Deep Learning Based Pneumonia Detection

Kaho Fardin Hasib
ID: 170204063
170204063@aust.edu

Ramisa Maliyat
ID: 180104079
180104079@aust.edu

Fatema Sultana
ID: 180104085
180104085@aust.edu

Sadia Sultana
ID: 180104094
180104094@aust.edu

***Abstract*—The application of machine learning and artificial intelligence in medicine is expanding. his is especially true in the medical field, which makes extensive use of biometric imaging. and where the gathering and processing of a lot of digital images is necessary for the diagnostic processes. The analysis of medical images using machine learning increases the farmness and validity of reporting. This study reviews the use of machine learning algorithms to analyze chest radiographs to improve decision-making and reach the correct diagnosis. This work mainly focuses on developing processing models using deep learning techniques such as convolutional neural networks, VGG16, ResNet-50, and InceptionNet. These models aim to help with the classification problem of determining whether there are pneumonia-related abnormalities in the chest x-ray and classifying this x-rays into two categories according to the detection results.**

***Index Terms*—VGG16, ResNet-50, InceptionNet, deep learning, convolutional neural networks, image preprocessing, and pneumonia.**

I. Introduction

A pulmonary infection called pneumonia can inflame the tiny air sacs in the lungs. Severe breathing difficulties and congestion, fever, chest pain, colds, or exhaustion are all signs of this illness. Infections like COVID-19 or the flu, as well as common colds and those caused by bacteria, fungi, and other organisms, can cause pneumonia. In developing nations, it is the main factor in child mortality. Lung X-rays are highly helpful in identifying diseases and abnormalities of the respiratory system in youngsters. For a full recovery from pneumonia, early diagnosis is crucial. The most prevalent method of diagnosis is the examination of X-ray scans, although this method depends on the radiologist's skill to interpret the images and is frequently not agreed upon by other radiologists. Therefore, a generalizing automatic method is needed to identify the illness. By identifying pneumonia in its early stages, this initiative can help individuals treat it. For a good recovery from pneumonia, early detection is essential. Discovering pneumonia in people infected with the 2019-nCoV coronavirus prompted further research on this subject. Determining whether changes in the chest X-ray are consistent with pneumonia and classifying the X-rays into two groups, based on the findings, helps classify the problem. The Guangzhou Women and Children's Medical Center contributed to this study's dataset, which is freely accessible on Kaggle. It contains 5947 photographs and classifies subjects as normal or with pneumonia. In this dataset, we used CNN, ResNet-50, VGG16, and InceptionNet as classification models. Spatial dropout is used to avoid overfitting.

II. Related Works

A. Pneumonia Detection Using Deep Learning Based on Convolutional Neural Network[1]

Deep learning and convolutional neural networks, one of the most recent methods to expedite the diagnosis of pneumonia, were reported in this research. They utilized a dataset made accessible on Kaggle by the Guangzhou Women and Children's Medical Center. Using CNN, they correctly identified X-rays with pneumonia in 334 out of 381 images, and X-rays without pneumonia in 187 out of 205 photos. They have an accuracy rate of 88.90

B. Pneumonia detection in chest X-ray images using compound scaled deep learning model[2]

In this paper, they have shown scaled ResNet50, which is the modified version of ResNet50. They used the same dataset that we have used. They obtained optimum results with 0.001 in the learning rate with SGD as the optimizer. They obtained a testing accuracy of 97.85% and a testing loss of 0.070.

C. Transfer Learning with Deep Convolutional Neural Network (CNN) for Pneumonia Detection using Chest X-ray[3]

They have provided a deep CNN-based transfer learning method for automatic pneumonia and its class recognition in this work. They have employed four different deep convolutional neural networks (CNNs) that have already been trained: Alex Net, ResNet18, DenseNet201, and Squeeze Net. They have made use of the pneumonia from chest X-rays dataset from Kaggle. For DenseNet20 training and testing, they have attained the best accuracy of 98 percent.

D. Pneumonia Detection from Chest X-ray Images Based on Convolutional Neural Network[4]

In this paper, they have presented a CNN-based model VGG16 model architecture with fewer layers to diagnose pneumonia on a chest X-ray image set. This model has contained only six layers combining the ReLU activation function, drop operation, and max-pooling layers. They have used Chest X-Ray images for the Kaggle competition to evaluate the model with a total of 5786 X-ray images. They have archived 96.07% accuracy with the VGG16 model.

III. Dataset

We have chosen a dataset which is provided by Guangzhou Women and Children's Medical Center, Guangzhou for this project which is available in Kaggle[1] .This dataset consists of 5856 chest X-ray images in JPEG format that are organized into three folders: train, test, and val. Each of these files has two subfolders that each include pictures of pneumonia affected or healthy people.

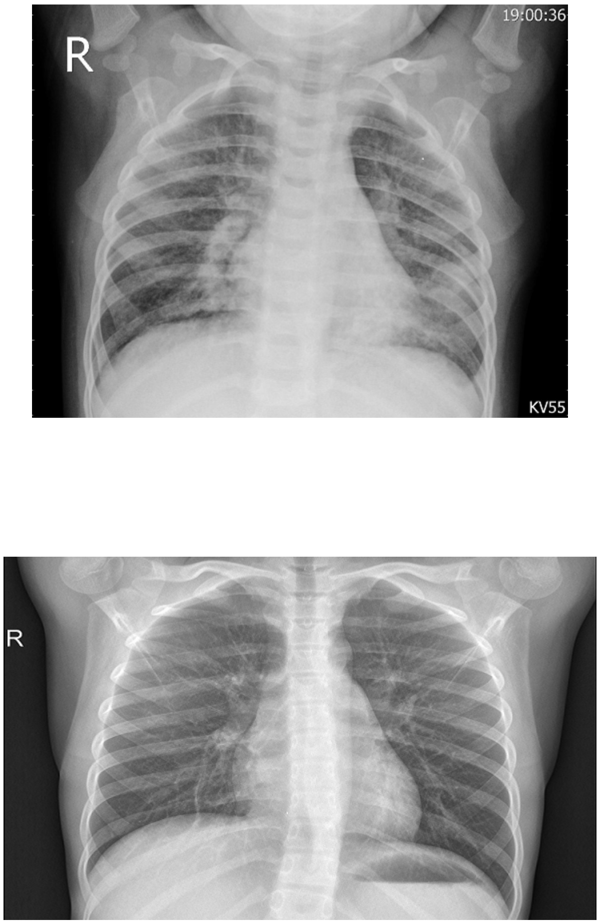

Fig. 1. An image of pneumonia affected chest(A), and a normal chest X-ray(B) from the dataset.

Before being eliminated from the analysis of the chest x-ray images, each chest radiograph was first examined for quality control. The photographs were graded by two qualified doctors prior to the diagnosis being utilized to train the AI system. As a result, the images are of a high caliber and come in various sizes. Table I displays the distribution of the images among the folders in our dataset.

IV. Background Study

A. Convolutional Neural Network (CNN)

Convolutional neural networks are deep learning models used to analyse input with a grid pattern, such as photographs

[1]https://www.kaggle.com/datasets/paultimothymooney/chest-xray-pneumonia!

TABLE I
DATASET DETAIL

	Train	Test	Validation
Normal	1299	234	50
Pneumonia	3833	390	50

(CNN). The top layer of a convolutional network is the convolutional layer. The fully connected layer is not followed by additional convolutional layers, other convolutional layers, or pooling layersWith proper filters, convolutional neural networks can accurately capture the spatial and temporal dependencies of images. As the visual data passes through the CNN layers, the larger features or shapes of the object are recognized first, and the target object last.Figure 2 shows CNN architecture.

B. Residual Networks(ResNet-50)

A traditional neural network, known as ResNet, stands for Residual Networks. In extremely deep CNNs, this model aids in solving the vanishing gradient problem. Considering that ResNet skips several layers, it trains deep neural networks. The outputs from earlier layers are added to the outputs of stacked layers via the skip connections between layers. A convolutional neural network called ResNet-50 may operate with 50 neural network layers. With 48 Convolution layers, 1 MaxPool layer, and 1 Average Pool layer, this model is a variation of the ResNet model.

C. Visual Geometry Group-16(VGG16)

VGG-16 or Visual Geometry Group-16 is the name of a convolutional neural network with 16 layers. The architecture of VGG16 is very simple and conventional, consisting of a block of 2 or 3 convolutional layers, followed by a pooling layer, 2 hidden layers of 4096 nodes each and an output layer (of 1000 nodes). followed by the final high density network. The final dense layer is the output layer with softmax activated. According to this concept, the entire network should be covered by 3 3 small receive fields (filters) with a step size of 1 pixel. As a result, the network would have the ability to converge more quickly.

D. InceptionNet

An architecture called InceptionNet, sometimes known as GoogleNet, provides inception modules—sub-networks that enable rapid training computation. InceptionNet has nine horizontally arranged inception components. It has 22 layers (27, including the pooling layers). Inception's last module makes use of global average pooling. As with any extremely deep network, it is vulnerable to the vanishing gradient problem. To avoid the network "dying out" in the middle, the authors added two additional classifiers. The total loss

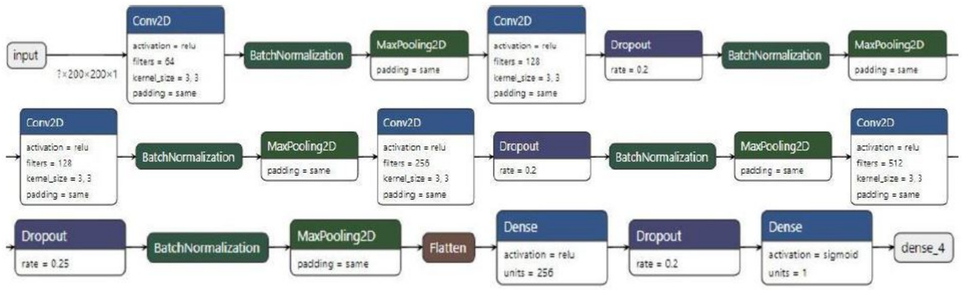

Fig. 2. Architecture of the CNN [1]

function is obtained by summing the auxiliary loss and the real loss.

Performance Matrices:As performance measurements, we have employed accuracy, loss, precision, recall, and f1 score. While fitting the models, validation split separates the data into training and validation data for each epoch. Typically, accuracy and loss rise with the number of epochs. Val accuracy starts to decline as Val loss increases (means model is cramming values not learning.) The ratio of True Positives to all the positives predicted by the model is known as precision. The precision decreases as the model predicts more false positives. Recall is the proportion of real positives to all the dataset's positives. The recall decreases as the model predicts more false negatives. The harmonic mean of recall and precision is known as the F1-score.

V. Methodology

A. *Image Pre-processing and Dataset Preparation*

The initial step before building a model is often preprocessing the input data. We initially altered photos to make them more suitable for training a convolutional neural network before training. We preprocessed and enhanced the data for this challenge using the Keras ImageDataGenerator function. Additionally, this class supports fundamental data enhancements such image flipping at random horizontal angles. In order to standardize the pixel intensity levels, we divided them by 255. Instead of using integer values from 0 to 255, the pixels in the image are now represented by floating point integers between 0 and 1. The models' performance ought to be enhanced as a result.

As previously mentioned, an imbalance in the proportion of training instances with and without images depicting pneumonia led to the data augmentation. When a certain form of data isn't anymore available,we can generate it artificially by rotating, cropping or zooming the existing images. All of this is feasible by utilizing the preprocessing tools in Keras.

B. *Selection of Models and Tools*

We have mainly used pandas,matplotlib, seaborn, numpy, keras etc. The models utilized in this experiment were the convolutional neural network (CNN), ResNet-50, VGG16, and InceptionNet. These models are a considerable breakthrough and are typically used for picture recognition and categorization. These models use x-ray images as input,after that bias and weight were added to different regions of the image, and then classifying the images into several groups. Here, we distinguish between two classes. These models consist of a few layers, such input and output, with hidden layers sandwiched in the between. These buried layers are where the majority of the calculation effort is done. Hidden layers may also contain pooling layers and completely connected layers. Each neuron in a convolutional layer is one of the most important model building blocks because it is only connected to some neurons (receptive fields) in the next convolutional layer. This form of structure permits the network to pay attention on a single low-level function in the first hidden layer and combine that function with other low-level functions to produce high-level functions in his second hidden layer. I can. Layer pooling is a technique for enhancing visual attributes while minimizing the size of the input image without significant loss of detail. The purpose of this is to reduce the amount of memory used and computational cost. This approach also reduces the possibility of overfitting by reducing the number of parameters.We have used both maximum pooling and average pooling in our work. In our experiments, the activation function of the ReLU was used. ReLU performs well in deep neural networks primarily because to its rapid computation and lack of saturation for positive values. Throughout the model construction phase, a dropout method was also used to enhance performance and prevent overfitting in neural networks. In the dropout method, certain neurons are randomly turned off so they are not used in that iteration. Simply adding a dropout can boost accuracy in the network by 1-2%. The neural network establishes which class the image belongs to using sigmoid activation.

Our full working process is depicted in Figure 3. We construct three distinct generators for train, test, and validation data after obtaining input images. We then performed image prepossessing. Changing the image's shape, standardizing the pixel values, etc. After that, four different binary classifiers received this image data. These models classified the images as normal or pneumonia. We then evaluate each model's accuracy, precision, recall, and f1 score. The outcomes of our models were then contrasted with those of the cited studies.

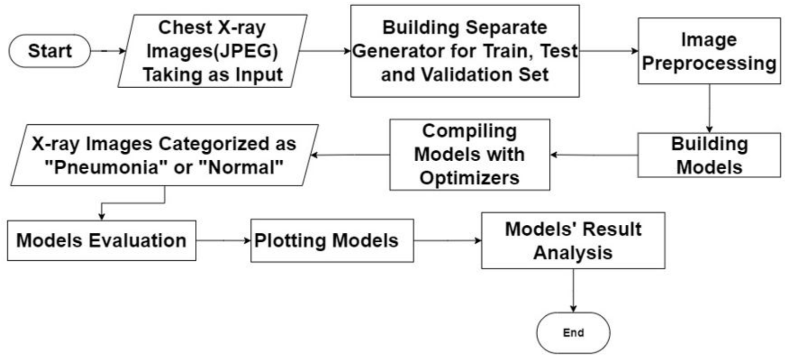

Fig. 3. Working Procedure

VI. Experimental Result

A. CNN

Our CNN or model uses 7 convolution and pooling layers along with spatial dropout to prevent overfitting, followed by 3 Dense and dropout layers to classify the images as having pneumonia or not having pneumonia.We got 92.31% accuracy after 25 epochs with a learning rate of 0.01.This model has 95.65% F1-score and loss is 15.09%.We have used adam optimizer here. In figure 4, the CNN accuracy curve is rising.

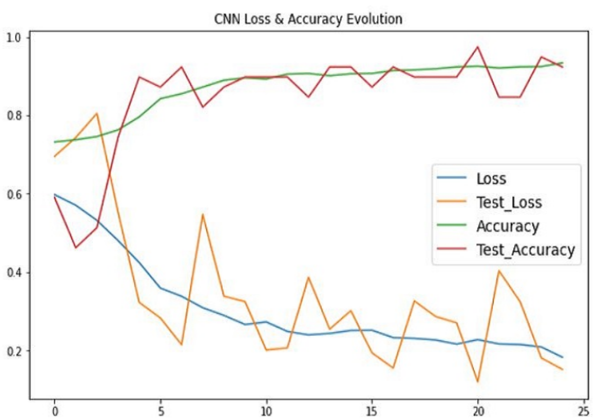

Fig. 4. Loss And Accuracy Evolution of CNN

It can be seen from the test loss curve how well the model fits the fresh data. At the end of the figure, the test and training loss is declining . This suggests that the model can continue to learn.

B. ResNet-50

Our ResNet-50 model uses 6 convolution layers followed by an activation layer. ResNet is a pre-trained model, which indicates that it was developed and trained by another person to address a related issue. Due to some of the instances' classification being arbitrary, the loss curve in figure 5 using RESNET is quite stable. The accuracy of this model is not

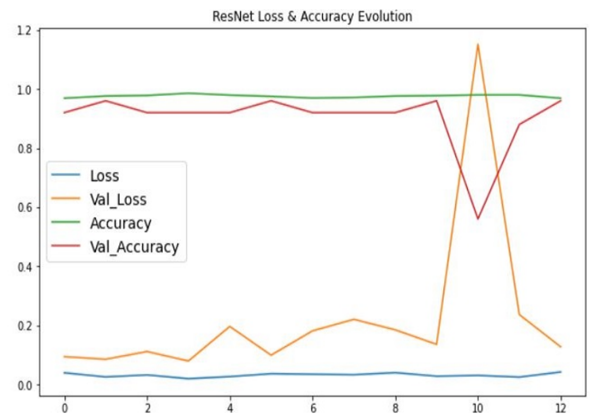

Fig. 5. Loss And Accuracy Evolution of ResNet-50

increasing as a result of loading pre-trained weight.In this model we got highest accuracy among four models and it is 96% after 13 epochs with a learning rate of 0.01.This model has 92.31% F1-score and the loss is 12.72%.We have reached to this accuracy by tuning the hyperparameters several times.When we trained that model with 0.001 learning rate the accuracy was below 50%.We have used stochastic gradient descent as optimizer here. With the aid of this model, the vanishing gradient issue can be solved.

C. VGG16

16 layers with weights are used in our VGG16 model. Thirteen convolutional layers, five Max Pooling layers, and three thick layers total 21 layers. The VGG16 model concentrated on having 3x3 convolution filter layers because it had a lot of hyper-parameters. The loading of pre-trained VGG-16 weights causes the accuracy curve, as shown in figure 6, to be flattened.In this model,we have used stochastic gradient descent as optimizer and learning rate was 0.01.After 15 epochs we got 92.00% accuracy and 100% F1-score.

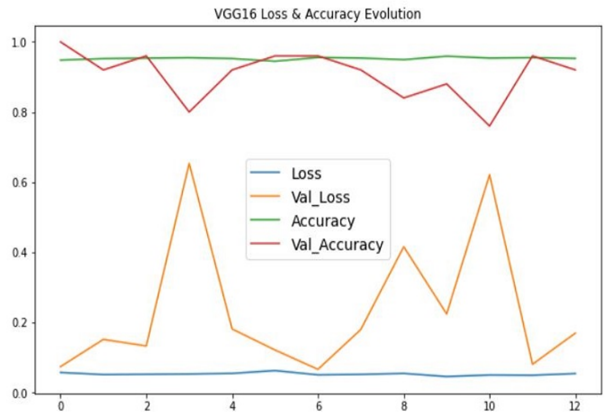

Fig. 6. Loss And Accuracy Evolution of VGG16

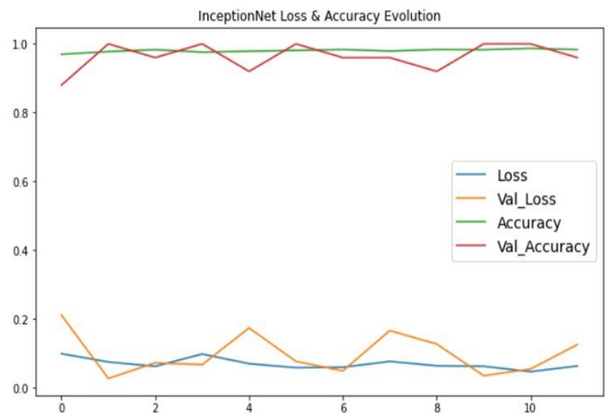

Fig. 7. Loss And Accuracy Evolution of Inception Net

D. Inception Net

Inception Net has 22 layers in total (27, including the pooling layers). The accuracy was improved and the computational complexity was decreased by the complete loss of the inception net during training. The progress of this model's loss and accuracy is shown in Figure 7.After 12 epochs with a learning rate of 0.01 we got 96.00% accuracy here and f1-score was 91.67%.

The findings of our four models are displayed in Table II. The highest level of accuracy is 96.00 percent for our ResNet-50 model and Inception Net models. The second-best score is held by CNN which is 92.31%.VGG-16 has quite similar accuracy as CNN.

TABLE II
PERFORMANCE METRICES OF MODELS

	Accuracy	Precision	Recall	F1 score	Loss
CNN	92.31%	91.67%	95.65%	95.65%	15.09%
ResNet-50	96.00%	100%	92.31%	92.31%	12.72%
VGG16	92.00%	83.33%	100%	100%	16.91%
Inception Net	96.00%	100%	91.67%	91.67%	12.55%

Figure 8 is the confusion matrix of our resnet-50 model.This confusion matrix shows that 99 negative examples and 243 positive examples are classified accurately. On the other hand, 147 actual positive examples are classified as negative and 135 actual negative examples are classified as positive.

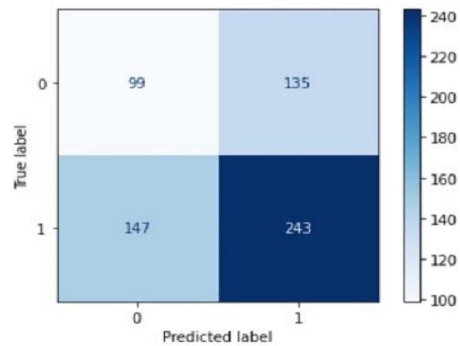

Fig. 8. Confusion Matrix of ResNet-50

VII. RESULT ANALYSIS

TABLE III
COMPARISON BETWEEN OUR MODELS TO THE CITED PAPER.

	Reference Paper	Our Model
CNN	88.90[1]	92.31
Resnet-50	97.85[2]	96.00
VGG-16	93.07[3]	92.00

The comparison between our three models and the reference publications is shown in Table III. The CNN model from Racic et al. (2021) has less accuracy than our CNN. For VGG16, we are virtually as accurate as Hashmi et al. (2021).Our Resnet model has an accuracy less than 1.85 of the reference paper.

VIII. CONCLUSION

In order to identify Pneumonia disease for this research, we employed Convolutional Neural Networks (CNN), ResNet-50, VGG16 (Visual Geometry Group From Oxford), and Inception Net. We avoided overfitting by using spatial dropout, which helps to improve the accuracy. This project gets more successful at identifying pneumonia and non-pneumonia cases by utilizing the dataset from the Guangzhou Women and Children's Medical Center, which is freely accessible on Kaggle. By identifying pneumonia at an early enough stage, individuals can help treat it.

REFERENCES

[1] Račić, L., Popović, T. and Šandi, S., 2021, February. Pneumonia detection using deep learning based on convolutional neural network. In 2021 25th International Conference on Information Technology (IT) (pp. 1-4). IEEE.

[2] Hashmi, M.F., Katiyar, S., Hashmi, A.W. and Keskar, A.G., 2021. Pneumonia detection in chest X-ray images using compound scaled deep learning model. Automatika: časopis za automatiku, mjerenje, elektroniku, računarstvo i komunikacije, 62(3-4), pp.397-406.
[3] Rahman, T., Chowdhury, M.E., Khandakar, A., Islam, K.R., Islam, K.F., Mahbub, Z.B., Kadir, M.A. and Kashem, S., 2020. Transfer learning with deep convolutional neural network (CNN) for pneumonia detection using chest X-ray. Applied Sciences, 10(9), p.3233.
[4] Zhang, D., Ren, F., Li, Y., Na, L. and Ma, Y., 2021. Pneumonia detection from chest X-ray images based on convolutional neural network. Electronics, 10(13), p.1512.