Classification and Detection of Chest Abnormalities from Chest Radiography Images: CXR\_SPNet an Efficient Deep Learning Approach

Shuvo Biswas 1, Bikash Kumar Paul 1,∗, FNU Shariful 2,3 and Ramkrishno Majumder 1

1 Department of Information and Communication Technology (ICT), Mawlana Bhashani Science and Technology University (MBSTU), Santosh, Tangail-1902, Bangladesh; [shuvo.ict13@gmail.com](mailto:shuvo.ict13@gmail.com) (S.B.); [ramkrishno2k20@gmail.com](mailto:ramkrishno2k20@gmail.com) (R.M.)

2 Department of Software Engineering (SWE), Daffodil International University, Asulia, Bangladesh.

3 School of Computing, University of North Florida, Jacksonville, Florida, USA.

\* Correspondence: [bikash@mbstu.ac.bd](mailto:bikash@mbstu.ac.bd)

**Abstract:** The whole world is the witness of the deadliest COVID 19 pandemic which unceasingly runs two years and make it more conspicuous. Thousands of peoples are sentenced to death because of this. At the early stage it affects the lungs by proliferating of the virus resulting the damages of the air sacs. After that the patient feel the short of breath because of low level of oxygen in blood. The Chest X-ray imaging is a salient way to observe the condition of the lung. To reduce the exposure of virus automatic, reliable and accurate system is essential. In this study, a novel deep learning-based approach, CXR\_SPNet, is proposed for the automatic classification of normal chest versus abnormal (COVID-19) chest. The novelty of this proposed CXR\_SPNet network is built to widen the network while maintaining its depth in comparison to existing networks. In this investigation a dataset called CXR has been opted which contains 2070 abnormal (COVID-19) and 1130 normal chest X-ray images. Here, the CXR\_SPNet as well as four pre-trained CNN models (VGG19, MobileNet, AlexNet, and GoogLeNet) on the same dataset to identify COVID-19-infected chests have been utilized carefully. Observing the confusion matrix demonstrates that, the proposed CXR\_SPNet successfully gained utmost accuracy (99.06%), specificity (99.5%), and low mean square error (0.00938) among all the pretrained models. In addition, by implementing various interpretable AI algorithms to classify X-ray images as well as find the specific location of COVID-19. Hence, it can be asserted that this research study will facilitate precisely to localize disease in X-ray images, interpretable artificial intelligence (XAI) techniques can advance medical science. Nevertheless, this model will be a potential candidate for fast screening of COVID-19 patients if this model is installed practically.

**Keywords:** Chest X-ray image; deep learning; classification; CXR\_SPNet; convolutional neural networks.

**1. Introduction**

Chest X-ray images have proven to have the ability to monitor and diagnose several lung disorders, including pneumonia, atelectasis, hernias, TB, and infiltrates. It also proved that chest x-ray images may provide a clear idea of the location and effect of COVID-19 on lung tissue [[1](#_bookmark20)]. As a result, images taken from chest X-rays can also be used to detect COVID-19. Chest X-ray is a rapid, cost-effective [[2](#_bookmark21)], and comprehensive clinical technique [[3](#_bookmark22)–[5](#_bookmark23)] that can be easily utilized in the diagnosis of important diseases such as COVID-19 and pneumonia. Compared to magnetic resonance imaging (MRI) and computed tomography (CT), chest X-rays deliver a lower radiation dose to the patient [[5](#_bookmark23)]. However, accurate ray image diagnosis requires special knowledge and skills [[6](#_bookmark24)]. Diagnosing COVID-19 and pneumonia using chest X-rays is much more difficult than using advanced imaging methods such as MRI or CT [[2](#_bookmark21)].

Only specialist doctors can diagnose COVID-19 by researching chest X-rays. There are very few specialist doctors who can diagnose this deadly disease. In most countries of the world, there are not enough doctors per population. As of 2017 records, Greece has 607 doctors per 100,000 inhabitants, ranking first in the world so far [[7](#_bookmark25)]. COVID-19 is a pandemic disease, and those who take care of patients or those sick with this disease, especially doctors, caregivers, and nurses, are at high risk, and many have even died. Concomitant health services, inadequate health workers, and hospital beds are contributing to the spread of these epidemics.

Computer-aided diagnosis (CAD) is a digital medical procedure that allows doctors to diagnose pneumonia precisely and quickly from chest X-rays [[8](#_bookmark26)]. Medical care is being made easier by managing large datasets using artificial intelligence methods that exceed human capabilities [[9](#_bookmark27)]. When CAD methods are integrated into radiology based diagnostic systems, both the workload and time of doctors are significantly reduced while increasing

the quantitative analysis and reliability of medical procedures [[10](#_bookmark28)]. Deep learning-based medical imaging and CAD systems have made the research field more popular [[9](#_bookmark27)–[11](#_bookmark29)]. Deep learning methods have demonstrated significant results over the past few years to automate the analysis of multimodal medical images in radiological applications [[12](#_bookmark30)–[14](#_bookmark31)]. One of the most reliable deep learning architectures is deep convolutional neural networks (DCNNs), which have been successfully used in a large range of real-world tasks like pattern recognition and image classification [[15](#_bookmark32)].

The following four ways can be handled by DCNN [[16](#_bookmark33)]: 1) training very large datasets using neural network weights 2) tuning on small datasets using the network weights of a fine-trained DCNN; 3) employing unsupervised fine-training to apply network weights to an application before using the DCNN model; and 4) using a fine-trained DCNN that acts as a feature extractor from the image, also known as an "off-the-shelf" CNN. DCNN has been effectively used in previous studies to detect common chest diseases, including mediastinal lymph nodes [[17](#_bookmark34)], tuberculosis screening [[18](#_bookmark35)], and CT images.

Inspired by the above discussion, we proposed a stacked parallel unit-based deep CNN framework for chest X-ray classification named CXR\_SPNet, which helps radiologists effectively and automatically identify normal and COVID-19-infected chests from X-ray images. The following steps are a summary of this study’s contributions:

1. We created a dataset of 3200 images for the classification of chest X-ray images. This dataset can be used as a benchmark by the scientific community. An expert radiologist labels the images in the Covid-19 class and the dataset is used to examine only the covid and normal images;
2. Conducting an experimental study using conventional chest X-rays, to classify normal and COVID-19-affected chests with the proposed deep learning framework;
3. The parallel stacked convolution unit located in CXR\_SPNet can provide superior results with a limited amount of input, whereas traditional CNN requires a large amount of data;
4. The compilation time and performance accuracy of our CXR\_SPNet model are compared with those of four traditional deep learning networks (such as AlexNet [[19](#_bookmark36)], VGGNet [[20](#_bookmark37)], GoogleNet [[21](#_bookmark38)], and MobileNet [[22](#_bookmark39)]);
5. In **section** [3](#_bookmark1), we examine the stacked parallel unit-based deep learning technique as well as more fine-tuned techniques for classifying chest X-ray images.;
6. Finally, our best CXR\_SPNet model obtained an accuracy of 99% from other models while having a precision of 99%.

The remaining portion of this study is as follows: Section [2](#_bookmark0) is the related work, Section [3](#_bookmark1) is the methodology, Section [4](#_bookmark7) is the result in part, Section [5](#_bookmark16) is the discussion, and lastly, Section [6](#_bookmark18) offers conclusions.

**2. Related Work**

The deep learning technique has been successfully used in the medical sciences with remarkable performance and encouraging outcomes when compared to human interaction in a variety of difficult tasks, including lung segmentation [[23](#_bookmark40)], various types of pneumonia detection and classification in chest X-ray images [[24](#_bookmark41)], skin cancer classification [[25](#_bookmark42)], brain disease classification [[26](#_bookmark43)], nasopharyngeal carcinoma identification [[27](#_bookmark44),[28](#_bookmark45)], and breast cancer detection [[29](#_bookmark46)]. Deep learning-based medical imaging algorithms have also been improved to aid clinicians and radiologists in the early identification, treatment, follow-up examination, and classification of COVID-19 [[30](#_bookmark47)].

Explainable artificial intelligence techniques have consistently produced reliable and trustworthy results in applications that utilize medical image-based data. In the last few years, researchers and scientists have researched and analyzed CX-ray images to classification COVID-19 utilizing deep learning methods.

The authors of [[31](#_bookmark48)] examined seven popular CNN-based models using a corpus of 25 COVID-19 patients and 50 controls. The best outcomes were generated by the DenseNet and VGG19 CNN models, with F1 scores for patients and controls of 0.91 and 0.89, respectively. Utilizing a chest X-ray database of only 50 COVID-19 cases and 50 controls, three deep CNN models (Inception V3, Inception Res-Net V2, and ResNet 50) [[32](#_bookmark49)] were chosen to detect COVID-19. The best accuracy, 98%, was attained by ResNet50.

A novel COVID-19 diagnostic algorithm called BMO-CRNN was introduced in the paper [[33](#_bookmark50)] and consists of two algorithms: (i) cascaded recurrent neural network (CRNN) and (ii) barnacle mating optimization (BMO). This approach uses chest X-ray images and a collection of CRNN hyper parameters, including batch size, learning rate, epoch, and activation function, to evaluate the optimal CRNN parameter. After executing the novel BMO-CRNN algorithm, it exhibited higher results, with a sensitivity of 97.01%, an F-value of 97.53%, and an accuracy of 97.31%.

A special CNN algorithm with unique image training was proposed for this study [[34](#_bookmark51)]. The research experiment used a dataset of 200 normal (healthy) and 180 COVID-19 images. According to this study, this algorithm achieved a classification accuracy of 91.6% Eight different popular deep-learning approaches were used in [[35](#_bookmark52)] to identify people with COVID-19 symptoms from 400 X-ray pictures, including ResNet15V2, VGG16, ResNet50, VGG19, DenseNet201, NasNetMobile, MobilenetV2, and InceptionResNetV2. On this dataset, NasNetMobile outperforms all pre-trained algorithms by achieving 93.94% accuracy.

For the binary image classification of controls and COVID-19 as well as the multiclass image classification of controls, pneumonia, and COVID-19, the authors of [[36](#_bookmark53)] made use of a chest X-ray database of 500 controls, 500 pneumonia patients, and 127 COVID-19 patients collected from different online sources. The darknet algorithm obtained an outstanding accuracy of 98% for binary image classification and 87% for multiclass image classification using a five-fold cross-validation method and transfer-learning techniques.

The dataset used in [[37](#_bookmark54)] consisted of a total of 1428 chest radiographs divided into three categories: healthy individuals, confirmed COVID-19 positives, and normal bacterial pneumonia. In this study, relatively small chest radiographs are trained and tested using the fine-tuned VGG16 architecture for classification tasks. The model tested accuracy rates of 92.5% and 96% for three classes (non-covid pneumonia, normal, and covid) and two classes (non-covid and covid), respectively. Four fine-trained CNN networks (SqueezeNet, DenseNet-121, ResNet50, and ResNet18) were used in [[38](#_bookmark55)] for image prediction. Experiments were developed using a dataset of 5000 pneumonia and no-detection images and a database of 184 COVID-19 photos. As reported in that paper, the specificity value was 92.9% and the sensitivity value was about 98%.

A new CNN model called multi-filter depthwise convolution (MD-Conv) was proposed in [[39](#_bookmark56)], which increased the training speed of the model by converting the depthwise convolution module composed of multiple filters into a single convolution module. Tests of the model were performed using 7,470 chest X-ray images of people. The model provided an accuracy of 93.4% and a maximum AUC of 98.3%.

Another study [[40](#_bookmark57)] showed that the proposed architecture of COVID-CAPS exhibits better performance than previous CNN-based architectures. The COVID-CAPS architecture has fewer trainable parameters, yet it achieved 95.7% accuracy and 95.8% specificity.

Deep CNNs demonstrated a remarkable ability to successfully and accurately differentiate 21,152 abnormal and normal chest radiographs, as shown in [[41](#_bookmark58)], which is a successful performance in medical science. For the classification of pneumonia versus normal, the CNN model had an accuracy of 94.64%, a specificity of 92.86%, and a sensitivity of 96.5%, after being pre-trained and fine-tuned on adult and pediatric patient datasets.

A modified version of the MobileNetV2 CNN model with a higher accuracy rate was proposed by the authors in [[42](#_bookmark59)]. In this study, to achieve the best results for testing and training the model, the authors created a large dataset of 52,000 images by augmenting chest X-ray images of 13,808 healthy and COVID-19 patients. The accuracy and sensitivity of the modified MobileNetV2 were both 98%.

Chest X-ray imaging has been used in most published studies to detect COVID-19, leading to tremendous improvements in treatment. However, numerous studies have shown that to be tested and trained, CNN algorithms require a large dataset. In studies [[31](#_bookmark48)–[37](#_bookmark54)], researchers used a small dataset (75–1428 chest X-rays) for testing and training the COVID-19 classification and attained varying degrees of accuracy (91.6–98%). However, its accuracy results are questionable because the models were not properly trained due to the use of small datasets. Additionally, the models used in the research [[38](#_bookmark55)–[42](#_bookmark59)] were trained on a huge dataset (5184–52,000 chest X-ray data), so the compilation time was very high, which means low accuracy, with most of the accuracy falling between 93.4–98%. From the above discussion, we can say that a model must be specific, fast to compile, and accurate for rapidly identifying people with COVID. To solve this problem, a parallel unit-based CNN model called CXR\_SPNet is proposed, which outperforms all currently used CNN models by increasing the accuracy rate and reducing the processing time. Also, we used a dataset with neither too many nor too few images. We created a medium range dataset called the CXR dataset, which contains 3200 chest X-ray images.

**3. Materials and Methods**

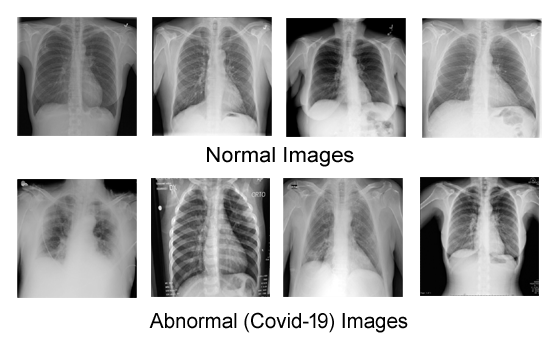
There are three components to this section: (i) dataset collection; (ii) fine-tuned CNNs models (iii) our newly proposed CNN-based model called CXR\_SPNet, which will improve chest X-ray image classification performance.

**3.1 Dataset collection**

This study used a database of X-rays [[43](#_bookmark60)] to train and test the CNN model. This database consists of four classes of datasets consisting of 6012 lung opacities, 1345 viral cases of pneumonia, 10,192 healthy individuals, and 3616 Covid-19 images, where each image size was 299x299. However, to facilitate research, we created a new dataset called CXR by randomly selecting images from only the healthy and COVID-19 X-ray sets among the four image sets, including 3200 X-ray images (Figure. [1](#_bookmark3) shows some examples of images from the CXR dataset). The CXR dataset is composed of 2070 abnormal (Covid-19) and 1130 normal X-ray images. Then, the images were resized for the classifier algorithms (for example, AlexNet was 227 × 227 pixels, while other models used 224 × 224 pixels). We divided the CXR dataset into three parts based on the train-test partition. The ratios of the three parts are 80:20, 70:30, and 90:10, respectively. Table. [1](#_bookmark2) lists the specific features of the CXR dataset.

**Table 1.** Dataset description.

|  |  |
| --- | --- |
| **Features** | **Parameters** |
| Total Images | 3200 |
| Disease Types | 2 |
| Disease Name | Abnormal (Covid), Normal |
| Image Size in Pixel | AlexNet is 227 × 227 pixels and others are 224 × 224 pixels |
| Abnormal (Covid) Images | 2070 |
| Normal Images | 1130 |
| Train-Test Splitting Ratio | 70:30, 80:20, 90:10 |
| Training Images | 2880 (for 90:10) |
| Testing Images | 320 (for 90:10) |



**Figure 1.** Few samples of CXR dataset images.

**3.2 Pre-trained CNN architectures**

In this section, our proposed CNN and previously trained networks are trained and tested on large image datasets to determine how efficient and effective they are at classifying modern medical image datasets [[44](#_bookmark61)–[46](#_bookmark62)]. In this regard, we have taken the help of four well-known pre-trained networks such as AlexNet [[19](#_bookmark36)], VGG19 [[20](#_bookmark37)], GoogleNet [[21](#_bookmark38)], and MobileNet [[22](#_bookmark39)], which play an important role in transfer learning techniques. Through these models, we have been able to easily distinguish normal and COVID-infected chests from chest X-ray images, which has created a great stir in medical science.

**3.2.1 AlexNet**

In 2012, AlexNet won the ILSVRC, developed by Krzyzewski et al. [[19](#_bookmark36)], demonstrating better results than previous models. Although it has more filters with deeper structures and more layered convolutional layers than LeNet, it provides better accuracy than LeNet. It is trained to classify 1,000 different classes of images using about 650,000 neurons. It adds ReLU to each convolutional and fully connected layer and has different kernel sizes (3x3, 5x5, and 11x11), data augmentation, dropout, and maximum pooling. This proves that nonlinear ReLU can help CNN train faster than Tanh or Sigmoid.

**3.2.2. VGG19**

The VGGNet was created by Andrew Zisserman and Karen Simonyan of the Oxford Robotics Institute based on CNN architecture [[20](#_bookmark37)]. On the big image dataset, the VGGNet did exceptionally well. Compared to the most popular AlexNet’s 11x11 filter, the VGGNet uses smaller filters (3x3) to have better image extraction performance. Both VGG16 and VGG19 architectures have different levels and depths, but VGG19 is deeper than VGG16. However, VGG19 has more parameters than VGG16, which makes network training more expensive.

**3.2.3 GoogLeNet**

Sejedi, et al. [[21](#_bookmark38)] in 2015, GoogleNet, a deep CNN model, was first trained and implemented on large datasets such as ImageNet. Compared to the Inception model, GoogleNet has about 12x fewer parameters and 22 layers, which makes the model more efficient and dynamic than other models. However, by combining more additional layers, the network may overfit and the number of parameters may increase. It takes 224 × 224 resolution images as input and processes them through three main units: (i) the pooling unit; (ii) the convolution unit; and (iii) the inception module. The architecture also includes activation, dense, and average pooling 2D layers. Currently, it is also known as the InceptionV1 deep CNN model.

**3.2.4. MobileNet**

The model was published by Sandler et al. [[22](#_bookmark39)] as a CNN model for devices with limited processing power, such as smartphones. The addition of fewer numerical learning parameters and inverted linear bottleneck modules significantly reduces the memory usage of the MobileNet model. The pre-trained MobileNet implementation is also readily accessible in several well-known deep-learning networks. Both the accuracy and training speed of the architecture are maximized by using 3.4 million parameters and removing the six-number activation function ReLU from each bottleneck unit.

**3.3. Proposed approach CXR\_SPNet model**

**Figure.** [3](#_bookmark6) illustrates the novel CXR\_SPNet architecture with details of each operation, which enhances the extraction of key features for classification. Our model combines two crucial components of CNN’s architecture: (i) the depth of the model, and (ii) the parallel unit. To learn additional feature mappings and enhance the model’s efficiency, we reduced the network’s depth and increased its width. A parallel unit is formed by combining multiple parallel convolutions and a single max pooling layer. The CXR\_SPNet is built to detect and classify which chest X-rays have COVID-19 and which do not. The CXR\_SPNet architecture has three main stages. These stages are (i) the initial stage; (ii) the intermediate stage; and, lastly, (iii) the final stage. The detailed operations of each stage of the CXR\_SPNet model are described in **Table** 2.

The parameters utilized for CXR\_SPNet training are 50 epochs, an 85% training dataset, a batch size of 32, a 5% validation dataset, and an Adam optimizer (learning rate of 0.001). To decrease the learning rate, factors = 0.4, patience = 2, and min\_delta = 0.001 are set. **Figure** [2](#_bookmark5) depicts the general pipeline of CXR\_SPNet.

**Table 2.** Proposed Model Architecture with various parallel and reduction unit. Here, S is Stride, BN is Batch Normalization Layer, FC is Fully Connected Layer, and SM is Softmax.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **SI** | **Operator** | **Kernel Size** | **S** | **No. of kernels** | **FC** |
| 1 | Conv2d, BN, Relu | 3x3 | 2x2 | 32 | - |
| 2 | Conv2d, BN, Relu | 3x3 | 1x1 | 128 | - |
| 3 | Conv2d, BN, Relu | 3x3 | 1x1 | 192 | - |
| 4 | Max Pooling2d | 3x3 | 2x2 | - | - |
| 5 | Conv2d, BN, Relu | 1x1 | 1x1 | 256 | - |
| 6 | Conv2d, BN, Relu | 3x3 | 1x1 | 512 | - |
| 7 | Max Pooling2d | 3x3 | 2x2 | - | - |
| 8 | Parallel Unit A |  | Four Units |  | - |
| 9 | Reduction Unit A |  | One Unit |  | - |
| 10 | Parallel Unit B |  | Five Units |  | - |
| 11 | Reduction Unit B |  | One Unit |  | - |
| 12 | Parallel Unit C |  | Three Units |  | - |
| 13 | Average Pooling2d | - | - | - | 2048 |
| 14 | FC, SM | - | - | - | 2 |

**Input image:** The CXR input images contain training and validation images, which are labeled into two classes: normal chest and abnormal chest. There are three channels in it. The image size of each channel is 224x224.

**Initial stage:** This layer is also called the "traditional convolutional layer," which contains different types of convolutional and max-pooling layers. This layer consists of the first seven layers, as shown in **Table.** [2](#_bookmark4). This stage consists of five convolutional layers with two different kernel sizes (1x1 and 3x3) and two max-pooling layers with kernel size (1x1). At the end of each convolution layer in this step, a rectified linear unit and a batch normalization layer are added. This stage is crucial to ensuring that any huge crude input images are reduced in dimension before moving on to the next stage.

**Convolutional layer:** This layer is coupled with a filter set [[47](#_bookmark63)] that learns the results of the previous layer since the weights are capable of identifying each convolutional filter. The CXR\_SPNet network’s first, second, third, fourth, and fifth convolution layers, which speed up the image filtration method, are made up of 32, 128, 192, 256, and 512 kernels, respectively.

**Batch normalization (BN) layer:** This layer applies a mini-batch to normalize each input channel. This layer plays an effective role in reducing the sensitivity of the CNN algorithm and increasing the training speed [[48](#_bookmark64)]. This stage has five batch normalization layers, which are placed after each convolution layer but before the ReLU layers.

**Rectified Linear Unit (ReLU):** ReLU is the non-linear activation function, which filters data at this level using the max (0, x) function [[49](#_bookmark65)], where x denotes the input to the neuron.

**Max pooling layer:** This layer increases the efficiency of the CNN model by selecting the largest elements from the activation map area covered by the filter. At this stage, two max-pooling layers of 3x3 kernel and 2x2 stride are used, where the first reduces the size of the input image from 109x109 to 54x54 and the second from 52x52 to 25x25.

**Intermediate stage:** This layer consists of two types of units, such as parallel units and reduction units. A "parallel unit" is a special type of unit that combines multiple convolutional filters. This special type of unit is used to extract multi-level features from a single input and combine sparse clusters. The parallel and reduction units can also be called the core of the CXR\_SPNet model. The parallel convolution unit can extract features from different convolution layers with different size filters, which makes CXR\_SPNet more efficient. Each of the three parallel units A, B, and C has one max pooling layer and seven, ten, and nine convolutional layers, respectively. On the other hand, reduction units A and B have one max pooling layer and, respectively, four and six convolutional layers. We are trying to find the best architecture for CXR\_SPNet by testing it using a train-test splitting ratio. In **Table.** [3](#_bookmark9), we construct three different variations of CXR\_SPNet according to the train-test splitting ratio. According to **Table.** [3](#_bookmark9), the CXR\_SPNet with a train-test ratio of 90:10 exhibits better performance.

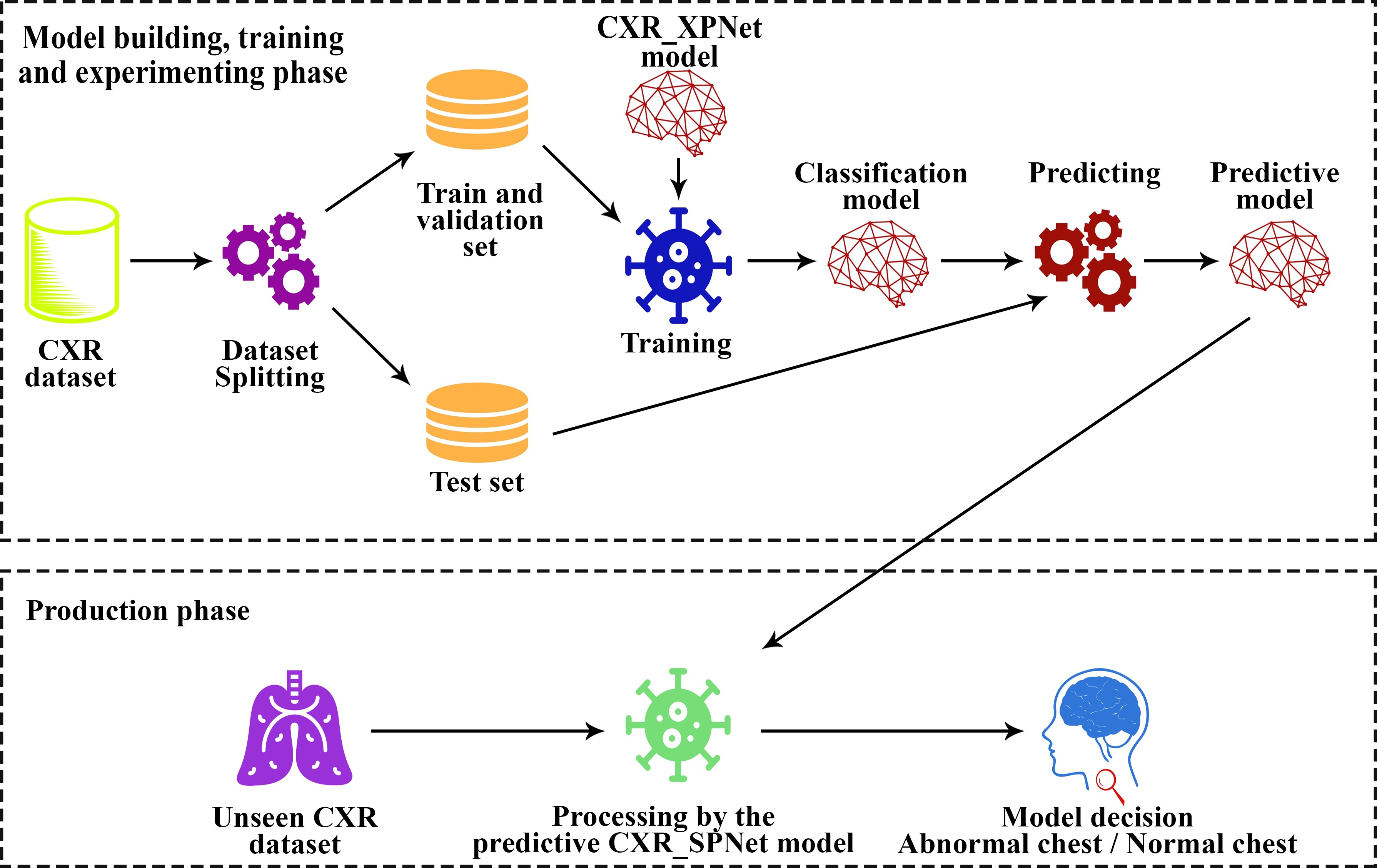
**Final stage:** This layer is also called the fully connected layer (FC). This layer is composed of a fully connected layer, a global average pooling layer, and a drop-out layer. In **Table.** [2](#_bookmark4) , the last three layers make up the final layer.

**Global average pooling layer:** Conventional CNNs use a globally averaged pooling layer instead of a fully connected layer. This is a special pooling technique because it combines the average of each activation map from an image and sends it directly to the softmax layer. In this study, the output of this layer is 2048.

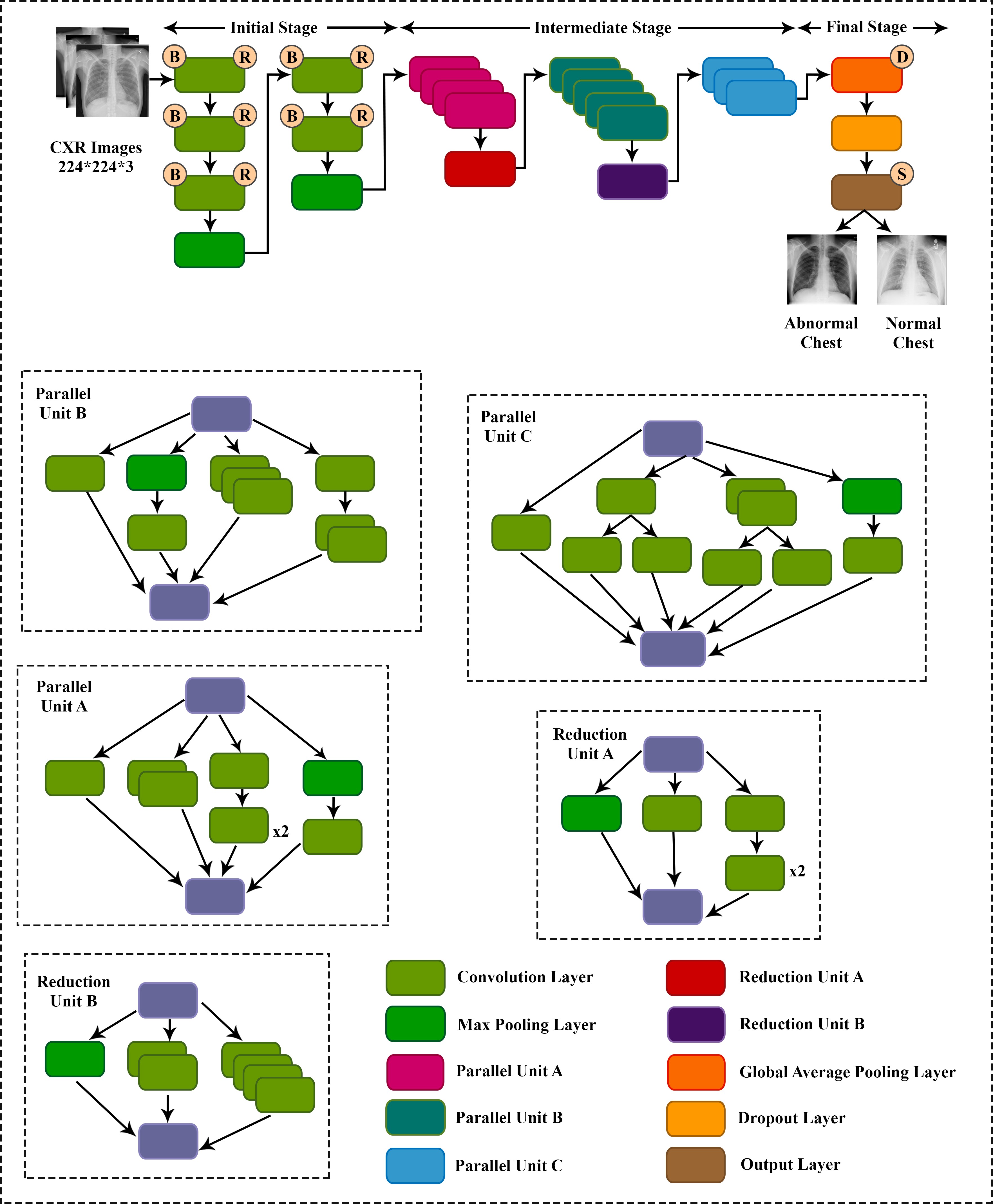
**Fully connected layer (FC):** All neurons in the preceding layer are concatenated by this layer [[50](#_bookmark66)] whose output is 2 and has one softmax activation function. This layer’s main function is to classify images into two categories: abnormal (COVID-19) and normal, by combining image features.

**Dropout Layer:** After the first FC layer, a dropout layer was used to increase model performance and avoid overfitting. The probability of dropping out in our model was 0.5.

**Softmax:** It is an output classifier that predicts what the model’s class labels will be. It determines how close the evidence data labels are to the training data.



**Figure 2.** The CXR\_SPNet model workflow for distinguishing normal and abnormal chest.



**Figure 3.** CXR\_SPNet Architecture. Here, B is Batch normalization, R is ReLU, D is Dense layer and S is Softmax.

More discriminative features are present in each convolution layer. Different levels of activation maps for both normal and abnormal images of CXR\_SPNet are shown in **Figure.** [4](#_bookmark8) and **Figure.** [5](#_bookmark10).

**4. Results**

**4.1. Evaluation metrics**

On our highly qualified dataset, we ran many experiments to gauge how well our network and specially pre-trained networks performed at classifying data. The different evaluation metrics are defined as follows:

**Accuracy (ACC):** The most crucial metric for deep learning algorithm outputs is accuracy. It is just the total values of the components of the confusion matrix divided by the sum of confusion matrix’s the true negative (T*−*) and true positive (T+). The ACC is calculated by the equation

**Precision (PRE):** It is determined by dividing the total true positives (T+) by the sum of the confusion matrix’s true positive (T+) and false positive (F+) and is defined as

**Recall (REC)/Sensitivity (SEN):** The ratio of true positive (T+) prediction values to the sum of the confusion matrix’s false negative (F*−*) and true positive (T+) prediction values is known as recall or sensitivity and is measured as

**F1-Score (FS):** It is a general indicator of the CNN algorithm’s accuracy which com- bines precision and recall and is defined as

**Specificity (SPE):** It is the ratio of true negative (T*−*) prediction values to the sum of the confusion matrix’s true negative (T*−*) and false positive (F+) prediction values. The specificity is calculated by the equation



**Figure 4.** Output of feature map for abnormal chest image (**a**) First convolution2d layer (**b**) First max\_pooling2d layer (**c**) Convolution2d\_25 layer.

## **Performance measures of CXR\_SPNet based on dataset splitting method**

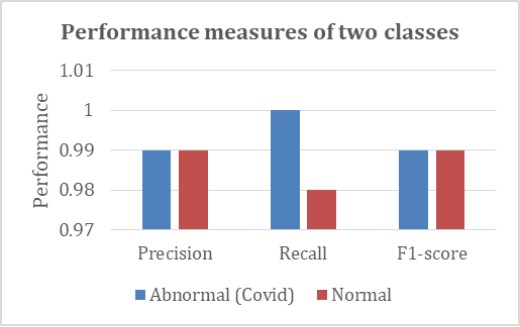
# The performance metrics analysis of CXR\_SPNet using the dataset splitting technique is described in **Table.** [3](#_bookmark9).

**Table 3.** performance analysis based on the dataset splitting technique.

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **Train Test ratio** | **Accuracy** | **Precision** | **Recall / Sensitivity** | **Specificity** | **AUC** | **F1-score** |
| 70:30 | 0.97917 | 0.97143 | 0.97143 | 0.98361 | 0.97752 | 0.97143 |
| 80:20 | 0.97813 | 0.96396 | 0.97273 | 0.98095 | 0.97684 | 0.96833 |
| **90:10** | **0.99062** | **0.99083** | **0.98182** | **0.99524** | **0.98853** | **0.98630** |

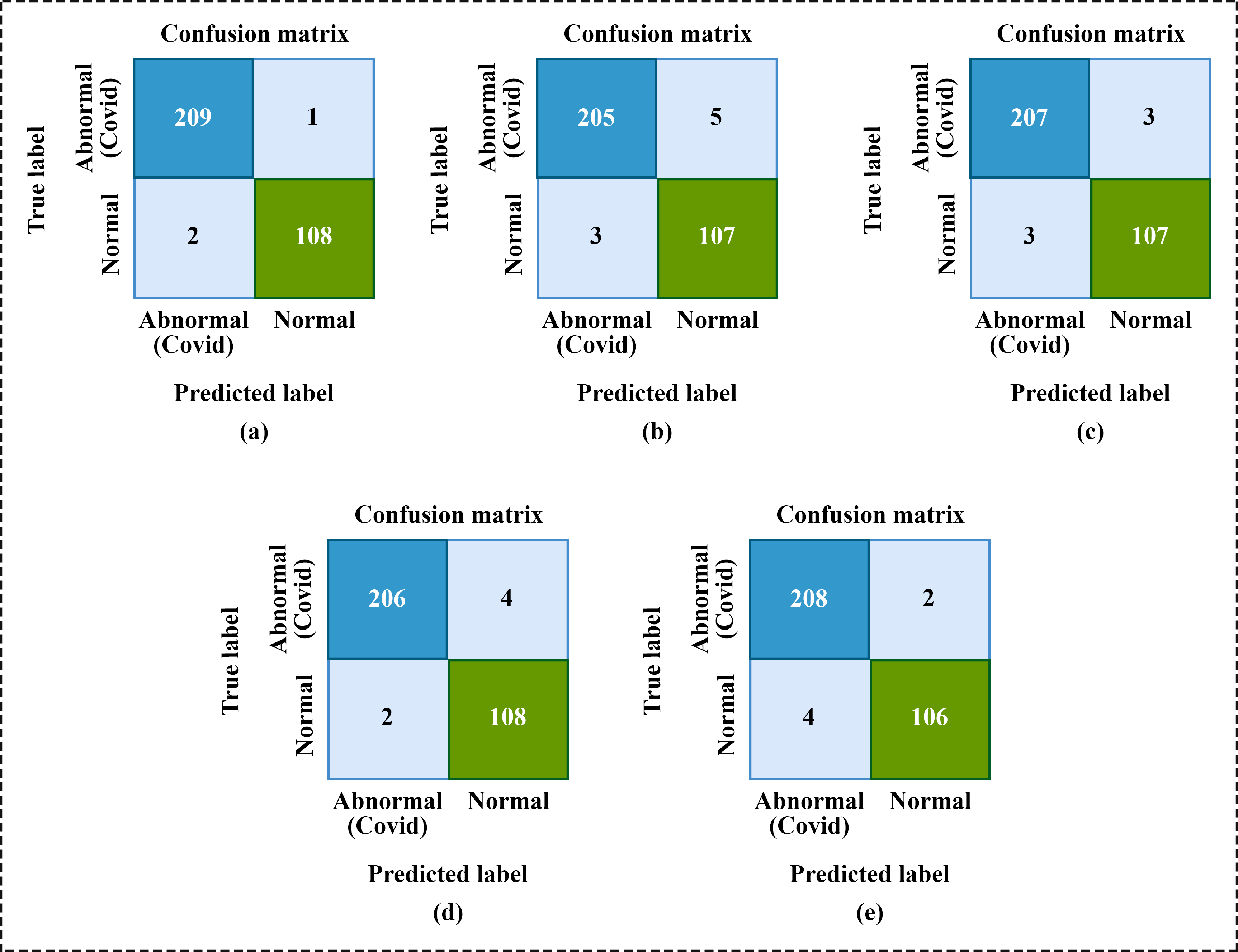


**Figure 5.** Output of activation map for normal image (**a**) First convolution2d layer (**b**) First max\_pooling2d layer.



**Figure 6.** Performance comparison of two classes (abnormal COVID and normal) for the proposed scheme. The horizontal axis represents the number of classes, and the vertical axis represents the performance of these classes. The recall score of abnormal COVID images is higher than that of normal x-ray images, but the precision and F1-score are the same for both classes.

To compare our proposed CXR\_SPNet with four ultra-modern models, we trained, evaluated, and compared our DFU dataset. The reports give in **Table.** [4](#_bookmark13) AlexNet [[19](#_bookmark36)] achieved the accuracy, precision, specificity, F1-score, AUC, and recall scores of 0.98125, 0.97273, 0.98571, 0.97273, 0.97922, and 0.97273, and the mean squared error rate is 0.01875. GoogLeNet [[21](#_bookmark38)] achieved the minimum accuracy, precision, specificity, F1-score, AUC, and recall scores of 0. 97500, 0. 95536, 0.97619, 0. 96396, 0. 97446, and 0. 97273, and the very high mean squared error rate is 0. 02500. VGG19 [[20](#_bookmark37)] has higher precision, and specificity than AlexNet and GoogLeNet but is lower than our proposed CXR\_SPNet. The precision, accuracy, specificity, recall, AUC, and F1-score of MobileNet [[22](#_bookmark39)] are respectively 0.96429, 0.98125, 0.98095, 0.98182, 0.98139, and 0.97297, and the mean squared error rate is 0.01875. CXR\_SPNet achieved the highest accuracy, precision, specificity, F1-score, AUC, and recall scores of 0.99062, 0.99083, 0.99524, 0.98630, 0.98853, and 0.98182, and the very low mean squared error rate is 0.00938. In **Table.** [5](#_bookmark15), a comparison of our proposed CXR\_SPNet model with other existing models is shown, where our model shows remarkable results. Performance measures for two classes in the proposed scheme are shown in **Figure.** [6](#_bookmark11).

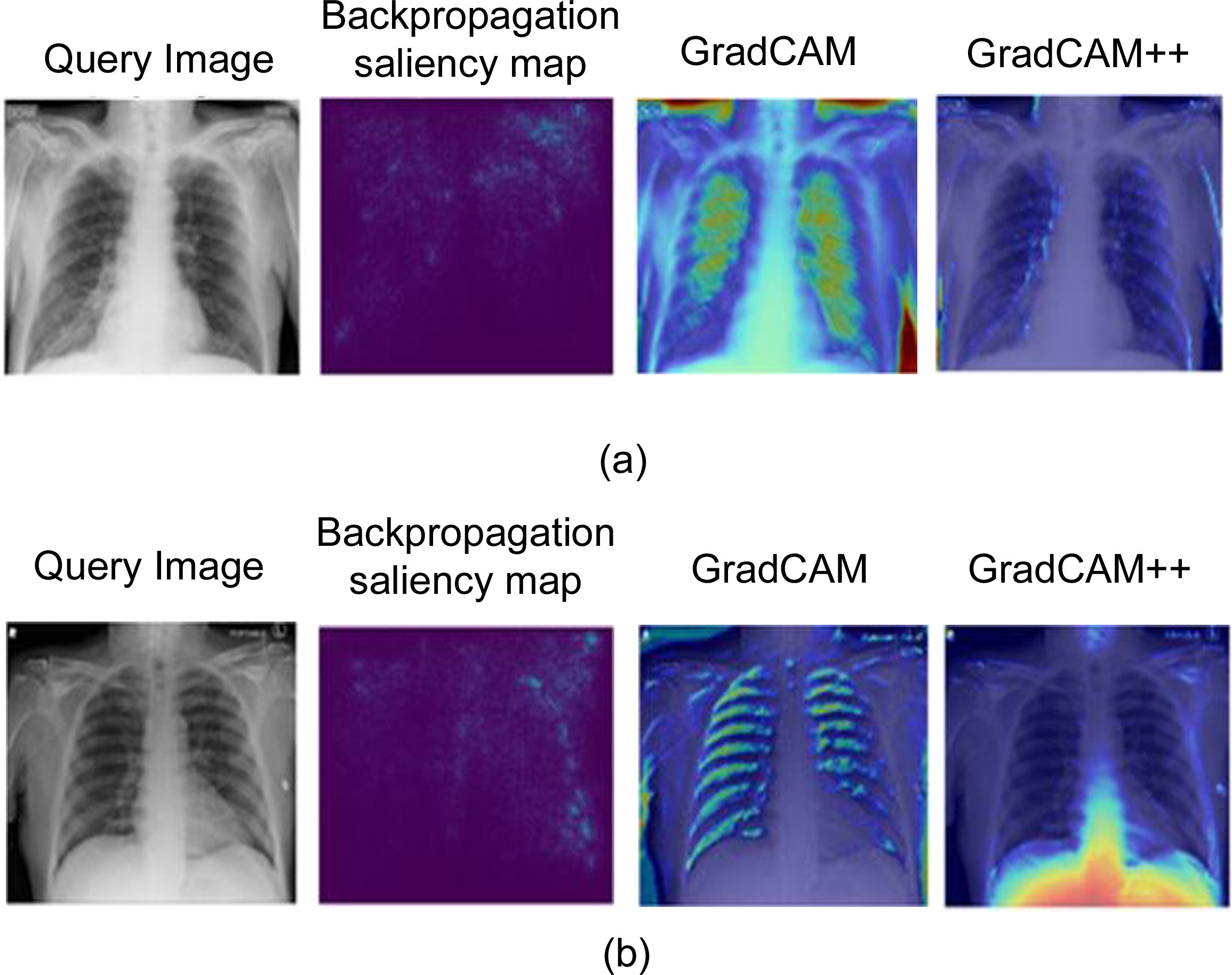


**Figure 7.** Confusion matrix of (**a**) CXR\_SPNet, (**b**) GoogLeNet, (**c**) AlexNet, (**d**) MobileNet, and (**e**) VGG19.

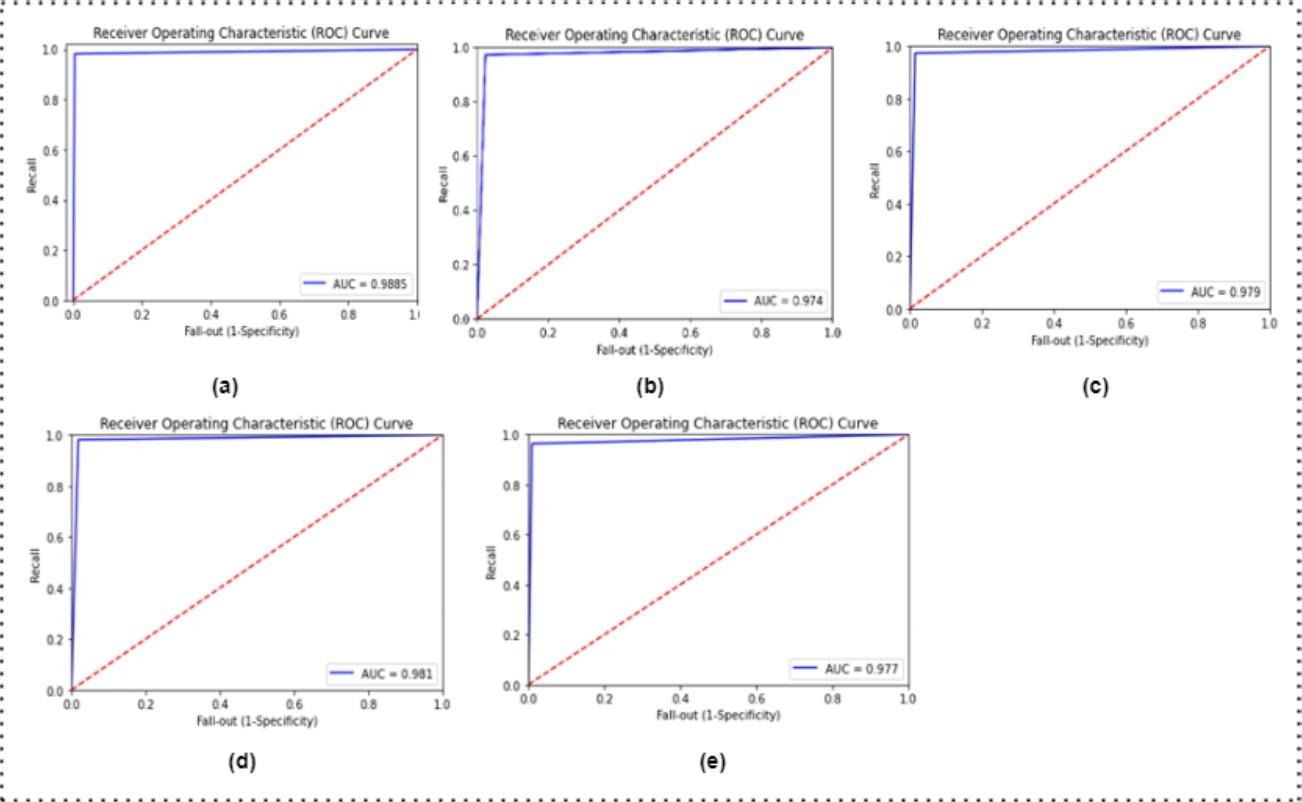
**4.3. Results comparison between CXR\_SPNet and the fine-tuned CNNs architecture**

**Table 4.** Comparison of CXR\_SPNet with ultra-modern Networks. Here, MSE is Mean Squared Error.

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| **Model** | **ACC** | **PRE** | **REC / SEN** | **FS** | **AUC** | **MSE** | **SPE** |
| AlexNet [19] | 0.98125 | 0.97273 | 0.97273 | 0.97273 | 0.97922 | 0.01875 | 0.98571 |
| VGG19 [20] | 0.98125 | 0.98148 | 0. 97364 | 0. 97248 | 0.97706 | 0. 01875 | 0.99048 |
| GoogLeNet [21] | 0.97500 | 0.95536 | 0.97273 | 0.96396 | 0.97446 | 0.02500 | 0.97619 |
| MobileNet [22] | 0. 98125 | 0.96429 | **0.98182** | 0.97297 | 0.98139 | 0.01875 | 0.98095 |
| **CXR\_SPNet** | **0.99062** | **0.99083** | **0.98182** | **0.98630** | **0.98853** | **0.00938** | **0.99524** |

Figure. [7](#_bookmark12) shows the confusion matrix of all CNN architectures in this study. The CXR\_SPNet model’s true positive (TP) = 108, false positive (FP) = 1, true negative (TN) = 209, and false negative (FN) = 2. Figure. [9](#_bookmark14) shows the ROC curves of all CNN architectures in this study. The CXR\_SPNet model produced a higher AUC score of 0.9885.

**Figure 8.** The first convolution layer output for (**a**) abnormal (Covid) and (**b**) normal images using an explainable AI algorithms.



**Figure 9.** ROC curves of (**a**) CXR\_SPNet, (**b**) GoogLeNet, (**c**) AlexNet, (**d**) MobileNet, and (**e**) VGG19.

**4.4. The result comparison study to classify normal vs. abnormal (Covid)**

**Table 5.** We contrast the outcomes of proposed CXR\_SPNet with those of other existing models.Here, SPE is Specificity, PRE is Precision, REC is Recall, ACC is Accuracy, and FS is the F1-Score.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Ref. | Data type | Total Samples | Classifier | Classes | Accuracy/ Performance (%) |
| [31] | X-ray | 75 | DenseNet , VGG19 | 2 | FS = 0.91 |
| [32] | X-ray | 100 | ResNet50 | 2 | ACC = 96.1 |
| [33] | X-ray | 247 | BMO-CRNN | 2 | SEN = 97.01, ACC = 97.31, FS = 97.53 |
| [34] | X-ray | 380 | Novel CNN Model | **2** | ACC = 91.6 |
| [35] | X-ray | 400 | NasNetMobile | 2 | ACC = 93.94 |
| [35] | X-ray | 1127 | Modified Darknet | 2 | ACC = 98 |
| [37] | X-ray | 1428 | VGG16 | 2 | ACC = 96 |
| [38] | X-ray | 5184 | SqueezeNet | 2 | SEN = 98, SPE = 92.9 |
| [39] | X-ray | 7470 | MD-Conv | 2 | ACC = 93.4 |
| [40] | X-ray | 13,975 | COVID-CAPS | 2 | ACC = 95.7 |
| [41] | X-ray | 21,152 | CNN | 2 | ACC = 94.64 |
| [42] | X-ray | 52,000 | Modified MobileNetV2 | 2 | ACC = 98, SEN = 98, PRE = 97 |
| **This Study** | **X-ray** | **3200** | **CXR\_SPNet** | **2** | **ACC = 99.06, PRE = 99.08, SPE = 99.5** |

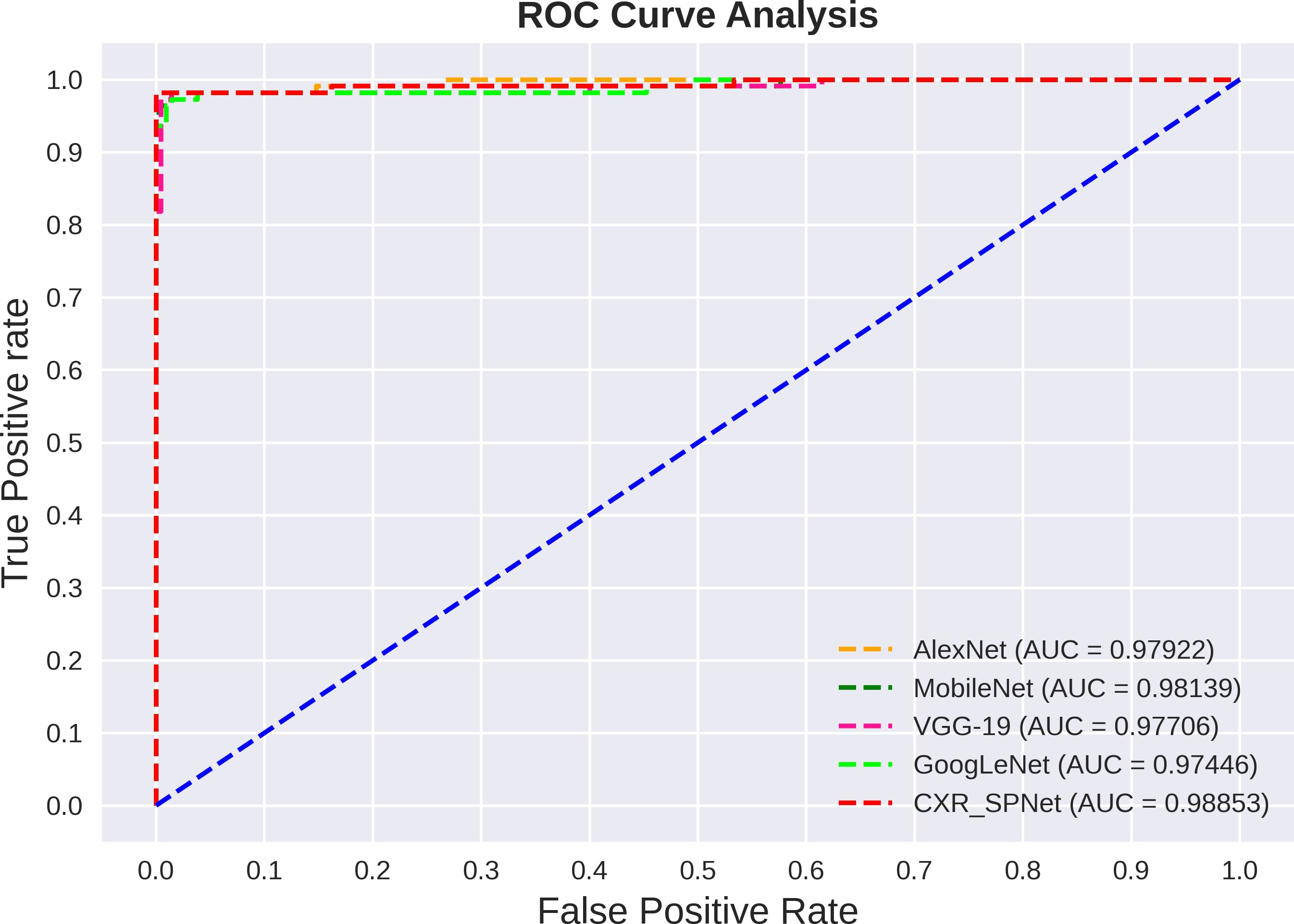
**5. Discussion**

The automatic classification of normal and abnormal chest X-ray samples using a novel approach is discussed in this paper. As we use a more generalizable Adam optimizer and fewer epochs for the network to converge, that reduces the training process. We test our model with some unknown images to check whether it can correctly classify chest X-ray images. A few samples of predictions obtained from CXR\_SPNet for both classes are provided in **Figure**. [11](#_bookmark19). Due to CXR\_SPNet’s great accuracy, precision, and specificity, it performed exceptionally well in correctly predicting both classes.

Additionally, the proposed CXR\_SPNet is compared with a few fine-tuned CNNmodels (VGG19, AlexNet, GoogLeNet, and MobileNet) on the same dataset in **Figure**. [10](#_bookmark17)’s ROC curve. It demonstrates that in the abnormal (Covid) vs. normal classification challenge, the CXR\_SPNet outperformed the conventional CNN designs. The proposed strategy is superior for many reasons, some of which are listed below:

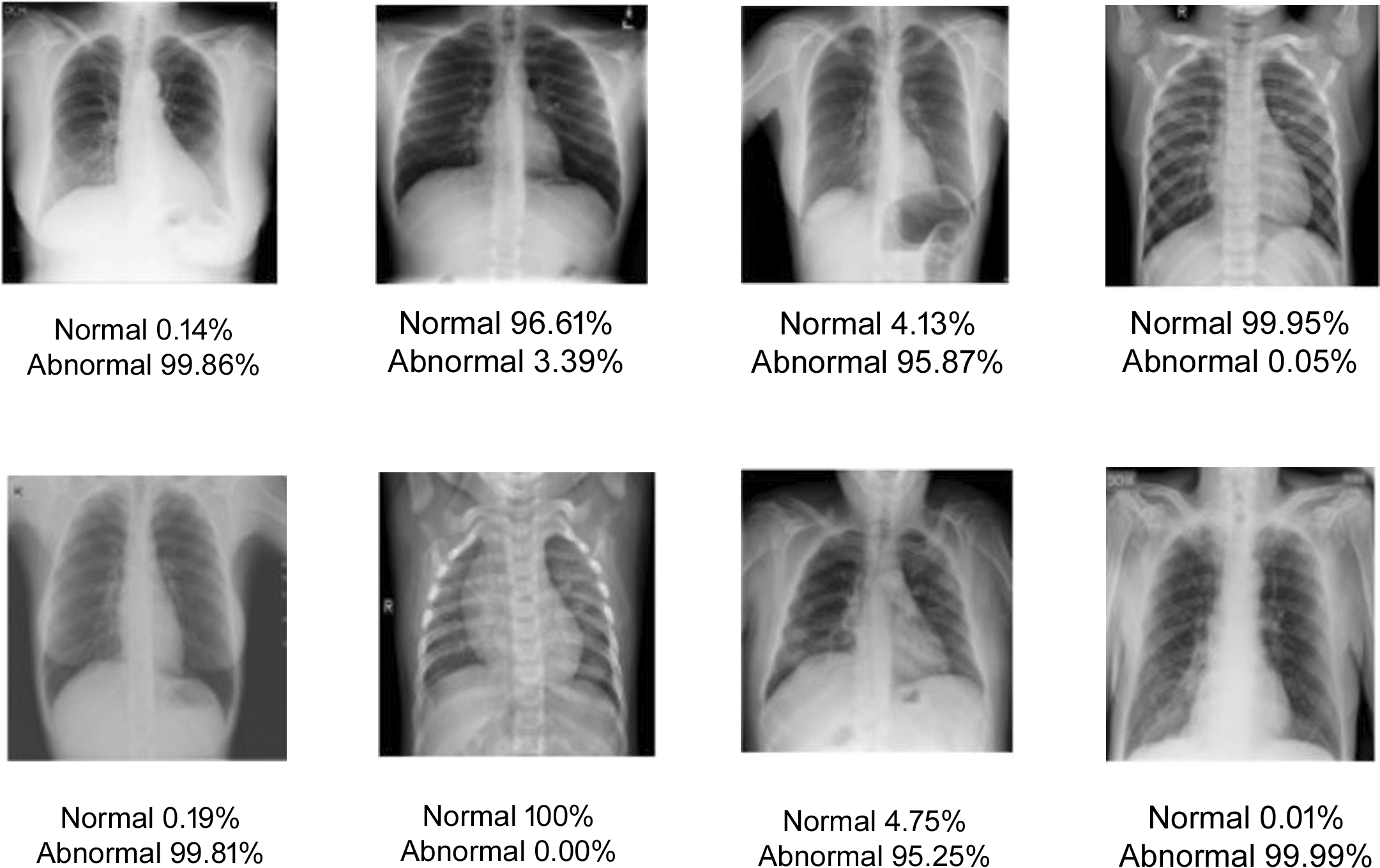
1. By using appropriate filters and depth in different convolution layers at the same level, this model captures all the information from the chest X-ray image and gives the model a fairly good receptive field.;
2. The CXR\_SPNet model uses 1x1 convolution inside each parallel unit that reduces the number of computational resources and training parameters;
3. In addition, the more generalizable Adam optimizer and the use of fewer epochs strengthen the predictive efficiency of CXR\_SPNet.

According to the findings in **Figure**. [7](#_bookmark12), the CXR\_SPNet can identify chest X-ray features at various phases of the COVID disease. The suggested method can therefore be used to automatically and objectively diagnose COVID. On the other hand, the maximum number of activation maps in the parallel unit can detect only the COVID-affected areas using new interpretable AI algorithms.



**Figure 10.** ROC curves for proposed CXR\_SPNet and others pre-trained model. The orange, green, deep pink, lime, and red colors show the “AlexNet,” “MobileNet,” "VGG19," “GoogLeNet” and “CXR\_SPNet,” respectively.

**6. Conclusions**

The introduced CXR\_SPNet is a lightweight, robust CNN model that can play a pivotal role in COVID-19 detection in medical science or clinical diagnostic by classifying CXR datasets into two classes (i.e. normal and abnormal (COVID)). Therefore, we have precisely demonstrated how the custom CXR\_SPNet architecture successfully attains ameliorated accuracy, precision, mean squared error, and specificity by reducing the processing time, number of neurons, and fully connected layers. These outstanding findings entail that the CXR\_SPNet can serve X-ray specialists in striving assessment more apace. To measure performance, the comparison among CXR\_SPNet with four fine-tuned architectures such as MobileNet, GoogleNet, VGG19, and AlexNet rigorously run where CXR\_SPNet outperformed all others. Throughout the simulation the CXR\_SPNet model performed at its best, with the supreme accuracy of 99.06 percent. It can be undoubtedly said that the proposed model will best suited for real time automatic detection for lung abnormalities or related lung diseases.

**Figure 11.** CXR\_SPNet based some prediction samples.

**References**

[1] Radiology Assistant. (2020). X-ray Chest images [dataset]. Retrieved March 23, 2020 from <https://radiologyassistant.nl/chest/lk-jg-1>.

[2] Antin B, Kravitz J, Martayan E (2017) Detecting pneumonia in chest X-rays with supervised learning. http://cs229.stanford.edu/ proj2017/fnal-reports/5231221.pdf

[3] Narayan Das N, Kumar N, Kaur M, Kumar V, Singh D (2020) Automated deep transfer learning-based approach for detection of COVID-19 infection in chest X-rays. IRBM. https://doi.org/ 10.1016/j.irbm.2020.07.001

[4] Ayan E, Ünver HM (2019) Diagnosis of pneumonia from chest X-ray images using deep learning. In: Scientifc meeting on electrical-electronics and biomedical engineering and computer science (EBBT), Istanbul, Turkey, pp 1–5. https://doi.org/10.1109/ EBBT.2019.8741582

[5] ref-3 [6] Jaiswal AK, Tiwari P, Kumar S, Gupta D, Khanna A, Rodrigues JJ (2019) Identifying pneumonia in chest X-rays: a deep learning approach. Measurement 145:511–518

[7] Euronews is a European television news network, headquartered in Lyon, Franc

[8] Antin B, Kravitz J, Martayan E (2017) Detecting pneumonia in chest X-rays with supervised learning. http://cs229.stanford.edu/ proj2017/fnal-reports/5231221.pdf

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

[10] Gaál G, Maga B, Lukács A (2020) Attention U-Net based adversarial architectures for chest X-ray lung segmentation. arXiv:2003.10304

[11] Jaiswal A, Gianchandani N, Singh D, Kumar V, Kaur M (2020) Classifcation of the COVID-19 infected patients using DenseNet201 based deep transfer learning. J Biomol Struct Dyn. https://doi.org/10.1080/07391102.2020.1788642

[12] L. Chen, P. Bentley, K. Mori, K. Misawa, M. Fujiwara, and D. Rueckert, "Self-supervised learning for medical image analysis using image context restoration," Medical Image Analysis, vol. 58, p. 101539, 2019/12/01/ 2019.

[13] F. Gao, H. Yoon, T. Wu, and X. Chu, "A feature transfer enabled multi-task deep learning model on medical imaging," Expert Systems with Applications, vol. 143, p. 112957, 2020/04/01/ 2020.

[14] M. Kim, C. Yan, D. Yang, Q. Wang, J. Ma, and G. Wu, "Chapter Eight - Deep learning in biomedical image analysis," in Biomedical Information Technology (Second Edition), D. D. Feng, Ed.: Academic Press, 2020, pp. 239-263.

[15] D.-X. Zhou, "Theory of deep convolutional neural networks: Downsampling," Neural Networks, vol. 124, pp. 319-327, 2020/04/01/ 2020.

[16] T. K. K. Ho, J. Gwak, O. Prakash, J.-I. Song, and C. M. Park, "Utilizing Pretrained Deep Learning Models for Automated Pulmonary Tuberculosis Detection Using Chest Radiography," in Intelligent Information and Database Systems, Cham, 2019, pp. 395-403: Springer International Publishing.

[17] Y. Miki et al., "Classification of teeth in cone-beam CT using deep convolutional neural network," Computers in Biology and Medicine, vol. 80, pp. 24-29, 2017/01/01/ 2017.

[18] F. Pasa, V. Golkov, F. Pfeiffer, D. Cremers, and D. Pfeiffer, "Efficient Deep Network Architectures for Fast Chest X-Ray Tuberculosis Screening and Visualization," Scientific Reports, vol. 9, no. 1, p. 6268, 2019/04/18 2019.

[19] Krizhevsky A, Sutskever I, Hinton GE (2012) Imagenet classification with deep convolutional neural networks. In Advances in neural information processing systems (pp. 1097–1105)

[20] Simonyan K, Zisserman A (2014) Very deep convolutional networks for large-scale image recognition.arXiv preprint arXiv: 1409.1556

[21] C. Szegedy et al., “Going deeper with convolutions,” in Proc. IEEE Conf. Comput. Vis. Pattern Recognit., 2015, pp. 1–9.

[22] M. Sandler, A. Howard, M. Zhu, A. Zhmoginov, and L. C. Chen, "MobileNetV2: Inverted Residuals and Linear Bottlenecks," Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition, pp. 4510-4520, 2018.

[23] C. L. Van, V. Puri, N. T. Thao and D. Le, “Detecting lumbar implant and diagnosing scoliosis from Vietnamese X-ray imaging using the pre-trained API models and transfer learning, Computers,Materials & Continua, vol. 66, no. 1, pp. 17–33, 2021.

[24] P. Rajpurkar, J. Irvin, K. Zhu, B. Yang, H. Mehta et al., “Chexnet: Radiologist-level pneumonia detection on chest X-rays with deep learning,” arXiv Prepr. arXiv: 1711.05225, pp. 3–9, 2017 (Pre-print).

[25] K. M. Hosny, M. A. Kassem and M. M. Foaud, “Skin cancer classification using deep learning and transfer learning,” in Proc. 9th Cairo International Biomedical Engineering Conference, Cairo, Egypt, IEEE, pp. 90–93, 2019.

[26] M. Talo, O. Yildirim, U. B. Baloglu, G. Aydin and U. R. Acharya, “Convolutional neural networks for multi-class brain disease detection using mri images,” Computerized Medical Imaging and Graphics, vol. 78, pp. 101673, 2019.

[27] M. K. A. Ghani, M. A. Mohammed, N. Arunkumar, S. A. Mostafa, D. A. Ibrahim et al., “Decisionlevel fusion scheme for nasopharyngeal carcinoma identification using machine learning techniques,” Neural Computing and Applications, vol. 32, no. 3, pp. 625–638, 2020.

[28] M. A. Mohammed, M. K. Abd Ghani, N. Arunkumar, S. A. Mostafa, M. K. Abdullah et al., “Trainable model for segmenting and identifying nasopharyngeal carcinoma,” Computers & Electrical Engineering, vol. 71, pp. 372–387, 2018.

[29] A. Cruz-Roaa, A. Basavanhally, F. González, H. Gilmore, M. Feldman et al., “Automatic detection of invasive ductal carcinoma in whole slide images with convolutional neural networks,” Medical Imaging 2014: Digital Pathology, vol. 9041, no. 216, pp. 904103, 2014.

[30] J. De Fauw, J. R. Ledsam, B. Romera-Paredes, S. Nikolov, N. Tomasev et al., “Clinically applicable deep learning for diagnosis and referral in retinal disease,” Nature Medicine, vol. 24, no. 9, pp. 1342– 1350, 2018.

[31] 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. 2021, 24, 1207–1220. [CrossRef]

[32] Tang, Y.-X.; Tang, Y.-B.; Peng, Y.; Yan, K.; Bagheri, M.; Redd, B.A.; Brandon, C.J.; Lu, Z.; Han, M.; Xiao, J.; et al. Automated abnormality classification of chest radiographs using deep convolutional neural networks. NPJ Digit. Med. 2020, 3, 1–8. [CrossRef]

[33] Minaee, S.; Kafieh, R.; Sonka, M.; Yazdani, S.; Soufi, G.J. Deep-COVID: Predicting COVID-19 from chest X-ray images using deep transfer learning. Med. Image Anal. 2020, 65, 101794. [CrossRef]

[34] Afshar, P.; Heidarian, S.; Naderkhani, F.; Oikonomou, A.; Plataniotis, K.N.; Mohammadi, A. COVID-CAPS: A capsule networkbased framework for identification of COVID-19 cases from X-ray images. Pattern Recognit. Lett. 2020, 138, 638–643. [CrossRef]

[35] Ahsan, M.; Gupta, K.; Islam, M.; Sen, S.; Rahman, L.; Hossain, M.S. COVID-19 Symptoms Detection Based on NasNetMobile with Explainable AI Using Various Imaging Modalities. Mach. Learn. Knowl. Extr. 2020, 2, 490–504. [CrossRef]

[36] Hemdan, E.E.; Shouman, M.A.; Karar, M.E. COVIDXnet: A framework of deep learning classifiers to diagnose COVID-19 in X-ray images. arXiv 2020, arXiv:2003.11055. Available online: http://arxiv.org/abs/2003.11055 (accessed on 19 August 2021).

[37] 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. 2020, 121, 103792. [CrossRef]

[38] Pandit, M.; Banday, S.; Naaz, R.; Chishti, M. Automatic detection of COVID-19 from chest radiographs using deep learning. Radiography 2020, 27, 483–489. [CrossRef]

[39] Hu, M.; Lin, H.; Fan, Z.; Gao, W.; Yang, L.; Liu, C.; Song, Q. Learning to Recognize Chest-Xray Images Faster and More Efficiently Based on Multi-Kernel Depthwise Convolution. IEEE Access 2020, 8, 37265–37274. [CrossRef]

[40] Ismael, A.M.; ¸Sengür, A. Deep learning approaches for COVID-19 detection based on chest X-ray images. Expert Syst. Appl. 2020, 164, 114054. [CrossRef]

[41] Shankar, K.; Perumal, E.; Díaz, V.G.; Tiwari, P.; Gupta, D.; Saudagar, A.K.J.; Muhammad, K. An optimal cascaded recurrent neural network for intelligent COVID-19 detection using Chest X-ray images. Appl. Soft Comput. 2021, 113, 107878. [CrossRef]

[42] Akter, Shamima, et al. "COVID-19 detection using deep learning algorithm on chest X-ray images." *Biology* 10.11 (2021): 1174.

[43] Kaggle. Available online: https://www.kaggle.com/tawsifurrahman/covid19-radiography-database?fbclid=IwAR0rw\_prTvf9 R0zInrJQkTFazeBaESxh3rB6otdrPdAWJDonEbIl2Nf6epk (accessed on 10 September 2021).

[44] Cheng PM, Malhi HS (2017) Transfer learning with convolutional neural networks for classification of abdominal ultrasound images. J Digit Imaging: 234–243.

[45] Ravishankar H, Sudhakar P, Venkataramani R, Thiruvenkadam S, Annangi P, Babu N, Vaidya V (2016) Understanding the mechanisms of deep transfer learning for medical images. In: Deep learning and data labeling for medical applications. Springer, Cham, pp 188–196

[46] Yu Y, Lin H, Meng J, Wei X, Guo H, Zhao Z (2017) Deep transfer learning for modality classification of medical images. Information 8(3):91

[47] Vedaldi A, Lenc K (2015) Matconvnet: convolutional neural networks for matlab. In: Proceedings of the 23rd ACM international conference on multimedia (pp. 689–692). ACM

[48] Ioffe S, Szegedy C (2015) Batch normalization: accelerating deep network training by reducing internal covariate shift. arXiv preprint arXiv:1502.03167

[49] Dahl GE, Sainath TN, Hinton GE (2013) Improving deep neural networks for LVCSR using rectified linear units and dropout. In: 2013 IEEE international conference on acoustics, speech and signal processing (pp. 8609–8613). IEEE

[50] Koushik J (2016) Understanding convolutional neural networks. arXiv preprint arXiv:1605.09081