Task 1: Chest X-Ray Images (Pneumonia) Classification

A labelled dataset of 5,232 chest X-ray images from children, including 3,883 characterized as depicting pneumonia (2,538 bacterial and 1,345 viral) and 1,349 normal, from a total of 5,856 patients to train the AI system.

Model 1

Network Detail(VGG16):

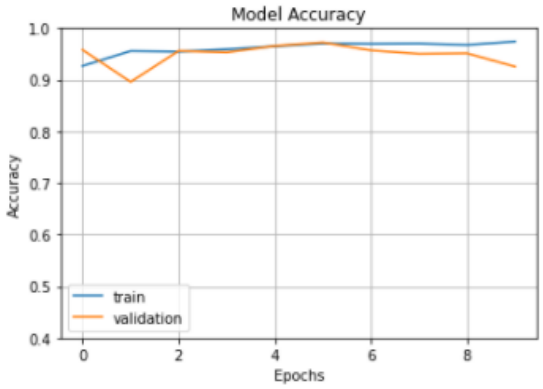
VGG16 is a widely used Convolutional Neural Network architecture with 16 layers. It is easy to implement and can be used for classification problems. Furthermore, the pre-trained model of vgg16 can also be used which is trained on Imagenet dataset.

In this task, we are using VGG16 pre-trained model. The benefit of this is that the weights are not randomly initialized rather they are fine-tuned weights. Furthermore, adam optimizer is used, with the batch size of 16, along with 10 epochs for training. Lastly, softmax has been used for classification.

Dataset Splits:

The dataset was already divided into train and test. The data was classified into two categories; Normal and Pneumonia. Moreover, the train set was further divided into validation set and training set with the ratio of 0.2. This means that 80% of the data will be used for training and 20% for validation.

Training Graphs:



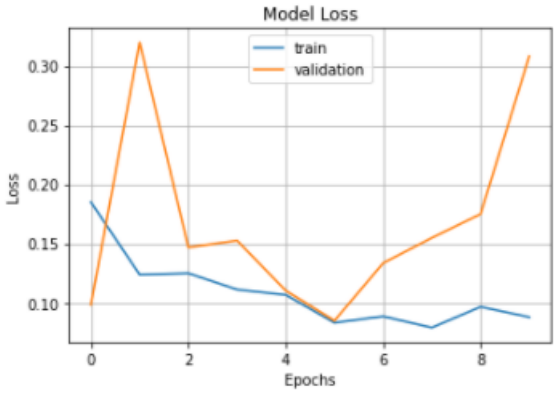


Figure 1. Model Training and Validation Accuracy

Figure 2.Model Training and Validation Loss

Figure 1 demonstrates the training and validation accuracy of the model. It can be observed that as the number of epochs increase, the accuracy of the training increases, however, the validation accuracy decreases after 8th epoch. Figure 2 demonstrates the training and validation loss of the model. The loss decreases as the number of epochs increases. There is a spike in validation loss after the 5th epoch, contradicting the loss of training.

Performance measures:

The performance measure used is Accuracy. Following table illustrates the quantitative measures.

	Training	Validation	Testing
Loss	0.0878	0.3082	0.2756
Accuracy	0.9742	0.9255	0.9263

Model 2

Network Detail(ResNet50):

ResNet stands for Residual Network. The authors resNet introduced the concept of skip connections. The skip connections skips some of the layers in the neural networks and feeds that output of one layer as the input to the next. The skip connections, solve the problem of vanishing gradient which mostly occurs because of the deep layers in convolutional networks. The variant of Resnet; Resnet50 uses 3-layer bottleneck blocks to ensure improved accuracy and lesser training time.

Further addition in the model is a flatten layers, a dense layer along with relu activation functions. The classifier used is softmax, along with Adam optimizer. Moreover, categorical cross entropy is used for loss. Lastly, the dataset is train over 10 epochs with batch size of 16.

Dataset Splits:

The dataset was already divided into train and test. The data was classified into two categories; Normal and Pneumonia. Moreover, the train set was further divided into validation set and training set with the ratio of 0.2. This means that 80% of the data will be used for training and 20% for validation.

Training Graphs:

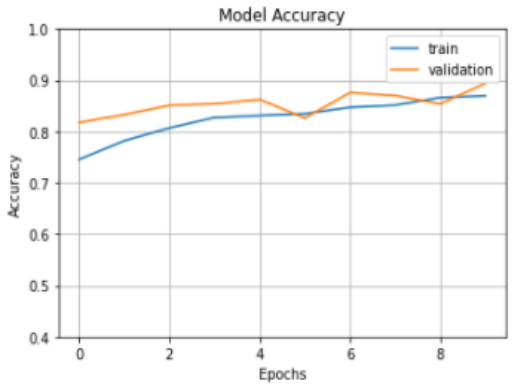


Figure 3. Model Training and Validation Accuracy

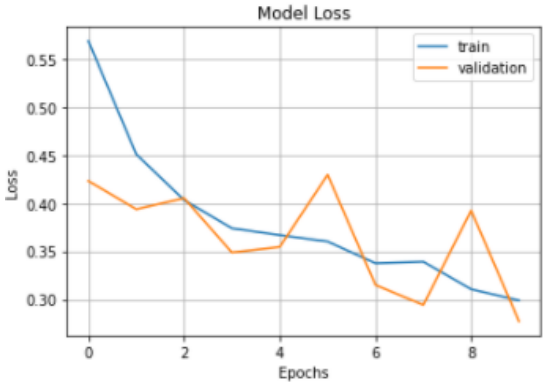


Figure 4. Model Training and Validation Loss

Figure 3 demonstrates the training and validation accuracy of the model. It can be observed that as the number of epochs increase, the accuracy of the training increases, however, the validation accuracy increases after 8^{th} epoch. Figure 4 demonstrates the training and validation loss of the model. The loss decreases as the number of epochs increases.

Performance measures:

The performance measure used is Accuracy. Following table illustrates the quantitative measures.

	Training	Validation	Testing
Loss	0.2990	0.2774	0.4416
Accuracy	0.8695	0.8940	0.7772

Performance comparison of the CNN architectures. [Training and inference Accuracy and Time]

	Training Accuracy	Testing Accuracy	Time (seconds)
VGG16	0.97	0.92	1147s
ResNet50	0.86	0.77	1177s

Conclusion:

The performance of VGG16 is better than ResNet50, even though there is no difference in time. There is a significant difference in the accuracy of both the models. Furthermore, both the models have higher testing accuracy than testing accuracy which indicates overfitting as well.

Task 2: 315 Bird Species - Classification

Data set of 315 bird species.45980 training images, 1575 test images (5 images per species) and 1575 validation images (5 images per species). All images are 224 X 224 X 3 color images in jpg format. Data set includes a train set, test set and validation set. Each set contains 315 sub directories, one for each bird species.

Model 1

Network Detail(VGG16):

VGG16 is a widely used Convolutional Neural Network architecture with 16 layers. It is easy to implement and can be used for classification problems. Furthermore, the pre-trained model of vgg16 can also be used which is trained on Imagenet dataset.

In this task, we are using VGG16 pre-trained model. The benefit of this is that the weights are not randomly initialized rather they are fine-tuned weights. Furthermore, adam optimizer is used, with the batch size of 16, along with 10 epochs for training. Lastly, softmax has been used for classification.

Dataset Splits:

The dataset was already divided into train and test. The data was classified into two categories; Normal and Pneumonia. Moreover, the train set was further divided into validation set and training set with the ratio of 0.2. This means that 80% of the data will be used for training and 20% for validation.

Training Graphs:

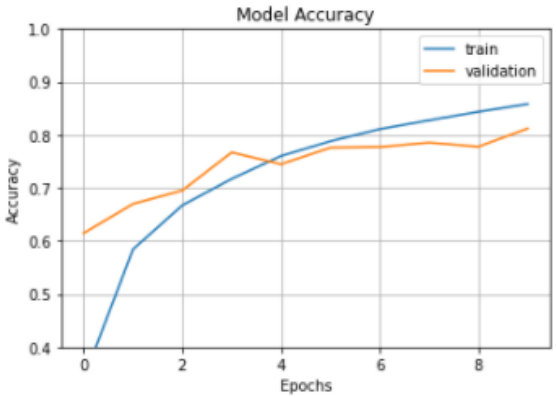


Figure 3. Model Training and Validation Accuracy

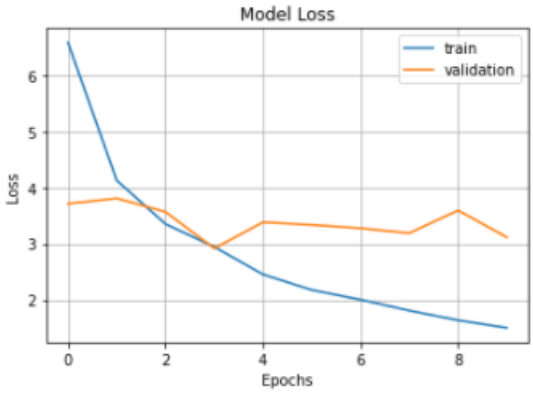


Figure 4.Model Training and Validation Loss

Figure 1 demonstrates the training and validation accuracy of the model. It can be observed that as the number of epochs increase, the accuracy of the training increases, however, the validation accuracy decreases after 7^h epoch and then increases. Figure 2 demonstrates the training and

validation loss of the model. The loss decreases as the number of epochs increases. However, the validation loss does not decrease in comparison with the training set.

Performance measures:

The performance measure used is Accuracy. Following table illustrates the quantitative measures.

	Training	Validation	Testing
Loss	1.5012	3.1149	2.7671
Accuracy	0.8583	0.8121	0.8267

Model 2

Efficient Net:

EfficientNet, introduced by Tan et al, is a convolutional neural network architecture and scaling method that uniformly scales all dimensions of:

- Depth
- Width
- Resolutions

The compound scaling method is justified by the intuition that if the input image is bigger, then t he network needs more layers to increase the receptive field and more channels to capture more f ine-grained patterns on the bigger image.
Further addition in the model is Global Average pooling and, a dense layer. The classifier used is softmax, along with Adam optimizer. Moreover, categorical cross entropy is used for loss. Lastly, the dataset is train over 5 epochs with batch size of 16.

Dataset Splits:

The dataset was already divided into train and test. The data was classified into two categories; Normal and Pneumonia. Moreover, the train set was further divided into validation set and training set with the ratio of 0.2. This means that 80% of the data will be used for training and 20% for validation.

Training Graphs:

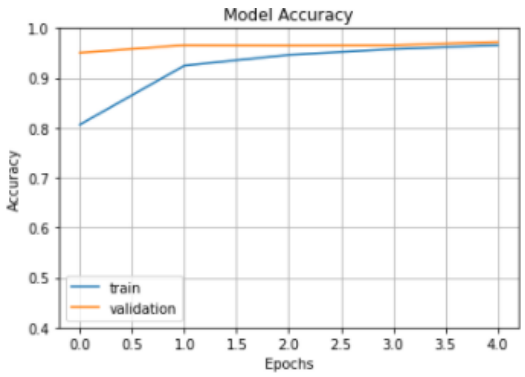


Figure 3. Model Training and Validation Accuracy

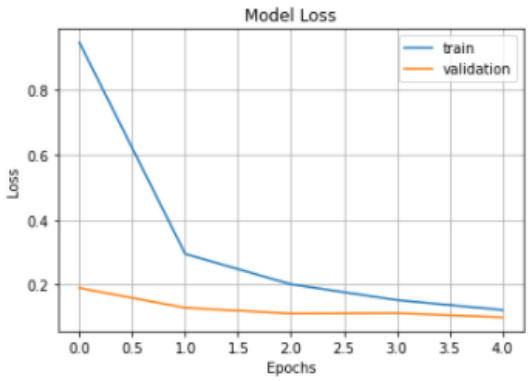


Figure 4. Model Training and Validation Loss

Figure 3 demonstrates the training and validation accuracy of the model. It can be observed that as the number of epochs increase, the accuracy of the training increases, and, the validation accuracy coincides with training after 3.5 epochs. Figure 4 demonstrates the training and validation loss of the model. The loss decreases as the number of epochs increases.

Performance measures:

The performance measure used is Accuracy. Following table illustrates the quantitative measures.

	Training	Validation	Testing
Loss	0.1218	0.0994	0.0682
Accuracy	0.9657	0.9721	0.9803

Performance comparison of the CNN architectures. [Training and inference Accuracy and Time]

	Training Accuracy	Testing Accuracy	Time (seconds)
VGG16	0.85	0.82	6919s/10 epochs
EfficientNetB5	0.96	0.98	4164s /5 epochs

Conclusion:

The performance of Efficient Net is significantly better than VGG16.There is a significant difference in the accuracy of both the models. Efficient-NetB5 has performed better in testing time, hence, the model does not indicate overfitting.