SJIF Impact Factor 2022: 8.197| ISI I.F. Value:1.241| Journal DOI: 10.36713/epra2016 ISSN: 2455-7838(Online)
EPRA International Journal of Research and Development (IJRD)
Volume: 7 | Issue: 5 | May 2022 - Peer Reviewed Journal

DETECTION OF COVID FROM CHEST X-RAYS USING GAN

Unnikrishnan M[1], Gontu Hemant Lakshman[2], KhwairakpamBidyanandaSingh[3], Sagar Pande[4]

[1]*Student, Department of CSE, LPU ,Jalandhar,Punjab,India*
[2]*Student, Department of CSE, LPU,Jalandhar,Punjab,India*
[3]*Student, Department of CSE, LPU, Jalandhar, Punjab, India*
[4]*Professor, Department of CSE, LPU ,Jalandhar,Punjab,India*

Article DOI: https://doi.org/10.36713/epra10226
DOI No: 10.36713/epra10226

ABSTRACT

A great deal of lung-related infections including COVID-19 can be diagnosed in the early stages by examining the x-ray images. With the advancement in deep learning techniques, we can back this up to a great extent. For the formulation of such models, the complexion of the dataset used for them plays a key role. An inadequate dataset with an imbalanced class of samples adversely affects the performance of such models. Through this research work, we introduce an improved version of Auxiliary Classifier GAN (ACGAN) which is computationally much more efficient than the prior model for the augmentation of the dataset. The dataset used in this research work is highly imbalanced due to the lack of availability of covid-19 infected samples. The proposed GAN architecture has done a great job in solving this issue. After the augmentation phase, the dataset including the spawned new samples is fed into the detection network to evaluate the performance of the model. The use of pre-existing ImageNet weights and Adam as an optimizer helped to train the model at a quicker rate and the overall accuracy achieved was 95.67 and validation accuracy of 92.57 with a loss value reduced to 0.675.

KEYWORDS: *Transfer Learning, Convolutional Neural Network (CNN), Data Augmentation, Deep Learning, Multiple perceptrons(MLP), ACGAN*

2. INTRODUCTION

Pneumonia is a dreadful respiratory septicemia which feigns the human lungs. Once a person got infected with the same his/her air sacs present in the lungs get filled with discharge and other fluids. The condition of the patient may get severe within few days and eventually dies if not backed up with proper medication. We can mainly classify pneumonia into 2 categories: viral and bacterial. Both are differed in the way in which it affects the lungs and the way in which the infection is treated. The bacterial infection proved to be more fatal and requires more medication. Usually the treatment includes antibiotic therapy. Meanwhile a good immune body can resolve the infection caused by virus and gets better by itself; this may vary person to person. It has been observed that pneumonia is more prevalent in children. The root cause of the infection comes from the pollution around us hence in accordance with our current lifestyle it is only going to increase. It is the top row of the most deadly diseases and is ranked 8th in case of mortality rate in United States. The chance of recovery diminishes as the detection of infection prolongs hence detection in earlier stages is crucial. Because of all these factors there is a desperate need of research and development of a automated procedure to diagnosis the infection with the help of deep learning models which can aid in the decline of mortality rate of pneumonia especially

SJIF Impact Factor 2022: 8.197| ISI I.F. Value:1.241| Journal DOI: 10.36713/epra2016 ISSN: 2455-7838(Online)
EPRA International Journal of Research and Development (IJRD)
Volume: 7 | Issue: 5 | May 2022 - Peer Reviewed Journal

in children [1].

For the successful detection of pneumonia in the lungs mainly the following tests are conducted: chest X-rays, chest MRI, needle biopsy of the lungs, ultrasound of the chest and CT scan of the lungs. Among these, chest X-rays examining method is more prevalent than others as other methods like CT scan imaging takes considerably more time. The scarcity of CT scan facilities in remote areas also facilities X-ray imaging. That's the primary reason why the dataset availed in this task is a cluster of X-ray images. Professionally qualified radiologist plays a decisive role in this process. In remote areas of our country this may be a cause of concern. So there is urgent need to back our medical field with a computer-aided diagnosis so that this expertise can be made accessible to a large population at a minimal cost. Convolution neural network has done a tremendous job in the image detection and segmentation. Biomedical diagnosis with the help of CNN has proven to work better than all the existing methods [1].

COVID has been a worldwide pandemic over the past few years (started in December 2019) and has caused panic with thousands of deaths and infections in every corner of the world. Although there are multiple challenges to stopping the virus, early detection and precautions can be taken using the help of deep learning.

Currently, Covid like most pandemic prone diseases is diagnosed using polymerase chain reaction (PCR). Although PCR is accepted as the gold standard for diagnosis it is not easily accessible in large amounts in highly contaminated areas. Many attempts have been made in the past to detect Covid through a CT scan. Algorithms based on deep learning were also used to detect anomalies on CT scans. CT scan has many drawbacks: it takes a long time to produce, considerably longer than X-ray imaging. High-quality CT scans are not available in many underdeveloped regions, which makes timely Covid detection impossible. On the other hand, X-rays are the most common and widely available diagnostic imaging technique, playing a crucial role in speeding up the detection process of Covid. X-ray units are a basic feature in most underdeveloped regions, even those in rural areas.

For this reason, our new model is trained to spot anomalies in chest X-ray images of suspected Covid patients, enabling fast and reliable detection.

Training the new model proved to be very challenging due to the lack of data-set and also the imbalanced data-set which was currently present during this project. There were more Covid cases than that of non-Covid cases in the data-set. Deep learning is very data-hungry and takes a considerable amount of training. We propose our new model of detecting anomalies in chest X-ray images using deep learning.

3. PROPOSED WORK

3.1 DATASET DESCRIPTION

The dataset embodies of CXR images distributed across 2 distinct classes: Covid and non-covid. The Covid genre consist of only 184 samples due to the lack of availability of open-source chest X-ray samples infected by the Covid. The primary reason of the scarcity is because normally the people diagnose the infection with the help of RT-PCR test and the doctors also won't care to examine the X-ray unless the health deteriorates severely whereas the collected sample of normal CXR counts to 3700 samples. The dataset is collected from multiple open sources like Kaggle and IEEE covid chest x-ray dataset. The images are of dimension 112*112*3 pixels. The dataset is highly imbalanced as it is evident from the stats. Training the model purely on this dataset may show bias towards the normal CXR class so there is an urging need of populating the dataset to a balanced one.

3.2 DATA AUGMENTATION

For the methodical framing of the model architecture, flawlessly balanced unambiguous dataset plays a vital role. Here in this case the dataset is highly imbalanced. So to resolve this dilemma, synthetic image generation of specified sparse class is mandatory. To enact this objective, a variation of generative adversarial network called Auxiliary Classifier Generative Adversarial Network (ACGAN) is implemented.

SJIF Impact Factor 2022: 8.197| ISI I.F. Value:1.241| Journal DOI: 10.36713/epra2016 ISSN: 2455-7838(Online)
EPRA International Journal of Research and Development (IJRD)
Volume: 7 | Issue: 5 | May 2022 - Peer Reviewed Journal

ACGAN

The classical generative model has fizzled to spawn high resolution samples due to the drawbacks like mode collapse and failing to attain Nash equilibrium. To address this issue, lot of variants were introduced in the later phases; one among them is ACGAN. As we have already discussed the typical GAN architecture is composed of two models, a generator for spawning realistic samples resembling an existing distribution and a classifier network called discriminator for segregating the fake images generated by the generator from the original ones. The two models are trained simultaneously like a min-max game to attain the desired objective.

ACGAN is an extension to a variant of GAN namely Conditional GAN. Earlier, we couldn't opt the class of images generated by the generator in case of traditional GAN. But with the introduction of cGAN, we attained more command and precision in the spawning of new samples. while training of the model, the generator is fed with the class label of the desired sample along with the latent vector and the discriminator is also provided with the class label of the image the generator is expected to spawn. Now, it is the task of the discriminator to segregate the fake from original and slowly promote the generation of specified labeled samples. In short, the cGAN model helps generate the desired class of synthetic samples from the latent noise space by providing class labels to both the generator and the discriminator during the training phase. But in case of ACGAN, only the generator is fed with the class label and the discriminator job is to predict the class of the images spawned by the generator model besides segregating the fake from the original. Hence the model has 2 outputs: first one is the probability of image being fake or real which is determined using sigmoid activation function and for optimizing the result binary cross entropy is used while the second one output is the probability of the image belonging to each known class in the dataset implemented using SoftMax activation function and categorical cross entropy is used as the loss function. The objective of the discriminator is to maximize the probabilities of precisely classifying the spawned images (Ls) and labeling the image with correct class (Lc) while the generator try to minimize Ls and maximize Lc. So overall the objective function narrows down to;

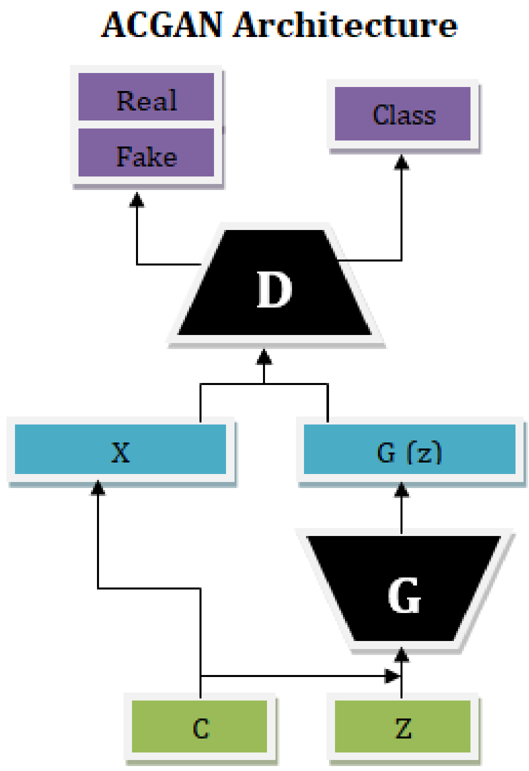

Figure 1: ACGAN Architecture

Discriminator: to maximize Ls+Lc& Generator: to maximize Lc – Ls

The log likelihood loss function that we utilize to optimize the model during the training procedure is as follows

Lc = E [log P(C = c | X real)] + E [log P(C = c | X fake)]

Ls = E [log P(S = real | X real)] + E [log P(S = fake | X fake)]

SJIF Impact Factor 2022: 8.197| ISI I.F. Value:1.241| Journal DOI: 10.36713/epra2016 ISSN: 2455-7838(Online)
EPRA International Journal of Research and Development (IJRD)
Volume: 7 | Issue: 5 | May 2022 - Peer Reviewed Journal

ACGAN GENERATOR

The generator model as mentioned in the previous section takes the latent vector along with the class label to spawn an image of dimension 112*112*3. It is fed with a noise latent vector of 20000 dimensions along with the class label. The latent point in the space is then construed by the fully connected dense layer to construct 1024 feature maps of dimension 7*7. The dense layer has relu activation function which makes the entire training procedure faster and reliable. The class label here is not in one hot encoded form as we usually seen, but is interpreted as an additional feature map in the initial phase of the model. a 7*7 feature map is obtained from a fully connected dense layer of 49 units which is fed with a learned embedding of the 50 dimensions noise vector. The obtained 7*7 feature map which represents the class label is concatenated with the reshaped noise latent vector obtained from the dense layer which makes the dimension of the input image 7*7*1025. The principal constituent of the generator model is transposed convolutional layers which are primarily used for up sampling the input that is, the dimension of the generated feature map is higher when compared to the input feature maps. The up sampled output of each transposed convolution layer is fed to a batch normalization layer followed by ReLU activation. The initial input image is up sampled from 7*7*1025 to 14*14*512 using trans conv layer of 512 filters, then to 28*28*256 with 256 filters, then to 56*56*128 with 128 filters, and finally to 112*112*3 dimensioned image with 3 filters. The filters used in each layer is of dimension 5*5 which is drifted through the input images and dot product of the square pixels is taken to obtain a single cell value of the output feature map. Strides of (2,2) is taken which specifies the square pixels skipped in between 2 consecutive dot products. Padding also added in the input image to preserve the information present in the edges of the image. In the final layer of the generator network, tanh activation is used while in all other layers ReLU activation is implemented.

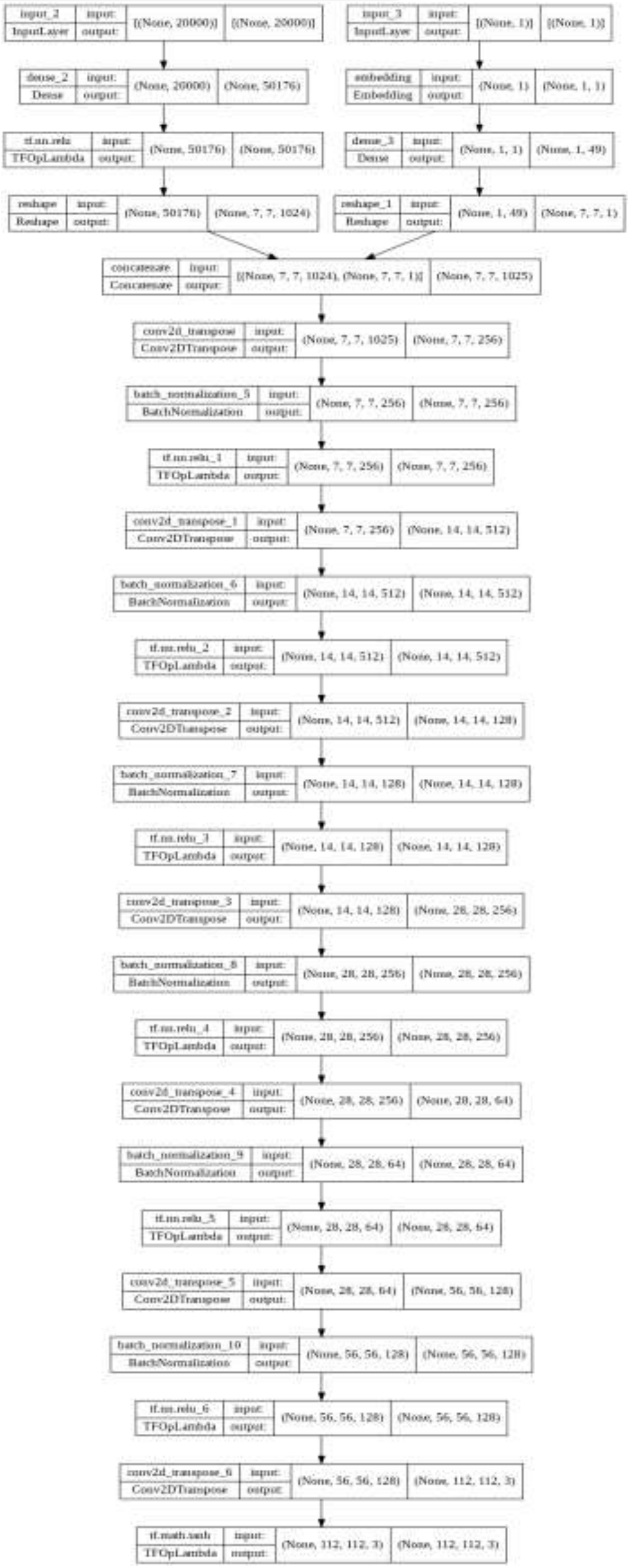

Figure 2: optimized computationally efficient ACGAN generator

SJIF Impact Factor 2022: 8.197| ISI I.F. Value:1.241| Journal DOI: 10.36713/epra2016 ISSN: 2455-7838(Online)
EPRA International Journal of Research and Development (IJRD)
Volume: 7 | Issue: 5 | May 2022 - Peer Reviewed Journal

To reduce the computational complexity and space complexity of the above discussed architecture, a profound optimization is done in the computation of the feature maps in each layer. The modification includes the introduction of 1*1 convolution layers in between the convolutional layers to reduce the dimension of the input to the later phases of the model and hence to drastically reduce the computational complexity. 1*1 convolutions simply take the dot product of values across the channel of each square pixel and hence can't recognize any spatial patterns present in the images. Despite of this, it can learn the features present across the channels of the image. After applying 1*1 convolution, no of channels present in feature maps can be reduced significantly which then fed as an input to later layers reduces the computation required.

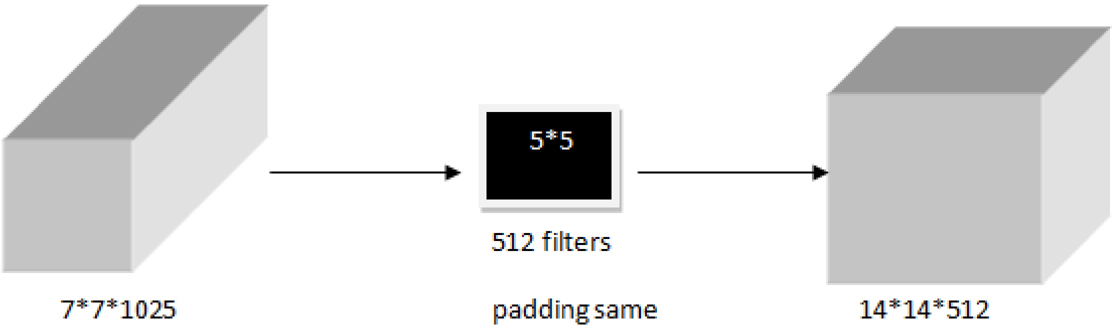

Figure 3: ***traditional convolution operation***

For obtaining the pixel value of each cell in the output, we need to take the dot product of the filters with the input pixel matrix across all the channels. So total number of multiplication operation required for determining the output feature map is as follows:
= Output image dimension * size of the filter * depth of the input channels
= (14*14*512)*(5*5)*(1025) = *2,571,520,000*
This comes around 2571+ million computation.

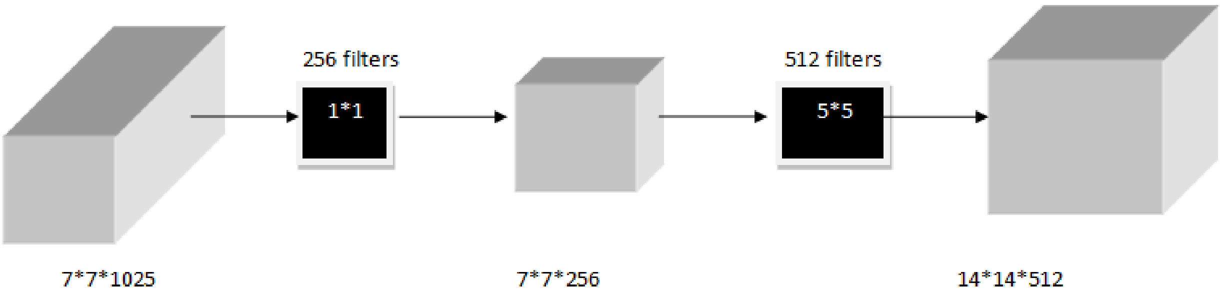

Figure 4: optimized computationally efficient convolution operation.

After applying the optimization discussed, the computation required reduced drastically as follows:
= operation required to find the intermediate output + operation required to find the final output
= (7*7*256)*(1*1)*(1025) + (14*14*512)*(5*5)*(256)
= 655,110,400 ~ 650+ million computation.

SJIF Impact Factor 2022: 8.197| ISI I.F. Value:1.241| Journal DOI: 10.36713/epra2016 ISSN: 2455-7838(Online)
EPRA International Journal of Research and Development (IJRD)
Volume: 7 | Issue: 5 | May 2022 - Peer Reviewed Journal

ACGAN DISCRIMINATOR

Similar to all other discriminator network, the prime objective of our discriminator model is also to facilitate the training of the generator model by detecting whether the image fed into it as input is real or not. Besides this goal, the discriminator is also assigned with a task of finding the label of the input image whether it is normal or covid-19 infected. The building block of this architecture is convolutional neural layers which is well known for its capability of image detection and hence is very handy in this classification task. It takes image of dimension 112*112*3 as input either from the generated samples or from the original dataset and predicts two probabilistic outputs: one using sigmoid and the other using SoftMax function.

The model constitutes of 5 layers of convolutional layers where each is followed by a batch normalization, a Leaky ReLU activation and by a dropout unit. 1st layer of the convolution has 32 filters, mainly for the detection of basic feature likes edges and borders, which takes the 112*112*3 image and extract a feature map of dimension 112*112*32 which is further fed to the next layer of 64 filters to give an output of 56*56*64. The extraction progress as follows: 56*56*64 to 28*28*128(128 filters) then to 14*14*256(256 filters) and finally to 7*7*512(512 filters). In the end the image obtained is flattened before applying the sigmoid and SoftMax unit functions. The filters used in each layer is of dimension 3*3 and strides of dimension (1, 1) is utilized in the initial layer followed by the (2,2) strides in the subsequent layers. To prevent the model from overfitting to the given dataset, dropout units are added after each layer which self recognize the dumb neurons during the training phase and drops it. This regularization technique modifies the model architecture itself for the prevention of overfitting unlike the L1 and L2 regularization which modifies the cost function to achieve the same.

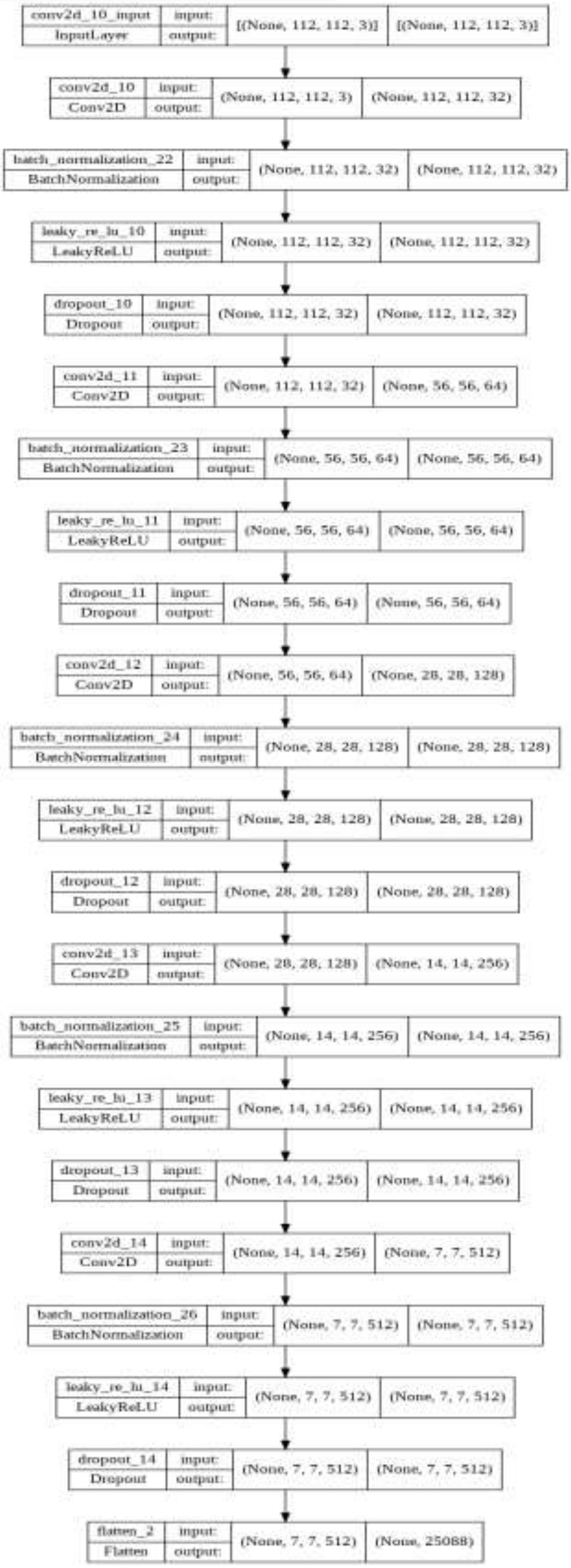

Figure 5: ***Discriminator Architecture***

SJIF Impact Factor 2022: 8.197| ISI I.F. Value:1.241| Journal DOI: 10.36713/epra2016 ISSN: 2455-7838(Online)
EPRA International Journal of Research and Development (IJRD)
Volume: 7 | Issue: 5 | May 2022 - Peer Reviewed Journal

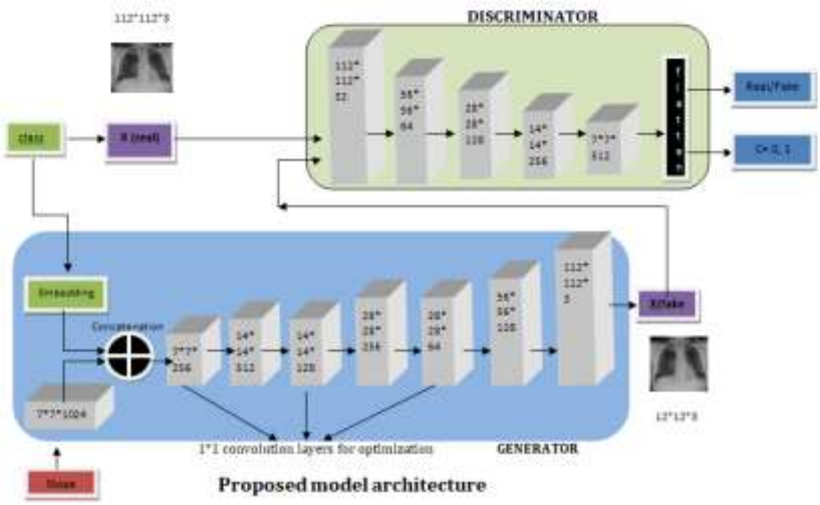

Figure 6: Proposed model Architecture

IMPORTANCE OF DATA AUGUMENTATION AND AC_GAN

Data Augmentation is one of the key steps which had a great impact on our results and training time. Reading over 10 similar research paper on the image detection topic there were some significant learnings and findings which were very helpful, one such key insight was that each model had a particular image size which brings the best from the model here image matrix size of [299,299] was perfect for Xception model.

One other learning was the models were not trained perfectly most of the dataset has more images with non-covid images and model was trained to those images yielded more accuracy as of the training set also contained more non-covid images and this all wouldlead to a false learning situation. This is where the data augmentation plays a major role, we balanced out the dataset with having images augmented and the number of non-covid and covid images were nearly equal our original dataset had 3000 non-covid images and some 700 covid images but to balance out the dataset what we did was generate 2400 covid images through data augmentation process using AC-GAN as the technology it was achieved.

3.3. DETECTION MODEL

XCEPTION MODEL:

As defined earlier convolution is the overlap of one function over the other where we perform the multiplication sum of the filter matrix over the input matrix to get the values in a new matrix but the convolution is always applied on to the kernel of the layer and there can be many kernels so the cost of the convolution is the major concern, more the number of multiplications more the cost of the computation. Xception solves that by using depthseparable convolution as the approach.

Depth Separable Convolution: approach taken here is to separate the work into phases

1.Depthwise convolution: stage where filtering is done

2.Pointwise convolution: where combination is done

This also is applied to only a single channel and since the depth is 1 which would decrease the overall multiplication cost

second phase is the combination phase where N filters are applied for m levels each of unit size. Let's compare the mode consider an input matrix of size a*a and depth d so the volume of the input is:

a*a*d

The filter size is k*k

So, for once we would multiply k*k*d times since the depth is 'd'.

SJIF Impact Factor 2022: 8.197| ISI I.F. Value:1.241| Journal DOI: 10.36713/epra2016 ISSN: 2455-7838(Online)
EPRA International Journal of Research and Development (IJRD)
Volume: 7 | Issue: 5 | May 2022 - Peer Reviewed Journal

But we stride the filter through the input let the number of strides be 'g' along the width and height so total multiplication for the kernel wouldbe:

g^2 *k^2 * d

Considering there are N kernels we would go up to N*(g^2 *k^2 * d)

But using Depth wise Separable Convolution we always consider only one channel at a time in the filtering stage.

In the Depth wise convolution stage consider the input volume be ='F*F*m' where m is the depth now the kernels of the size k*k would work only in a single channel then there would be total 'm' such kernel multiplications

k*k*1 and total of m such operations. And considering the strides the output volume would be:

Og*Og*m

Pointwise convolution: here the linear combination of each of these layers is considered which is nothing but a 1*1 convolutional operation on the output matrix of the first phase and as there are total of m channels so after having N such filters, we would have a total volume of Og*Og*N;

Here Og = No of strides in the input layer.

Now total multiplication here are: M *Og^2*k^2 which is significantly less costly than the normal convolution.

There are three flows Entry Flow, Middle Flow and Exit Flow. And all convolution and separable convolution are batch normalized. The flow starts from the entry then middle which is repeated 8 times and then the exit flow.

Xception is faster than inception even though it has the same parameters. Also provides high speed calculation with less complexity due to the separable Convolutions.

Provides top-1 accuracy of 0.790 and top-5 accuracy of 0.945.

Optimizers used here are Adam and SoftMax is the activation functionfor the last layer.

SoftMax

Most of the commonly used optimizers aresigmoid, RELU, TanH which all are mathematical functions having graphs but SoftMax is the optimizer which works on probability. SoftMax is generally used as the optimizer for the output layer but then it is useful to predict which input belongs to which class and the number of SoftMax units should be equal to the number of output classes.

f(y) = e^yi/Σke^yk

Each unit holds the probability of the class based on the concept of total probability that is probability of the class by sum of the probability of all the classes.

And the formula above shows the probability of a class given the total probability. SoftMax converts the real-world output numbers in probability. SoftMax is very helpful when it comes to multiclass classification and we could always know the higher probability output would be of the target class.

SJIF Impact Factor 2022: 8.197| ISI I.F. Value:1.241| Journal DOI: 10.36713/epra2016 ISSN: 2455-7838(Online)
EPRA International Journal of Research and Development (IJRD)
Volume: 7 | Issue: 5 | May 2022 - Peer Reviewed Journal

ADAM:

When we consider normal gradient descent algorithm the curve for the descent of the loss function is rather a zig-zag path for every epoch we travel upward vertically and then downward and it slows down our overall horizontal movement to the minima and to prevent that zig-zag flow of the optimizers are used to aid the training process and make it quicker.

ADAM is the combination of 2 optimizers that is RMSProp and Moment Gradient descent it is one of the RMS (Root Mean Square prop)

4. RESULT

The above techniques were implemented on covidX dataset which was further augmented using a variation of AC-GAN a total of 6000 images were generated which further divided into the train and test folders with 80 and 20 percent division ratio. The model was trained for 4 times with 5 ,15 25, and 35 epochs and after 35 epochs no significant improvement in accuracy was noticed and running for more epochs was a waste of computation power given the improvement. There was a total of 158 batches and each epoch took 158-160s to complete.

Use of pre-existing ImageNet weights and Adam as an optimizer helped to train the model at a quicker rate and the overall accuracy achieved was 95.67 and validation accuracy of 92.57 with loss value reduced to 0.675.

The charts and the confusion matrix for the above implementation:

```
Confusion Matrix for AUG_Xcep
[[279 333]
 [255 317]]
Classification Report for AUG_Xcep
              precision    recall  f1-score   support

    CovidCXR       0.52      0.46      0.49       612
 NonCovidCXR       0.49      0.55      0.52       572

    accuracy                           0.50      1184
   macro avg       0.51      0.51      0.50      1184
weighted avg       0.51      0.50      0.50      1184
```

Figure 7: Confusion Matrix of the detection model

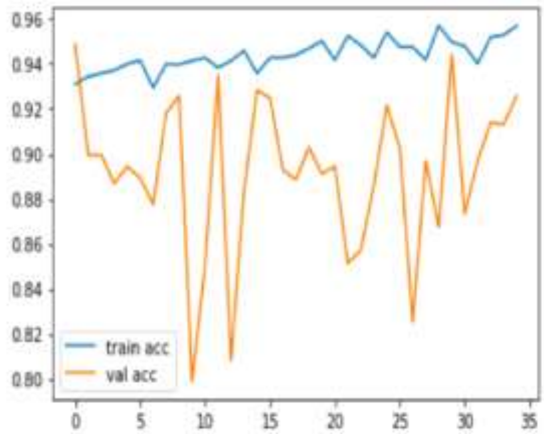

Figure 8: ***Training Accuracy and Validation Accuracy***

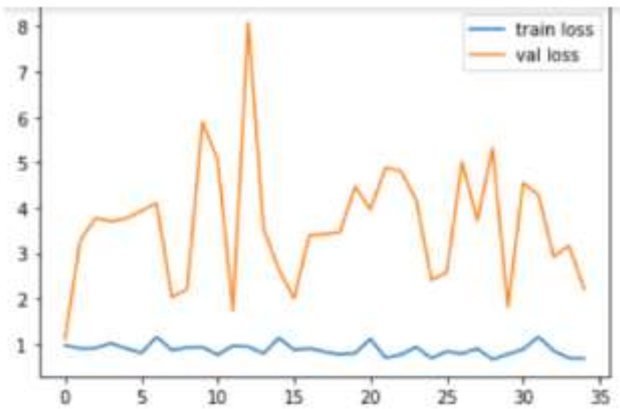

Figure 9***: Training Loss and Validation Loss***

.

SJIF Impact Factor 2022: 8.197| ISI I.F. Value:1.241| Journal DOI: 10.36713/epra2016 ISSN: 2455-7838(Online)
EPRA International Journal of Research and Development (IJRD)
Volume: 7 | Issue: 5 | May 2022 - Peer Reviewed Journal

5. CONCLUSION

The very purpose for working in field of chest x-ray detection was the situation and the chaos covid-19 as a pandemic caused in daily lives of citizens. People were afraid for visiting the covid centers for testing and detecting covid through x-rays was a better alternative and will also decrease the congestion at the RT-PCR centers. Having said that if we want this to be a alternate and a option for detection infections and diseases through x-rays a robust model with high classification accuracy is required. Lot of significant improvements have been made in this field since engineers started detecting pneumonia through x-rays and new paper related to covid -19 gave us some beautiful insights and information to carry out the research we wanted to conduct.

After reviewing more than 30 papers we set had set some requirements which we wanted our model to achieve those were decreased learning time and creating a balanced dataset which would allow us train the model perfectly. To achieve these things our solutions were GAN and Transfer Learning. GAN is a data augmentation technique which overall takes input a class of images and morphs those images to the desired output class. ACGAN was selected as a base model and some modification was done to it. After building the model and compiling it we then ran it over our initial dataset containing 3700 images 3000 non-covid and 700 covid images respectfully. To create a balanced dataset, we generated 2400 covid images and created a balanced mixture.

After that the augmented dataset was fed to the detection algorithm which in this case was a transfer learning model known as Xception. Transfer learning algorithms are trained on ImageNet dataset and have pretrained weights for all the scenarios, since our weights were already initialized this helped us to decrease the training time for the model which was nearly 160sec on avg for an epoch. In total we ran our model for more than 35 epochs but at epoch 35 the performance started to bottleneck and no such significant improvement was noticed after that considering how valuable those resources are we finished our training and got a accuracy of 95.67 with validation accuracy of 92.57.

6. REFRENCES

1. *Hashmi, Mohammad Farukh, et al. "Efficient pneumonia detection in chest xray images using deep transfer learning." Diagnostics 10.6 (2020): 417.*
2. *Bozorgtabar, Behzad, et al. "Informative sample generation using class aware generative adversarial networks for classification of chest Xrays." Computer vision and image understanding 184 (2019): 57-65.*
3. *Ravichandiran, Viveak, Aditya Srivastava, and Ying Chen. "Evaluation Metrics."*
4. *Loey, Mohamed, Florentin Smarandache, and Nour Eldeen M Khalifa. "Within the lack of chest COVID-19 X-ray dataset: a novel detection model based on GAN and deep transfer learning." Symmetry 12.4 (2020): 651.*
5. *Menon, Sumeet, et al. "Generating realistic covid-19 x-rays with a mean teacher+ transfer learning gan." 2020 IEEE International Conference on Big Data (Big Data). IEEE, 2020.*
6. *Liang, Zhaohui, et al. "Enhancing automated COVID-19 chest X-ray diagnosis by image-to-image GAN translation." 2020 IEEE International Conference on Bioinformatics and Biomedicine (BIBM). IEEE, 2020.*
7. *Khalifa, Nour Eldeen M., et al. "Detection of coronavirus (COVID-19) associated pneumonia based on generative adversarial networks and a fine-tuned deep transfer learning model using chest X-ray dataset." arXiv preprint arXiv:2004.01184 (2020).*
8. *Zhang, Longling, et al. "FedDPGAN: federated differentially private generative adversarial networks framework for the detection of COVID-19 pneumonia." Information Systems Frontiers 23.6 (2021): 1403-1415.*