# Pneumonia Detection from Chest X-ray: A CNN Approach

**Samruddhi Raut**
*Masters in Robotics*
*Northeastern University*

***Abstract*— Pneumonia is the leading cause of death in children under the age of five. It accounted for approximately 16% of the deaths of children under the age of five, killing around 880,000 children in 2016 according to a study conducted by UNICEF. Affected children were mostly less than two years old. Timely detection of pneumonia in children can help to fast-track the process of recovery. Chest X-ray imaging is the most frequently used method for diagnosing pneumonia. However, the examination of chest X-rays is a challenging task and is prone to subjective variability. In this paper, I present convolutional neural network models to accurately detect pneumonic lungs from chest X-rays, which can be utilized in the real world by medical practitioners to treat pneumonia. Also, I employed pretrained deep transfer learning models to compare the results with self-built convolution neural network. Experimentation was conducted on Chest X-Ray Images (Pneumonia) dataset available on Kaggle. The proposed model performance (accuracy: 92.64%, recall: 94%, F1 score: 93%) is superior to the pre-trained transfer learned models DenseNet121 (accuracy: 92.64%, recall: 94%, F1 score: 93%), VGG-16(accuracy: 92.64%, recall: 94%, F1 score: 93%) Res-50(accuracy: 92.64%, recall: 94%, F1 score: 93%), InceptionV3(accuracy: 92.64%, recall: 94%, F1 score: 93%). Recall and F1 scores are calculated from the confusion matrix of each model for better evaluation. The model's performance in pneumonia detection shows that the proposed self-built CNN model effectively classifies normal and abnormal X-rays in practice, hence reducing the burden of radiologists.***

**Keywords—CNN, Transfer learning, Pretrained Models, Recall, F1-Score, Accuracy.**

## I. INTRODUCTION

One of the major factors associated with pneumonia in children is indoor air pollution. Apart from this, under-nutrition, lack of safe water, sanitation and basic health facilities are also major factors. Pneumonia is an interstitial lung disease caused by bacteria, fungi, or viruses. There are more than 1400 children infected with pneumonia per 100,000 children. The Global Burden of Disease Study reported that lower respiratory tract infections, including pneumonia, were the second largest cause of death in 2013. The large number of child deaths by pneumonia alarms scientists worldwide to propose more effective and acute methods to detect pneumonia. With technology developing, more and more measures are developed, in which radiology-based methods are most popular and useful. Diagnostic radiological techniques for pulmonary disease include chest X-ray imaging, computed tomography (CT), and magnetic resonance imaging (MRI), among which chest X-ray imaging is most effective and economical as it is more available and portable in hospital and has lower exposures of dose radioactivity for patients. Fig 1. shows an example shows an example of a pneumonic and a healthy lung X-ray. The white spots in the pneumonic X-ray (indicated with red arrows), called infiltrates, distinguish a pneumonic from a healthy condition. However, chest X-ray examinations for pneumonia detection are prone to subjective variability. Even for very professional and experienced doctors, the diagnosis of pneumonia through X-ray images is still a tremendous task because X-ray images have similar region information for different diseases, such as lung cancer. Therefore, it is very time consuming and energy-consuming to diagnose pneumonia through traditional methods and impossible to diagnose whether a patient suffers pneumonia through a standardized process. Thus, an automated system for the detection of pneumonia is required.

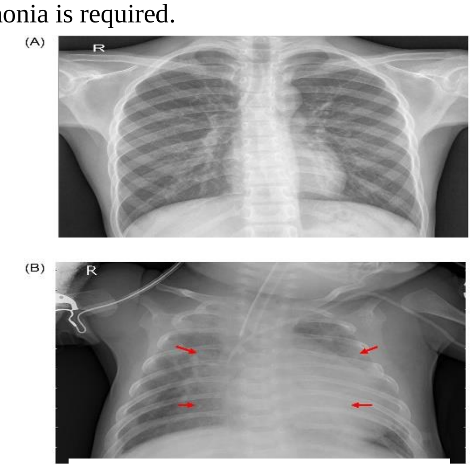



**Fig 1. Examples of two X-ray plates that display** (a) a healthy lung and (b) a pneumonic lung. The red arrows in (b) indicate white infiltrates, a distinguishing feature of pneumonia

Hence, in this study, I propose a convolutional neural network to accurately detect pneumonic lungs from chest X-rays, which can be utilized in the real world by medical practitioners to treat pneumonia. Also, I employed the transfer learning and used various pretrained models like DenseNet121, VGG16, Resnet50 and InceptionV3 to compare the results with the proposed CNN. These models have been trained to classify chest X-ray images into normal and pneumonia in a few seconds, hence serving the purpose of early detection of pneumonia. The main objective of the paper is to develop CNN models from scratch which can classify and thus detect pneumonic patients from their chest X-rays with high validation accuracy, recall and F1 scores. Recall is often favored in medical imaging cases over other performance evaluating parameters, as it gives a measure of false negatives in the results.

The rest of this paper is organized as follows: Section 2 explores the work related to this field that has been accomplished till now. Section 3 explains the methodology of the paper, explaining the architecture of the models, and the dataset used to train and test the models. Section 4 presents the results achieved by the various CNN models and compares the performance of each model using accuracy and loss graphs and confusion matrices. Section 5 provides a brief conclusion to the paper and delivers the best-suited model. Furthermore, the future scope of this research work has also been discussed. All the references which are cited in the paper have been listed in the end.

## II. Literature Review

Many researchers have tackled the problem of classifying images with high accuracy. Here are some citations:

Pneumonia detection using chest X-rays has been an open problem for many years [1, 2], the main limitation being the scarcity of publicly available data. Traditional machine learning methods have been explored extensively. Chandra et al. [3] segmented the lung regions from chest X-ray images and extracted eight statistical characteristics from these regions, which they used to classify them. They implemented five traditional classifiers: multi-layer perceptron (MLP), random forest, sequential minimal optimization (SMO), classification via regression, and logistic regression. They evaluated their method on 412 images and achieved a 95.39% accuracy rate using the MLP classifier. Kuo et al. [4] used 11 features to detect pneumonia in 185 schizophrenia patients. They applied these features in many regression and classification models, such as decision trees, support vector machines, and logistic regression, and compared the results of the models. They achieved the highest accuracy rate, 94.5%, using a decision tree classifier; the other models fell short by large margins. Similarly, Yue et al. [5] used 6 features to detect pneumonia in chest CT scan images of 52 patients; the best AUC value they achieved was 97%. However, these methods cannot be generalized and were evaluated on small datasets.

Several methods have been introduced to describe a brief process in pneumonia detection using chest X-ray images in recent years, especially some deep learning methods. Rubin et al. [6] developed a CNN model to detect common thorax disease from frontal and lateral chest X-ray images. MIMIC-CXR dataset was used to perform large-scale automated recognition of these images. The dataset was split into training, testing and validation sets as 70%, 20% and 10%, respectively. Data augmentation and pixel normalization were used to improve overall performance. Their DualNet CNN model achieved an average AUC of 0.72 and 0.688 for PA and AP, respectively.

Also, A deep convolutional neural network to classify pulmonary tuberculosis was developed by Lakhani et al. [7]. Transfer learning models such as AlexNet and GoogleNet were also used to classify chest X-ray images. The dataset was split into training, testing and validation sets as 68%, 14.9% and 17.1%, respectively. Data augmentation and pre-processing techniques were employed to get the best performing model achieving an AUC of 0.99. Precision and recall of the model were 100 and 97.3%. An AG-CNN model was developed by Guan et al. [8] to recognize thorax disease. ChestX-ray14 dataset was used to detect thorax disease from chest X-ray images. Global and local branch attention-guided CNN was used for classification purposes. Their model was better than other models mentioned in their research paper, achieving an AUC of 0.868. A deep convolutional neural network model was developed by Rajpurkar et al. [9] to classify chest X-ray images into pneumonia and other 14 diseases. ChestX-ray14 dataset was used for training the model. They compared their ChXNet model (121 layered model) with practicing academic radiologists. Their ChXNet model achieved an F1 score (95% CI) of 0.435 outperforming radiologists which achieved an F1 score (95% CI) of 0.387. A deep convolutional neural network model having five convolutional layers some followed by max-pooling layers, having three fully connected layers was trained by Krizhevsky et al. [10]. This network contained 60 million different parameters. By employing dropout, this model achieved a top-five error percent of 17%. Simonyan et al. [11] developed a highly accurate model employing multiple small kernel-sized filters to achieve top-five test accuracy 92.7%. This model was trained on the ImageNet dataset and submitted to the ILSVRC 2014 competition. A convolution neural network for classification and segmentation of brain tumor MRIs was developed by Xu et al. [12]. Multiple techniques such as data augmentation, feature selection and pooling techniques were employed in this model. The validation accuracy for classification achieved by this model is 87.5%, and validation accuracy of segmentation is 84%, 256 × 256 pixels sized frontal chest radiographs which were fed to a deep convolution neural network to detect abnormalities. A convolutional neural network with five convolution layers employing leaky ReLU, average pooling and three fully connected layers was developed by Anthimopoulos et al. [13] to detect interstitial lung disease patterns in a dataset containing 14,696 images belonging to seven different classes. This model achieved a classification accuracy of 85.5%. He et al. [14] developed a residual neural network (RNN) to classify images present in the ImageNet dataset. RNN introduced the concept of shortcut connections to tackle the problem of vanishing gradients. This model which was submitted to ILSVRC 2015 attained state-of-the-art classification accuracy. A transfer learning model, extension of AlexNet using data augmentation techniques, was developed by Glozman et al. [15]. This model was trained on ADNI database. Two neural network models were presented by Hemanth et al. [16] which are MCPN and MKNN. These models classified MRIs with high accuracies and tackled high convergence time for Artificial Neural Networks.

## III. METHODS

CNN models have been created from scratch and trained on Chest X-Ray Images (Pneumonia) dataset on Kaggle. Keras neural network library with TensorFlow backend has been used to implement the models. Numpy and Pandas is used to manipulate the data Matplotlib.pyplot and seaborn is used to produce plots for visualization, util will provide the locally defined utility functions. Also, several modules from the Keras framework are being used for building deep learning models. The training and testing of the model were performed on GoogleColab. Dataset consists of 5216 training images, 624 testing images and 16 validation images. Data augmentation has been applied to achieve better results from the dataset. The self-built convolution model is trained on varying Epoch and learning rate. The four pretrained models have been trained on the training dataset, each with different number of convolutional layers. Each pretrained model was trained for 10 epochs and 100 steps per epochs. The following sub-headings further explain the above stages in depth.

### A. *The Data Set*

The dataset is organized into 3 folders (train, test, val) and contains subfolders for each image category (Pneumonia/Normal). There are 5,863 X-Ray images (JPEG) and 2 categories (Pneumonia/Normal) Fig.3. Chest X-ray images (anterior- posterior) were selected from retrospective cohorts of pediatric patients of one to five years old from Guangzhou Women and Children's Medical Center, Guangzhou. All chest X-ray imaging was performed as part of patients' routine clinical care. For the analysis of chest x-ray images, all chest radiographs were initially screened for quality control by removing all low quality or unreadable scans. The diagnoses for the images were then graded by two expert physicians before being cleared for training the AI system. In order to account for any grading errors, the evaluation set was also checked by a third expert. Each of these three folders contains two subfolders containing images diagnosed as pneumonia or normal. The subfolder names represent the data labels. The images are high-quality and come in various sizes, but they were subsequently resized for training the model. Because there were more images of X-rays labeled as "Pneumonia" in the training dataset, in order to increase the number of training examples of "Normal" images, the data augmentation was used. A couple of examples of X-ray images, with their corresponding labels are shown in Fig. 2.

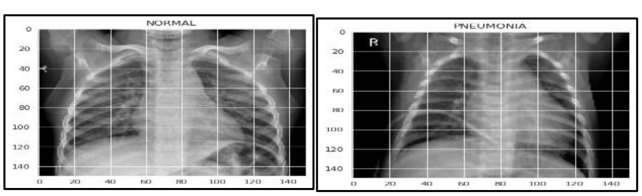



Fig.2

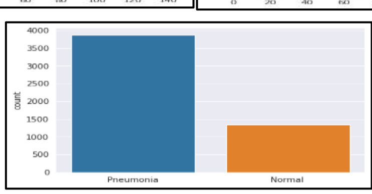



Fig.3

### B. *Pre-processing of Images*

Dataset Preparation Most of the time, the first step before building a model is the preprocessing of the imported data. The original images are RGB, but for the purpose of this experiment, they were imported as grayscale and resized. After that, the pixel intensity values were normalized by dividing the pixel values with 255. This way the pixels in the image are represented with floating point numbers between 0 and 1, rather than integer numbers in the range 0 to 255. This should positively affect the performance of CNN [17]. As stated before, the data augmentation has been performed, because of the disbalance in the number of training examples of images showing pneumonia versus those that were normal. This was done because the degree of model overfitting is determined by both its power and the amount of training it receives, providing a CNN with more training examples of images diagnosed as normal, in order to reduce overfitting. As there is no more available data of certain type, the data is artificially created by randomly rotating some training images by 30 degrees, randomly zooming by 20% some training images and randomly shifting the images horizontally by 10% of the width, also randomly shifting images vertically by 10% of the height, and randomly flipping the images horizontally. All of this can be performed with Keras preprocessing tools as a tensor flow backend .

### C. *CNN Architecture*

CNN models are feed-forward networks with convolutional layers, pooling layers, flattening layers and fully connected layers employing suitable activation functions. Convolutional layer. It is the building block of the CNNs. Convolution operation is done in mathematics to merge two functions. In the CNN models, the input image is first converted into matrix form. Convolution filter is applied to the input matrix which slides over it, performing element-wise multiplication and storing the sum. This creates a feature map. $3 \times 3$ filter is employed to create 2D feature maps when images are black and white. Convolutions are performed in 3D when the input image is represented as a 3D matrix where the RGB color represents the third dimension. Several feature detectors are operated with the input matrix to generate a layer of feature maps which thus forms the convolutional layer

**Activation functions.** All the models presented in this paper use two different activation functions, namely ReLU activation function and sigmoid activation function. The ReLU activation function stands for rectified linear function. It is a nonlinear function that outputs zero when the input is negative and outputs one when the input is positive. The ReLU function is given by the following formula: This type of activation function is broadly used in CNNs as it deals with the problem of vanishing gradients and is useful for increasing the nonlinearity of layers. ReLU activation function has many variants such as Noisy ReLUs, Leaky ReLUs and Parametric ReLUs. Advantages of ReLU over other activation functions are computational simplicity and representational sparsity. Sigmoid activation function is used in all the models presented in this paper. This broadly used activation function is employed in the last dense layer of all the models. In this activation function the input to the function is transformed into a value between 0.0 and 1.0. binary cross-entropy loss function is mostly used with this type of activation function.

**Pooling layer:** Convolutional layers are followed by pooling layers. The type of pooling layer used in the self-built CNN model is max-pooling layers. The max-pooling layer having a dimension of 2 × 2 selects the maximum pixel intensity values from the window of the image currently covered by the kernel. Max-pooling is used to down sample images, hence reducing the dimensionality and complexity of the image.

**Flattening layer and fully connected layers:** After the input image passes through the convolutional layer and the pooling layer, it is fed into the flattening layer. This layer flattens out the input image into a column, further reducing its computational complexity. This is then fed into the fully connected layer/dense layer. The fully connected layer has multiple layers, and every node in the first layer is connected to every node in the second layer. Each layer in the fully connected layer extracts features, and on this basis, the network makes a prediction. This process is known as forward propagation. After forward propagation, a cost function is calculated. It is a measure of performance of a neural network model. The loss function used in all the models is binary cross-entropy. After the loss function is calculated, back propagation takes place. This process is repeated until the network achieves optimum performance. Different optimization algorithm such as Adam, RMS Prop etc. has been used in the models.

**Reducing overfitting:** Dropout technique helps to reduce overfitting and tackles the problem of vanishing gradients. Dropout technique encourages each neuron to form its own individual representation of the input data. This technique on a random basis cuts connections between neurons in successive layers during the training process. Data augmentation technique has also been employed to reduce overfitting. As there is no more available data of certain type, the data is artificially created by randomly rotating some training images by 30 degrees, randomly zooming by 20% some training images and randomly shifting the images horizontally by 10% of the width, also randomly shifting images vertically by 10% of the height, and randomly flipping the images horizontally. Learning rate of models was also modified, to reduce overfitting.

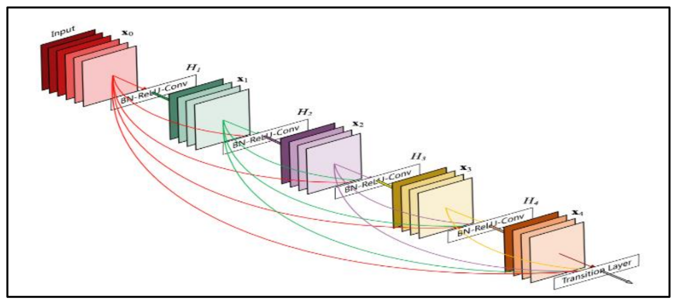
### D. *Transfer Learning and Pretrained Models*

Transfer Learning is a machine learning method where we reuse a pre-trained model as the starting point for a model on a new task. To put it simply—a model trained on one task is repurposed on a second, related task as an optimization that allows rapid progress when modeling the second task. By applying transfer learning to a new task, one can achieve significantly higher performance than training with only a small amount of data. Transfer learning is so common that it is rare to train a model for an image or natural language processing-related tasks from scratch.

Instead, researchers and data scientists prefer to start from a pre-trained model that already knows how to classify objects and has learned general features like edges, shapes in images. ImageNet, AlexNet, and Inception, VGG16, ResNet50, DENSEnet121 are typical examples of models that have the basis of Transfer learning. Following pretrained models are being used in this paper.

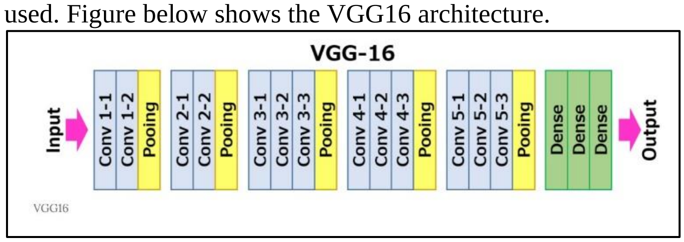
*i. DenseNet121*

DenseNet is an architecture that focuses on making the deep learning networks go even deeper, but at the same time making them more efficient to train, by using shorter connections between the layers. DenseNet is a convolutional neural network where each layer is connected to all other layers that are deeper in the network, that is, the first layer is connected to the 2nd, 3rd, 4th and so on, the second layer is connected to the 3rd, 4th, 5th and so on. This is done to enable maximum information flow between the layers of the network. Below figure shows the architecture of DenseNet-121. It has 120 Convolutions and 4 AvgPool. All layers i.e., those within the same dense block and transition layers, spread their weights over multiple inputs which allows deeper layers to use features extracted early on. In this paper we use the Adam as an optimizer and binary-cross entropy as a loss function for this model.



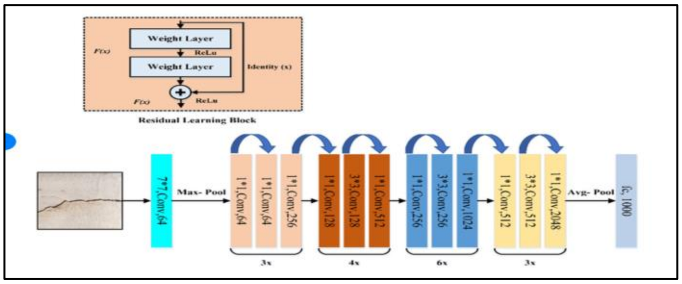

*ii. VGG16*

VGG16 is a convolutional neural network model proposed by K. Simonyan and A. Zisserman from the University of Oxford in the paper "Very Deep Convolutional Networks for Large-Scale Image Recognition". The model achieves 92.7% top-5 test accuracy in ImageNet, which is a dataset of over 14 million images belonging to 1000 classes. VGG16 was trained for weeks and was using NVIDIA Titan Black GPU's. VGG16 has a very simple and classical architecture, with blocks of 2 or 3 convolutional layers followed by a pooling layer, plus a final dense network composed of 2 hidden layers (of 4096 nodes each) and one output layer (of 1000 nodes). Only 3x3 filters are used. Figure below shows the VGG16 architecture.



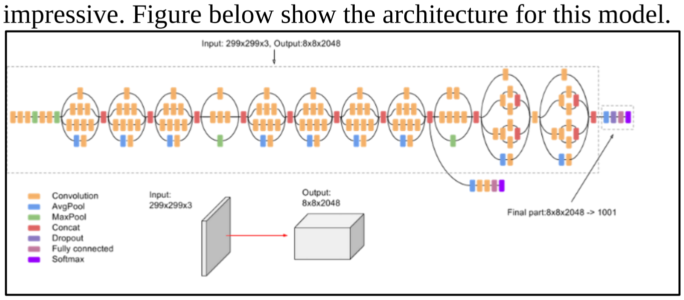

*iii. ResNet50*

ResNet50 is a variant of ResNet model which has 48 Convolution layers along with 1 MaxPool and 1 Average Pool layer. It has 3.8 x 10^9 Floating points operations. The Result were pretty good on the ImageNet validation set, The ResNet 50 model achieved a top-1 error rate of 20.47 percent and and achieved a top-5 error rate of 5.25 percent, this is reported for single model that consists of 50 layers not an ensemble of it. below is the table given if you want to compare it with other ResNet's or with other models. Figure shows ResNet50's architecture.




*iii. InceptionV3*

Inception v3 was trained on ImageNet and compared with other contemporary modelsInception-v3 is a convolutional neural network architecture from the Inception family that makes several improvements including using Label Smoothing, factorized 7 x 7 convolutions, and the use of an auxiliary classifier to propagate label information lower down the network (along with the use of batch normalization for layers in the side head). The major modifications done on the Inception V3 model are Factorization into Smaller Convolutions Spatial Factorization into Asymmetric Convolutions Utility of Auxiliary Classifiers Efficient Grid Size Reduction. In total, the inception V3 model is made up of 42 layers which is a bit higher than the previous inception V1 and V2 models. But the efficiency of this model is really impressive. Figure below show the architecture for this model.




***E. Algorithm of CNN classifier***

Fig.4 shows the flowchart of the overall schema of research. The number of epochs for all the classifier models presented in this paper was alter between range 6 to 48 and after training and testing several CNN models over the course of research, the best parameters are obtained at epochs 12. Classifier models trained for a greater number of epochs than 16 have showed overfitting. Several optimizer functions were also trained and studied. RMSprop optimizer function was finalized to be used for the self-built CNN classifier after it gave the best results. A simple classifier model shown in the below figures has been built. Also, I used ReduceLROnPlateau class to reduce learning rate when a metric has stopped improving. Models often benefit from reducing the learning rate by a factor of 2-10 once learning stagnates. This callback monitors a quantity and if no improvement is seen for a 'patience' number of epochs, the learning rate is reduced. This help in reduce the overfitting. The results have been summarized in the subsequent section of this paper.

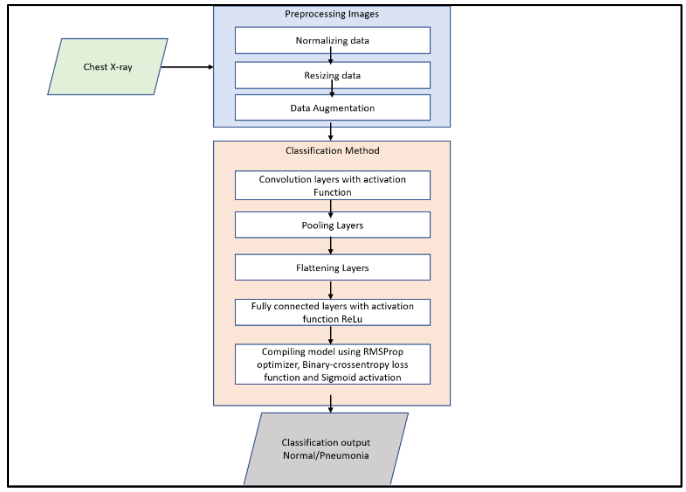
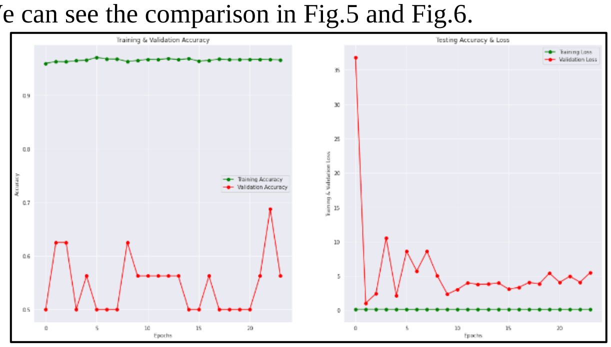


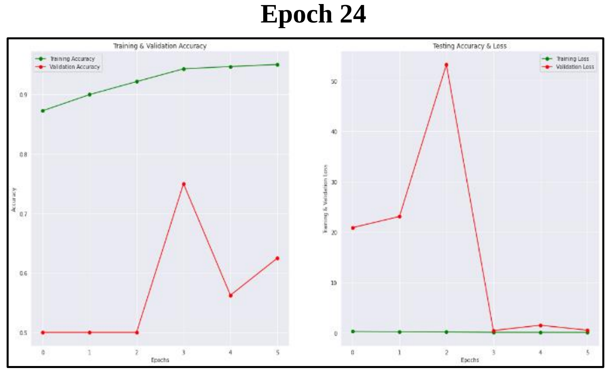

Fig.4.

## IV. Experimental Results and Discussion

Evaluation of the results can be done by analyzing the various parameters such as training accuracy and loss, validation accuracy and loss and of course model accuracy for self-built CNN model. The batchsize and the learning rates are 32 and 0.000001, respectively. I tried varying the epochs range form 6-48 to obtain the better results. It has been observed that once the epochs range start exceeding beyond the epoch 24 the model starts overfitting with training accuracy 97% and validation accuracy 56%. Also, in case of epoch 6 we get very less accuracy compared to the number of epochs equal to 12 and 16. We can see the comparison in Fig.5 and Fig.6.




**Epoch 24**




**Epoch 06**

The final graphs for Epochs 12 and learning rate 0.000001 are shown below. The progress of the training and validation accuracy and loss during time is shown in the Fig.7. They show that over time, both the training and validation accuracy were getting better, especially after the 8th epoch. Also, as the number of epochs reaches near 12 the validation loss comes to a good level of near 0. The loss is a way to analyze how well the model is doing on both the training and validation set. It calculated how well the model is doing on each example in these sets and calculates the sum of errors made on them.

One additional way of representing the results of the model is to build a confusion matrix. The Y-axis of the confusion matrix holds the predicted values, while the X-axis holds the true values. The confusion matrix for our latest experiment is illustrated in Fig. 8. With the trained model, 373 out of 402 were accurately predicted as images of X-rays with pneumonia, while 205 out of 222 were accurately predicted as X-rays without pneumonia. This gives us a model accuracy of 92.64%.

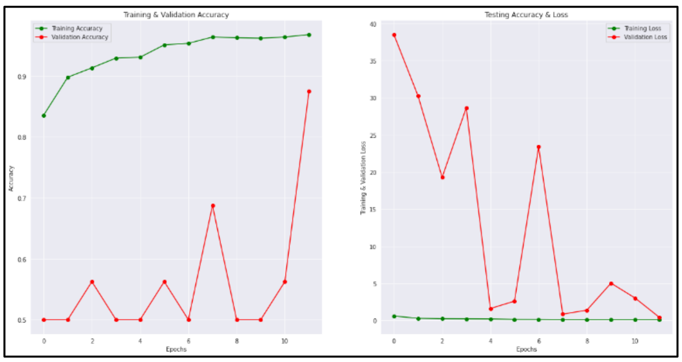



Fig.7.

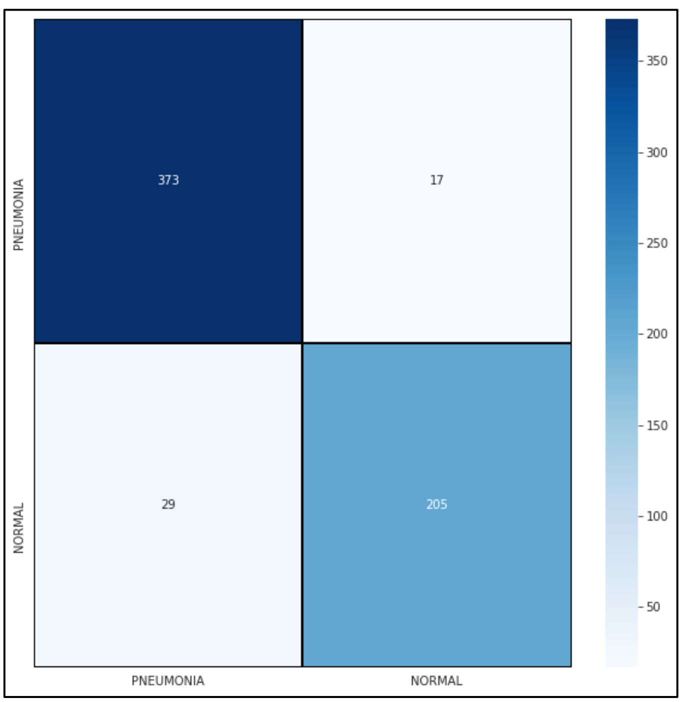
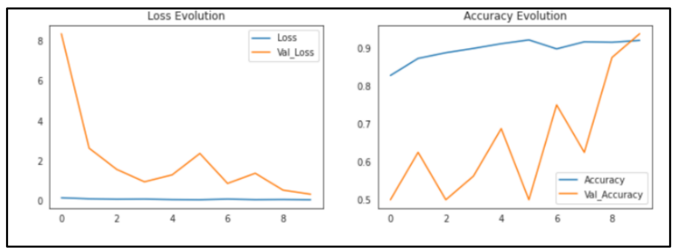
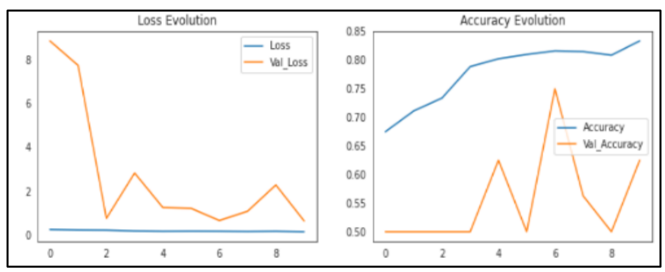



Fig.8.

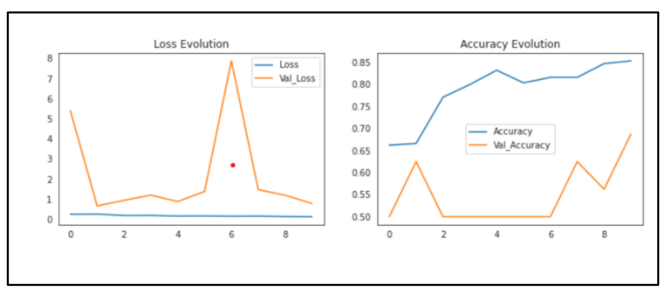
To study the performance of each classifier model, validation accuracy, recall and F1 score were evaluated as the performance measures. Accuracy and loss graphs were also studied. The confusion matrix is also computed for each model. Fig.9. show the confusion matrices, accuracy graphs and loss graphs of all CNN classifier models. For all the transfer learned pretrained model the No.of epochs, Steps per epochs and the validation steps are set as 10,100 and 25, respectively. Also, the dropout rates for ResNet50, VGG16, and InceptionV3 are set as 0.3,0.4 and 0.6. Table 1 and Fig.9 show that classifier model ResNet50 and VGG16 significantly underperformed compared to the self-built model and DenseNet121(428 layers). The accuracy graphs and loss graphs show overfitting. Accuracy, recall and F1 scores are also low. In case of self- built CNN, in addition to extra convolution layer, employing dropout and lowering the learning rate of optimizer improved the performance considerably. It achieved the least overfitting along with highest accuracy and recall. Several attempts were made to better the performance by adding more convolutional layers and changing the parameters. Also, I used ReduceLROnPlateau class to reduce learning rate when a metric has stopped improving. Models often benefit from reducing the learning rate by a factor of 2-10 once learning stagnates. This callback monitors a quantity and if no improvement is seen for a 'patience' number of epochs, the learning rate is reduced.



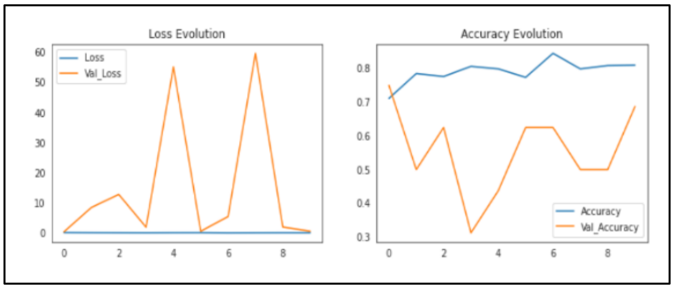

**DenseNet121**




**VGG16**




**ResNet50**




**InceptionV3**

Table 1 Performance comparison of different CNN models

| Classifier Model | Validation Accuracy (%) | Validation Loss (%) | Recall (%) | F1 Score (%) |
|---|---|---|---|---|
| Self-Built CNN | 92 | 8 | 94 | 93 |
| DenseNet121 | 88 | 12 | 91 | 88 |
| VGG16 | 85 | 15 | 87 | 84 |
| ResNet50 | 83 | 17 | 85 | 83 |
| Inception V3 | 87 | 13 | 90 | 89 |

## V. Conclusion

This paper describes the use of deep learning in order to classify digital images of chest X-rays according to presence or absence of changes consistent with pneumonia. The implementation was based on CNN model using Python programming and scientific tools. Also, this paper shows a comparison between results obtained from self-built CNN model and Transfer learn pretrained models. Pretrained models used for this paper are: DenseNet121, VGG16, ResNet50 and InceptionV3.

The validation accuracy, recall and F1 score of CNN classifier self -built model are 92%, 94% and 93%, respectively, which are quite high compared to other transfer learned pretrained models. DenseNet121 classifier model also comes very close in performance with 88% validation accuracy, 91% recall and 93% F1 score. For this dataset, self-built CNN outperformed the transfer learned pretrained models DenseNet121, VGG16, ResNet50 and Inception V3. The paper by Chakraborty [18] achieved the overall accuracy of 93.62% and recall of 91% trained on the same dataset. The paper by Liang [19] achieved recall of 93.7% on the same dataset. The models presented by me at best could achieve 92.64% accuracy which is lower, but 94% recall has been achieved. Recall is often favored in medical imaging cases over other performance evaluating parameters, as it gives a measure of false negatives in the results. High recall values will ensure that the number of false-negative instances is lower, hence lowers the risk to the patient's life. Thus, it is concluded that self-built CNN classifier model can, therefore, be effectively used by medical officers for diagnostic purposes for early detection of pneumonia in children as well as adults. A large number of X-ray images can be processed very quickly to provide highly precise diagnostic results, thus helping healthcare systems provide efficient patient care services and reduce mortality rates. This convolutional neural network model was successfully achieved by employing various methods of parameter tuning like adding dropout, changing learning rates, changing the batch size, number of epochs, adding more complex fully connected layers and changing various stochastic gradient optimizers. Further research steps will include experimenting with various preprocessing and CNN configurations, data augmentation techniques, as well as using additional X-ray datasets with additional data labels showing other pathologies. It is also expected that neural network models based on GAN [20], generative adversarial networks, would also be trained and compared with the existing models.

### Acknowledgment

I would like to thank Professor Bruce Maxwell for constant guidance and support and for providing me with the opportunity to work on this project as the final project for CVPR course.

### References

[1] Albahli S., Rauf H., Algosaibi A. & Balas V. AI-driven deep CNN approach for multi-label pathology classification using chest X-Rays. *PeerJ Computer Science*. 7 pp. e495 (2021) pmid:33977135

[2] Albahli S., Rauf H., Arif M., Nafis M. & Algosaibi A. Identification of thoracic diseases by exploiting deep neural networks. *Neural Networks*. 5 pp. 6 (2021)

[3] Chandra T. & Verma K. Pneumonia detection on chest X-Ray using machine learning paradigm. *Proceedings Of 3rd International Conference On Computer Vision And Image Processing*. pp. 21-33 (2020)

[4] Kuo K., Talley P., Huang C. & Cheng L. Predicting hospital-acquired pneumonia among schizophrenic patients: a machine learning approach. *BMC Medical Informatics And Decision Making*. 19, 1–8 (2019) pmid:30866913

[5] Yue H., Yu Q., Liu C., Huang Y., Jiang Z., Shao C., et al. & Others Machine learning-based CT radiomics method for predicting hospital stay in patients with pneumonia associated with SARS-CoV-2 infection: a multicenter study. *Annals Of Translational Medicine*. 8 (2020) pmid:32793703

[6] Rubin, J., Sanghavi, D., Zhao, C., Lee, K., Qadir, A., Xu-Wilson, M.: Large Scale Automated Reading of Frontal and Lateral Chest X-Rays Using Dual Convolutional Neural Networks (2018). arXiv preprint arXiv:1804.07839

[7] Lakhani, P., Sundaram, B.: Deep learning at chest radiography: automated classification of pulmonary tuberculosis by using convolutional neural networks. Radiology **284**(2), 574–582(2017)

[8] Guan, Q., Huang, Y., Zhong, Z., Zheng, Z., Zheng, L., Yang, Y.: Diagnose Like a Radiologist: Attention Guided Convolutional Neural Network for Thorax Disease Classification (2018). *arXiv preprint* arXiv:1801.09927

[9] Rajpurkar, P., Irvin, J., Zhu, K., Yang, B., Mehta, H., Duan, T., Ding, D., Bagul, A., Langlotz, C., Shpanskaya, K., Lungren, M.P.: Chexnet: Radiologist-Level Pneumonia Detection on Chest X-rays with Deep Learning (2017). arXiv preprint arXiv:1711.05225

[10] Krizhevsky, A., Sutskever, I., Hinton, G.E:. Imagenet classification with deep convolutional neural networks. In: Advances in Neural Information Processing Systems, pp. 1097–1105 (2012)

[11] Simonyan, K., Zisserman, A.: Very Deep Convolutional Networks for Large-Scale Image Recognition (2014). arXiv preprint arXiv:1409.1556

[12] Xu, Y., Jia, Z., Ai, Y., Zhang, F., Lai, M., Eric, I., Chang, C.: Deep convolutional activation features for large scale brain tumor histopathology image classification and segmentation. In: 2015 international conference on acoustics, speech and signal processing (ICASSP), pp. 947– 951 (2015)

[13] Anthimopoulos, M., Christodoulidis, S., Ebner, L., Christe, A., Mougiakakou, S.: Lung pattern classification for interstitial lung diseases using a deep convolutional neural network. IEEE Trans. Med. Imaging **35**(5), 1207–1216 (2016)

[14] He, K., Zhang, X., Ren, S., Sun, J.: Deep residual learning for image recognition. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 770–778 (2016)

[15] Glozman, T., Liba, O.: Hidden Cues: Deep Learning for Alzheimer's Disease Classification CS331B project final report (2016)

[16] Hemanth, D.J., Vijila, C.K.S., Selvakumar, A.I., Anitha, J.: Performance improved iteration-free artificial neural networks for abnormal magnetic resonance brain image classification. Neurocomputing **130**, 98–107 (2014)

[17] Chakraborty, S., Aich, S., Sim, J.S., Kim, H.C.: Detection of pneumonia from chest x-rays using a convolutional neural network architecture. In: International Conference on Future Information & Communication Engineering, vol. 11, no. 1, pp. 98–102 (2019)

[18] Liang, G., Zheng, L.: A transfer learning method with deep residual network for pediatric pneumonia diagnosis. In: Computer Methods and Programs in Biomedicine (2019)

[19] Radford, A., Metz, L., Chintala, S.: Unsupervised Representation Learning with Deep Convolutional Generative Adversarial Networks (2015). arXiv preprint arXiv:1511.06434

[20] https://journals.plos.org/plosone/article?id=10.1371/journal.pone.02566 30

[21] https://www.researchgate.net/publication/349408817_Pneumonia_Detec tion_Using_Deep_Learning_Based_on_Convolutional_Neural_Network

[22] Zhang, D.; Ren, F.; Li, Y.; Na, L.; Ma, Y. Pneumonia Detection from Chest X-ray Images Based on Convolutional Neural Network. Electronics 2021, 10, 1512. https:// doi.org/10.3390/electronics10131512

[23] KaggleDataset: https://www.kaggle.com/datasets/paultimothymooney/chest-xray-pneumonia?datasetId=17810&sortBy=voteCount