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

A. Introduction and Problem Statement

Deep learning has advanced greatly and provides many benefits over conventional ML techniques. The primary goal of our project is to develop a CNN classification model for three different datasets. Using different experimental settings comparative experiments will be carried out, and the effects of the various parameters on classification will be examined.

Problem and its applications: COVID-19 patients can also be 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].To diagnose it correctly by identifying it from numerous similar cases of pneumonia, Tuberculosis, etc. [2]. Hence, it is our model topic to distinguish between the chest X-ray images of COVID and other diseases related to chest.

General Strategy: Our proposed ML and CNN based model, will aim to differentiate the X-ray images into Normal, COVID-19, Pneumonia and few other classes. For this, we are using 3 datasets of different sizes to train our CNN models - AlexNet, VGG, ResNet.

Challenges: The challenges that arised with our problem initially was choosing a proper CNN model for classification. We have chosen AlexNet, VGG and ResNet. The dataset consists of images of different format and hence the pre-processing of data was a challenge which we dealed by manipulating the images using transforms to resize, random horizontal flip, and using tensors to normalise the images with the mean and standard deviation to force the network to be in range of 0-1 rather than 0-255. Also, after convolution operation we performed max pooling to reduce the dimensionality which enabled us to reduce the number of parameters and helped in shortening the training time. Further, training the model with huge datasets which consists of 21164, 6939 and 2482 images respectively was very time consuming and required us to use online GPU offered by Google Collab to train faster. Another challenge we faced was comparing the results of different models which was done by generating confusion matrix.

Expected results: 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. We want to evalute which CNN model will provide more accurate results on each dataset.

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 CNN 8 layers deep. The first 5 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 is to distinguish between COVID-19 and Pneumonia; images that which are a bit plain will lead to the complexity of network in finding features if we had not used CNN. Hence, we decided to use AlexNet as our first approach.

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. [6]. VGG came into picture as it addresses the depth of CNNs. VGG takes a 224x224 pixel RGB image as input. 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.

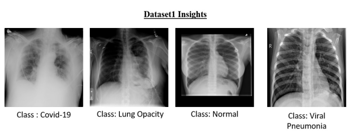
For the next step we decided to go with VGG which has all the capabilities of AlexNet and it goes further and tries to get better results, via ReLU activation function and other changes in its layers. Also it is the easiest to implement and it will form the basis for other configurations and training for other VGG models as well.

ResNet50 : ResNet is a gateless or open-gated variant of the HighwayNet,the first working very deep feedforward NN with hundreds of layers, much deeper than previous NNs. [7] Typical ResNet models are implemented with double or triple layer skips that contain non-linearities (ReLU) and batch normalization in between. Models with several parallel skips are referred to as DenseNets. ResNet50 is a widely used variant of ResNet model which has 48 Convolution layers along with 1 MaxPool and 1 Average Pool layer.

The reason we used ResNet is because ResNets are one of the most efficient NN 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 chose ResNet50 as it has 50 convolutional layers which could get deeper features from the images and since our photos are X-Ray images we needed this kind of depth in layers.

Datesets:

Dataset 1: [6] "COVID-19 Radiography Database" This dataset consists of 21,164 images which are categorised into 4 classes namely - COVID Positive(3616 images), Normal(10,192 images), Non-COVID lung infection(6012 images), Pneumonia(1345 images). All the images are in 'PNG' file format and 299x299px shape. 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.



Dataset 2 : [7] "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 contains PNG format for all images and average shape of 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.

Dataset 3 : [8] "X-ray Body Parts-fastai" This particular dataset contains largest number of classes-22 with 2,482 images when compared to our other datasets. The train class has 1,738 images & test class has 743 images. In this dataset, relevent class type to our project was class Chest with 724 Images of train and 49 of test. Other 1,014 images are distributed among rest of the 21 classes. The image type for this dataset is PNG and the shape is 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. Unbalances in data was observed in this dataset due to less number of images per class, which we will plan to deal with further.

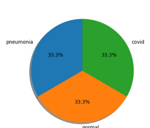


Image Distribution in Dataset2

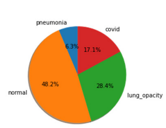


Image Distribution in Dataset1

C. Attempts at solving the problem

For all of our models we used batch size of 32, Cross Entropy Loss as a loss function, learning rate of 0.01 and 5 epochs. Our failed attempts were: Giving input size as (256x1x1) and getting our calculated output size as (256x0x0). Output size was too small and caused loss function; Setting learning rate to lr=0.1 and getting 48.41 accuracy with 30 epochs, which takes time but the result was not satisfying; Using pre-trained VGG11 for dataset 2 gives 32.39 accuracy with 5 epochs which indicates that more epochs is need for dataset or batch size should be reduced to train the model properly. Some models are not ready yet and these results 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	0.9210	0.1192
Dataset2	0.1051	0.1064	0.1137
Dataset3	0.3462	-	0.3599

Recall:

DS/Model	AlexNet	VGG11	ResNet50
Dataset1	0.7099	0.8620	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 the results shown above, all the measurements are not even close to what we expected because of two reasons. Firstly, 5 epochs is not enough for such big tasks and secondly, size of datasets are too small.

D. Future Improvements

Further, we will do more training of network with increased epochs as 5 seems to not be enough. We try to have more precision and recall. For instance, ResNet50 problem will be solved with this approch. Also, in order to get better results in short period of time, it is better to use pre-trained models, For which we are going to use transfer learning for AlexNet and VGG11 with dataset 2, since these two models were not accurate. We will tune the hyper-parameters of the CNN to get better result. For instance, for dataset 2, with batch size as 32, a lot of features are missed because of less number of images. Other than that, learning rate could change the output accuracy if it is not defined properly. Lastly, loss functions other than Cross Entropy Loss will be used to learn 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*. 1, 2

[7] Dataset 2: *https://www.kaggle.com/datasets/amanullahasraf/covid19-pneumonia-normal-chest-xray-pa-dataset*. 1, 2

[8] Dataset 3: *https://www.kaggle.com/code/alexanderyyy/xraybodyparts-fastai/notebook*. 2

[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*