**Lung X-ray Classification CNN Model Report**

May 5, 2024

**Group 1 Authors:** Chris Why | Faeez Aroos | Mihir Sabnis

## **COVID-19 Medical Crisis and Administrative Problem**

During the COVID-19 pandemic, healthcare systems across the globe were overwhelmed with a surge in patients requiring critical care from COVID-19, often leaving hospitals in a situation where they were over capacity. During the peak of the pandemic, hospitals operated at or above capacity. Data from the COVID Tracking Project indicated that ICU capacity was a critical issue, with many states reporting over 70% ICU occupancy during the height of the pandemic.[[1]](#footnote-0) Accurate diagnostic tools could have reduced this burden by ensuring only the patients who needed critical care were admitted to ICUs. Hospitals faced a dual challenge: identifying and treating COVID-19 patients and efficiently managing limited resources such as ICU beds and ventilators.

Accurately classifying patients into diagnosis categories: COVID-19, viral pneumonia, non-COVID or pneumonia lung issues, or normal healthy lungs is crucial for multiple reasons:

* COVID patients require isolation and specific treatment protocols to prevent the spread and address the respiratory issues that arise. In extreme cases, this can require the use of a respirator.
* Viral pneumonia needs treatment that differs from bacterial pneumonia, often involving supportive care and antiviral medications
* Non-COVID and non-viral pneumonia lung issues encompass a range of conditions that may require different treatments and may not be as severe
* Normal lungs can be identified and discharged, reducing pressure on hospital resources.

False positives for COVID-19 place an unnecessary strain on healthcare resources. During critical times, it's essential to allocate ICU beds, ventilators, and medical personnel to patients critically in need. Reducing false positives could have redirected these lifesaving resources to those in critical condition.

## **Dataset**

A collaborative team of researchers from Qatar University in Doha, and the University of Dhaka in Bangladesh, along with collaborators from Pakistan and Malaysia, and medical doctors, have developed a significant database of chest X-ray (CXR) images. The dataset used for the analysis is specifically designed to facilitate the study of COVID-19, as well as other normal and viral pneumonia lung infections. The dataset contains **3,616 COVID-19-positive cases, 10,192 normal images, 6,012 images of lung opacity (non-COVID lung infections), and 1,345 viral pneumonia images**.

## **Data Preprocessing and Cleaning**

**1. Data Ingestion and Directory Structure Analysis:** We created a function that analyzes the dataset directory to collect image paths and assign labels based on folder names.

**2. Dataframe Creation:** All valid image paths and their corresponding labels are compiled into a Pandas DataFrame by the **define\_df** function.

**3. Data Splitting:** We used the **split\_data** function to divide up the dataset into training, validation, and test sets to facilitate robust training and ensure the model's generalization capability. It employs stratification to maintain a balanced representation of each class across the data splits, addressing potential class imbalances.

**4. Image Data Generators:** We used the **create\_gens** function from Keras's **ImageDataGenerator** to generate data loaders for training, validation, and test datasets. These generators perform real-time image loading and preprocessing, including horizontal flipping for data augmentation and a simple scalar function for normalization. A batch size = 32 was set on generation.

**5. Visualization of Data Samples:** We visually confirmed the proper preprocessing and labeling of images by displaying a batch from both generators to ensure data integrity and accuracy.

## **Model Selection**

**1. ResNet101:** a 101-layer residual network that employs skip connections to enable smooth gradient flow and avoid the vanishing gradient problem in deep networks. Its residual blocks support deep feature learning, making it suitable for vision tasks requiring rich representations, maintaining performance even with increased network depth.

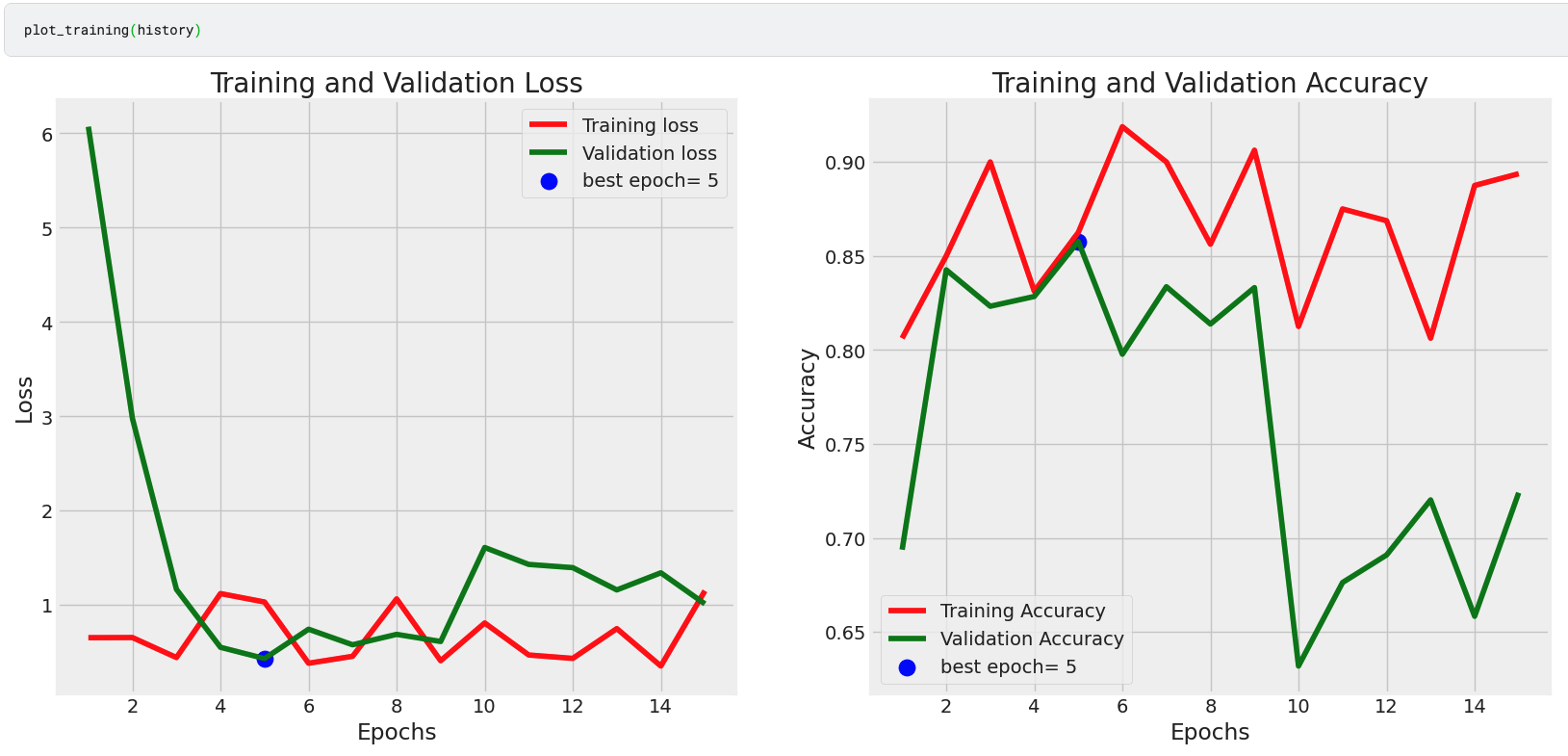
**2. VGG16:** A 16-layer convolutional neural network known for its simplicity, it stacks 3x3 convolutional layers in a uniform architecture. Max pooling reduces feature volume, while fully connected layers handle classification. Its straightforward design captures vital patterns, making it ideal for medical image analysis.

**3. KerasEfficientNetB1** Part of the EfficientNet family, it scales CNNs uniformly using a compound coefficient. MBConv and SE optimizations balance latency and accuracy, achieving high precision with fewer parameters. Its efficiency makes it well-suited for fast inference in clinical settings where rapid decision-making is crucial.

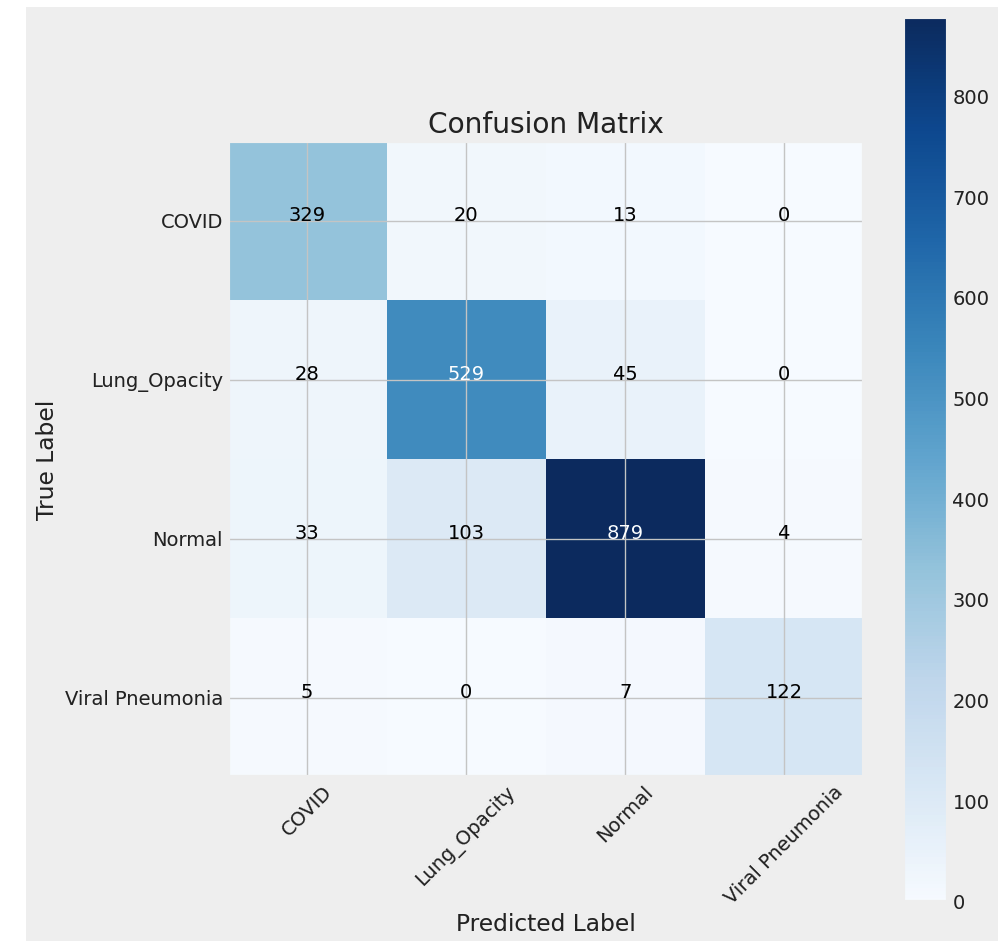
## **Model 1 - ResNet101**

As a base model,a pre-trained ResNet101 architecture is used as the feature extractor with its final classification layers removed. Weights are initialized from a pre-trained model file (resnet101\_weights\_tf\_dim\_ordering\_tf\_kernels\_notop.h5). Training is set to 20 epochs, with each epoch consisting of 5 steps for faster computation due to time constraints. A batch size of 64 ensures efficient data loading.

The model achieves a **training accuracy** of approximately **89.1%**. With a **validation accuracy** of around **86.2%,** the model generalizes well to unseen data during training. The validation loss shows some fluctuations but remains relatively stable throughout the epochs.

****

The model demonstrates strong classification ability across all four categories. Some degree of confusion exists between similar conditions (e.g., COVID-19 and lung opacity), which is a known challenge due to overlapping visual features.



The model consistently maintains precision above **80%** across all four classes. Recall is also relatively high across the board, ranging from **88% to 91%** in each class. The F1-score, representing the harmonic mean of precision and recall, remains consistently strong. The model achieves a **test accuracy of approximately 87.8%**, demonstrating strong predictive performance on unseen data. The macro average **F1-score stands at 88%** and the **weighted average F1-score** is also **88%**, reflecting the model's overall efficacy in accurately identifying various lung conditions.

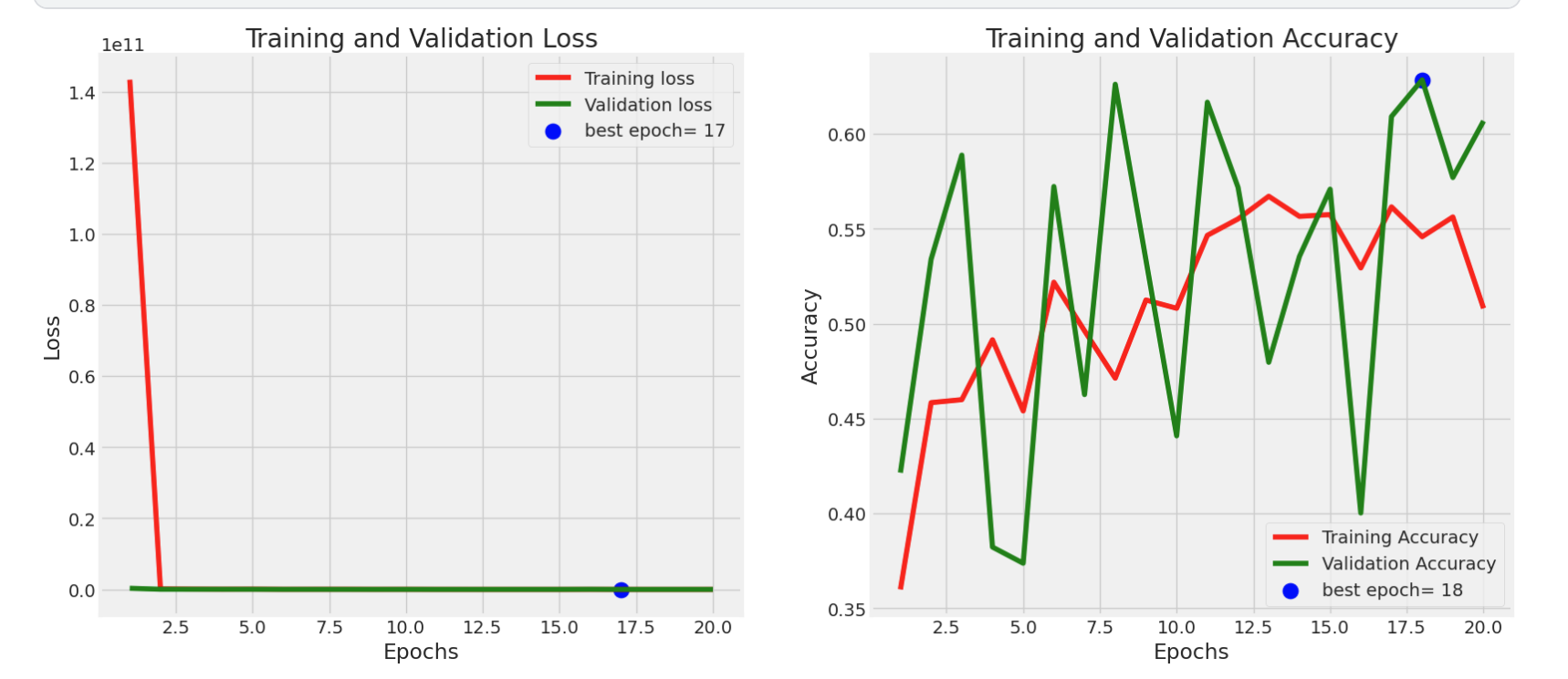
In conclusion, The ResNet101-based model provides a robust classification of COVID-19 and other lung infections. Slight discrepancies in the confusion matrix can be mitigated by increasing steps per epoch for greater data exposure. Overall, the model is effective for its purpose in aiding the diagnosis and classification of lung infections to efficiently allocate medical resources.

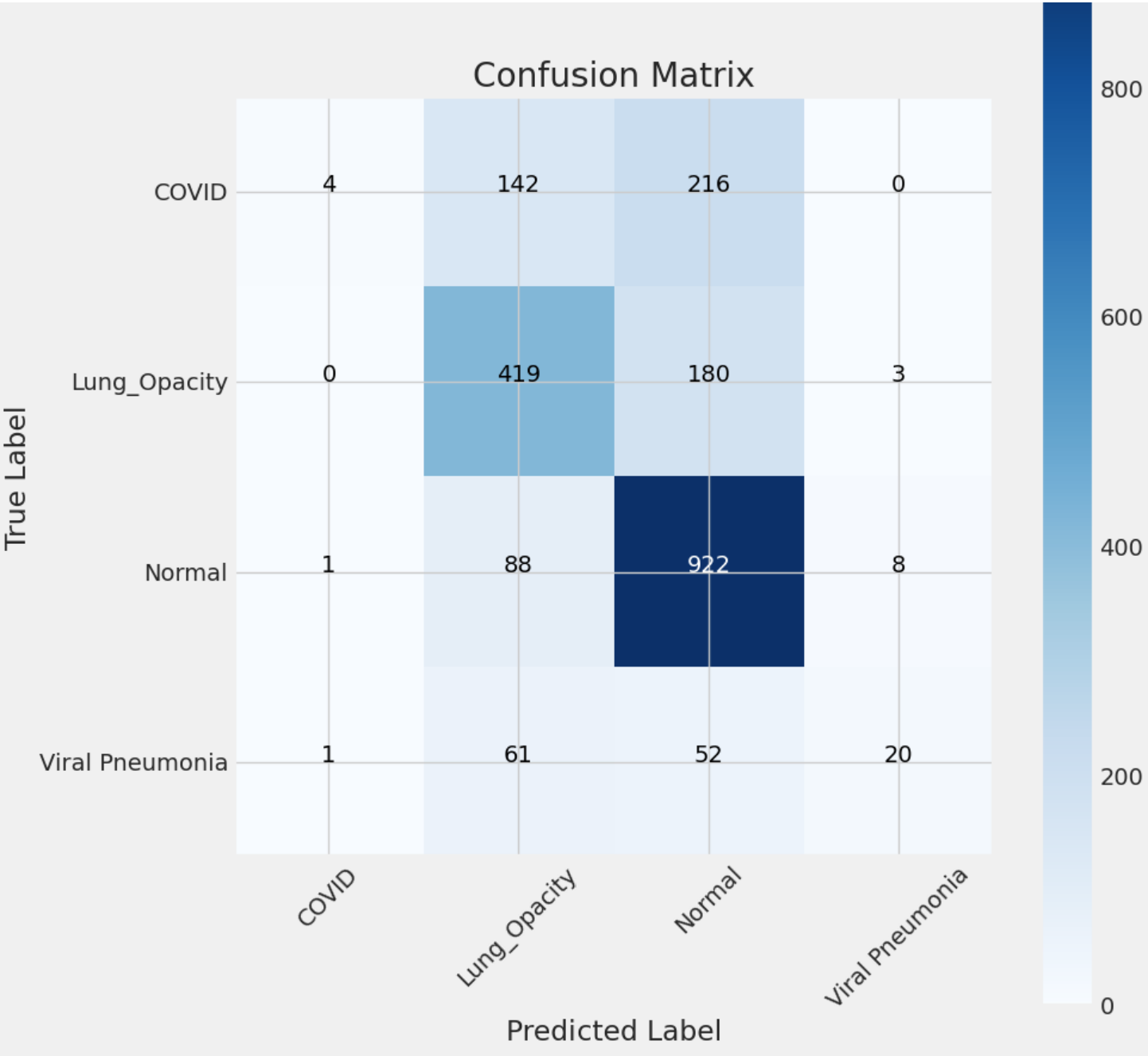
## **Model 2 - VGG16**

Using the same train, test, and validation data as used in Model 1, the weights of a VGG16 architecture were trained. Similarly, the output layer is a dense layer with a softmax activation function to generate class probabilities across four categories (COVID-19, Normal, Lung Opacity, and Viral Pneumonia).

Similar to Model 1, the model is trained with several specific hyperparameters to optimize its performance. The input image size was adjusted to 50x50 to reduce computational resources and accelerate training speed. The training process is configured to run for 20 epochs, with each epoch comprising 100 steps to ensure a balanced training dataset along with a batch size of 32.

The training and validation results provided illustrate the effectiveness of the model in classifying lung infections across four classes: COVID-19-affected lungs, normal lungs, lung opacity (non-COVID lung infections), and viral pneumonia. The model achieves a **training accuracy of approximately 61.7%** with a **validation accuracy of around 63.4%**.





Unlike Model 1, Model 2 has lower predictive accuracy as highlighted by the confusion matrix. In many cases, COVID-19 was mispredicted as a normal or lung opacity diagnosis. For images where the true label was lung opacity or normal, the model was fairly accurate in predicting the diagnosis. Images from viral pneumonia patients were most commonly misclassified as lung opacity or normal.

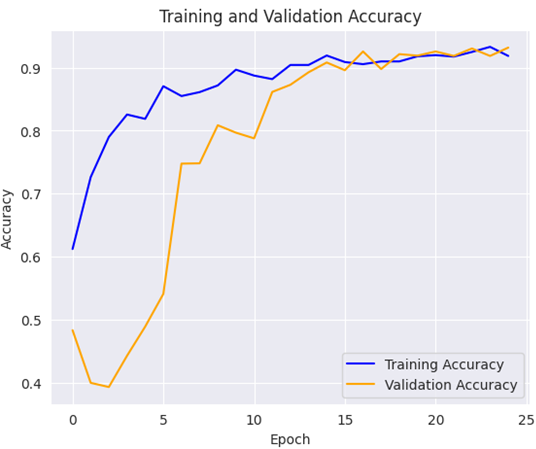
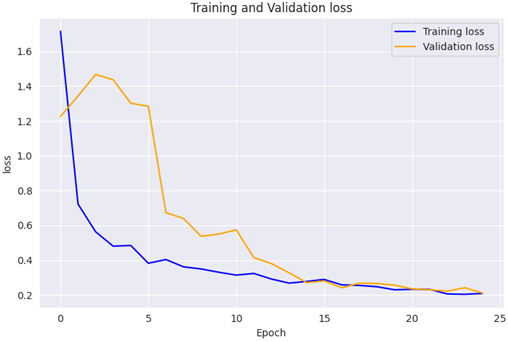
The evaluation metrics and final results reveal the model's limited capability in classifying lung infections, including COVID-19. The model maintains a **precision above 59%** across all classes, yet recall varies significantly between 1% and 90%. The F1-scores, both macro and weighted averages, are **42%** and **57%** respectively, indicating unbalanced and generally weak predictive performance. Despite achieving a **test accuracy of about 64.5%**, the model shows discrepancies that could potentially be improved with further tuning and additional training data. Overall, the model's performance suggests it is not ideally suited for effectively aiding in the detection and classification of various lung infections.

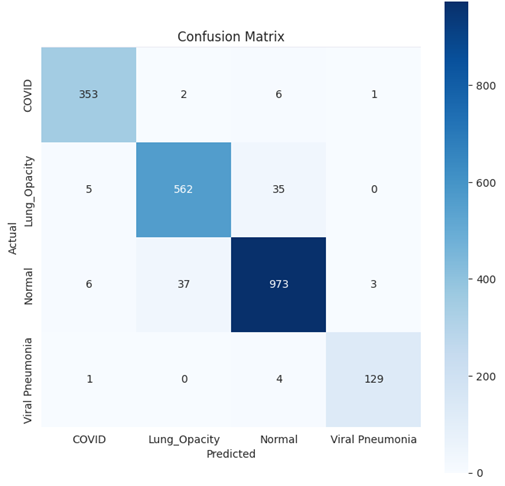
## **Model 3 - EfficientNetB1**

The core of the model is based on the **EfficientNetB1 architecture,** adapted from TensorFlow's **Keras** applications. It's configured without the top layer (classification layers) to act as a sophisticated feature extractor. The model weights are loaded from a pre-trained EfficientNetB1 model and then trained on the same train, test, and validation split. The output feeds into a dense layer with **128 neurons with 'relu' activation** that handles the activation of learned features. A **Dropout layer with a rate of 0.5** reduces the risk of overfitting. A final Dense layer with **4 neurons and a 'softmax' activation** function classifies the inputs into four categories.

This model is created with several specific hyperparameters to optimize performance. The model utilizes the **Adamax optimizer**, which tends to have a consistent convergence in scenarios with noisy gradients. The **learning rate is 0.00055** with a **loss function of categorical cross-entropy**. The **metric used is Accuracy** to optimize for the proportion of accurately classified images.

The model was trained for **25 epochs** to avoid excessive training time and was configured to **50 steps per epoch**, to ensure that each epoch processes a significant portion of the data, enhancing learning outcomes without overfitting. The model achieved a **training accuracy of 93.62%** and a **validation accuracy of 93.19%**. We can see in the graph that both training and validation loss stabilized and achieved convergence.



Model 3 was the best-performing model on test data. It achieved a **test accuracy of 95.61%**. With reference to the confusion matrix, we can see that the model is the best at **predicting COVID-19 lungs (97.51% accuracy)** and **Viral pneumonia (96.12% accuracy)**. It seems to confuse normal lungs and lungs with other problems probably due to the varying opacity of the issues which makes it more difficult to classify. The model is fantastic at reducing the number of normal lungs misclassified as COVID or pneumonia (99.12% accurately predicting normal lungs as not COVID or pneumonia).

| **Class** | **True Positives** | **False Positives** | **False Negatives** |
| --- | --- | --- | --- |
| **COVID** | 353 | 3 | 6 |
| **Lung Opacity** | 562 | 5 | 35 |
| **Normal** | 973 | 46 | 0 |
| **Pneumonia** | 129 | 1 | 4 |

## **Conclusion and Evaluation** Based on the analysis and model performance, it is evident that advanced imaging analysis models can play a crucial role in the accurate classification of lung-related illnesses, particularly in the context of COVID-19. The successful implementation of these models can significantly enhance diagnostic accuracy and efficiency, leading to better management of medical resources. With accurate diagnosis, we can reduce the time to treatment and better manage hospital resources to only those who critically need it in a time of crisis like the COVID-19 pandemic. The pandemic boosted attention and funding in the medical deep learning space as well as the demand for clean and consistent datasets for training and testing purposes. The most critical part of this data pipeline is ensuring that the dataset is clean and accurate to real life as these models will be used to help determine the treatment for patients, which if wrong can be fatal. There are serious ethical implications with using such models and therefore we must ensure utmost accuracy and compliance with privacy practices. The model we chose is the Keras EfficientNetB1 as it accurately classifies critical illnesses (COVID, pneumonia) and rarely misclassifies healthy lungs as critical illnesses. This allows us to better allocate medical resources to the critically ill. The model also trains faster and more accurately on a smaller dataset. The ResNet101 model, with its high precision and recall across various lung conditions, demonstrates strong predictive performance, making it a valuable tool in healthcare settings. However, the VGG16 model shows room for improvement, particularly in reducing misclassifications. In training our models, we were heavily limited by the computation power and long training times. We chose hyperparameters that were manageable in the time given. These models have a high potential for further optimization.

## Overall, the development and refinement of these models offer substantial benefits for future healthcare scenarios, emphasizing the need for continued improvement and adaptation of such technologies to ensure effective response in medical crises. The models were trained using limited computational resources. Retraining the models with a larger number of epochs, batch size, steps per epoch, and a lower learning rate would likely further improve the accuracy of the models. This approach not only meets immediate healthcare needs but also sets a precedent for managing future pandemics through rapid diagnosis and efficient resource allocation.

## **References**

*The COVID Tracking Project,*[**https://covidtracking.com/**](https://covidtracking.com/), (accessed 3rd May, 2024)

1. *The COVID Tracking Project,*[**https://covidtracking.com/**](https://covidtracking.com/), (accessed 3rd May, 2024) [↑](#footnote-ref-0)