Classification of Chest X-Ray images to distinguish between COVID-19 and Pneumonia
COMP 6721 Applied Artificial Intelligence - Proposal - By Group D

A. Problem Statement and Application

COVID-19 is an infectious illness and an pandemic brought on by the coronavirus strain. SARS-CoV-2 responsible for the severe acute respiratory syndrome. Patients suffering from COVID-19 can also present with abnormalities on chest X-ray images that are characteristic of infection [1]. According to WHO, the most common diagnosis in severe COVID-19 patients is severe Pneumonia. [2]
With people all round the world suffering from COVID, it is important to diagnose it correctly by identifying it from numerous similar cases of pneumonia, Tuberculosis, lung infection etc. [2]. Hence, it is our project topic to distinguish between the chest X-ray images of COVID and other diseases related to chest as it can be helpful in the early diagnosis of COVID-19.
With our proposed ML and CNN based model, we aim to differentiate the X-ray images into Normal, COVID-19, Pneumonia and few other classes.
The challenges that might arise with our problem is training the model with similar chest X-ray images where pneumonia can be mistaken with COVID. As the dataset consists of images of different format and hence the pre-processing of data would be a challenge.
The goal throughout developing the application is to study the impact of different training models in our application by interchanging the datasets, CNN architectures, hyper parameters etc.

B. Proposed Methodologies

For the implementation of the model, we have selected 3 CNN architecture which are VGG11, ResNet50 and AlexNet.

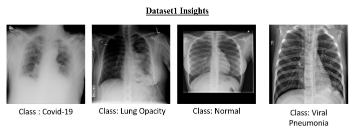
AlexNet : AlexNet is a convolutional neural network that is 8 layers deep. The first five were convolutional layers, some of them followed by max-pooling layers, and the last three were fully connected layers. It used the non-saturating ReLU activation function, which showed improved training performance over tanh and sigmoid. [5] In our project we used vanilla version of the AlexNet. According to our use case which distinguish between COVID-19 and Pneumonia, a networks with less number of layers would not find many features so we decided to use AlexNet as our first approch in Convolution Neural Network.

VGG11 : AlexNet is primarily capable of object-detection. It takes into account over-fitting by using data augmentation and dropout. VGG replaces tanh activation function with ReLU by encapsulating its distinct features for over-pooling. VGG came into picture as it addresses the depth of CNNs.Input. VGG takes in a 224x224 pixel RGB image. The convolutional layers in VGG use a very small receptive field . There are also 1x1 convolution filters which act as a linear transformation of the input, which is followed by a ReLU unit. VGG has three fully-connected layers: the first two have 4096 channels each and the third has 1000 channels, 1 for each class. All of VGG's hidden layers use ReLU. To go to the next step we decided go to VGG which has all the capabilities of AlexNet and it goes next step and tries to get better result, via ReLU activation function and other changes in its layers. the main reason that we used VGG11 is because, it is the easiest to implement and will form the basis for other configurations and training for other VGG models as well.

ResNet50 : ResNet It is a gateless or open-gated variant of the HighwayNet,[2] the first working very deep feedforward neural network with hundreds of layers, much deeper than previous neural networks. Typical ResNet models are implemented with double- or triple- layer skips that contain nonlinearities (ReLU) and batch normalization in between. Models with several parallel skips are referred to as DenseNets. ResNet50 is a variant of ResNet model which has 48 Convolution layers along with 1 MaxPool and 1 Average Pool layer. It is a widely used ResNet model. The reason that we used ResNet is that because ResNets are one of the most efficient Neural Network Architectures, as they help in maintaining a low error rate much deeper in the network. Moreover, since we wanted to use the basic version on ResNets, we choose ResNet50, due to several reasons, first is the most popular ResNets, and also it has 50 convolutional layers which could get deeper features from the photo and since our photos are X-Ray we need this kind of layers.

Datesets:
Dataset 1:"COVID-19 Radiography Database" This dataset consists of 21,164 images which are categorised into 4 classes [COVID Positive, Normal, Non-COVID lung infection, Pneumonia]. There number of images per classes is as follows: COVID 19: 3616 Images, Normal: 10,192 Images, Lung Opacity: 6012 Images (This class represents the infection of lungs which is not caused by Covid-19) & Viral Pneumonia: 1345 Images. All the images are in Portable Network Graphics (PNG) file format and shape of the image is 299x299px. This data is collected from Germany medical schools, GitHub, Kaggle, Radiological Society of North America(RSNA) and various other public datasets. The preview of the images of each classes from Dataset1 is as shown in the below image.

CVPR #*****



Dataset 2 : [6] ”COVID19, Pneumonia, Normal Chest X-ray Dataset” This dataset has 3 classes containing 6939 multi-type images. Number of images per class is 2313 Images. Each class contain file format of PNG for all images and average shape of images in all classes is: 1024x1024px. We have collected this datset from Kaggle but it was originally gathered using GitHub, the Radiopaedia, Italian Society of Radiology (SIRM) , Figshare data repository websites. Then,912 augmented images were collected from Mendeley.

Dataset 3 : [7] ”Xray Body Parts-fastai” This particular dataset contains largest number of classes with compared to our other image dataset. Number of Classes are 22. This dataset contains a total of 2,482 images. And Images per class are: Train class: 1,738 images & Test class: 743. Among the dataset the ralevent class type to our project was class Chest, this class has 724 Images from train and 49 for test. Other 1,014 images are distributed among rest of the 21 classes. The image type for this dataset is PNG and the shape of all images are 420x512 px.

The average images per class is less than 49 images. Some of the classes have very few images.For ex Clavicles class there are only 9 images only.

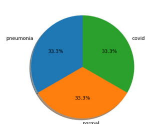


Image Distribution in Dataset2

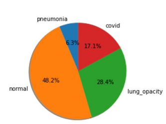


Image Distribution in Dataset1

C. Attempts at solving the problem

Parameters and Hyper parameters : For all of the bellow models we used batch size of 32, Cross Entropy Loss as a loss function, and learning rate of 0.01 and 5 epochs (the reason that we used this low number of epochs was 1- to compare how our networks work with less epochs and 2- time limitation). if any parameters changes during any training it will be mnetioned.

Attempts : We tried 3 models AlexNet, VGG11 and ResNet50. Accuracy, F1 Score, Recall and precision for each model is available bellow. (some models are not ready yet and these result are just preliminary)

Accuracy:

DS/Model	AlexNet	VGG11	ResNet50
Dataset1	77.5	88.80	47.69
Dataset2	33.10	31.92	34.10
Dataset3	0.6924	-	71.98

Precision:

DS/Model	AlexNet	VGG11	ResNet50
Dataset1	0.8440	88.80	0.1192
Dataset2	0.1051	0.1064	0.1137
Dataset3	0.3462	-	0.3599

Recall:

DS/Model	AlexNet	VGG11	ResNet50
Dataset1	0.7099	88.80	0.2500
Dataset2	0.3333	0.3333	0.3333
Dataset3	0.5000	-	0.5000

F1 score:

DS/Model	AlexNet	VGG11	ResNet50
Dataset1	0.7751	0.8880	0.4769
Dataset2	0.3152	0.3192	0.3410
Dataset3	0.6924	-	0.7198

As all the result shows all the measurement are not even close to what we expected. this is because of two reasons, 1- 5 epochs is not enough for such big task and 2- size of datasets are too small. our first failed attempts was, given input size: (256x1x1). Calculated output size: (256x0x0). Output size is too small and causes loss function comes Nan after some time. other failed approach that we used is, set learning rate to lr=0.1 and get 48.41 accuracy with 30 epochs, which take time but the result was not satisfying. another failed approaches, is we used pre-trained VGG11 for dataset two gets 32.39 accuracy with 5 epochs which indicate that the more epochs is need for dataset or reduce the batch size. to train the model properly.

D. Future Improvements

There are several things which needs to be done in the future. First of all train our networks more, 5 epochs are not enough for our application as we can see that the result is not satisfying. We need to have more precision and recall regarding to our use case. For instance, ResNet50 problem will be solved with this approch. Secondly, in order to get better result in short period of time, it is better to use pre-trained models. For example, we are going to use transfer learning in for AlexNet and VGG11 for second dataset, since these two models were not accurate. Third, we will tune the hyper-parameters of the cnn model to get better result. For example, for dataset two the amounts of image it is not enough so with 32 of batch size, a lot of features will be missed. Other than that, learning rate could change the output accuracy if it is not defined properly, for example in VGG model as it was mentioned in the attempts it could cause huge impact. Lastly, loss functions are also have huge rule for the model, since we used in all of our model we used Cross Entropy Loss, we do not know the impact of other loss functions on our models.

E. Bibliography

References

[1] Ng Ming-Yen, et al. *Imaging profile of the COVID-19 infection: radiologic findings and literature review.* Radiology: Cardiothoracic Imaging 2.1 (2020). 1

[2] *https://www.who.int/docs/default-source/coronaviruse/clinical-management-of-novel-cov.pdf.* 1

[3] M.E.H. Chowdhury, T. Rahman, A. Khandakar, R. Mazhar, M.A. Kadir, Z.B. Mahbub, K.R. Islam, M.S. Khan, A. Iqbal, N. Al-Emadi, M.B.I. Reaz, M. T. Islam *"Can AI help in screening Viral and COVID-19 pneumonia?"* IEEE Access, Vol. 8, 2020, pp. 132665 - 132676.

[4] Rahman, T., Khandakar, A., Qiblawey, Y., Tahir, A., Kiranyaz, S., Kashem, S.B.A., Islam, M.T., Maadeed, S.A., Zughaier, S.M., Khan, M.S. and Chowdhury *Exploring the Effect of Image Enhancement Techniques on COVID-19 Detection using Chest X-ray Images.* arXiv preprint arXiv:2012.02238, M.E., 2020.

[5] Krizhevsky, Alex and Sutskever, Ilya and Hinton, Geoffrey E. *ImageNet Classification with Deep Convolutional Neural Networks.* Association for Computing Machinery, 2017. 1

[6] Dataset 1: *https://www.kaggle.com/datasets/tawsifurrahman/covid19-radiography-database*.

[7] Dataset 2: *https://www.kaggle.com/datasets/amanullahasraf/covid19-pneumonia-normal-chest-xray-pa-dataset*.

[8] Dataset 3: *https://www.kaggle.com/code/alexanderyyy/xraybodyparts-fastai/notebook*.

[9] *https://github.com/agchung*

[10] *https://radiopaedia.org/?lang=gb*

[11] *https://www.sirm.org/en/category/articles/covid-19-database/*

[12] *https://figshare.com/articles/dataset/COVID-19ChestX-RayImageRepository/12580328/2*

[13] *https://www.kaggle.com/datasets/ibombonato/xray-body-images-in-png-unifesp-competion*