iJRASET
International Journal For Research in Applied Science and Engineering Technology

INTERNATIONAL JOURNAL FOR RESEARCH
IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 11 Issue: X Month of publication: October 2023

DOI: https://doi.org/10.22214/ijraset.2023.56233

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 11 Issue X Oct 2023- Available at www.ijraset.com

Detecting Pneumonia Using Deep Learning and Computer Vision

Grace Iordanov
Bellevue Digital Discovery

Abstract: Pneumonia is one of the most common causes of death of kids under the age of 5 and it currently stands as one of the top reason for hospitalization in the US for people over the age of 65, but with early identification and treatment the person who has the disease can survive. Multiple deep learning models for image classification including CNNs, RNNs can allow rapid identification of Pneumonia to begin treatment of the patient to increase the chance of survival. Pneumonia is usually detected through X rays. Formal detection of the disease through this method is very timely and can cause human errors. During this research I am going to use the VGG 19 model, the Resnet 50 model and Xception model to detect pneumonia from pictures of Lung x rays to see if it is possible to be able to detect accurately instances of pneumonia in x rays using convolutional neural networks. I will do analysis on the models and find which model is the most accurate for this classification task.
Keywords: Convolution, Depthwise, CNN, Overfit,

I. INTRODUCTION

Pneumonia causes the bronchioles in one or both of the lungs to swell and to fill up with fluid and not allow any air to pass through them. There are two types of pneumonia, the first is viral pneumonia which is caused by viruses such as COVID-19 and the common cold. The second type of pneumonia is called bacterial pneumonia and it is caused by a bacterial infection in the lungs by bacteria such as by streptococcus pnumonmaine or mycoplasma pnumonmaine. Bacterial pneumonia is more dangerous than viral pneumonia and is more likely to cause death.
Bacterial pneumonia can be prevented from getting worse with anti biotics anti fungal medications and viral pneumonia can be managed until it goes away on its own. It is important for doctors to be able to identify pneumonia as minimized error as possible to be able to give someone the treatment they need to prevent them from dying from this condition. The current methods used to diagnose pneumonia leave room for error and can require a lot of time that the patient with the symptoms cannot wait for. Using the output of deep learning technologies along with the review of a medical expert can minimize as much as possible the chance of a false negative and can save the patients life.

II. METHODS

I am going to be using the Kaggle Chest X Ray Images (Pneumonia) dataset which contains to 5,863 images identify true positive and true negative cases of pneumonia. The dataset is organized into folders "normal" for images that are benign for pneumonia and "pneumonia" for images that contain malignant instances of pneumonia in the images for the training and the test set. Each of the images in this dataset were scaled to the size of 224 by 224 pixels to be able to be fit into the VGG-19 model and were given a zoom range of 0.5 to reduce noise.

Here are some examples of the images

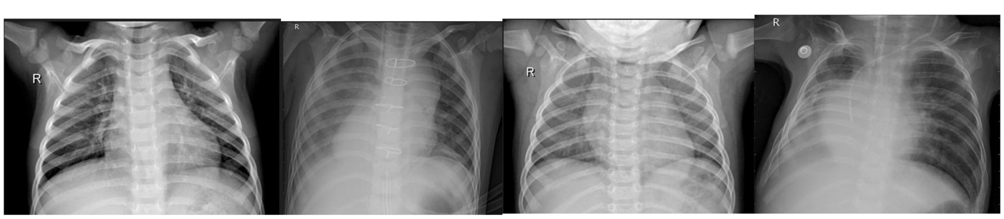


Fig I. Images before standardized to target size

A. Convolutional Neural Networks

The VGG-19, Resnet 50 model and Xception model are all models based on convolutional neural networks. Convolutional neural networks are neural networks include a convolutional layers, pooling layers and a flattening layer in addition to dense layers. In the convolutional layer of the neural network the images are separated into different fields and depending on the color channels in these fields each part of the field with a specific value for a color channel will be given a certain value such as a 1 for a part of a field with a white color or a zero for a part of the field with a black color channel.

This will allow a small portions of the entire image to be represented by numerical values, after the convolution in the max pooling layers, the max value of groups of the values will be placed into an even smaller field of numbers. The flattening layer is the final layer of the network before the dense layers. In the case of the models in this paper I am going to use 1 dense layer after the flattening layer with a softmax activation function.

B. VGG-19

I will begin by using the VGG-19 model on the dataset. The VGG-19 model is a purely convolutional network model with 19 hidden layers. To use the model in the predetermined state I will leave the trainable parameter set to false for the model and I will process the output of the model into a flatten layer which will be the input to a Dense layer. The model bellow shows how the input moves through the model and is convoluted at each step. This can be shown by the decreasing dimensions of input by a factor of two for each convolution. At the end the input is arranged into a matrix of 4096 units.

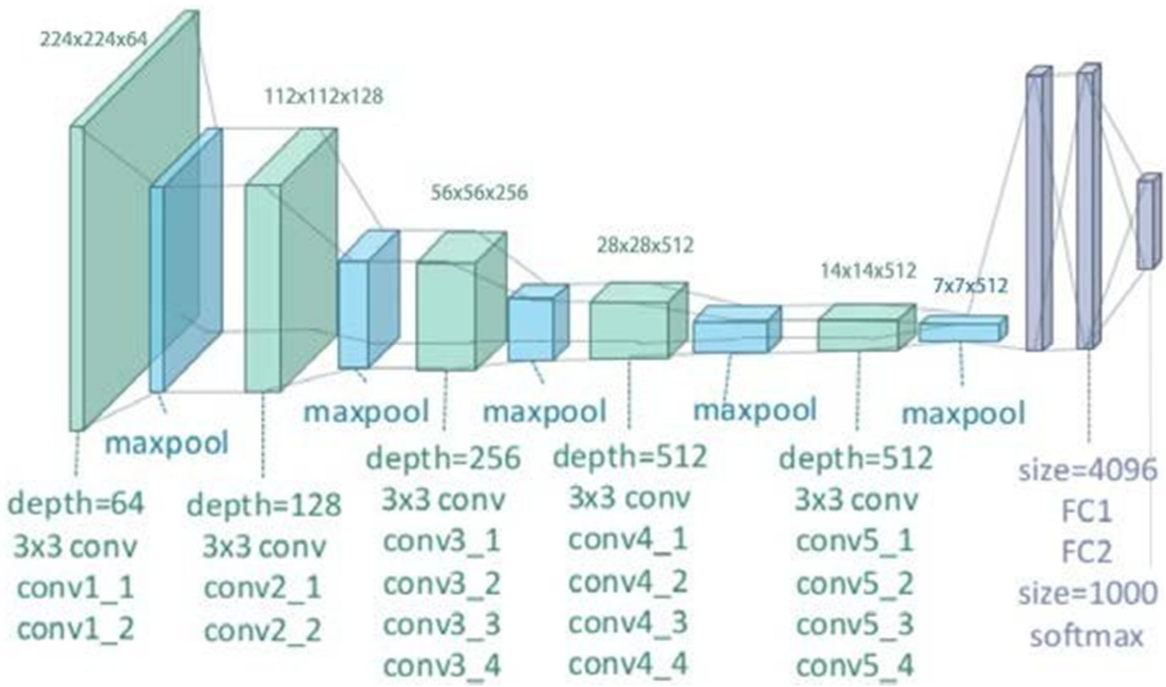


Fig II. VGG-19 Network Architecture

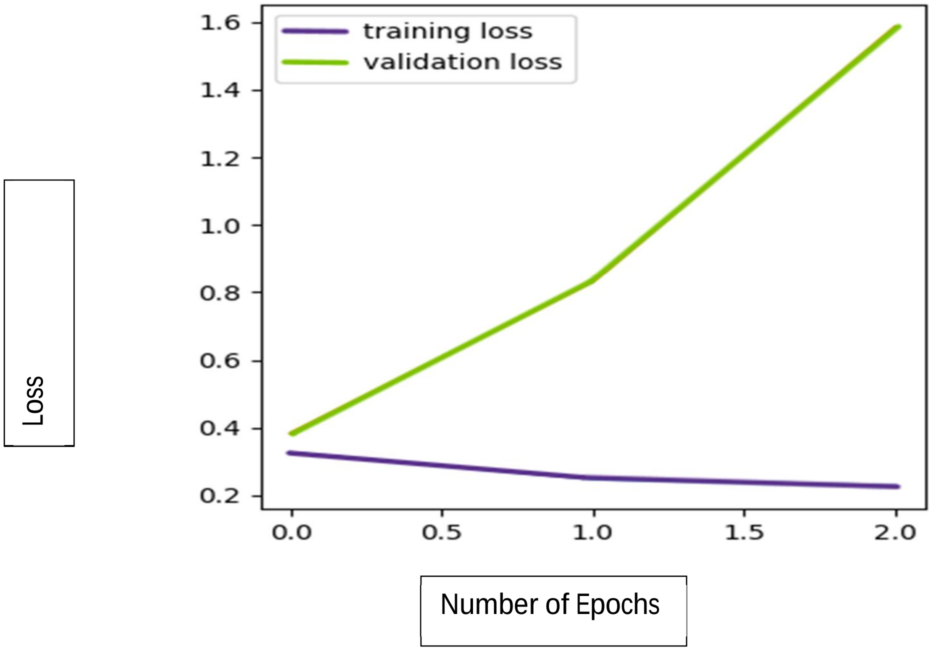
I will only be training this model for 3 epochs due to the speed it takes for the model to train given the amount of images in the dataset. To make sure that the model does not over fit I am going to apply a zoom ratio of 0.5, a horizontal flip. Here Is the accuracy that I got after training this model:

Here is the accuracy of the model

Table 1.
Accuracy For VGG-19 Model

Model	Accuracy	Loss
VGG-19	83.48%	1.79%

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 11 Issue X Oct 2023- Available at www.ijraset.com



The accuracy of 83.48% is impressive for a dataset with over 5200 images but given by this graph it looks like the training loss is low and relatively stable while validation loss is very high indicating potential overfitting of the model. Therefore I am going to next try out the resnet50 model on the data to see if the overfitting can be minimized with this model.

C. Resnet 50

Resnet 10 is another deep convolutional network with blocks that take the input from three layers back and insert it along with the input from the previous layer to the current layer into it. This eliminates the problem of the vanishing gradient over time. The vanishing gradient is a problem which takes place from loss functions in the neural network that receive a gradient that keeps shrinking per each layer of the network that it passes through the loss functions when the network backpropagates the gradient. With a gradient that doesn't change the weights of each of the layers in the network do not change indicating the network has already found the local optimum. Just like the VGG-19 model the Resnet 50 is a pretrained model, the model is pretrained on the Image Net dataset which contains 7500 various images of different objects.

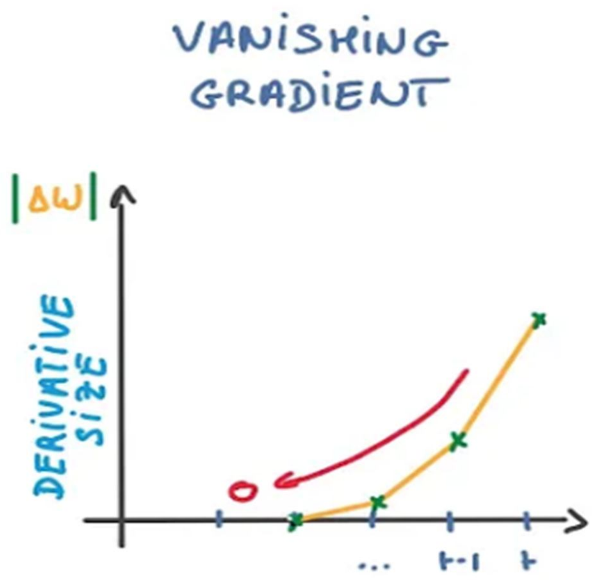


Figure III. Vanishing Gradient Derivative Visualization

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 11 Issue X Oct 2023- Available at www.ijraset.com

Therefore, I will run the images through the resnet-50 model to see if there are any improvements in the accuracies with the absence of the vanishing gradient problem.

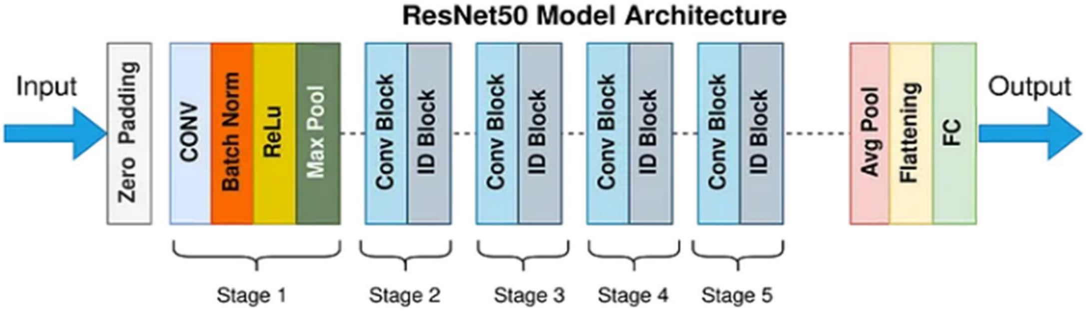


Fig IV. Resnet50 Model Architecture

To enter in the images to this model I will need to resize them to an input shape of 180 x 180 by 3. When using this model I will continue to leave the trainable parameter to false and I will leave the number of epochs at 3.

Here are the results:

Table 2.
Accuracy For Resnet 50 Model

Model	Accuracy	Loss
Resnet 50	77.08%	46.07%

The accuracy has decreased by 6.4% but the value loss has decreased by 133%. The resnet 50 allowed for the model to become less fitted to the model thereby decreasing the accuracy but decreasing the loss of the model. 179% from the previous model is a stark indication of overfitting.

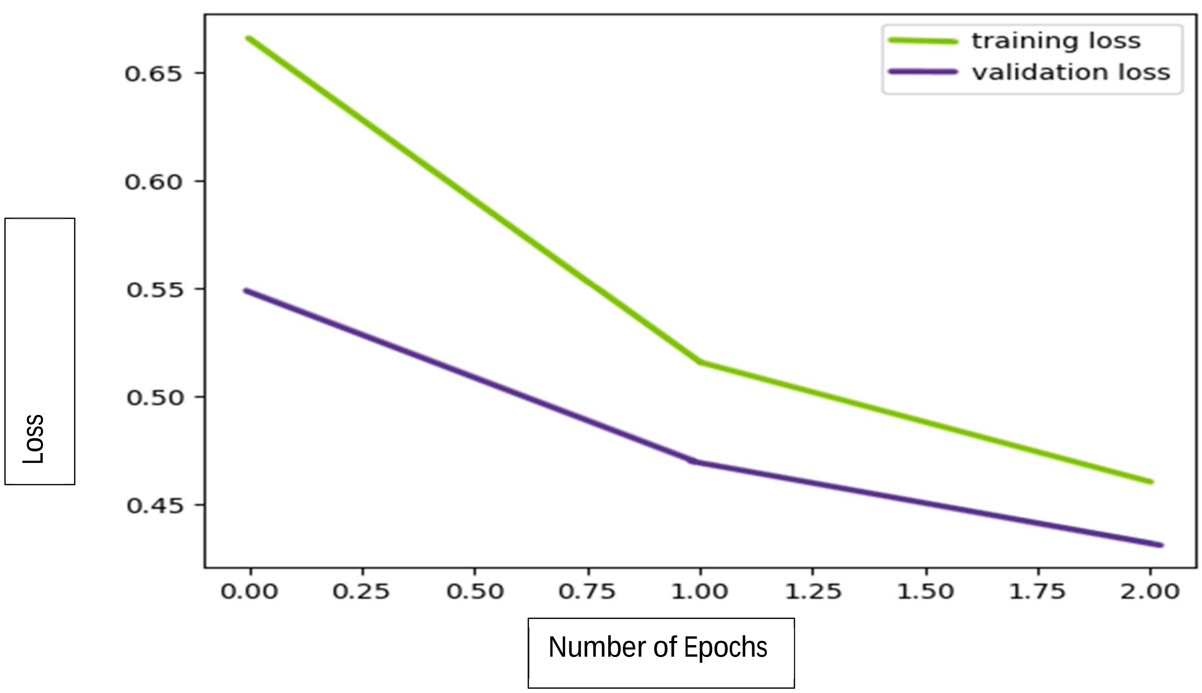


Fig V. Resnet50 Model Loss Data

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 11 Issue X Oct 2023- Available at www.ijraset.com

As we can see from the model shown the slope of the line is higher for the loss for the training and test sets indicating less overfitting over the time the model has been trained. It looks like though this model at first has been overfitting but has slowly become adapted to the diversity of images in the dataset hinting the importance of using a large dataset when training a model.

D. Xception

Now I am going to train the xception model on the data to see if it can outperform the VGG-19 model and the Resnet 50 model while maintaining a low validation loss.

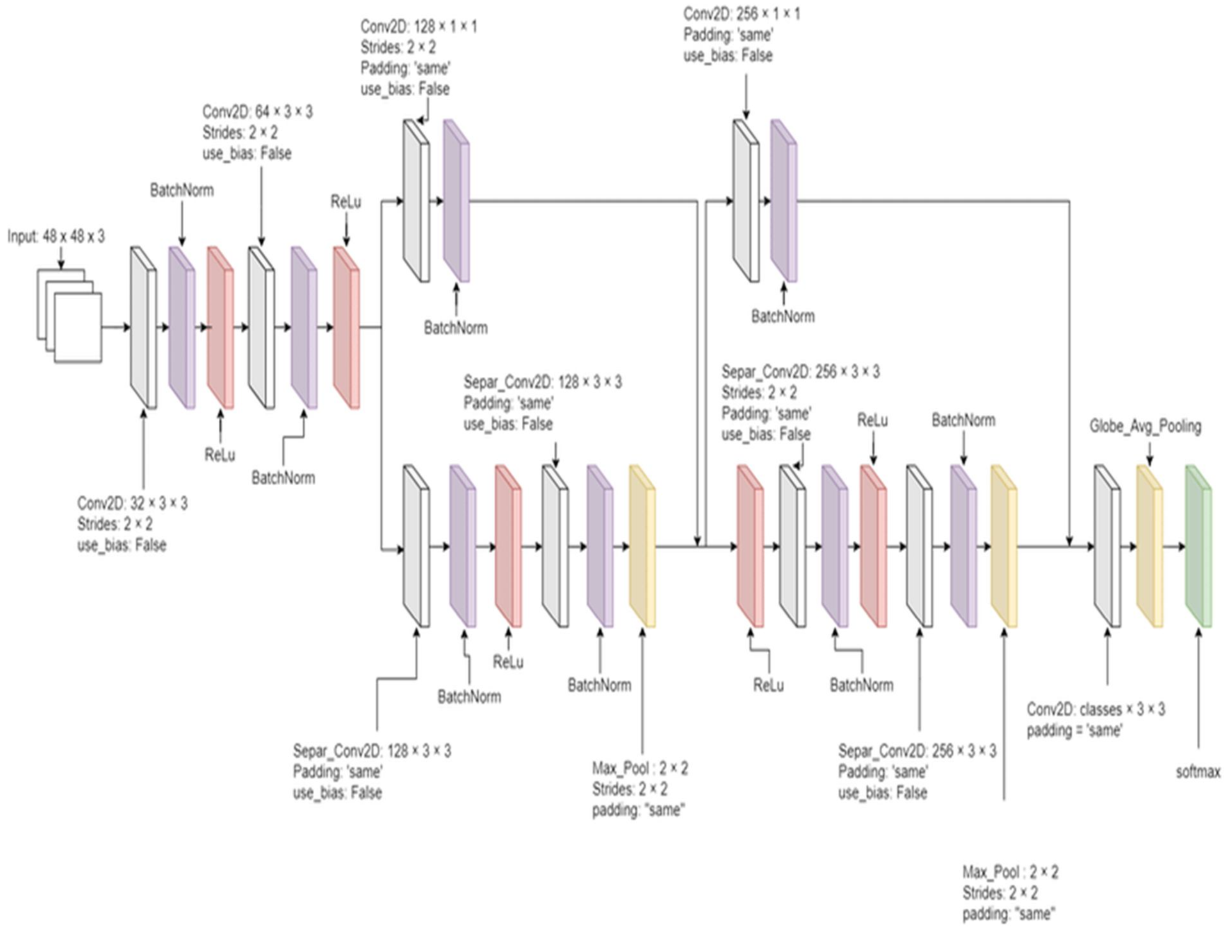


Fig VI. Xception Model Architecture

In a similar way to the resnet architecture this model also has a method of preventing the vanishing gradient issue, but this time it is through splitting the convolution up into two depthwise convolutions which can be shown by the top and bottom processes in Figure 6. In this process the input is combined and a pointwise convolution which is illustrated as the top process of the image. The xception model has 71 layers. To be able to train the model on the images I will need to rescale them to the input shape of 299 x 299 x 3. Now I am going to train this model on the dataset.

Here are the results

Table 3.
Accuracy For Xception Model

Model	Accuracy	Loss
Xception	62.66%	44.20%

For this dataset the xception model did poorer than the resnet 50 by 14.42%, but the value loss has decreased by 1.87%.

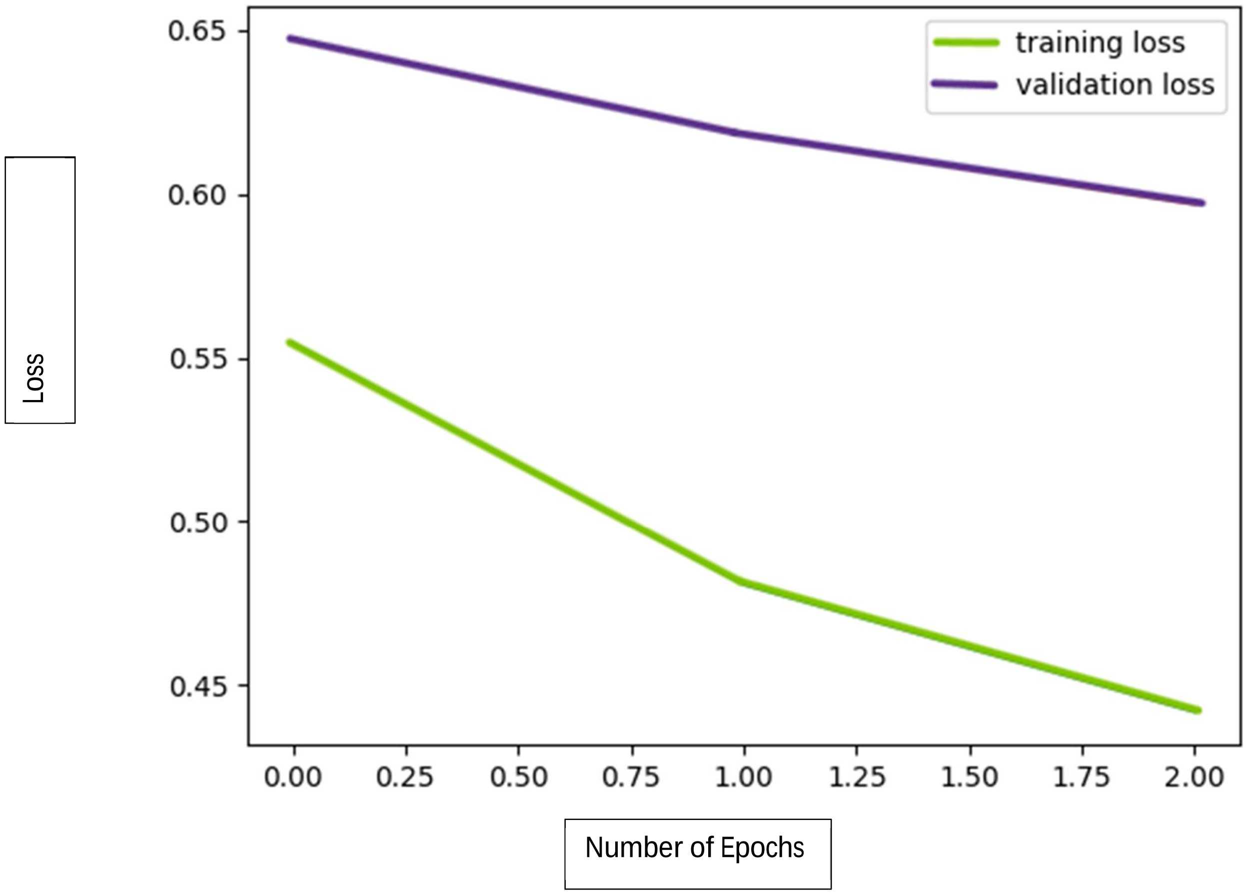


Fig VI. Xception Model Loss Data

III. CONCLUSIONS

By using VGG-19, Resnet 50 and Xception I have been able to accurately detect pneumonia from images of x-rays. I have explored the models of VGG-19, Resnet 50 and Xception and I have analyzed how they differ in structure, function and in accuracy for detecting malignant or benign cases of pneumonia from the images. From this research I have found that the VGG-19 model is the most prone to overfitting as can be seen with the high levels of value loss, and the Resnet 50 model has the best accuracy and value loss out of all of the models scoring 77.08% for accuracy and 44.02% for loss.

This shows that it can be possible to use deep learning to detect diseases at high accuracy rates such as pneumonia from x ray images. To better improve the accuracies of the models I can increase the epochs to 5 to give the model more time to train or increase the learning rate by 1% to allow the model to train faster. Also adding rotations, zooms and flips to the images can help to increase the accuracy rate by adding more diversity to the images in the dataset.

REFERENCES

[1] (3/2021) The Extensive Usage of Facial Image Threshing machine for facial emotion recognition performance Alwin Poulos Jung Hwan Kim Dong Seong Han
[2] (5/2018) Breast Cancer Screening using convolutional neural network and follow up digital mammography Clifford K Yang Research Gate
[3] (8/18/2019) The Annotated Resnet 50 Suvaditya Mukherjee
[4] (4/4/2020) Ayush Choudhary Vanishing Gradient And Exploding Gradient Problems Medium