Shape

Description automatically generated with medium confidence



**Essay / Assignment Title:** AI-Driven Pneumonia Diagnosis: Chest X-ray Classification

**Programme title:MSc of Data Analytics**

**Name:Turan Can Gün (Q1055616)**

**Year:2024 (2023 October Intake)**

**CONTENTS**

[INTRODUCTION 3](#_Toc1988586108)

[CHAPTER ONE DATASET PREPARATION 4](#_Toc474453680)

[CHAPTER TWO NEURAL NETWORK ARCHITECTURE DESIGN 8](#_Toc976025806)

[2.1 CONVOLUTIONAL NEURAL NETWORK (CNN) 9](#_Toc587541108)

[2.1.1 CNN WORKING PRINCIPLE 10](#_Toc1428438827)

[2.2 Pre –Trained Models 11](#_Toc596029796)

[2.2.1 MobileNet and ResNET 11](#_Toc2097114615)

[CHAPTER THREE CODE IMPLEMENTATION 12](#_Toc1693920293)

[3.1 MobileNet 12](#_Toc368537729)

[3.2 ResNET 18](#_Toc1179762273)

[CHAPTER FOUR MODEL EVALUATION 19](#_Toc396018648)

[4.1 Definition of Metrics 19](#_Toc1223313201)

[4.2 Comparision Results of MobileNet and ResNET 20](#_Toc797541013)

[CONCLUDING REMARKS 23](#_Toc1659379560)

[BIBLIOGRAPHY 24](#_Toc971747970)

[APPENDIX (if necessary) 26](#_Toc1045181180)

**Statement of compliance with academic ethics and the avoidance of plagiarism**

I honestly declare that this dissertation is entirely my own work and none of its part has been copied from printed or electronic sources, translated from foreign sources and reproduced from essays of other researchers or students. Wherever I have been based on ideas or other people texts I clearly declare it through the good use of references following academic ethics.

(In the case that is proved that part of the essay does not constitute an original work, but a copy of an already published essay or from another source, the student will be expelled permanently from the postgraduate program).

Name and Surname (Capital letters):Turan Can Gün

.......................................................................................................................................

Date: 31/10/2024

# INTRODUCTION

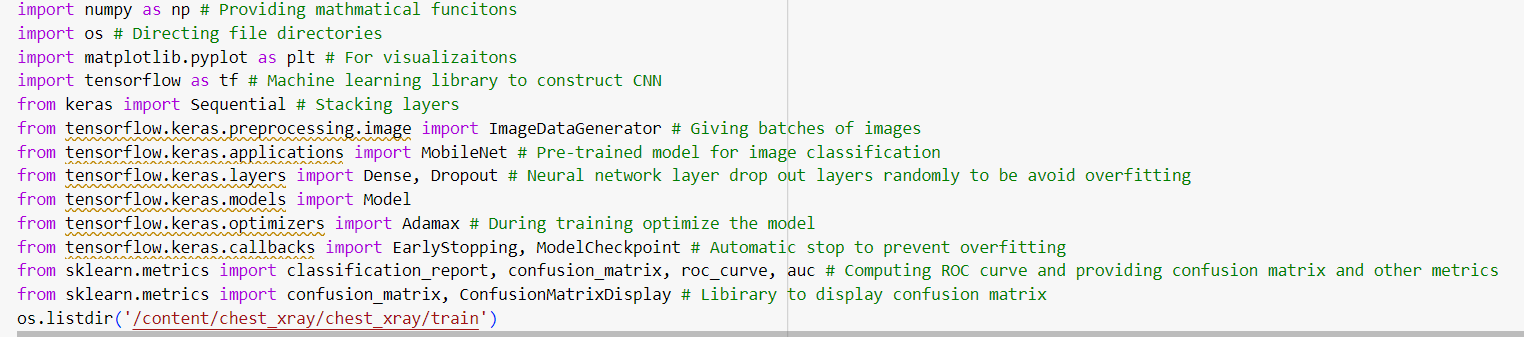
Computer vision is a subfield of artifitical intellligence (AI) also it has revolutionary impact in healthcare industry.Furthermore computer vision has a background from machine learning (ML) algorithms.In practically as an example ML algorithms are able to provide trustful insights regarding to detecting the most prone prone people some chronic diseases such as diabetes , high blood pressure (C. Mosquera et. Al , 2023). Moreover which can provide tailored dietary to these people to recorever their health situation. Specifically computer vision field has wide range usages areas like pneumonia detection , tuberculosis detection , lung cancer detection , COVID-19 diagnosis.

Computer vision concept can be handled by different ways such as Convolutional Neural Netowrks (CNN) , Image Segmentation , Object Detection and Localization and Trasnfer Learning.The most popular one also used way in this study which is Convolutional Neural Network (CNN).

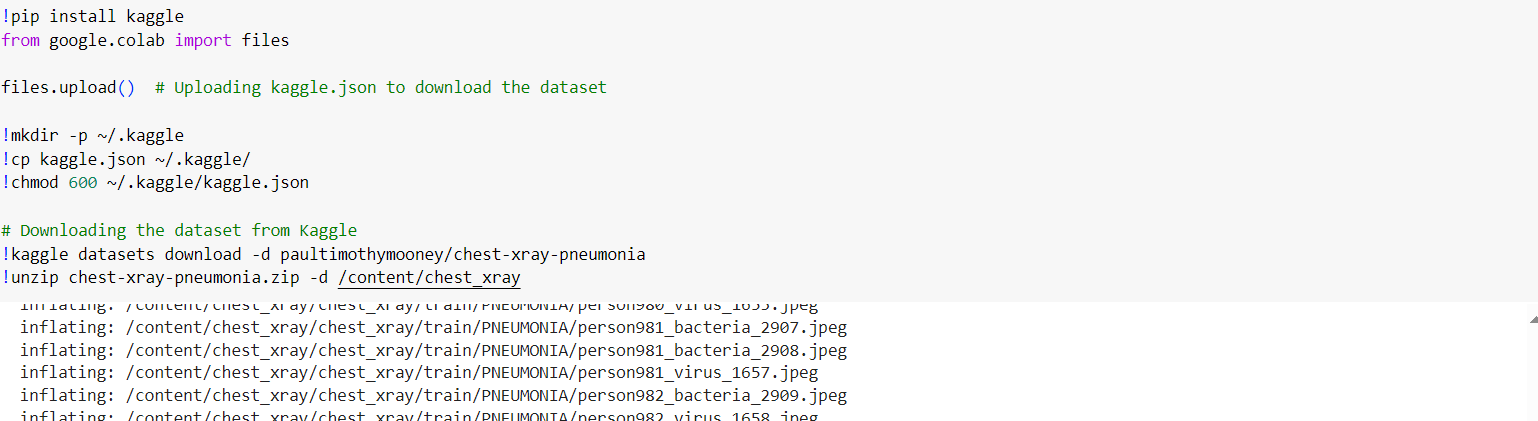
In that study various steps will be mentioned from dataset preparation to model construction and evaluation performance results.Moreover the study was performed on Google Colab environment.About hardware configuration T4 GPU is used due to parallel computation.It has 16 GB of GDDR6 memory and 2560 CUDA cores.

# CHAPTER ONE DATASET PREPARATION

The dataset has been obtained from Kaggle.In this there are some explanations regarding to the dataset such as implementation into the Google Colab then various visualizaitons about some facts across the dataset.

Figure 1.1 Explanations of Necessary Libraries

Before the implement the project various libraires must be uploaded into the Google Colab Libraries because which have pre-defined configurations so Pythpn is very popular around the world across machine learning algorithms due to user friendly.These libraries are essential for from mathematical calculations to optimizing the algorithm to get as highest as benchmark score at end of the day.

Figure 1.2 Downloading the Dataset from Kaggle Directly into Colab Environment

Instead of the uploading the dataset into Google Drive then get it back into Google Colab this study implement this step like downloading the dataset from Kaggle directly by using Kaggle Personal Account Token (.json file) because downloading speed could be reach up to 300 Mb/sn which means that this way is too much effective than other conservative method.

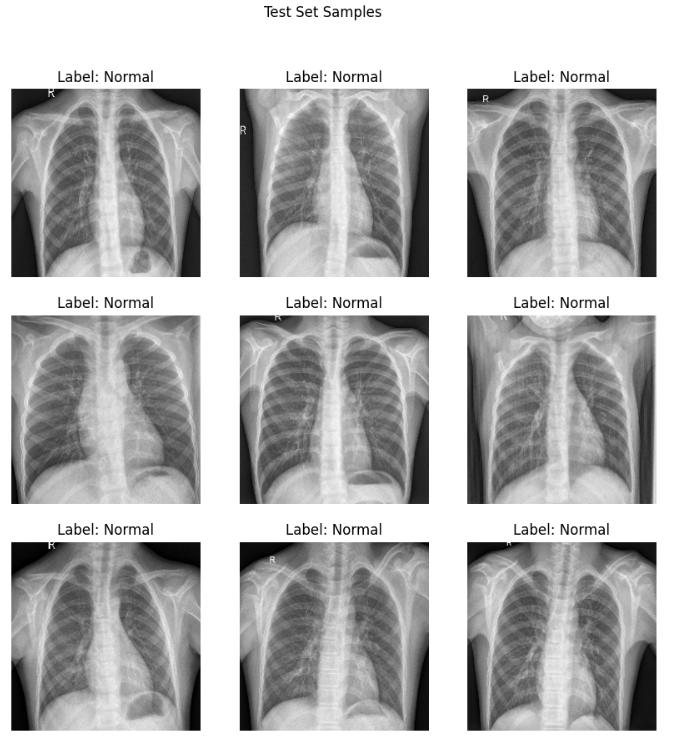
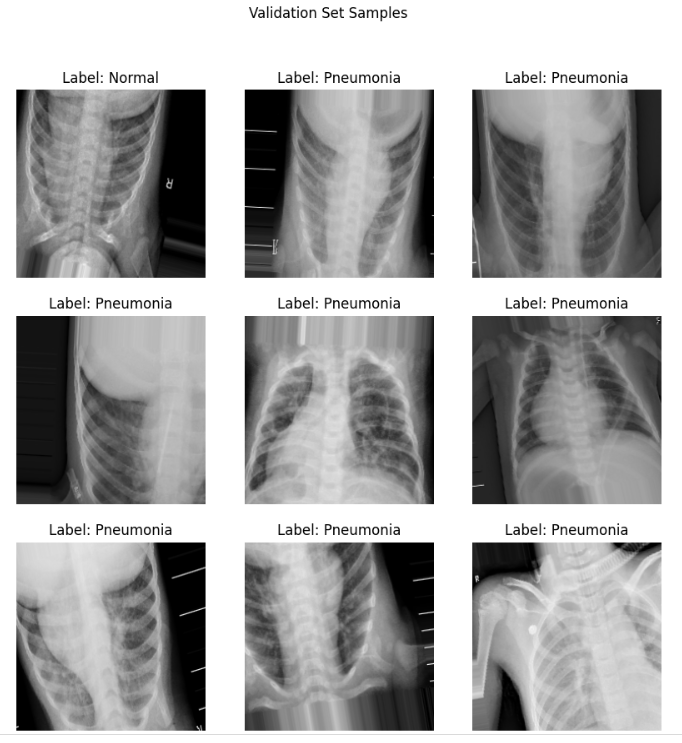
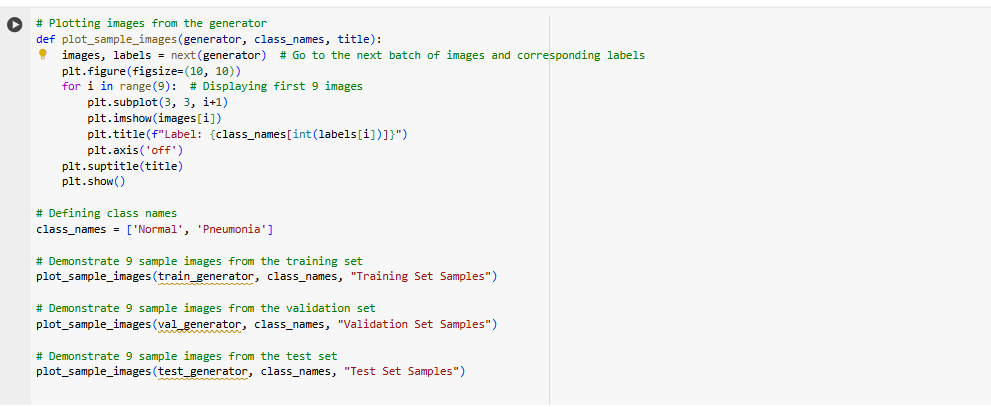


Figure 1.3 Samples from the Dataset

After downloading the dataset from Kaggle the code was used to demonstrate various samples from the dataset to check them.

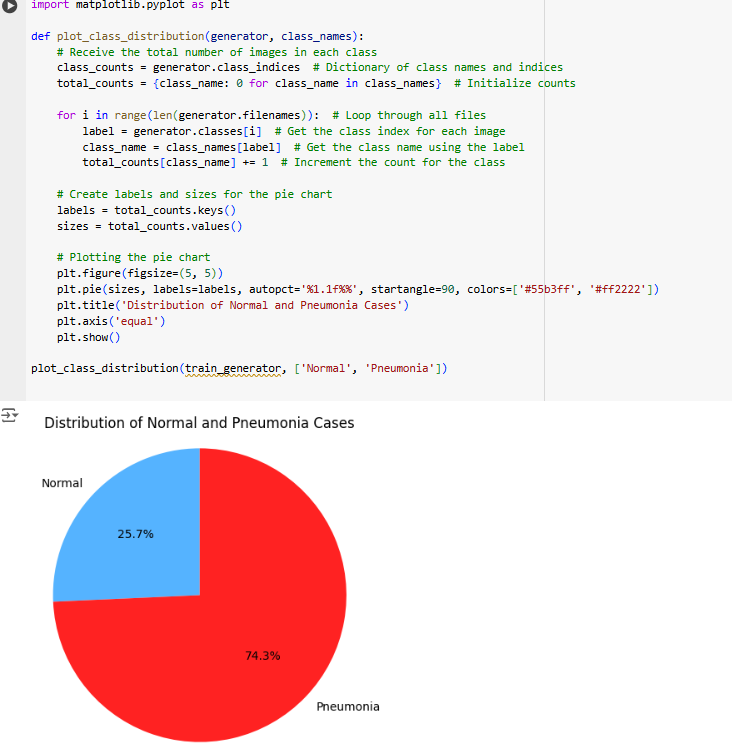


Figure 1.4 Distribution of Normal and Pneumonia Cases in the Dataset

In the dataset mainly there are two cases one of them is called as Normal that indicates health situation of Chest X-Ray images , about Pneumonia case which implies unhealthy situation of Chest X-Ray images then majority (%74.3) of the dataset is built up from Pneumonia cases as Figure 1.3 is reported.

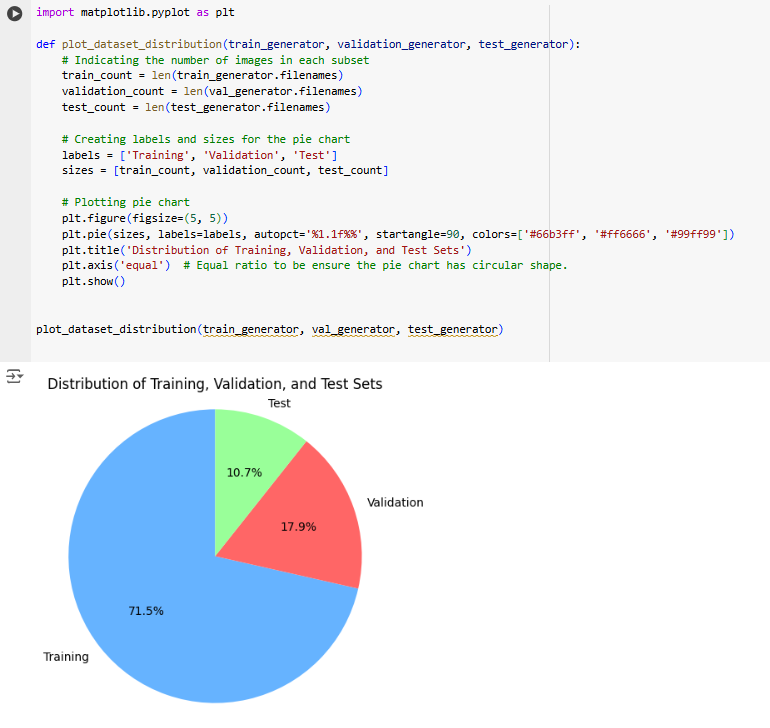


Figure 1.5 Distribution of Training-Validaiton-Test

Another analysis is regarding to the dataset is related to its sturcutre as Figure 1.4 shown training,validation and test parts are sharing the dataset by %71,5 , %17.9 and %10.7 respectively.

Training part is critical for training step of seleceted ML algorithms on the other hand validaiton part is essential to adjust vairous hyperparameters of the model and test part is important across unibiased evaluation of trained dataset.

# CHAPTER TWO NEURAL NETWORK ARCHITECTURE DESIGN

Recently, machine learning (ML) has become very widespread in research and has been incorporated in a variety of applications, including text mining, spam detection, video recommendation, image classification, and multimedia concept retrieval (Rozenwald MB. et al.,2020).Moreover Deep Learning (DL) is the most popular algorithm in the ML areas becuase it is able to provide excellent opportunuties across hardware technologies such as High Performance Computing (HPC) ( Potok TE. et al. , 2018).

From general perspective Computer Visison area is feed by deep – learning , machine learning and artifitical intelligence and also which can be illustrated like

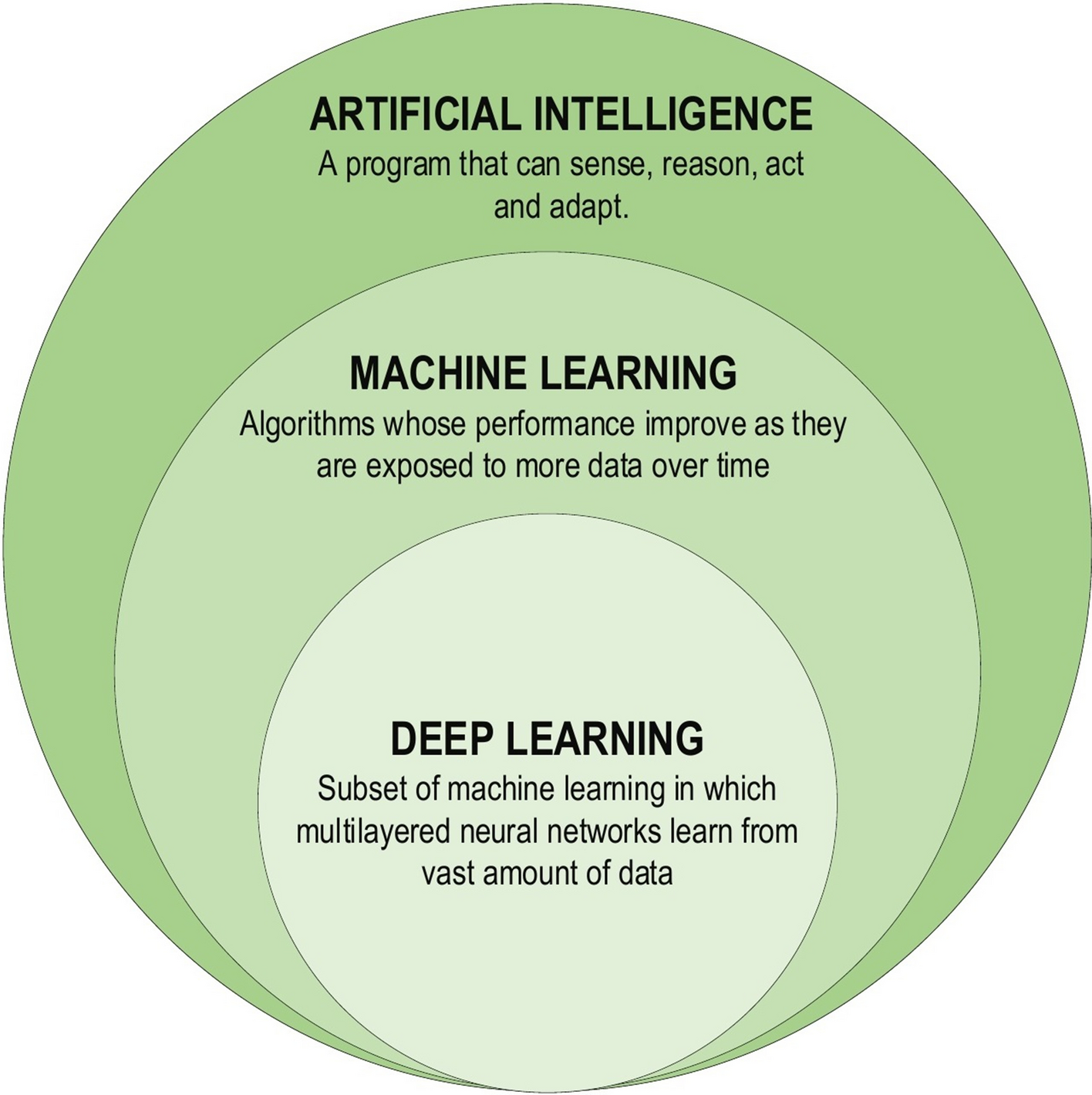


Fİgure 2.1 Relationship Between DL-ML-AI (Alzubaidi, L. et al. , 2021)

## 2.1 CONVOLUTIONAL NEURAL NETWORK (CNN)

Comvolutional Neural Network (CNN) is one of the most frequent utilized method within deep learning area becuase that is able to recognize and identitfy relationship between features wtihout any human any human intervention ( Gu J. et al. , 2018).Usage areas of CNN can be described as computer vision (Fang W. et al. , 2020) , speech recoginiton (Palaz D, Magimai-Doss M, Collobert R. , 2019) , face recognition etc. (Li HC. et al. , 2020).If there must be more specific example for that a visual cortex of cat’s brain can be simulated by CNN (Hubel DH, Wiesel TN , 1962).

## 2.1.1 CNN WORKING PRINCIPLE

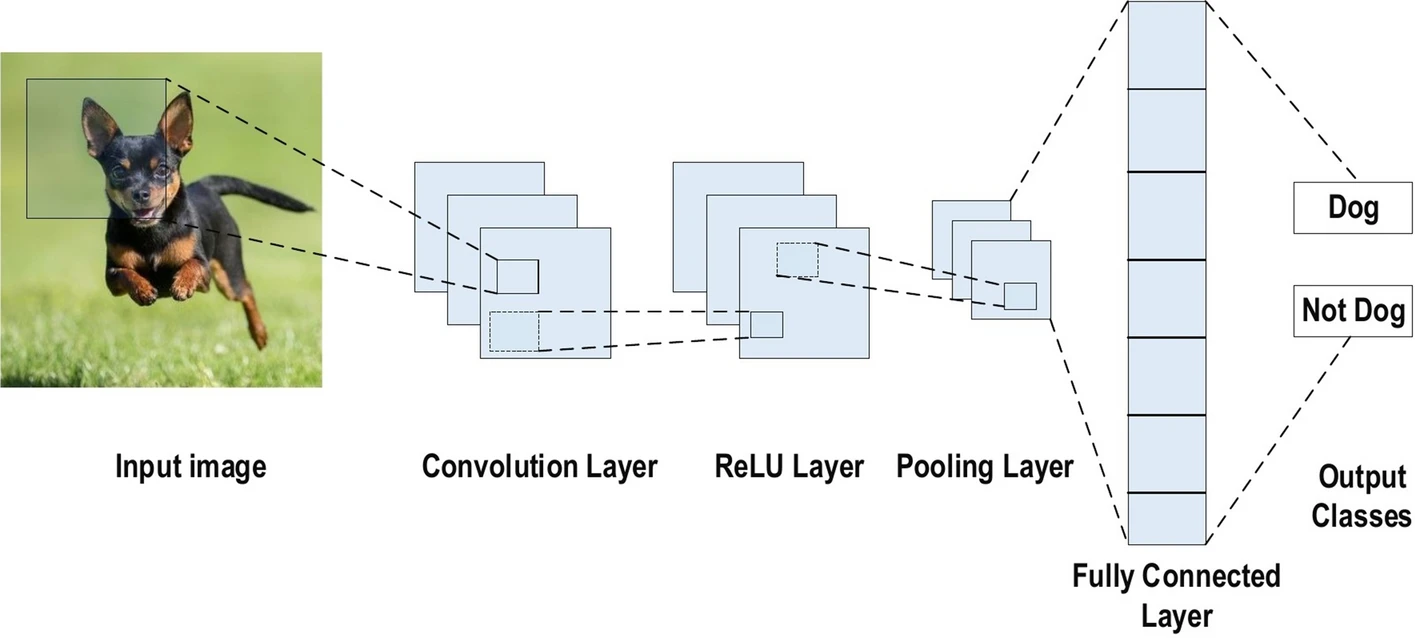


Figure 2.2 Illustration of CNN (Alzubaidi, L. et .al. , 2021)

Working principle of Convolution Neural Network is based on layers and each layer has three dimension they called as height (m) , width (m) and depth (r) then height and width must be equal to each other in the structure.Basically CNN process the input image on 5 step ;

1. Input layer receives image.
2. Convolution layer extract various features of input.
3. Pooling layer indicates the most infleunce features of input.
4. Fully connected layer classifies features.
5. Output gives as predicted results or classes.

There are various advantages to use CNN than other traditional methods ;

* CNN has less number of trainable netowrk due to weight sharing property this situation leads to enchance generalizaiton and avoid from overfiitng.
* Feature extraction and classification are of CNN causes highly organized output as well as reliant on the features.
* Large scale network implentation is more possible than other ways.

## 2.2 Pre –Trained Models

As asubtypes of deep learning networks Convolution Neural Netowrk (CNN) , Recurrent Neural Netowork (RNN) , Graph Neural Network (GNN) were utilized by scientist to handle different kind of Artifitical Intelligence (AI) tasks in recent years.Althoguh these netowrks are useful there is a big problem which is called as “Data Hungry”.To overcome this issue transfer learning has been used which aims that instead of fully train any model , humans started to learn with very sample at the beginning of specified problem then gained knowledge was trasnferred to selected transfer learning model (Thrun and Pratt , 1998 ; Pan and Yang ,2009).

After a while experts have been aware potential advantages of pre – trained models so they started to work on that (Qiu et al. , 2020).To develop pre – trained models they have been adopt self – supervised learning.For example the model should be predict “China” when masked last word of sentence of “Beijing is the capital of China”.However still pre-trained models have some problems based on meaning of words such as synonymous words.

There are many pre-trained based on CNN models available like ResNet , VGG , MobileNET etc.In this study ResNet and MobileNET utilized then comapred from various aspects.

## 2.2.1 MobileNet and ResNET

MobileNET is a pre – trained model which is based on Convlotion Neural Network (CNN) it is able to provide low energy consumption as well as shorter process time.Main usage areas are image classification , face rccognition and object detection.

ResNet is a pre-trained model which is developed by Microsoft.Especially it can achieve higher benchmark score for image classification taks.

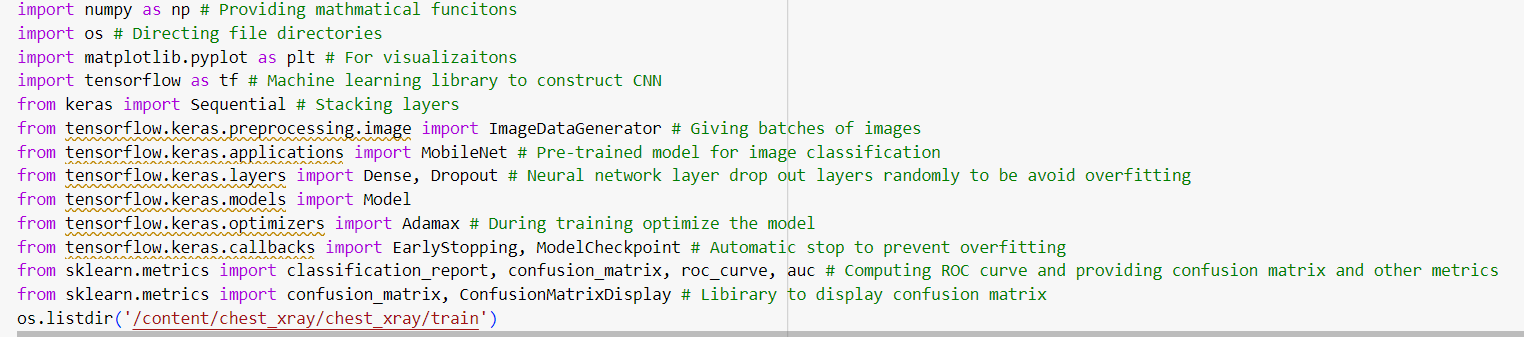
|  |  |  |
| --- | --- | --- |
| Feature | ResNET | MobileNet |
| Architecture | Very Deep Layers | More Superficial |
| Performance | Higher Accuracy | Faster but Less Accuracy |
| Computational Power | High | Low |
| Training Period | Long | Short |
| Usage Area | Large Scale Applicaitons | Suitable for Mobile Systems |
| Transfer Learning | Specialized on Higher Accuracy | Specialized on Higher Adoptation |
| Depth Issue | Fix Lost Gradient Issue | Restricted |

Table 2.1 Comparision of ResNET and MobileNet

# CHAPTER THREE CODE IMPLEMENTATION

## 3.1 MobileNet

In this study will be realized AI-Driven Pneumonia Diagnosis by using Chest X-Ray dataset.This part will explain each steps during the code implementation which includes various stages from library to performance results.

Figure 3.1 Uploading Necessary Libraries

Before start the implementation necessary libraries must be uploaded and they are needed because of mathematical functions , file operations , visualizations , implement ML algorithms (in study case pre-trained models) , avoid overfitting , early stopping and benchmark (performance) results.

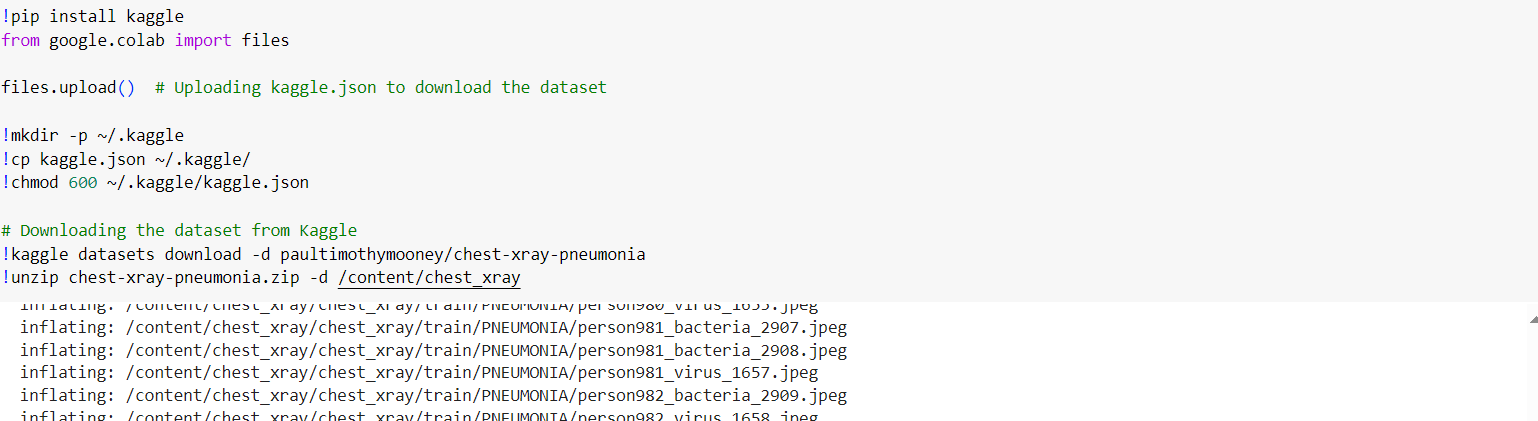
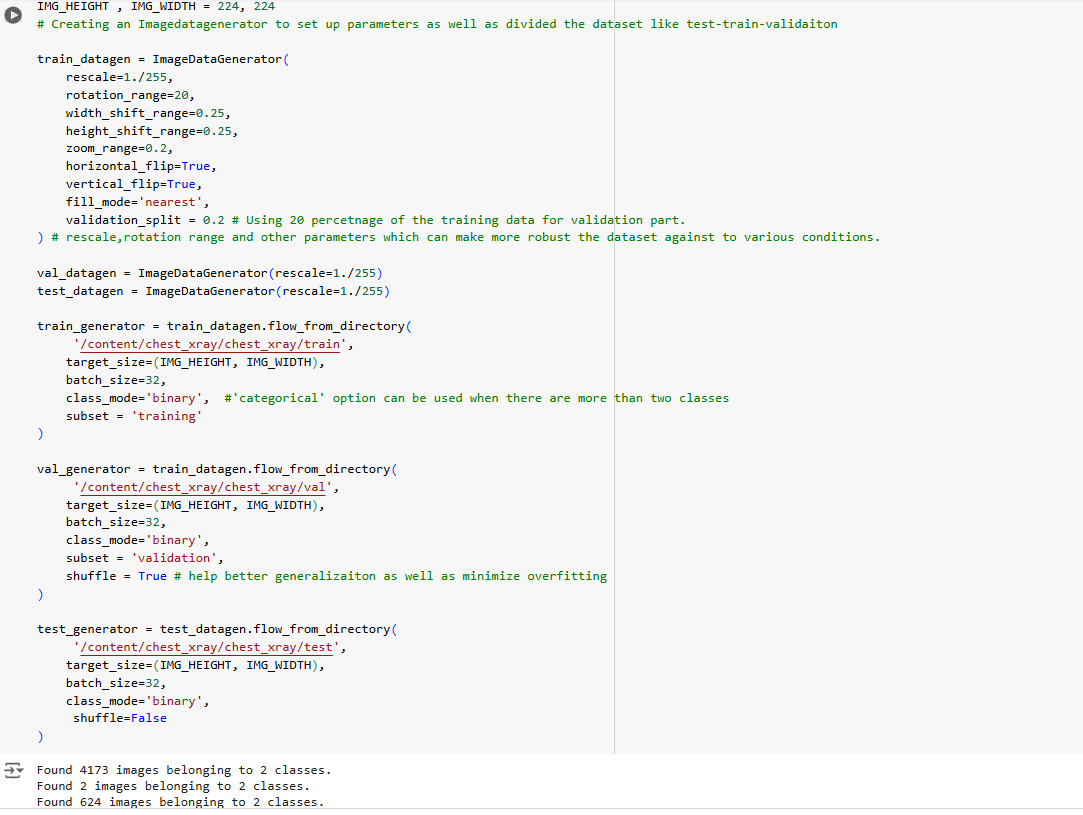


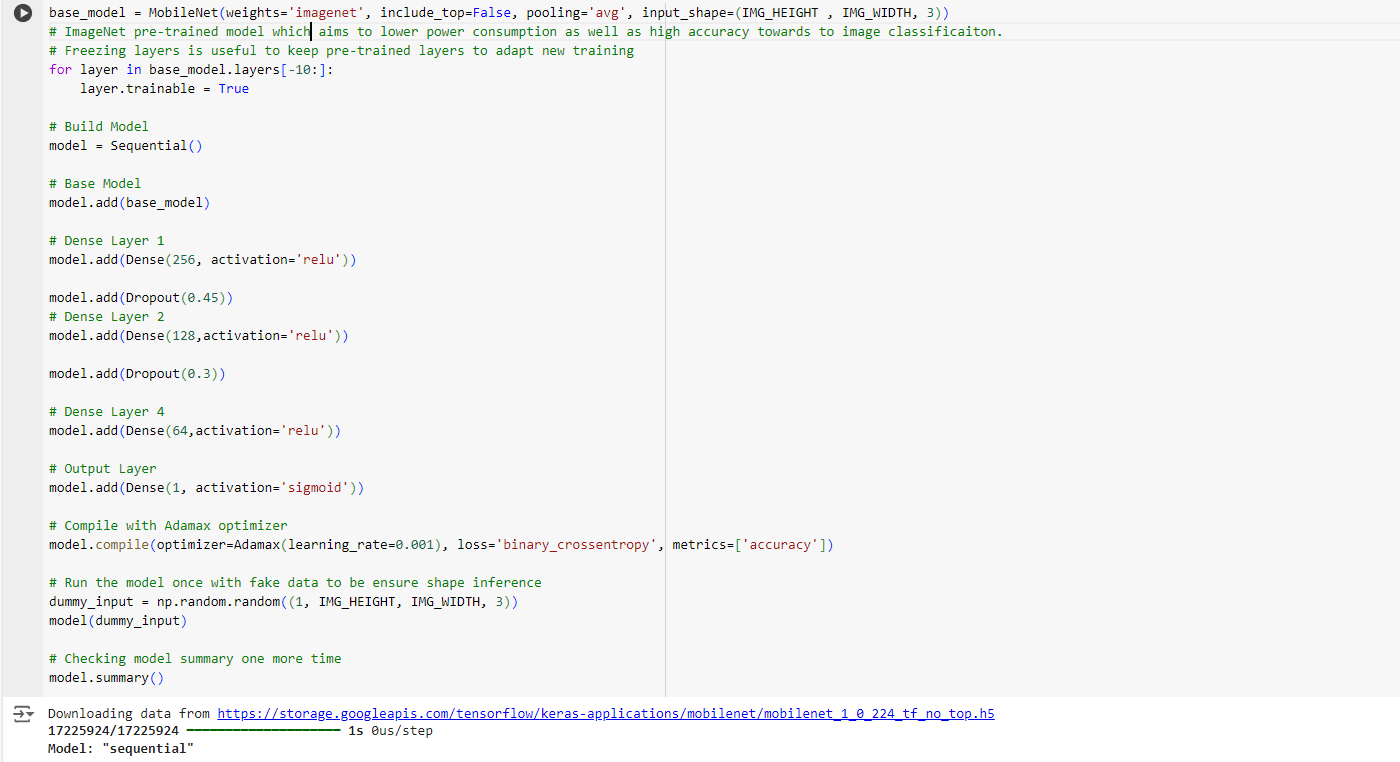
Figure 3.2 Downloading the Dataset from Kaggle

To download Chest X-Ray dataset from Kaggle directly by Kaggle token (.json file) into Google Colab environment is much easier than uploading to Google Drive – downloading from Google Drive circle due to high download speed (up to 300 MB/sn).

Figure 3.3 Set Up Train-Validation-Test Generator

Before the training step there are some sub stages first of them setting generator as well as their paramaeters.As Figure 3.3 indicates firstly an image data generator was created with its parameters like rescale , rotation range , horizontal flip and so on.These parameters are able to make the dataset more robust across to real life situations to get as much as possible accuracy score.

Train part and validation part of the dataset was used for training stage of the model then test part evaluated trained data unbiasedly.Moreover batch size is 32 for all of them which means that in each step the model will be trained by 32 batches.Lower number of batches leads to slower training but higher generalizaiton at the end of the day also “Shuffle=True” option helps to avoid overfitting.

Figure 3.4 Building MobileNet Pre-Trained Model

After setting up generators pre trained build was built with its some additional properties.For instance drop out funciton is used since to be avoid from overfiitng on the other side activation=relu command indicates that a neuron in the model fire or not baiscally also which is able to avoid vanishing graidents during training.

Sigmoid fucntion is related output of the training as binary classification (0 or 1).Furthermore it is less successfull than relu still sigmoid has smooth graident.

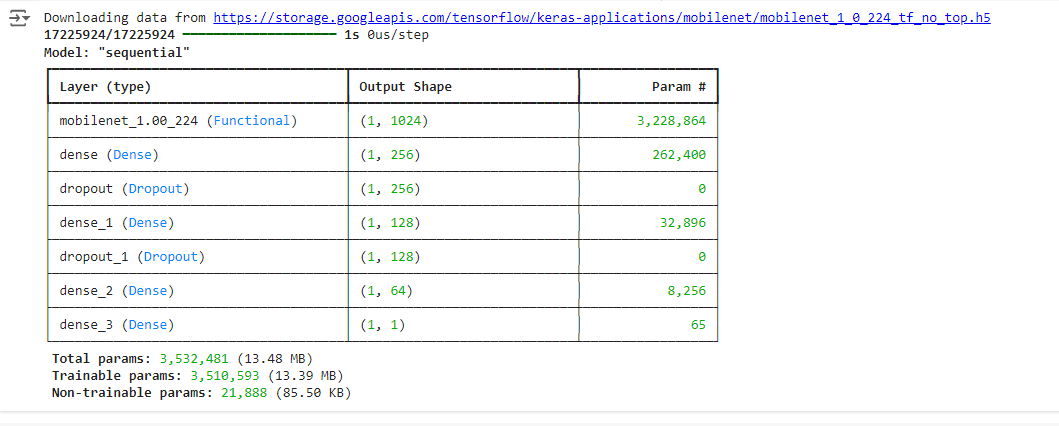
Figure 3.5 Parameters and Shape of Layers

Figure 3.5 shows that parameters as well as output shape of layers and over %99 of total parameters are able to trained by the model.

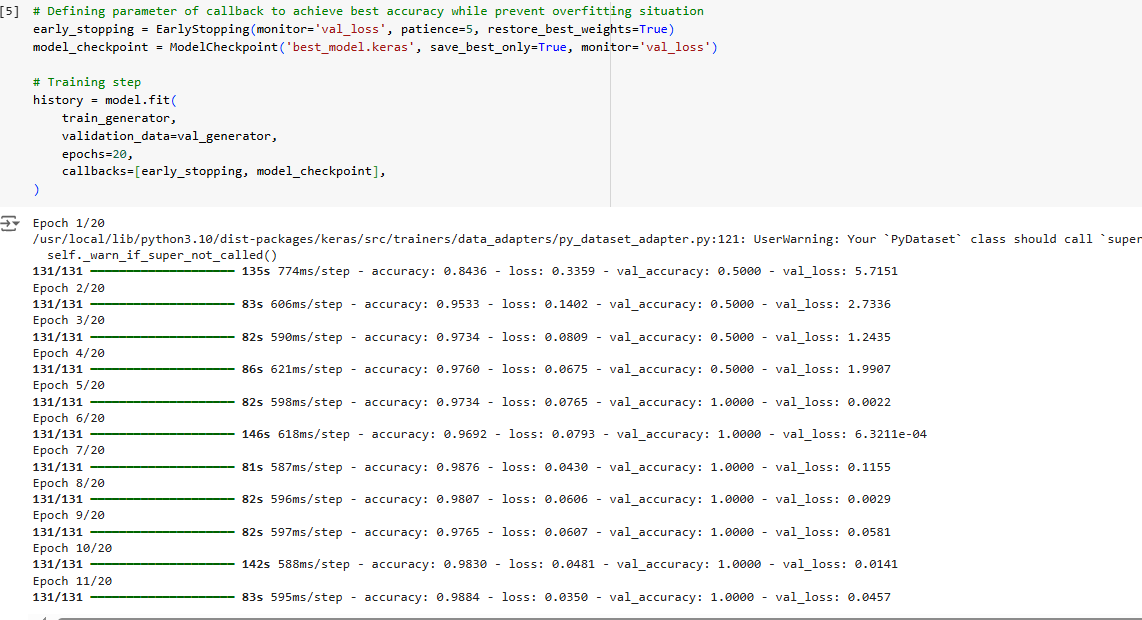
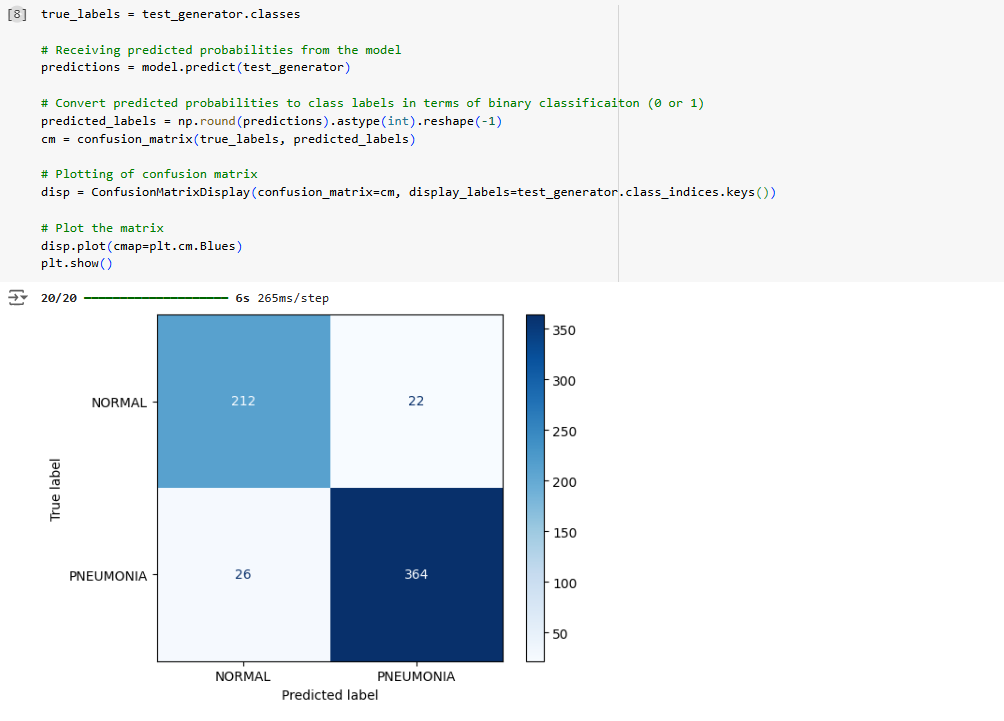
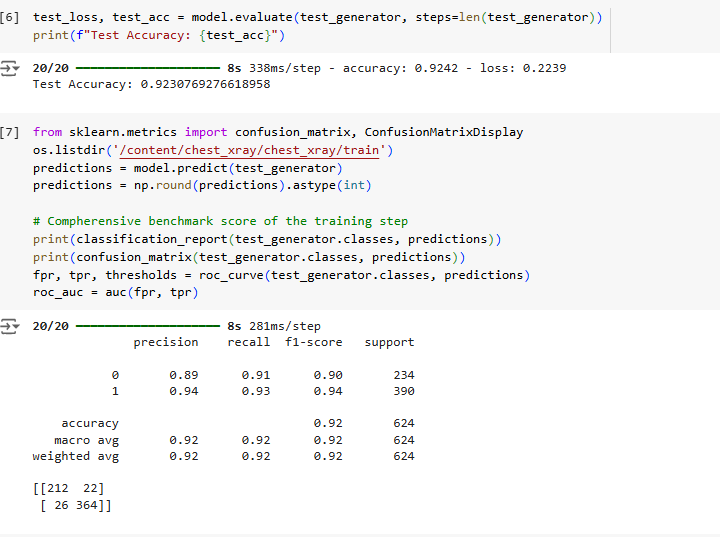


Figure 3.6 Training Step of the Model

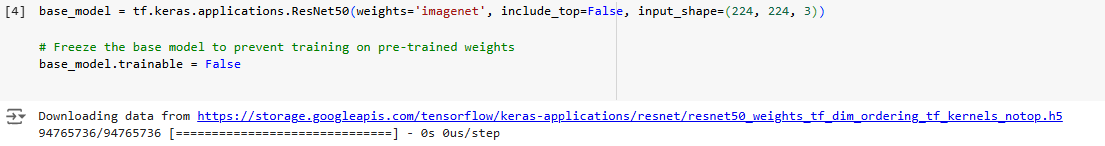
After complete pre-trained steps and check the parameters of model’s layer training step was handled.Furthermore in the training part which consists EarlyStopping method to be avoid from overfitting by monitoring validation loss.Number of epoch is selected 20 for training stage and each epoch is processed via 131 steps.Until epoch 11 the dataset is trained by MobileNet then the model was stopped by EarlyStopping to prevent overfitting and overfiitng can be described as the model learns very well than needed this situation leads to inaccurate results when the model face any data out of the dataset in real life.When loss – validation loss and test loss decreased wihch means overfitting occurs so lower performance across generalization.On the other side as complexity of the model increase at the same time possibility of overfitting increases so basic model might be good option to prevent overfitting.

Figure 3.7 Illustration of Confusion Matrix by Technically and Visually

Confusion matrix is a benchmark tool to evaluate the performance result of machine learning algorithms towards to classification problem across two or more classes then which was proposed in 1940s and 1950s for signal detection.In this code implementation which was visualized as numerically as well as visually as a result numerical result can be confirmed by calculating manually with using visually report (right side).

## 3.2 ResNET

In this study as mentioned before there are 2 pre-trained model was selected on of them is MobilleNet other is ResNET when one part of the all the code which is able to swtiched to another pre-trained model.

Figure 3.8 Switching Pre-Trained Model

From Tensorflow library by using this code part selected pre-trained model can be changed also all of other preprocess of training is same as MobileNet and still EarlyStopping feature was used to prevent overfitting.

# CHAPTER FOUR MODEL EVALUATION

## 4.1 Definition of Metrics

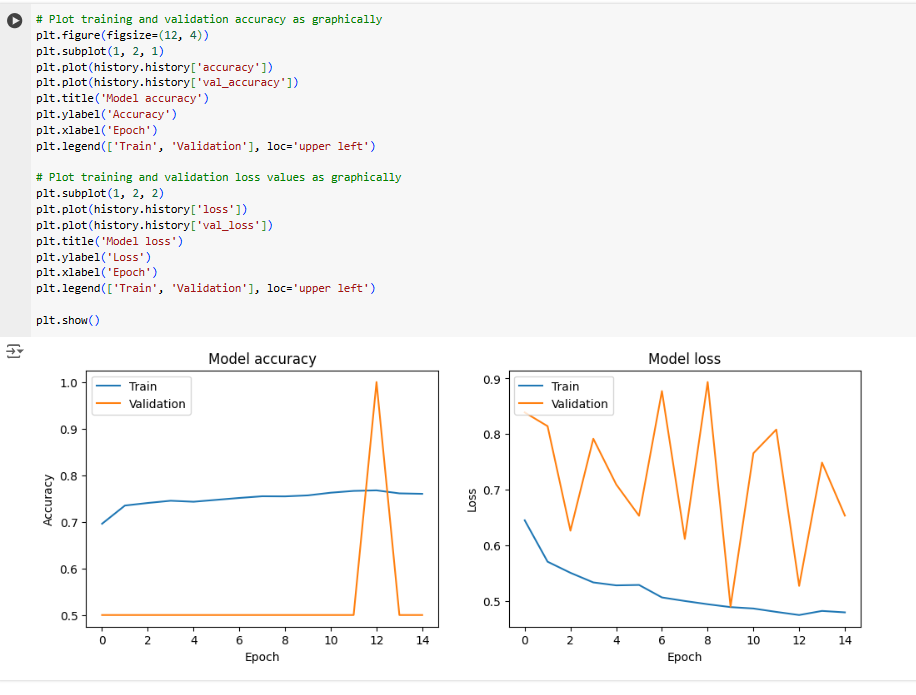
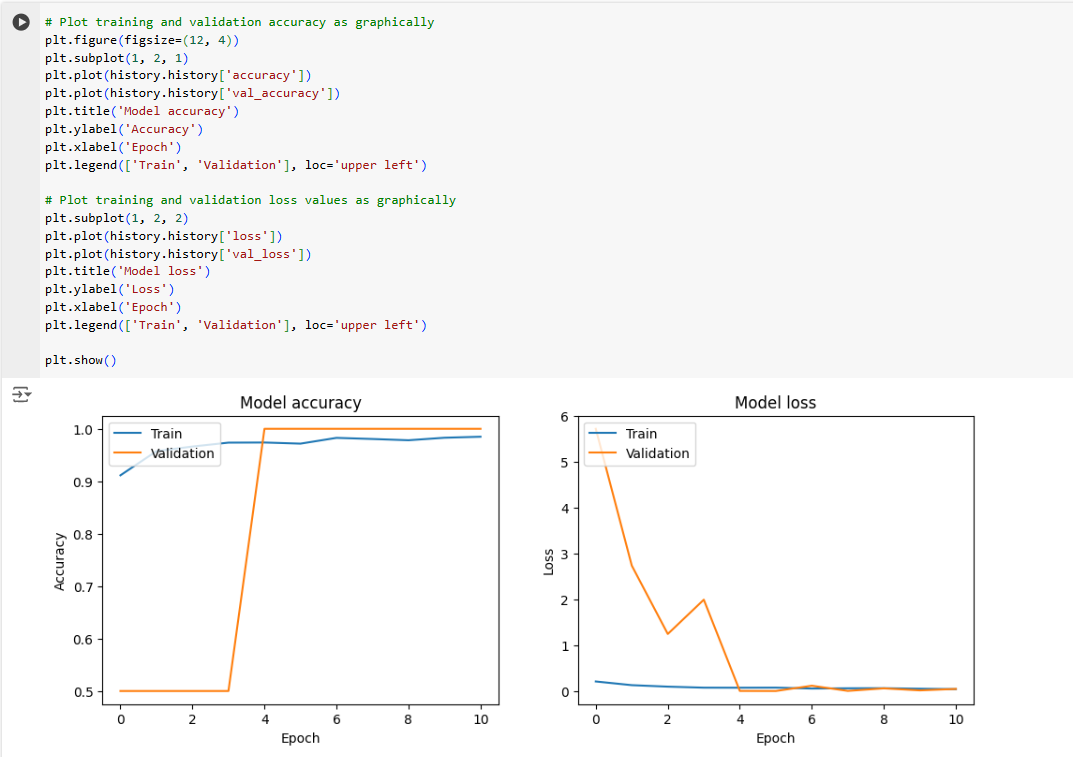
Evaluating performances of various machine learning algorithms depend on vairous benchmark scores such as accuracy , precision , recall , F-1 score.

* Accuracy is one of the most popular metrics to evaluate performance.Basically it computes all predicitons out of all predictions made.
* Precision computes all true positive predictions which means that main focus area quality of the positive predictions.
* Recall tries to find capacity of the model regarding to identify all positives.
* F-1 score can be considered as harmonic mean of recall and precision.When there is need of false positive and false negatives.

Although these metrcis are essential to evaluate the performance of the model still there are some considerations should be taike into account ;

* Accuracy might not be suitable when dataset has imbalanced structure.
* Precision is critical when there is scenario like cost of false positives high.
* Recall is crticial is when there is scenario like cost of false negatives is high.
* The best time to use F-1 when precision and recall both of them have importance.

## 4.2 Comparision Results of MobileNet and ResNET

Figure 4.1 Graphics of Accuracy and Loss Results of MobileNet – ResNET Respectively

* MobileNet shows more stable training results as well as validation accuracy on the other side ResNET isn’t stable as much as MobileNet so which causes overfitting situation.One of the major concern about overfitting is not generalize enoughly.

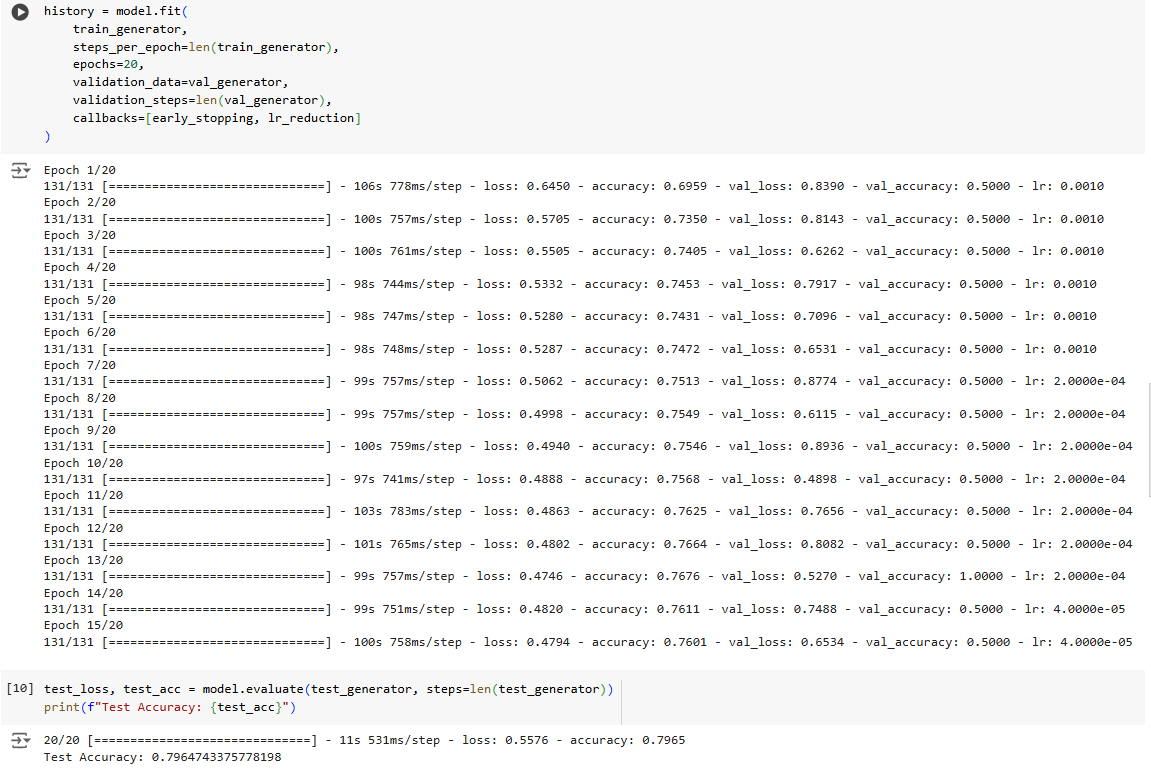
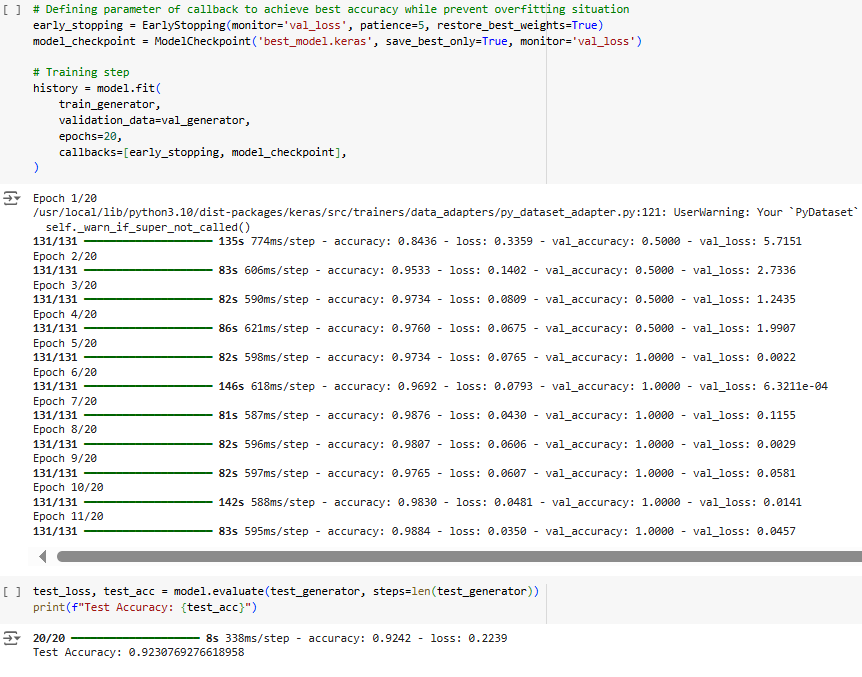


Figure 4.2 Test Accuracy Results of MobileNet – ResNET Respectively

* Both model stopped training before the final epoch to avoid overfitting and end of the day MobileNet has higher test accuracy (%92.30) than ResNET (%79.64).

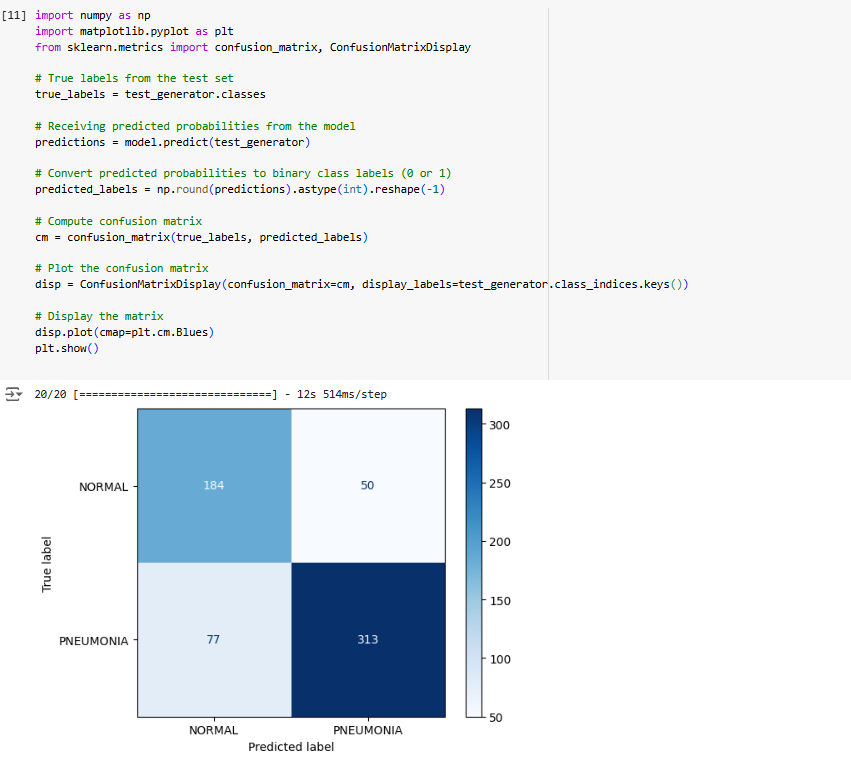
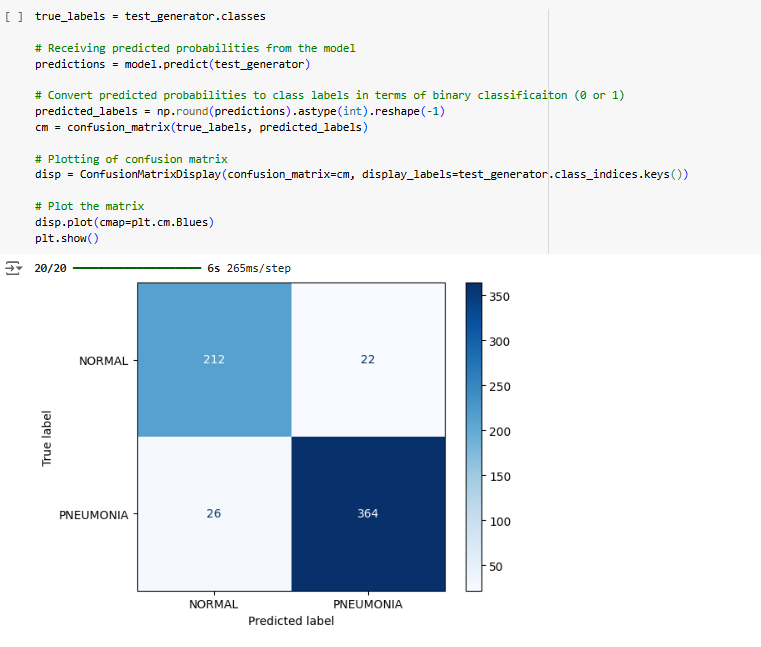


Figure 4.3 Confusion Results of MobileNet – ResNET Respectively

* MobileNet seems to have bigger advantage over ResNET in terrms of identify the more positive values at the same time decreasing number of FP and FN.

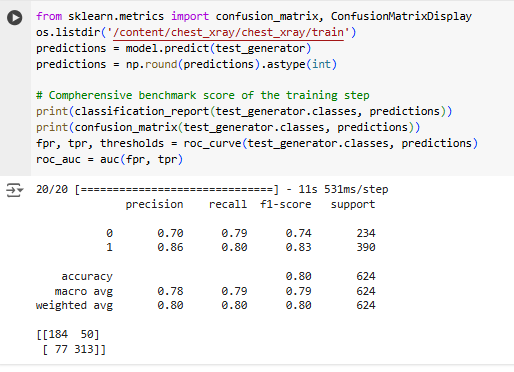
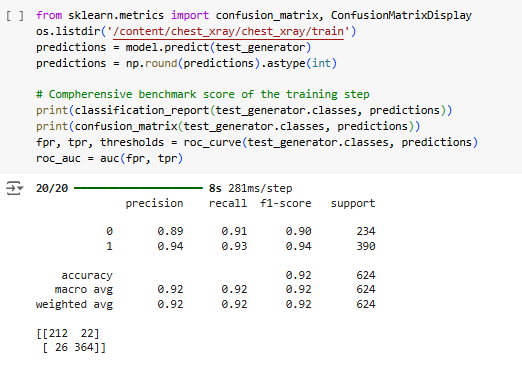


Figure 4.4 Confusion Matrix Display Results of MobileNet and ResNET Respectively

* Figure 4.4 implies that precision , recall and F-1 score with macro average of two models as well as their weighted average then as other previous figures support which is clear that MobileNet is able to demonstrate better performance than ResNET.

# CONCLUDING REMARKS

In that study two pre – trained models are handled across Chest X-ray dataset which are MobileNet and ResNET.During the pre- process of the model implementation there were no abnormal situation regarding to structure of the dataset also majority of the dataset is built up from pneumonia cases as well as traning part has highest share all whole of the dataset.

During the training stage to be avoid from overfitting situation EarlyStopping function is used.After this stage performance (benchmark) results are obtained.According to them MobileNet is able to demonstrate more satisfactory results than ResNET.On the other side there is a consideraiton about ResNET which is Colab environment has some issues related to ResNET algorithm in terms of hardware consistency since sometimes ResNET algorithm doesn’t work on Google Colab such as during the training stage epoch progress can’t continue properly.Moreover to overcome this situation local systems can be used with GPU but extra libraries are needed for that.Unfortanetly installing these libraries is complex because of different laptop configurations.

In a conclusion , as a sub-area of AI computer vision has critical role in heaşth industry as well as there are many potential huge impacts for the future and in academia there are many contiuned studies to either improve existed machine learning or developing new machine learning algorithms.

# BIBLIOGRAPHY

1. Alzubaidi, L., Zhang, J., Humaidi, A.J., Al-Dujaili, A., Duan, Y., Al-Shamma, O., Santamaría, J., Fadhel, M.A., Al-Amidie, M. and Farhan, L., 2021. *Review of deep learning: concepts, CNN architectures, challenges, applications, future directions.* *Journal of big Data*, *8*, pp.1-74.
2. Fang, W., Love, P.E., Luo, H. and Ding, L., 2020. *Computer vision for behaviour-based safety in construction: A review and future directions.* *Advanced Engineering Informatics*, *43*, p.100980.
3. Gu, J., Wang, Z., Kuen, J., Ma, L., Shahroudy, A., Shuai, B., Liu, T., Wang, X., Wang, G., Cai, J. and Chen, T., 2018. *Recent advances in convolutional neural networks.* *Pattern recognition*, *77*, pp.354-377.
4. Hubel, D.H. and Wiesel, T.N., 1962. *Receptive fields, binocular interaction and functional architecture in the cat's visual cortex*. *The Journal of physiology*, *160*(1), p.106.
5. Li, H.C., Deng, Z.Y. and Chiang, H.H., 2020. *Lightweight and resource-constrained learning network for face recognition with performance optimization.* *Sensors*, *20*(21), p.6114.
6. Mosquera, C., Lara, M.A.R., Díaz, F.N., Binder, F. and Benitez, S.E., 2023. *Introducing Computer Vision into Healthcare Workflows*. In *Digital Health: From Assumptions to Implementations* (pp. 43-62). Cham: Springer International Publishing.
7. Palaz, D., Magimai-Doss, M. and Collobert, R., 2019. *End-to-end acoustic modeling using convolutional neural networks for HMM-based automatic speech recognition.* *Speech Communication*, *108*, pp.15-32.
8. Potok, T.E., Schuman, C., Young, S., Patton, R., Spedalieri, F., Liu, J., Yao, K.T., Rose, G. and Chakma, G., 2018. *A study of complex deep learning networks on high-performance, neuromorphic, and quantum computers*. *ACM Journal on Emerging Technologies in Computing Systems (JETC)*, *14*(2), pp.1-21.
9. Rozenwald, M.B., Galitsyna, A.A., Sapunov, G.V., Khrameeva, E.E. and Gelfand, M.S., 2020. *A machine learning framework for the prediction of chromatin folding in Drosophila using epigenetic features*. *PeerJ Computer Science*, *6*, p.e307.
10. Thrun, S. and Pratt, L., 1998. *Learning to learn: Introduction and overview*. In *Learning to learn* (pp. 3-17). Boston, MA: Springer US.
11. Qiu, X., Sun, T., Xu, Y., Shao, Y., Dai, N. and Huang, X., 2020. *Pre-trained models for natural language processing: A survey.* *Science China technological sciences*, *63*(10), pp.1872-1897.
12. <https://github.com/TuranCANG/Turan-Can-G-n-BSBI-Projects.git> (codes are available there)

# APPENDIX (if necessary)

AI : Artfitical Intelligence

CNN : Convolutional Neural Networks

DL : Deep Learning

GNN : Graph Neural Netowrk

GPU : Graphics Process Unit

ML : Machine Learning

RNN : Recurrent Neural Network