Multiclass classification of Covid-19, Tb, Pneumonia, and health cases using Deep Learning.

MSc Research Project
Data Analytics

Aditya Pramod Shinde
Student ID: 20178883

School of Computing
National College of Ireland

Supervisor: Prof. Vikas Sahni

National College of Ireland
Project Submission Sheet
School of Computing

Student Name:	Aditya Pramod Shinde
Student ID:	20178883
Programme:	Data Analytics
Year:	2022
Module:	MSc Research Project
Supervisor:	Prof. Vikas Sahni
Submission Due Date:	15/08/2022
Project Title:	Multiclass classification of Covid-19, Tb, Pneumonia, and health cases using Deep Learning.
Word Count:	6231
Page Count:	15

I hereby certify that the information contained in this (my submission) is information pertaining to research I conducted for this project. All information other than my own contribution will be fully referenced and listed in the relevant bibliography section at the rear of the project.

ALL internet material must be referenced in the bibliography section. Students are required to use the Referencing Standard specified in the report template. To use other author's written or electronic work is illegal (plagiarism) and may result in disciplinary action.

Signature:	APShinde
Date:	15th August 2022

PLEASE READ THE FOLLOWING INSTRUCTIONS AND CHECKLIST:

Attach a completed copy of this sheet to each project (including multiple copies).	☐
Attach a Moodle submission receipt of the online project submission, to each project (including multiple copies).	☐
You must ensure that you retain a HARD COPY of the project, both for your own reference and in case a project is lost or mislaid. It is not sufficient to keep a copy on computer.	☐

Assignments that are submitted to the Programme Coordinator office must be placed into the assignment box located outside the office.

Office Use Only	
Signature:	
Date:	
Penalty Applied (if applicable):	

Contents

List of Figures

List of Tables

Multiclass classification of Covid-19, Tb, Pneumonia, and health cases using Deep Learning.

Aditya Pramod Shinde
20178883

Abstract

In recent times we have been made aware of the high infection rate and lethality that can be caused by the outbreak of pulmonary diseases like Covid-19. The need for an economical and quick diagnosis of such a disease is also increasing. It is important to consider that a disease like covid has a resemblance to some of the other pulmonary diseases like pneumonia and tuberculosis and has overlapping symptoms and characteristics with them. There have been many previous types of research that have used chest x-rays for the classification of covid and healthy cases or covid and pneumonia or covid and TB using binary classification. This research proposes a deep learning framework for multiclass classification of Covid-19, pneumonia, Tuberculosis, and normal cases with the use of chest X-rays. In this research, I have used a combination of three open datasets available on Kaggle to create my dataset for research which has a total of 4136 images. I have used image data augmentation and applied pre-trained models like VGG16 and DenseNet121. I have also created a CNN from scratch which can perform classification with 91 per cent overall accuracy and high precision and an F1 score which I have discussed further in the report. This investigation can help physicians and patients in early diagnosis and can act as a pre-diagnosis method to start early treatment or isolation of the patient. It can also act as a patient prioritization tool to help physicians focus on people that have a high probability to be infected in case of a high number of patients due to a sudden outbreak.

1 Introduction

The recent pandemic caused by Covid-19 has taught us the importance of precautions and early diagnosis. This has also created a requirement for affordable and alternate preliminary diagnosis methods for early detection to contain disease outbreaks. Chest x-rays are affordable and good diagnostic methods. However, diseases like Covid-19, Tuberculosis, and Viral pneumonia have close resemblance and overlapping characteristics and symptoms making it difficult even for physicians to differentiate between these diseases solely based on chest x-rays Bai et al. (2020). Covid -19 is the most infectious disease followed by Tuberculosis which infected around 10 million people in the year 2020. There have been many types of research that have done binary classification of covid and healthy cases, covid vs pneumonia or covid vs TB. I have decided to perform a multiclass classification of Covid-19, pneumonia, TB, and healthy cases to provide an ability to differentiate these diseases properly from others and provide a proper pre-diagnosis using chest x-rays. The factors that affect this research are availability of labelled Xray datasets and having balanced classes. The dataset used has balanced classes and utilises data augmentation. The research question posed in this research investigates how convolutional neural networks and pre-trained models can be employed for multiclass classification of Covid-19, Pneumonia, Tuberculosis, and Normal cases from chest X-rays. To address the research question I have developed a convolutional neural network from scratch and have used pre-trained models like VGG16 and DenseNet121 along with some modifications and have evaluated these models on metrics such as accuracy, precision, recall and f1 score. I have also derived a confusion matrix of predictions for all models to better understand the performance with respect to each class. The major contribution of this research is a convolutional neural network capable of performing multiclass classification of closely related and difficult-to-differentiate diseases which can be used as a pre-diagnosis and suggestive tool for physicians to start early treatment or for taking precautionary measures like isolation. This tool can also be used for the purpose of patient prioritization for further diagnosis by physicians so that the people in need can get timely treatment when there is an outbreak, which drastically increases the number of patients that a physician has to attend to. In this research 2 covers the study of related work, 3 covers the methodology of the research, 4 has the design specifications, 5 has project implementation, 6 has results and discussion of the research findings and 7 covers conclusion and future work.

2 Related Work

I have tried to study the related work necessary for my project and divided it into 3 section 5.1 Existing research related to this topic , 5.2 Pre-Trained Models and 5.3 Model Evaluation Techniques.

2.1 Existing research related to this topic

The paper proposed by Mahbub et al. (2022) has focused on creating a lightweight deep CNN model that can be used on chest x-ray to classify pulmonary diseases. The deep CNN created is capable of binary classification. This CNN was tested on 6 different scenarios which are covid vs healthy cases, pneumonia vs healthy, TB vs health, covid vs pneumonia, covid vs TB and pneumonia vs TB. The binary classification has provided great results with all scenarios having above 98% accuracy. The literature gave a detailed description and summary of the model used and the figures used in the paper are very informative and well structured. The problem undertaken by this research paper is very inspiring but the approach used is binary classification instead of multiclass classification which is the motive of our research. This research provides ample comparative metrics and assessment of all models to test the performance.

Automatic analysis of Xrays to classify covid, pneumonia and TB can directly reduce the efforts and time taken for diagnosis. The Paper written by N et al. (2022) has proposed a deep CNN for multiclass classification of covid, pneumonia, TB and normal condition with the help of chest Xrays. The authors have used two datasets available on Kaggle. However, the amount of training and testing data that is used is relatively less for building a CNN. The model was able to achieve an accuracy of 87.41 per cent which is very good when it comes to multi-class classification. The report put forward lacks a proper comparison of the model alongside alternatives and only focuses on CNN. The description of the model that is mentioned is just a higher-level view which is not sufficient to faithfully reproduce the model. This model can be tested on a larger dataset for its proper validation but its description makes it a difficult task for other researchers.

Yadav et al. (2021) have suggested a deep unsupervised framework for lung disease classification utilizing X-rays and CT images. They have proposed a novel framework that uses unlabeled data to detect lung diseases with the help of a multi-layer generative adversarial network which is called Lung-GAN. In the medical and healthcare industry procuring a high volume of labelled imaging data is the biggest challenge that Yadav et al. (2021) have decided to overcome with this technique. In this CNN is used for image representation and a stacking classifier is utilized alongside Linear SVC for the classification problem. This system can achieve 94-95 per cent accuracy by using unlabelled data. The availability of a labelled dataset allows us to explore other models however the base motive of classification of four closely resembling images and its feasibility is proven by this research.

A hybrid approach using fuzzy logic and modular artificial neural network is explored byVarela-Santos and Melin (2021). The chest x-rays of pneumonia and lung nodules are analyzed and their features are divided by the modular neural network. The images are first pre-processed using image segmentation to get the region of interest and then feature extraction is performed on these images which are fed to a feed-forward neural network which analyses its features. The fuzzy inference is connected to the output of the neural network which classifies the images. This hybrid model was able to achieve $95 - 98$ per cent accuracy. This model is very complex yet the research gave us insight into hybrid models and their use in our field of research.

The research by Ibrahim et al. (2021) makes use of x-rays along with CT scan images for multiclass classification of covid19, pneumonia, and lung cancer. The deep learning methods explored in this research are ResNet152V2 + Bidirectional GRU, VGG19-CNN, ResNet152V2, and ResNet152V2 + Gated Recurrent Unit. This research has used data augmentation to balance the classes in the dataset. The model that performed the best was the use of VGG19 along with CNN. The methodology of this research is computationally expensive for implementation.

A deep CNN-based architecture called CovXNet is used by Mahmud et al. (2020) which is previously trained on a large pneumonia dataset, and transfer learning is applied with some fine tuning is used on the covid x-ray dataset. A stacking algorithm is used for the optimization of the model. The images are resized and min-max normalization is performed on these images for pre-processing dilated depth-wise convolutions are used in this architecture for feature extraction purpose and the stacking algorithm boosted the optimization of the proposed architecture.

The SqueezeNet is used by Ucar and Korkmaz (2020) for Covid-19 diagnosis. This paper discussed how the learning abilities of CNN help achieve high accuracy in image classification problems. At the time of this research, Covid labelled dataset was not available easily and the researcher got to work with an imbalanced dataset to classify covid, pneumonia and healthy cases. Data augmentation was used to resolve class imbalance and SqueezeNet was finetuned with Bayesian optimization. I have a larger dataset available which makes it an advantage for my research.

The binary classification of Covid and healthy cases was performed by Jain et al. (2021) who compared the performance of multiple CNN models. This research made use of Inception V3, Xception, and ResNet. To prevent overfitting of any classes data augmentations like rotating, zooming and shearing have been used. The activation used is LeakyReLU which helps with the problem of dead neurons. The models were evaluated on the metrics of accuracy and Xception provided the best accuracy of 97.97 per cent.

An automated screening model for covid-19, pneumonia and healthy patients is suggested by Das et al. (2021) with the help of ResNet50 and VGG16. The researcher has also used the TLCoV-CNN model for covid detection and used transfer learning has been applied. In this research, VGG16 gave the highest accuracy of 97.67 per cent and an average ROC-AUC of 0.97 was obtained. This research shows the high efficiency of VGG16 used with transfer learning for chest x-ray classification which I have applied in my research as well.

The research performed by Demir et al. (2022) has utilized a Convolutional autoencoder for the extraction of features and an SVM classifier optimized by the Bayesian algorithm is used for the classification of Covid-19 pneumonia and healthy cases. The images were resized to 100 * 100 pixels and Laplacian-based gradient operations are applied to the image and a novel algorithm SDAR is used for feature selection. The model was evaluated on basis of precision, sensitivity, accuracy and F1 score and the model performed with an accuracy of 99.75 per cent. The approach of using an SVM classifier with a bayesian algorithm helped in optimizing the model but this also increased the complexity of this system, The research proposed by Al-antari et al. (2021) has used the YOLO predictor or classification of 8 different respiratory diseases which are pneumothorax, cardiomegaly, nodules, cardiomegaly, atelectasis, pneumonia, infiltration, effusion, and masses. This model was trained on over 50000 images and evaluated with the use of five-fold cross-validation. The model was able to predict covid in real time but the training resources required for this research are very high and it took 18 hours to train each fold. The main objective of this model is to focus on the speed of classification but this makes it very heavy computationally.

A modified ResNet50V2 has been developed by Ahamed et al. (2021) which makes use of CT scans and chest x-rays for the classification of covid, viral pneumonia, bacterial pneumonia and healthy patients. The dataset is acquired from multiple sources and all the images were sharpened using a filter All of the images used in Imagenet are 224*224 pixels which is the reason for resizing the images for use in pre-trained models. The ResNet50V2 has two newly added bases modified and fine-tuned. The model was able to achieve an accuracy of 96-99 per cent which is very high for this problem.

The research by Heidari et al. (2020) makes use of two image pre-processing techniques on chest x-ray images which are removal of diaphragm and histogram equalization. The images are transformed into pseudocolours using a bilateral low pass filter. These pre-processed images are passed to a CNN which is pre-trained to classify covid, pneumonia and normal cases. The removal of the diaphragm and use of a bilateral low pass filter was able to boost accuracy up to 94-98 per cent which is 88 per cent without pre-processing. This research shows methods of improving the accuracy of a CNN model without making changes to the model.

The issues faced because of blurry, low contrast and noisy images have a direct effect on the accuracy of the prediction of any classification model. The method proposed by Caseneuve et al. (2021) helps to overcome this issue by proposing a method for cleaning the dataset before model training. In this research, an operator-based edge detector detects the boundaries and other elements of chest x-ray images. This method aims to eliminate low-quality and blurry images which are unfit for training and can degrade the model training. With this approach, the researcher was able to obtain 95 per cent accuracy.

2.2 Pre-Trained Models

A combination of Xception and ResNet50V2 has been proposed by Rahimzadeh and Attar (2020) for the classification of normal, covid and pneumonia classes using chest x-rays. This research makes use of pre-trained networks trained on the ImageNet weights. The weakness of this research is the use of a very small dataset which only has 180 X-ray images and only 31 covid class images. I have utilized a much bigger dataset and used pre-trained models with ImageNet weights similar to this research.

The classification of covid, viral pneumonia, bacterial pneumonia and healthy cases is performed by Narin et al. (2021) using 5 pre-trained CNN models which are Inception-ResNetV2, ResNet50, ResNet152, InceptionV3, and ResNet101. The best performing model is ResNet50 which gave an accuracy of 96.5 -99.7 per cent accuracy. This experiment uses 3 binary datasets along with data augmentation and all the images were resized to 224 *224 pixels for use with the pre-trained models. This research helped me to understand the implementation and pre-processing required for pre-trained models.

The research by Ismael and Şengür (2021) has explored the binary classification of covid and healthy patients using CNN for the classification of chest x-ray images. This research has used 5 pre-trained models for feature extraction and SVM has been used for classification. The models used are VGG19, ResNet18, ResNet101, ResNet50, and VGG16. The researcher also created an end-to-end CNN model which gave an accuracy of 91 per cent but ResNEt50 used along with SVM gave the highest accuracy with 95 per cent accuracy of covid and 90 per cent accuracy for healthy cases.

The use of chest x-ray for the classification of covid and pneumonia has been proposed by Gupta et al. (2021). The models used for this research are ResNet101, NASNet, MobileNet, Xception, and InceptionV3 all of which are pre-trained and fine-tuned and blended with the help of an integrated stacking technique. The fuzzy colour technique is used in pre-processing for noise reduction and enhancement of images. The pre-processing of images contributed to increasing the efficiency of the model and an accuracy score above 99 per cent was acquired. However, this model was applied on a very small dataset and application on a larger dataset is needed to prove the reliability of this technique.

2.3 Model Evaluation Techniques

The research by KC et al. (2021) has evaluated 8 pre-trained models for the classification of 5 classes using chest x-rays. The dataset was first divided into a 90:10 split in which 90 per cent of data is for training and 10 is for testing. The training set was split further into 80:20 ratios in which 80 is reserved for training and 20 for validation. All the models were trained for 1000 epochs. The DenseNet121 outperformed all these models and it was evaluated on accuracy, precision, recall and F1 score.

A multiclass classification system for covid, normal and pneumonia is proposed by Choudhuri and Paul (2021). This research has used pre-trained models using VGG-16 and a classic CNN. The database is divided into train test split and a test set is used for evaluation of the model on parameters like sensitivity, specificity, accuracy, precision f1 score and recall. This method of evaluation is also used in my research.

3 Methodology

The research methodology consists of five steps namely data gathering, data preparation, data transformation, data modelling and evaluation of results as shown in Fig. 1.

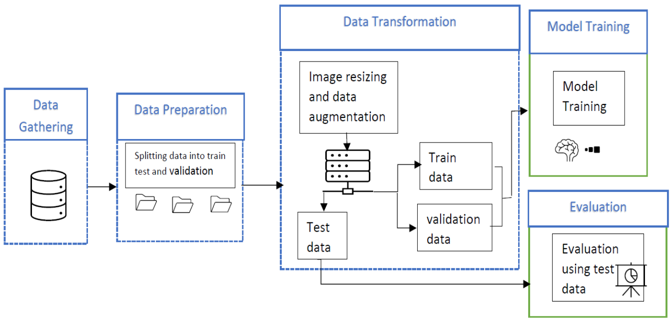

Figure 1: Research Methodology

3.1 Data Gathering

The datasets required for the classification problem are taken from 3 kaggle datasets The first dataset by Chowdhury et al. (2020) and Rahman et al. (2021) consists of Covid, pneumonia and healthy cases[1]. The second dataset by Rahman et al. (2020) consists of TB and normal cases Xrays [2] and the third dataset also contains some more TB and normal Xrays [3]. The dataset also had metadata and mask which I don't require for this research.

3.2 Data Preparation

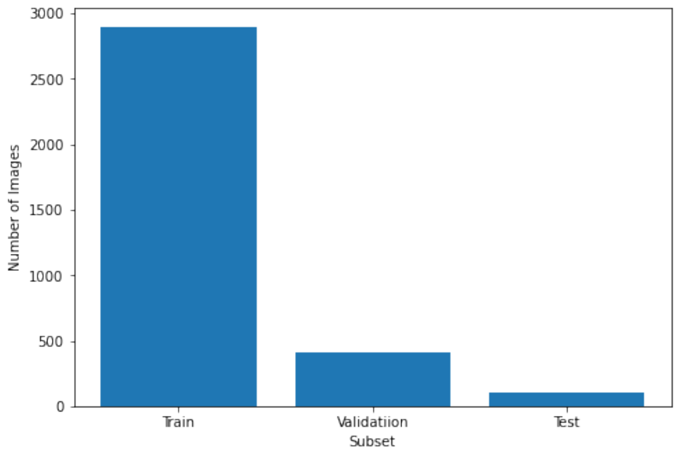

Figure 2: Folder Split ratio

[1]Dataset1:`https://www.kaggle.com/datasets/tawsifurrahman/covid19-radiography-database`
[2]Dataset2:`https://www.kaggle.com/datasets/tawsifurrahman/tuberculosis-tb-chest-xray-dataset`
[3]Dataset3: `http://https://www.kaggle.com/datasets/raddar/tuberculosis-chest-xrays-shenzhen`

In this step, I have combined the dataset and split the data into the train, test and val folder which is for training, testing and validation respectively as per KC et al. (2021). The train folder has 2900 images with 725 images of each class, the test folder has 828 images with 207 images of each class and the val folder have 408 images with 102 images for every class as represented in Fig. 2. I have decided to keep the classes balanced in the preparation itself to avoid the class imbalance problem.

3.3 Data Transformation

Figure 3: Data Augmentation

Data transformation involves resizing images that I have split into training and validation folders and data augmentation involves restructuring these images to form new images for model training. In the transformation of data, I have used Keras Image data generators to apply resizing of images to 224 *224 pixels as suggested by Narin et al. (2021) because I will be using pre-trained models trained on these dimensions. The process of data augmentation helps us to create more data by slightly altering the existing images in my dataset as suggested by Das et al. (2021) which can be seen in Fig. 3. The Image data augmentation we applied include rescale =1./255, zoom range = 0.4, rotation range = 10 and horizontal flip = True. We have also applied target size = (224, 224), batch size = 32, class mode = "categorical", shuffle = True and seed = 42 for training and validation ImagedataGenerators. This augmentation is applied to training and validation data and the images are shuffled to avoid overfitting. The training and validation data is used for model training and the testing ImagedataGenerators which has rescale=1./255, target size=(224, 224), batch size=1,class mode=None, shuffle=False and seed=42 is used for evaluation of my models.

3.4 Model Training

In this phase, I have used the data from train and validation ImageDataGenerators and have applied it for my model training. I have the VGG-16 pre-trained mode used by Das et al. (2021) for Covid, Pneumonia and health cases classification from chest x-ray in which the model performed with an accuracy of 97 per cent. I have also used pre-trained DenseNet121 used by KC et al. (2021) for chest x-ray classification. Both of these models have been trained on the imagenet dataset and have input shape of (224,224,3) along with the addition of some layers which is discussed in the design specification of these models. I have also created a CNN from scratch for this specific problem which proved to be a good option when explored by Choudhuri and Paul (2021) where they created a CNN from scratch and also used pre-trained models in which CNN made for their specific task outperformed pre-trained models.

3.5 Model Evaluation

For the evaluation of my models, I have used the testing set of my data and calculated the accuracy, precision, recall and f1 score alongside creating a confusion matrix of predicted and expected results from each model to better understand their performance and evaluate them properly on these criteria. This technique for evaluation was inspired by Choudhuri and Paul (2021) who have also evaluated the CNN model built from scratch alongside pre-trained models.

4 Design Specification

The models used for multiclass classification can be separated into two major categories CNN built from scratch and pre-trained models.The pre-trained models I used are VGG-16 and DenseNet121.

4.1 CNN Model

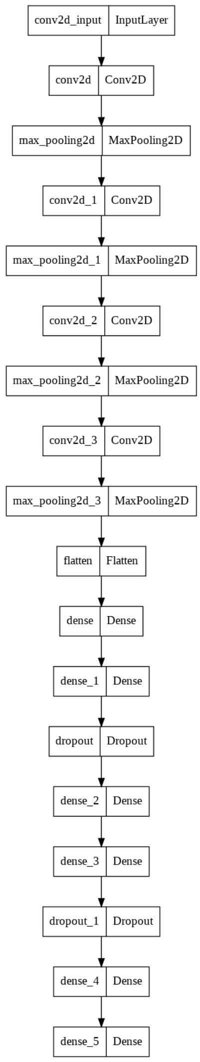

Figure 4: CNN model Architecture

For CNN I have used sequential model which has a convolutional layer with filters=128,strides= 1, kernel size = (5, 5), activation='relu' and input shape=(224, 224, 3) for the input layer followed by a max pooling layer with stride (3,3) followed by another convolutional layer which has filters=64,kernel size = (5, 5) and activation='relu' whose output goes to max pooling with stride (3,3). The output of the max-pooling layer is then given to a convolutional layer of filters=32, kernel size = (3, 3) and activation='relu' followed by the max pooling layer of stride (2,2) this goes to the convolutional layer which has filters=32, kernel size = (3, 3) and activation='relu' again followed by max pooling layer of stride (2, 2). Then I have added a flattened layer two dense layers first have parameters 4096 followed by 256 and both have activation as relu. These dense layers are followed by a dropout layer of 0.1. The output of dropout is followed by two more dense layers with parameters of 128 and 16 in that order, both having activation as relu followed by another dropout layer of 0.1. Then a dense Layer of 8 and activation relu followed by the output layer having parameter equal to the number of classes which is 4 and activation softmax which adds up to a total of 3,427,516 parameters. For compiling the model the parameters used are loss function set as 'categorical crossentropy', optimizer set to 'adam' and metrics used as 'accuracy'.

4.2 Pre-Trained models

For both VGG16 and densenet121 pre-trained models, the input shape is set as 224,224 with channel 3 which stands for RGB. The weights are trained on imagenet and include top is set as false. All the layers are frozen with layer trainable set as false. The VGG16 is added with some layers at the output which are an average pooling layer with pool size (4,4) followed by a flatten layer. The flatten layer output goes to the dense layer of unit 32 an activation set as relu followed by a dropout layer of 0.2. This is followed by a dense layer of units 512 and activation relu which goes to the final dense layer with the unit set as 4 and activation set as softmax which gives 14,750,052 total parameters and 35,364 trainable parameters. The DenseNet121 has output added as an average pooling layer with pool size (4,4) followed by a flatten layer. The flatten layer output goes to the dense layer of unit 16 an activation set as relu followed by a dropout layer of 0.2. This is followed by a dense layer of units 512 and activation relu which goes to the final dense layer with the unit set as 4 and activation set as softmax giving 7,064,660 total parameters out of which 27,156 are trainable parameters. Both of these models have been compiled with adam optimizer with loss function set as categorical cross entropy and metrics set as accuracy. The models have been trained for 100 epochs with callbacks given as early stopping and model checkpoint callback. The early stopping monitors validation accuracy with mode set to max having verbose 1 and patience 20 while the model checkpoint saves weights in chest-x-ray.h5 and monitors validation loss and saving best weights set as true with mode as a minimum and verbose set as 1.

5 Implementation

In this section, the details such as setup required, tools utilized data handling and model implementation are covered.

5.1 Environmental setup

For this research, I have used Google Colab Pro which has 25 GB RAM and 110 GB of disk space for saving my data This provided me with the high computational power and GPU which I required for the implementation of the models. Python language with version 3.7.13 is used as the primary programming language. For Deep Learning TensorFlow version 2.8.2 is used and Keras 2.8.0 is used for implementation.

5.2 Data Handling

The dataset procured from the three Kaggle public repositories is first unzipped and only useful data is stored and a new data is created which has three subfolders namely train, test and Val each of these folders have 4 folders named after the 4 classes of classification and each of these folders have images of that particular class. The train folder has 4 folders with 725 images in each folder the test folder has 108 images in each subfolder and the Val folders have 207 images each. The primary folder of my dataset is stored on google drive and this drive is mounted in my Google Colab notebook to perform data augmentation using Keras Image data generators, this is used for model training.

5.3 Model Implementation

The pre-trained models used in the research which are VGG16 and DenseNet121 are trained on the imagenet dataset and their weights have been frozen, additional layers have been added to the output layer to improve the performance of the model and the parameters are finalised after trial and error of different combinations. Both of these models have been set to run for 100 epochs but early stopping has been employed to stop the model if the validation accuracy is not improved for over 20 epochs. I have also used checkpoints and set save best only as true. The CNN model has been made from scratch with multiple trial and error attempts to increase the accuracy of the model. The CNN takes 224,224 as input size a total of 3,427,516 parameters. I have used early stopping which monitors validation loss and if the validation loss has no improvement for 20 epochs the model is stopped. I have set this model to run for 100 epochs.

6 Results and Discussion

The models that I have used in this research are all trained for 100 epochs with batch size 32 and the pre-processing techniques used also remain the same to maintain uniformity. I have made use of test data to run the predictions of all the models and create a confusion matrix for evaluation.

6.1 VGG16 Pre-Trained Model

The confusion matrix of the VGG16 model in Fig. 5 shows that it has the highest accuracy in the prediction of pneumonia cases compared to other models. This pre-trained model was able to perform fairly well in my classification problem and had an overall accuracy of 85%.

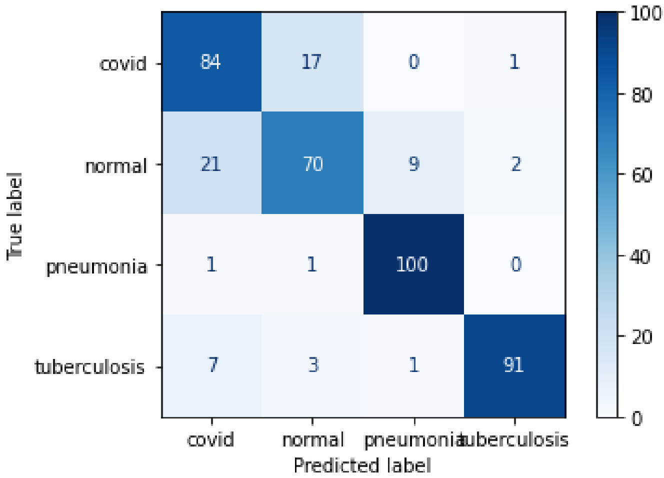

Figure 5: VGG16 Confusion Matrix

6.2 DenseNet121 Pre-Trained Model

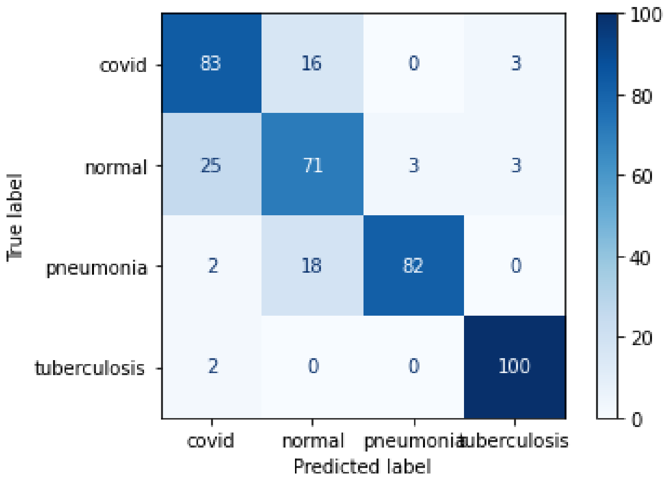

Figure 6: DenseNet121 Confusion Matrix

The Confusion matrix of the DenseNet121 model n Fig. 6 shows that it has the highest accuracy for the TB class only misclassifying 2 of the classes as Covid. It had the lowest accuracy for the normal class and also had difficulty in comparing Normal and Covid classes as seen in the confusion matrix.

6.3 End-to-End End CNN Model

The end-to-end CNN model as shown in Fig. 7 performed overall best compared to other models and had the highest accuracy in normal and covid cases. It also had above 90% accuracy for pneumonia and TB class. This model achieved 91% overall accuracy.

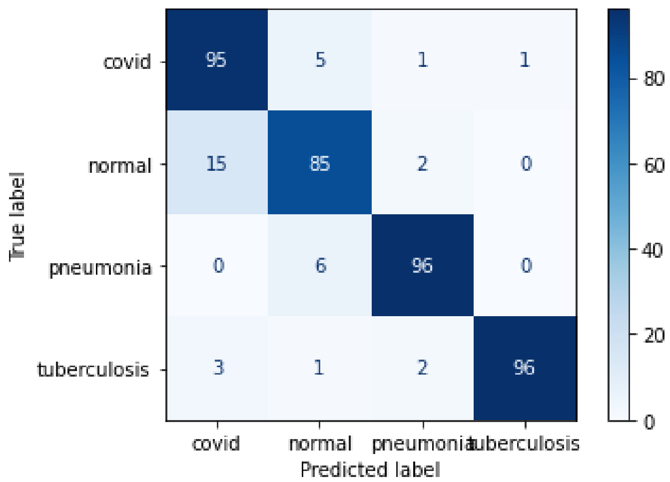

Figure 7: CNN Model Confusion Matrix

6.4 Discussion

The CNN model designed from end to end performed better than the pre-trained models VGG16 and DenseNet121 it was able to achieve the highest accuracy as well as highest precision, recall and F1 score as shown in Table 1. The CNN model that I have proposed achieved 91% accuracy as shown in Fig. 8 in this research has also performed better than the recent research in this field conducted by N et al. (2022) who have worked on a smaller dataset and have achieved 88% of accuracy. The limitation of this research is that the covid and normal classes identification can be improved further which can increase the accuracy above 95%. Overall the proposed CNN model has performed as per expectation and provided great results for the classification problem.

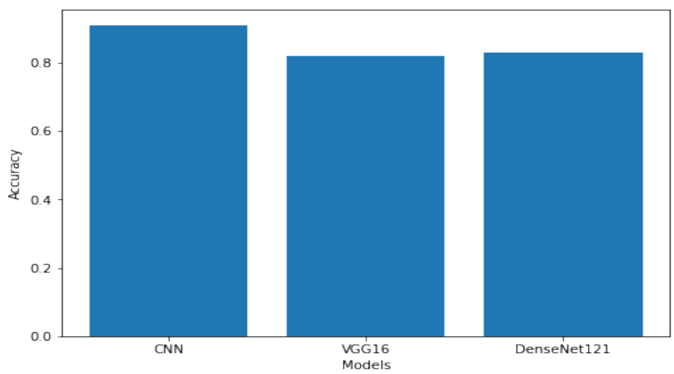

Figure 8: Model Accuracy Comparison

Table 1: Performance Metrics

Models	**Avg. Precision**	**Avg. Precision**	**Avg F1score**
CNN Model	91.50	13.65	13.65
VGG16 Model	84.75	0.01	13.65
DenseNet121 Model	83	92.50	13.65

7 Conclusion and Future Work

This research aimed to explore how convolutional neural networks and pre-trained models can be employed for multiclass classification of Covid-19, Pneumonia, Tuberculosis, and Normal cases from chest X-rays. The Results demonstrate that the end-to-end CNN model performed the best compared to VGG16 and DenseNet12 when compared based on accuracy, precision, recall and F1 score. The limitation of this study was that all individual models had certain strengths and certain weaknesses in prediction as seen VGG16 performed best for classification of TB and DenseNet121 performed best for pneumonia. The limited availability of labelled datasets in the medical field is also one of the challenges and limitations. This research can be used to create an early diagnosis system and a suggestive tool for a physician by creating a system that takes low-cost chest x-rays and checks which disease the patient has. All these diseases have similar symptoms and overlapping characteristics and this multi-class classification system can provide the physician with an edge for early treatment and taking precautionary measures to further reduce the infection rate of the diseases. For future work, I would suggest the use of the Ensemble model unifying all of these models and utilising a larger dataset to combine the strengths of each model and create a much better Multi-Class classification model for these diseases.

References

Ahamed, K. U., Islam, M., Uddin, A., Akhter, A., Paul, B. K., Yousuf, M. A., Uddin, S., Quinn, J. M. and Moni, M. A. (2021). A deep learning approach using effective preprocessing techniques to detect covid-19 from chest ct-scan and x-ray images, *Computers in Biology and Medicine* **139**: 105014.
URL: *https://www.sciencedirect.com/science/article/pii/S0010482521008088*

Al-antari, M. A., Hua, C.-H., Bang, J. and Lee, S. (2021). "fast deep learning computer-aided diagnosis of covid-19 based on digital chest x-ray images", *Applied Intelligence* **51**(5): 2890–2907.
URL: *https://doi.org/10.1007/s10489-020-02076-6*

Bai, H. X., Hsieh, B., Xiong, Z., Halsey, K., Choi, J. W., Tran, T. M. L., Pan, I., Shi, L.-B., Wang, D.-C., Mei, J., Jiang, X.-L., Zeng, Q.-H., Egglin, T. K., Hu, P.-F., Agarwal, S., Xie, F.-F., Li, S., Healey, T., Atalay, M. K. and Liao, W.-H. (2020). Performance of radiologists in differentiating covid-19 from non-covid-19 viral pneumonia at chest ct, *Radiology* **296**(2): E46–E54. PMID: 32155105.

Caseneuve, G., Valova, I., LeBlanc, N. and Thibodeau, M. (2021). Chest x-ray image preprocessing for disease classification, *Procedia Computer Science* **192**: 658–665. Knowledge-Based and Intelligent Information Engineering Systems: Proceedings of the 25th International Conference KES2021.
URL: *https://www.sciencedirect.com/science/article/pii/S1877050921015556*

Choudhuri, R. and Paul, A. (2021). Multi class image classification for detection of diseases using chest x ray images, *2021 8th International Conference on Computing for Sustainable Global Development (INDIACom)*, pp. 769–773.

Chowdhury, M. E. H., Rahman, T., Khandakar, A., Mazhar, R., Kadir, M. A., Mahbub, Z. B., Islam, K. R., Khan, M. S., Iqbal, A., Emadi, N. A., Reaz, M. B. I. and Islam, M. T. (2020). Can ai help in screening viral and covid-19 pneumonia?, *IEEE Access* **8**: 132665–132676.

Das, A. K., Kalam, S., Kumar, C. and Sinha, D. (2021). Tlcov- an automated covid-19 screening model using transfer learning from chest x-ray images, *Chaos, Solitons Fractals* **144**: 110713.
URL: *https://www.sciencedirect.com/science/article/pii/S0960077921000667*

Demir, F., Demir, K. and Şengür, A. (2022). Deepcov19net: Automated covid-19 disease detection with a robust and effective technique deep learning approach, *New Generation Computing* .
URL: *https://doi.org/10.1007/s00354-021-00152-0*

Gupta, A., Anjum, Gupta, S. and Katarya, R. (2021). Instacovnet-19: A deep learning classification model for the detection of covid-19 patients using chest x-ray, *Applied Soft Computing* **99**: 106859.
URL: *https://www.sciencedirect.com/science/article/pii/S1568494620307973*

Heidari, M., Mirniaharikandehei, S., Khuzani, A. Z., Danala, G., Qiu, Y. and Zheng, B. (2020). Improving the performance of cnn to predict the likelihood of covid-19 using chest x-ray images with preprocessing algorithms, *International Journal of Medical Informatics* **144**: 104284.
URL: *https://www.sciencedirect.com/science/article/pii/S138650562030959X*

Ibrahim, D. M., Elshennawy, N. M. and Sarhan, A. M. (2021). Deep-chest: Multi-classification deep learning model for diagnosing covid-19, pneumonia, and lung cancer chest diseases, *Computers in Biology and Medicine* **132**: 104348.

Ismael, A. M. and Şengür, A. (2021). Deep learning approaches for covid-19 detection based on chest x-ray images, *Expert Systems with Applications* **164**: 114054.
URL: *https://www.sciencedirect.com/science/article/pii/S0957417420308198*

Jain, R., Gupta, M., Taneja, S. and Hemanth, D. J. (2021). Deep learning based detection and analysis of covid-19 on chest x-ray images, *Applied Intelligence* **51**(3): 1690–1700.

KC, K., Yin, Z., Wu, M. and Wu, Z. (2021). Evaluation of deep learning-based approaches for covid-19 classification based on chest x-ray images, *Signal, Image and Video Processing* **15**(5): 959–966.
URL: *https://doi.org/10.1007/s11760-020-01820-2*

Mahbub, M. K., Biswas, M., Gaur, L., Alenezi, F. and Santosh, K. (2022). Deep features to detect pulmonary abnormalities in chest x-rays due to infectious diseasex: Covid-19, pneumonia, and tuberculosis, *Information Sciences* **592**: 389–401.
URL: *https://www.sciencedirect.com/science/article/pii/S0020025522001001*

Mahmud, T., Rahman, M. A. and Fattah, S. A. (2020). Covxnet: A multi-dilation convolutional neural network for automatic covid-19 and other pneumonia detection from chest x-ray images with transferable multi-receptive feature optimization, *Computers in Biology and Medicine* **122**: 103869.

N, H. V. R., Sreeja, A. N., Reddy, K. H., Hasmitha, V. and R, L. (2022). Multiclass deep model for diagnosis of covid-19 using chest x-ray, *2022 6th International Conference on Trends in Electronics and Informatics (ICOEI)*, pp. 1767–1771.

Narin, A., Kaya, C. and Pamuk, Z. (2021). Automatic detection of coronavirus disease (covid-19) using x-ray images and deep convolutional neural networks, *Pattern Analysis and Applications* **24**(3): 1207–1220.

Rahimzadeh, M. and Attar, A. (2020). A modified deep convolutional neural network for detecting covid-19 and pneumonia from chest x-ray images based on the concatenation of xception and resnet50v2, *Informatics in Medicine Unlocked* **19**: 100360.

Rahman, T., Khandakar, A., Kadir, M. A., Islam, K. R., Islam, K. F., Mazhar, R., Hamid, T., Islam, M. T., Kashem, S., Mahbub, Z. B., Ayari, M. A. and Chowdhury, M. E. H. (2020). Reliable tuberculosis detection using chest x-ray with deep learning, segmentation and visualization, *IEEE Access* **8**: 191586–191601.

Rahman, T., Khandakar, A., Qiblawey, Y., Tahir, A., Kiranyaz, S., Abul Kashem, S. B., Islam, M. T., Al Maadeed, S., Zughaier, S. M., Khan, M. S. and Chowdhury, M. E. (2021). Exploring the effect of image enhancement techniques on covid-19 detection using chest x-ray images, *Computers in Biology and Medicine* **132**: 104319.
URL: *https://www.sciencedirect.com/science/article/pii/S001048252100113X*

Ucar, F. and Korkmaz, D. (2020). Covidiagnosis-net: Deep bayes-squeezenet based diagnosis of the coronavirus disease 2019 (covid-19) from x-ray images, *Medical Hypotheses* **140**: 109761.
URL: *https://www.sciencedirect.com/science/article/pii/S0306987720307702*

Varela-Santos, S. and Melin, P. (2021). A new modular neural network approach with fuzzy response integration for lung disease classification based on multiple objective feature optimization in chest x-ray images, *Expert Systems with Applications* **168**: 114361.
URL: *https://www.sciencedirect.com/science/article/pii/S0957417420310423*

Yadav, P., Menon, N., Ravi, V. and Vishvanathan, S. (2021). Lung-gans: Unsupervised representation learning for lung disease classification using chest ct and x-ray images, *IEEE Transactions on Engineering Management* pp. 1–13.