Delta University Scientific Journal
Journal home page: https://dusj.journals.ekb.eg

# Diagnosis of Pneumonia from Chest X-Ray Images using Deep Learning

**Amany Sarhan*[1], Aml Alaa Khattab[1], Mahmoud Shaheen[2]**

[1]*Computer and Control Engineering Department, Faculty of Engineering, Tanta University, Tanta, Egypt;*
[2]*Faculty of Artificial Intelligence, Delta University for Science and Technology, Gamasa, Egypt*
Correspondence: Email : amany_sarhan@f-eng.tanta.edu.eg , pg_190285@f-eng.tanta.edu.eg

**ABSTRACT**
Pneumonia is a disease of the lungs caused by a bacterial infection. Early diagnosis is critical to the outcome of treatment. In general, a trained radiologist may identify the condition using chest X-ray pictures. Diagnoses might be subjective for a variety of reasons, including the appearance of disease, which may be ambiguous in chest X ray pictures or mistaken with other conditions. As a result, computer-aided diagnostic systems are required to advise practitioners. In this study, we employed two well-known convolutional neural network models, Xception and VGG16, and a custom CNN model to diagnose pneumonia. We employed transfer learning and fine-tuning throughout the training step. The test findings indicated that the custom CNN model outperformed VGG16 network and the Xception model with 93% accuracy.

***Keywords:*** Pneumonia; transfer learning; Xception; VGG16; deep learning.

## 1. Introduction

Pneumonia is inflammation of the tissues in one or both lungs that usually caused by a bacterial infection. In the USA annually more than 1 million people are hospitalized with the gripe of pneumonia. Unfortunately, 50.000 of these people die from this illness (CDC on Pneumonia statistics). Fortunately, pneumonia can be a manageable disease by using drugs like antibiotics and antivirals. However, early diagnosis and treatment of pneumonia is important to prevent some complications that lead to death (Aydogdu et al.). Chest X-ray images are the best known and the common clinical method for diagnosing of pneumonia. However, diagnosing pneumonia from chest X-ray images is a challenging task for even expert radiologists. The appearance of pneumonia in X-ray images is often unclear, can confuse with other diseases and can behave like many other benign abnormalities.

These inconsistencies caused considerable subjective decisions and varieties among radiologists in the diagnosis of pneumonia (Neuman et al.). Therefore, there is a need for computerized support systems to help radiologists for diagnosing pneumonia from chest X-ray images. Recent developments in deep learning field, especially convolutional neural networks (CNNs) showed great success in image classification (Krizhevsky et al.). The main idea behind the CNNs is creating an artificial model like a human brain visual cortex. The main advantage of CNNs, it has the capability to extract more significant features from the entire image rather than hand crafted features (LeCun et al.). Researchers developed different CNN based deep networks and these networks achieved state of results in classification, segmentation, object detection and localization in computer vision (He et al. and Ronneberger et al.).

Besides the natural computer vision problems, CNNs achieved very successful results in solving medical problems such as breast cancer detection (Ragab et al.), brain tumor segmentation (Gab Allah, et al.), alzheimer disease diagnosing (Arafa, et al.), Knee (Sarhan et al.) etc. Various studies about the detection of pneumonia using deep learning has been introduced using DenseNet 121 (Antin et al. and Rajpurkar, et al.) and CheXNet (Chollet). Other used custom CNN, Inception-V3, VGG16 (Mujahid et al., Sharma, S. and Guleria), Quaternion Channel-Spatial Attention Network (Singh et al.) or AlexNet (Ibrahim et al.). Based on this information, we modified and trained two well-known networks for classifying pneumonia from chest X-ray images. Our first network is based on the Xception model.

The second one is VGG16 based model. Besides, we utilized transfer learning, fine-tuning and data augmentation methods. For an objective comparison between them, we used same parameters when training both networks. Also, we compared the performance of two networks on the test data with different metrics. The results show that the Xception model outperforms Vgg16 model in diagnosing pneumonia. On the other hand, VGG16 model showed better performance in diagnosing normal cases.

## 2. Previous work

Deep learning (DL) is a subset of artificial intelligence (AI) that draws inspiration from the structure of the human brain. It is defined as a subfield of machine learning that operates similarly to the biology of human brains by collecting data and processing it using neural networks. Many biological health challenges, including cancer detection (brain tumor and breast cancer), are being addressed with computer-aided diagnostics based on AI models. DL models may discover hidden elements in photos that medical specialists cannot see. Convolutional neural network (CNN) is the major DL tool that is widely used in various subfields of the healthcare system due to its ability to draw features and develop the ability to differentiate between different classes. In addition, Transfer learning (TL) has made it simpler to swiftly retrain neural networks on specific datasets with good accuracy. Pneumonia is a frequent illness caused by several microbiological species, including bacteria, viruses, and fungus. The term "pneumonia" is derived from the Greek word "pneumon," which means "lungs." Thus, the term pneumonia refers to a lung ailment. Pneumonia is a medical condition that causes inflammation of one or both lung parenchymas.

Other causes of pneumonia include food aspiration and chemical exposure. Pneumonia occurs as a result of pathogen-induced inflammation, which causes the lung's alveoli to fill up with fluid or pus, resulting in a decrease in carbon dioxide ($CO_2$) and oxygen ($O_2$) exchange between blood and the lungs, making it difficult for infected people to breathe. Using DL to classify the illness early will help in assisting the patients to recover faster. Several work have been dedicated towards this solution. Mujahid et al. had used several pre-trained convolutional neural network (CNN), such as VGG16, Inception v3, and ResNet50. The dataset is obtained from Kaggle containing 7750 X-ray images. They preprocessed the dataset before transfer learning tasks. Ensembles are created by combining CNN, Inception-V3, VGG-16, and ResNet50. The experimental findings demonstrate that Inception-V3 with CNN achieved the greatest accuracy and recall scores of 99.29% and 99.73%, respectively.

Sharma, S. and Guleria introduced a deep learning (DL) model with VGG16 to identify and categorize pneumonia using two CXR image datasets. For the first dataset, the VGG16 using Neural Networks (NN) achieves 92.15% accuracy. Furthermore, the experiment using NN and VGG16 was carried out on another CXR dataset encompassing 6,436 pictures of pneumonia, normal, and covid-19 with 95.4% accuracy.

Singh et al. categorized chest X-Ray pictures for pneumonia diagnosis using a QCSA network (Quaternion Channel-Spatial Attention Network), which combines the spatial and channel attention mechanisms with a Quaternion residual network on Kaggle X-ray dataset. Their architecture obtained 94.53% accuracy. Ibrahim et al. suggested using a deep learning strategy based on a pretrained AlexNet model to classify COVID-19, non-COVID-19 viral pneumonia, bacterial pneumonia, and normal Chest X-ray images gathered from several public datasets. The model classified Chest X-ray images of COVID-19 pneumonia and non-COVID-19 viral pneumonia with 99.62% accuracy, 90.63% sensitivity, and 99.89% specificity. The three-way classification model obtained 94.00% accuracy, 91.30% sensitivity, and 84.78%. Finally, for the four-way classification, the model obtained 93.42% accuracy, 89.18% sensitivity, and 98.92% specificity.

## 3. Material and methods

### *3.1. Dataset Details*

In this study, a dataset consisting of 5856 frontal Grey-scaled images chest X-ray images provided on Kaggle. The images in the dataset are varying resolutions such as 712x439 to 2338x2025. There are 1583 normal case, 4273 pneumonia case images in the dataset. Fig. 1 shows some X-ray image samples from the dataset. Table 1 represents the distribution of the data when training, validating and testing phases of the proposed model. In our models 0 represents normal cases, 1 represents pneumonia cases.

### *3.2 Data per-processing*

In this study, the collected radiograph images have been fixed to the size of $200 \times 200$ resolution to improve the system stability. The data normalization process was also applied from range 0 to 1 on all of the datasets to prevent the model from overfitting. The dataset was divided into three portions such as training, validation, and testing splitted as 80:20 for Train-validation and testing respectively. The proposed model was trained using the training and validation sets of normal and pneumonia disease radio graphs.

**Table 1: Distribution of the dataset**

| | Train | Validation | Test |
|---|---|---|---|
| **Normal** | 1349 | 234 | 234 |
| **Pneumonia** | 3883 | 390 | 390 |
| **Total** | 5232 | 624 | 624 |

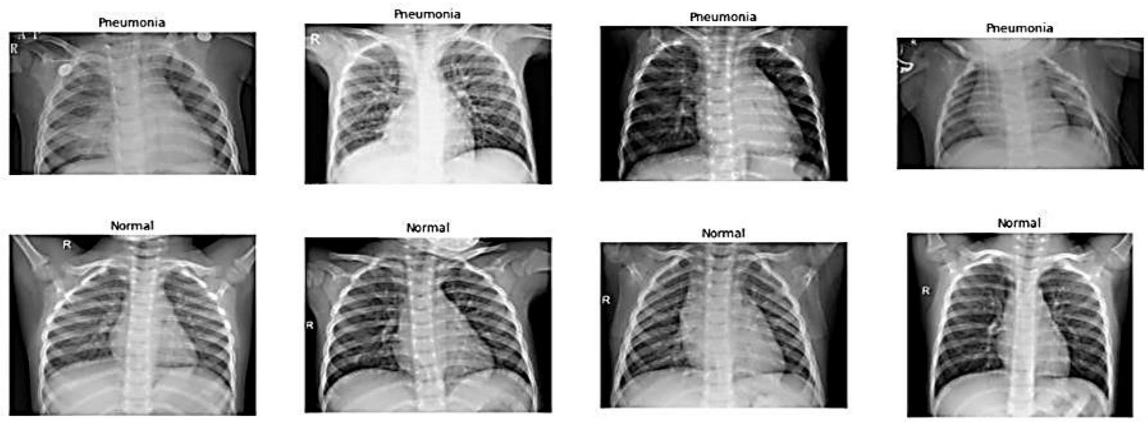


**Fig. 1. Sample of Chest X-ray images from the dataset**

*3. ٣Data augmentation and transfer learning*

One of the major problems of deep learning is that it requires a large quantity of data to get accurate results. However, there may be insufficient data in some cases. Obtaining and annotating data, particularly for medical conditions, is a costly and time-consuming operation. Fortunately, there are several answers to this problem. One of them is data augmentation, which prevents overfitting and increases accuracy (Gab Allah et al.). In this work, the training time data augmentation strategy was used. These techniques modifies the radio graphic images. We employed several augmentation methods, including rotation with 90 degrees, width shifting and height shifting, zooming with range 0.1, horizontal and vertical flipping, and rotating at 40-degree angles. Fig. 2 illustrates example of data augmentation.

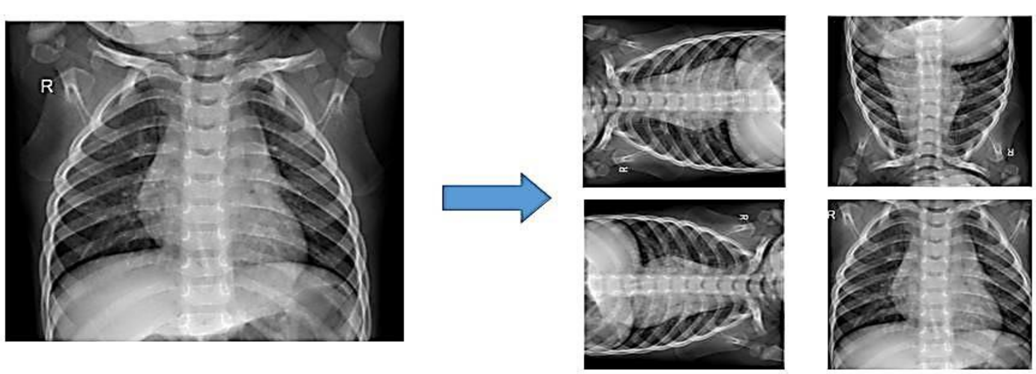

**Fig. 2. Some examples of Chest X-ray images data augmentation**

Other performance-enhancing strategy in deep models, particularly CNNs, is known as transfer learning (Iman et al.). Transfer learning is the concept of transcending the solitary learning paradigm and applying information gained from one job to tackle related ones. Today, just a few individuals train a full CNN from scratch. Because it requires a large volume of data. Instead, use pretrained CNNs on huge datasets like ImageNet, which has 1.2 million images and 1000 classes. There are three distinct transfer learning techniques in CNNs. These include feature extraction, fine-tuning, and pre-trained models. In this work, we adopted a fine tuning strategy inspired by the observation that early layers of CNNs contain more general characteristics such as edges, colors, and blobs. So, this layer should be useful for many other tasks. But last layers have more data specific features. Therefore, we fixed some early layers of our models and trained our models excluding fixed layers.

### *3.4 First Proposed model: Xception model*

Xception model stands for the extreme version of Inception model. It uses a modified depth wise separable convolutional layer. In the first step, it uses a 1x1 pointwise convolution and follows by a 3x3 depth wise convolution. Fig. 2 shows used separable convolution in Xception model. The Xception architecture has 36 convolutional layers for feature extraction. After convolutional layer followed by a logistic regression layer. If desired, fully connected layers can be used between convolutional layers and logistic regression layer. The convolutional layers structured as 14 modules and all of them has a linear residual connection between them. The base model Xception has archived better performance than Inception-V3 classification of ImageNet dataset.
We utilized transfer learning and fine-tuning on the Xception model, therefore used pretrained ImageNet weights before the start to train and the last 10 early layers were frozen. After the global average polling layer, we added two fully connected layers (1024, 512) and a two-way output layer with SoftMax activation function. Fig. 2 shows the proposed Xception model.

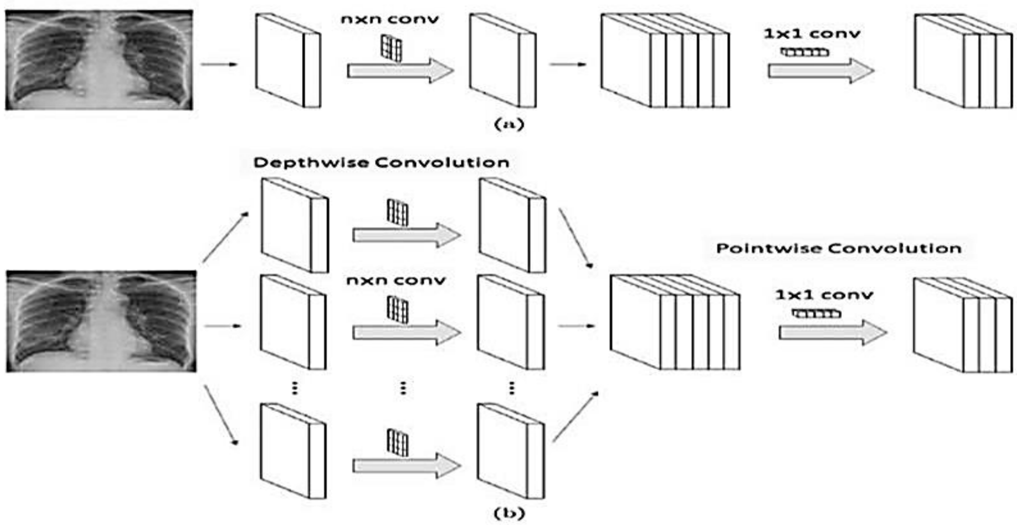


**Fig. 2. Separable convolution process used in the Xception model**

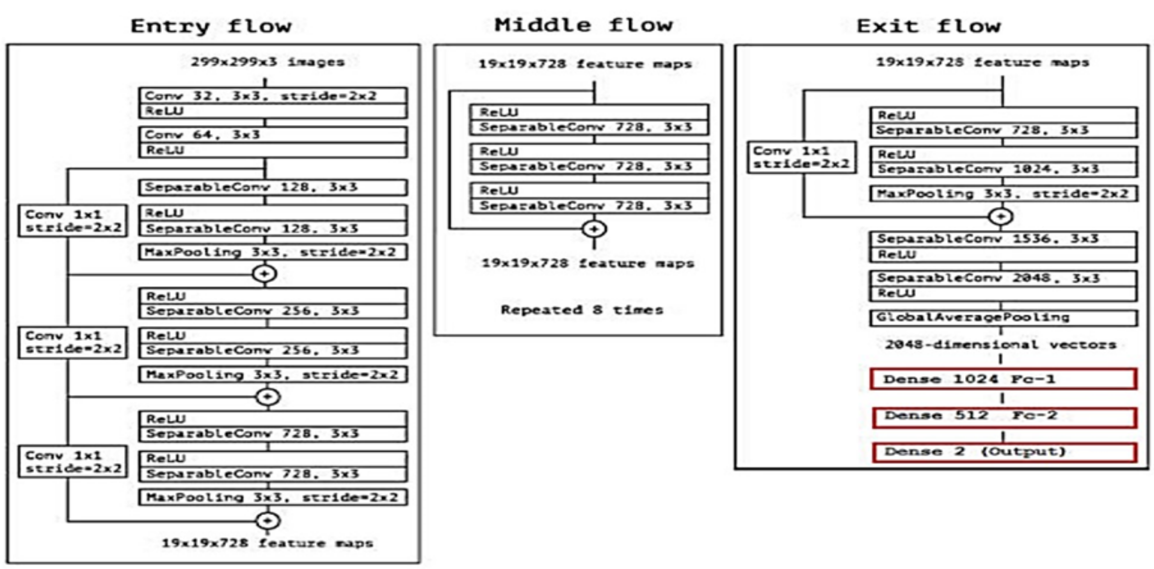


**Fig. 3. Fine-tuned Xception network**

*3.5 Second Proposed model: VGG16 model*

The model has 16 convolutional layers with small receptive fields (3x3), 144 million parameters, five maxpooling layers (2x2 size) and three fully-connected layers, with the final layer has a soft-max activation function. We used this model with pre-trained weights on ImageNet and modified fully connected layers of the network. Also, we fixed the last 8 early layers to avoid the train. Fig. 4 shows our modified network.

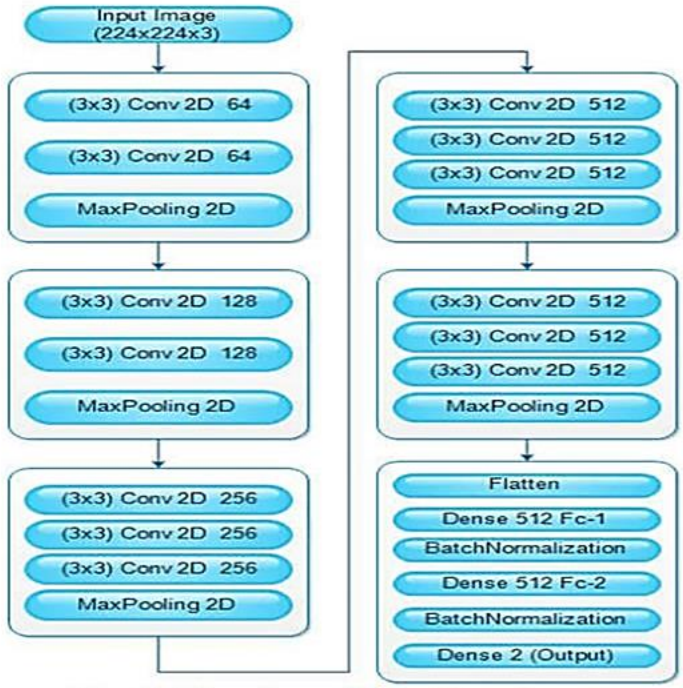


**Fig. 4. VGG16 model architecture**

*3.6 Third Proposed model: CNN Architecture*

In this study, we designed a convolutional neural network (CNN) to detect pneumonia from chest X-ray images. The CNN model consists of multiple layers, starting with an input layer that accepts images of size 200 x 200 (after resizing), where each image is a 3D tensor representing the image height, width, and channels. The core of the model comprises convolutional layers, each followed by Rectified Linear Unit (ReLU) activation to introduce non-linearity, and max-pooling layers to down sample the feature maps and reduce spatial dimensions. The first convolutional layer applies 256 filters with a kernel size of (3, 3), and this is followed by a max-pooling operation with a (2, 2) window. Subsequent convolutional layers use 64 and 16 filters, respectively, each accompanied by max-pooling layers. To improve convergence and prevent internal co variate shifts, each convolutional block is followed by batch normalization. Fig. 5 shows the model architecture.

The output from the final convolutional layer is passed through a flattening layer to convert the 3D feature maps into a 1D vector, which is then fed into fully connected layers. The first dense layer has 64 units with ReLU activation and is followed by a dropout layer with a dropout rate of 0.5 to mitigate overfitting. The final output layer consists of a single neuron with a sigmoid activation function, which outputs the probability of pneumonia presence, making the model suitable for binary classification. The sigmoid function, defined as:

$$\sigma(x) = \frac{1}{1+e^{-x}} \qquad (1)$$

produces a value between 0 and 1, representing the likelihood of pneumonia.

*3.7 Adam Optimizer*

After defining the architecture, we compiled the model using the Adam optimizer with a learning rate of 0.0001. Adam, short for Adaptive Moment Estimation, combines the advantages of both RMSProp and SGD with momentum, making it efficient in handling sparse gradients and ensuring faster convergence. The binary crossentropy loss function was used, which is appropriate for binary classification tasks. This loss function is computed as:

$$L(y,\hat{y}) = -\frac{1}{n}\sum_{i=1}^{n}[y_i \log(\hat{y}_i) + (1-y_i)\log(1-\hat{y}_i)] \quad (2)$$

where $yi$ is the true label and $\hat{yi}$ is the predicted probability. This model architecture allows for efficient learning from the chest X-ray data, enabling the accurate detection of pneumonia in patients.

| Layer (type) | Output Shape | Param # |
|---|---|---|
| conv2d_3 (Conv2D) | (None, 200, 200, 256) | 2,560 |
| activation_5 (Activation) | (None, 200, 200, 256) | 0 |
| max_pooling2d_3 (MaxPooling2D) | (None, 100, 100, 256) | 0 |
| batch_normalization_3 (BatchNormalization) | (None, 100, 100, 256) | 400 |
| conv2d_4 (Conv2D) | (None, 100, 100, 64) | 147,520 |
| activation_6 (Activation) | (None, 100, 100, 64) | 0 |
| max_pooling2d_4 (MaxPooling2D) | (None, 50, 50, 64) | 0 |
| batch_normalization_4 (BatchNormalization) | (None, 50, 50, 64) | 200 |
| conv2d_5 (Conv2D) | (None, 50, 50, 16) | 9,232 |
| activation_7 (Activation) | (None, 50, 50, 16) | 0 |
| max_pooling2d_5 (MaxPooling2D) | (None, 25, 25, 16) | 0 |
| batch_normalization_5 (BatchNormalization) | (None, 25, 25, 16) | 100 |
| flatten_1 (Flatten) | (None, 10000) | 0 |
| dropout_2 (Dropout) | (None, 10000) | 0 |
| dense_2 (Dense) | (None, 64) | 640,064 |
| activation_8 (Activation) | (None, 64) | 0 |
| dropout_3 (Dropout) | (None, 64) | 0 |
| dense_3 (Dense) | (None, 1) | 65 |
| activation_9 (Activation) | (None, 1) | 0 |

**Fig. 5. Custom CNN model architecture**

## Results

This system has been developed on a device with 16 GB RAM, Ryzen7, and GPU Nvidia GEFORCE RTX 3050 with 4 GB RAM connected to Kaggle environment on TPU T100. We evaluated the three deep learning models by using 624 frontal chest X-ray images. The test set contains 234 normal and 390 pneumonia cases. In this section, training strategies and test results were presented. Before the training phase, all images are resized for the target network model. Because, Xception network accepts images at 299x299x3 dimensions whereas VGG16 network accepts images at 224x224x3 dimensions. Also, all image pixels were normalized range in [0, 1]. We trained both networks by using the same parameters. These train parameters are, epoch size is set as 50, categorical cross entropy selected as loss function, ADAM is used as the optimizer, learning rate set as 1e-4, weight decay set as 0.9 and batch size set as 16. Different approaches were used for avoiding overfitting. Firstly, batch normalization was used after every convolutional layer. Secondly, the dropout method was used after fully connected layers with 0.5 rate. Lastly data augmentation was used for avoiding overfitting.

The accuracy, sensitivity and specific results of the three models are given in Table 2. Table 3 shows Xception, VGG16 and custom CNN models in terms of precision, recall, f1 score for the two classes: pneumonia and Normal. Finally, Table 4 gives the training and testing time of the two models (Xception and VGG16).

Precision (the ratio of true positive predictions to the total predicted positives) is crucial in this context as it indicates how many of the predicted pneumonia cases were actually positive.

$$\text{Precision} = \frac{True\ positives}{True\ positives + False\ positives} \quad (3)$$

Recall (the ratio of true positives to the total actual positives) is important to ensure that the model detects the maximum number of pneumonia cases, especially in critical medical applications.

$$\text{Recall} = \frac{True\ positives}{True\ positives + False\ Negatives} \quad (4)$$

The F1-score combines precision and recall, providing a balanced measure of both metrics. It is particularly useful when dealing with imbalanced datasets, as in the case of pneumonia detection.

$$\text{F1- score} = 2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} \quad (5)$$

Figures 7-9 shows the results of the custom CNN model as it achieved the best results of the three models. We can see that the custom CNN has the highest accuracy, precision, recall and F1-Score. It is also able to identify the image class with minimum number of false negative and positive as indicated by the confusion matrix.

**Table 2: Accuracy, sensitivity and specific results of the three models**

| | Accuracy | Sensitivity | Specificity |
|---|---|---|---|
| **Xception** | 0.82 | 0.85 | 0.76 |
| **VGG16** | 0.87 | 0.82 | 0.91 |
| **Custom CNN** | 0.93 | 0.90 | 0.95 |

**Table 3: Precision, recall, F1-Score results of the three models in the two classes**

| | Xception | | | VGG16 | | | Custom CNN | | |
|---|---|---|---|---|---|---|---|---|---|
| | Precision | Recall | F1-Score | Precision | Recall | F1-Score | Precision | Recall | F1-Score |
| **Normal** | 0.86 | 0.65 | 0.74 | 0.83 | 0.86 | 0.84 | 0.94 | 0.76 | 0.91 |
| **Pneumonia** | 0.82 | 0.94 | 0.87 | 0.91 | 0.89 | 0.90 | 0.95 | 0.79 | 0.92 |

**Table 4: Time of training and testing of the two models**

| | Training Time | Testing Time |
|---|---|---|
| **Xception** | 100 mins | 13 secs |
| **VGG16** | 83 mins | 10 secs |

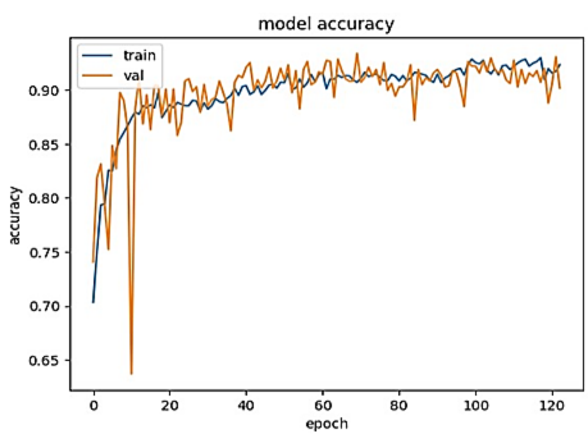


**Fig. 6. Custom CNN model accuracy vs no. of epochs**

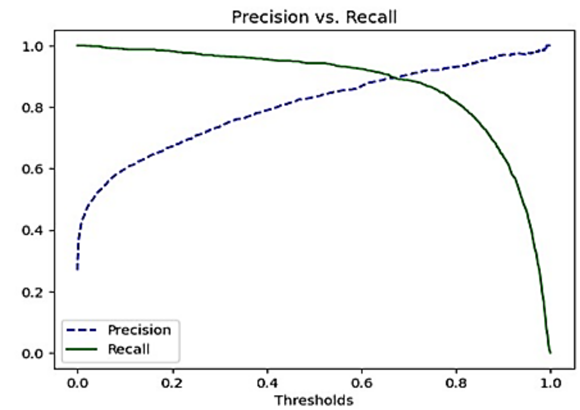


**Figure 7. Custom CNN model Precision vs Recall diagram**

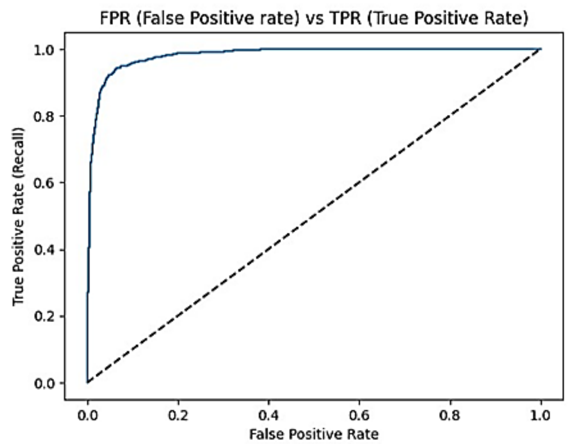


**Fig. 8. Custom CNN model Recall diagram**

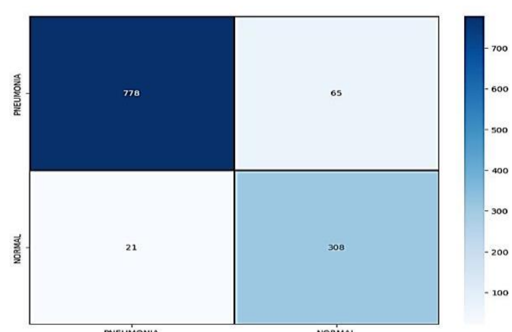


**Fig. 9. Custom CNN model Confusion Matrix**

## Conclusion

In this study, we compared two CNN network's performance on the diagnosis of pneumonia disease. While training our model we used from transfer learning and finetuning. After the training phase, we compared two network test results. The test results showed that Vgg16 network outperforms Xception network by accuracy 0.87 %, specificity 0.91%, pneumonia precision 0.91% and pneumonia f1 score 0.90%. Whereas Xception network outperforms Vgg16 network by sensitivity 0.85%, normal precision 0.86% and pneumonia recall 0.94%. According to the experimental results and confusion matrices in Fig. 7 every network has own detection capability on the dataset. Xception network is more successful for detecting pneumonia cases than Vgg16 network. At the same time Vgg16 network is more successful at detecting normal cases. In the future work we will ensemble of two networks. In this way we will combine strengths of two networks and will achieve more successful results on diagnosing of pneumonia from chest X-ray images.

**Disclosure**

The author reports no conflicts of interest in this work.

**References**

https://www.kaggle.com/datasets/paultimothymooney/chest-xray-pneumonia.

Arafa DA, Moustafa H, Ali HA, Ali-Eldin A and Saraya MS. A deep learning framework for early diagnosis of Alzheimer's disease on MRI images. Multimedia Tools Applications, vol. 83, pp. 3767–3799, 2024.

Aydogdu M, Ozyilmaz, E., Aksoy, H., Gursel, G. and Ekim, N. Mortality prediction in community-acquired pneumonia requiring mechanical ventilation; values of pneumonia and intensive care unit severity scores," Tuberk Toraks, vol. 58, no. 1, pp. 25-34, 2010.

CDC. (2019). Pneumonia Statistics. Available: https://www.cdc.gov/pneumonia/prevention.html

Chollet F, "Xception: Deep learning with depthwise separable convolutions," arXiv preprint, p. 1610.02357, 2017

He K, Zhang X, Ren S and Sun J. Deep residual learning for image recognition. Proceedings of the IEEE conference on computer vision and pattern recognition, 2016, pp. 770-778.

Gab Allah AM, Sarhan AM, Elshennawy NM. Classification of Brain MRI Tumor Images Based on Deep Learning PGGAN Augmentation. Diagnostics, vol. 11, 2343, 2021.

Ibrahim AU, Ozsoz M, Serte S, Al-Turjman F and Yakoi PS. Pneumonia Classification Using Deep Learning from Chest X-ray Images During COVID-19. Cognitive Computing, vol. 16, pp. 1589–1601, 2024.

Iman M, Arabnia HR, Rasheed K. A Review of Deep Transfer Learning and Recent Advancements. Technologies, 2023, vol. 11, no. 40.

Kaggle Dataset. Available online: https://www.kaggle.com/paultimothymooney/chest-xray-pneumonia (accessed on 28 January 2022).

Krizhevsky A, Sutskever I and Hinton GE. Imagenet classification with deep convolutional neural networks. Advances in neural information processing systems, pp. 1097-1105, 2012.

LeCun Y, Bengio Y and Hinton G. Deep learning. Nature, vol. 521, no. 7553, p. 436, 2015.

Mujahid M, Rustam F, Álvarez R, Mazón LV, Díez J, and Ashraf I. Pneumonia Classification from X-ray Images with Inception-V3 and Convolutional Neural Network. Diagnostics, vol. 12, no. 1280, 2022.

Neuman MI, Lee EY, Bixby S, Diperna S, Hellinger J, Markowitz R, Servaes S, Monuteaux MC, Shah SS. Variability in the interpretation of chest radiographs for the diagnosis of pneumonia in children. Journal of Hospital Medicine, vol. 7, no. 4, pp. 294-8, 2012

Ragab DA, Sharkas M, Marshall S and Ren J. Breast cancer detection using deep convolutional neural networks and support vector machines. PeerJ, vol. 7, p. e6201, 2019.

Rajpurkar P, Irvin J, Ball RL, Zhu K, Yang B, Mehta H. Deep learning for chest radiograph diagnosis: A retrospective comparison of the CheXNeXt algorithm to practicing radiologists. PLOS medicine, 2018.

Ronneberger O, Fischer P and Brox T. U-net: Convolutional networks for biomedical image segmentation. Proceedings of the International Conference on Medical image computing and computer-assisted intervention, pp. 234-241, Springer, 2015.

Sarhan AM, Gobara M, Yasser, S et al. Knee Osteoporosis Diagnosis Based on Deep Learning. Int J Comput Intell Syst 17, 241 (2024).

Sharma, S. and Guleria, K. A Deep Learning based model for the Detection of Pneumonia from Chest X-Ray Images using VGG-16 and Neural Networks, Procedia Computer Science, Volume 218, Pages 357-366, 2023, Singh S, Kumar M, Kumar A et al. Pneumonia detection with QCSA network on chest X-ray. Scientific Reports, vol. 13, no. 9025, 2023.

Tsang SH Review: Xception—With Depthwise Separable Convolution, Better Than Inception-v3, 2019.

Yang HM, Duan T, Ding D, Bagul A, and Langlotz C. Chexnet: Radiologist-level pneumonia detection on chest x-rays with deep learning. arXiv preprint arXiv:1711.05225, 2017.