International Journal of Engineering Trends and Technology
ISSN: 2231 – 5381 /doi:10.14445/22315381/IJETT-V70I2P227
Volume 70 Issue 2, 229-238, February, 2022

Original Article

HCUGAN: Hybrid Cyclic UNET GAN for Generating Augmented Synthetic Images of Chest X-Ray Images for Multi Classification of Lung Diseases

Swathi Buragadda[1], Kodali Sandhya Rani[2], Sandhya Venu Vasantha[3], M. Kalyan Chakravarthi[4]

[1]*Asst.Professor, Dept of CSE, Lakireddy Balireddy College of Engineering, Mylavaram*
[2]*Asst.Professor, Dept of CSE, Dhanekula Institute of Engineering and Technology, Gangur*
[3]*Associate Professor Dept of IT, Maturi Venkata Subba Rao Engineering, Hyderabad*
[4]*Assistant Professor (Senior) School of Electronics Engineering, VIT-Ap University, Amaravathi*

buragaddaswathi@gmail.com, kodalisandhyarani@gmail.com, vasantha_it@mvsrec.edu.in

Abstract - *In 2019, when COVID-19 started spreading as a virally infection across the globe, people started researching using the available chest x-ray images of lungs. But during the detection process, the research article found that the available images are less and most of the symptoms are similar to pneumonia disease. So in this article, before performing the multi-classification process, tries to enhance the size of the dataset by combining the UNET with cyclic GAN's. Most of the real-world data exist in an imbalanced form, which affects the overall architecture and performance of the design. Many of the researchers have implemented manipulation techniques like translation, rotation, and others for enhancement of dataset size, but due to the high dimensionality of the medical images, these basic and simple approaches doesn't have any impact on the model. So, the improved cyclic GAN's mechanism helps the research article to create a balanced dataset with more augmented or reconstructed CXR images by performing segmentation using the UNET operation.*

Keywords - *Cyclic GAN, Semantic Segmentation, UNET's, Up-Down sampling, Cycle Consistency, Contraction and Expansion Path.*

I. INTRODUCTION

Over the past few years, CNN has been the most audible term coined by most of the researchers, industry people, and other prominent persons of AI. CNN is a deep learning approach that is used to differentiate among the various objects of inputs based on the prioritized weights assigned to them. The main advantage of these neural networks lies in their recognition process of dependencies among the spatial and temporal features using different kinds of filters. The major goal of any network is to get understand the input images with less trainable parameters. In a real-time scenario, we can see images of different colour systems like RGB, CMYK and others, so there is a need for a network that can efficiently convert the image into an easy processing format, but still, it needs to maintain all its important features to get accurate prediction results. The major components of networks are:

i. Kernel: This is also known as "filter", which is used to read the pixel values of the image in stage by stage manner. The number of pixels it needs to process is dependent on the size of the kernel. Suppose we have K as 4*4*1 then, the reading of complete image needs 16 shift operations by reading 4 values horizontally and vertically. This can be pictorially represented, as shown in figure 1.

1	1	1	0	0	1
0	0	1	1	0	0
1	1	0	1	0	0
0	0	0	1	1	1
0	1	1	0	1	0
1	0	0	1	0	1

6*6 Input Image

Reading Operation

1	1	1	0
0	0	1	1
1	1	0	1
0	0	0	1

Kernel Operation-1

1	1	0	0
0	1	1	0
1	0	1	0
0	0	1	1

Kernel Operation-2

Fig. 1 Reading Operation by Kernel Component

The main aim of the convolution layers is to obtain the convoluted matrix, which contains information about the high-level features. With the increase in the number of layers, we can obtain low-level features like colour, gradients and others.

ii. Padding: It is an attribute that decides the size of the output obtained. If the padding attribute is considered valid, then it returns an output matrix with reduced dimensions. If the padding attribute is considered as same, then it returns the output matrix with the same number of dimensions as in the case of input.

iii. Pooling Layer: The time complexity of the system increases with the increase of features in the image. To preserve all the important features with high-speed processing, the CNN makes use of this pooling layer. It also reduces the computational power of the system. In the proposed system, it has implemented a max-pooling layer to extract the features from the image.

iv. Fully Connected Layer: Before sending the input to this layer, the system needs to flatten the matrix, to resize the matrix into necessary dimensions. After receiving the flatten values, the task of this layer is to identify the attributes that have non-linear relations between them. The proposed system used both feedforward and backpropagation technique along with softmax activation function to classify the age of the person.

II. LITERATURE SURVEY

Saleh Albahli [1] designed Efficient GAN’s using CXR images to balance the dataset. To create augmented images using a combination of an autoencoder with GAN’s during this process, the images are rated on the scale of 0 to 5 because the dataset consists of multiple classes starting from no disease to infiltration. If an image exceeds the threshold value, then it discards the image, and finally, this model has created 5000 images for each class label and made it balanced to design a proper detection system. The deep learning contains 1 pooling, 1 flatten and 2.

Fully-connected layers are designed to create fake images by involving simple image manipulations along with a pre-trained model known as “Inception”. The model also implements residual units to enhance the workflow of the model by simplifying the connections.

Mohamed Loey [2] proposed model-based GAN transfer learning to enhance the size of the dataset in which GAN’s are utilized to pre-process the images, and transfer learning involves the partition of data for achieving good validation results. In this process of transfer learning, applies three popular pre-trained models, known as “ALEX, GOOGLE, and RES” Nets to fine-tune the X-ray images by resizing all the images to 512*512. The main aim of this model is to reduce the loss function by defining the mathematical equation CONVET. After creating the augmented images, it computes the mean of each image and reutilizes it for performing transfer learning by defining the last layer with a customization size of 4*1during this process, the coefficients values of the transferred data updates and any values that exceed the threshold value gets automatically discarded from the mini-batch after each epoch. In this research, the model is designed using 3D CNN to produce a feature map vector in the form of a box layer, and the most important characteristic is defining the size of the kernel equal to the feature map and reducing the cost function by defining a non-linearity function which can perform the training process as fast as it can.

Pir Masoom Shah [3] created synthetic images using Deep Convolution GAN. In this research, the model works on three types of class labels, namely healthy, pneumonia, and covid. The major advantage of this research is before performing the classification. It checks the generality of the validation data by performing K-Means clustering on the three-class labels. To find the relevance between the features of the image, it performed an Attenuation map to check the confidence value of each image in the decision-making process. After this, the system creates synthetic data and trains it using the EfficientNETB4 to perform classification. This model uses activation maps to provide more information about the internal relationships that are hidden in the image.

Devansh Srivastav [4] combined GAN’s with a pre-trained model known as “VGG-16” the model used imbalanced data with high false-positive images, so it suffers from biased data, which results in poor accuracy results. The model, to reduce the bias, produced new images by applying image augmentation operations on original images but also by applying it on images generated using DCGAN’s. To design this, it implemented transposed convolution layers of 4 in size and generated random samples by using simple random distribution theory, whereas the Discriminator, which has to classify the images, works with 4 fully connected layers in the forward direction. To decide the number of normalization layers, the model applies statistical analysis, which tries to minimize the loss function and finds that it needs 2 layers of batch normalization and 1 layer of dropout.

Nour Eldeen M. Khalifa [5] designed a fine-tuned transfer learning model combined with GAN’s. In this model, it compared the combination of three pre-trained models with GAN’s, namely ALEX, GOOGLE, and RESNET-18 and found that RESNET-18 performs better than two other models. In this research, the generator model induced noise using the Gaussian mechanism to classify the model. The researchers designed a fine-tuned model by ensembling the Squeeze, ALEX and GOOGLE and augmentation is performed with 12 layers to fool the Discriminator as much as possible

Longing Zhang [6] proposed a novel framework known as “Federated GAN’s using differential mechanism”, whose major goal is “Privacy Preservation” of the patients whose records are stored across the hospital public clouds and also the original image data is not shared with the unofficial people by creating a mask image known as “CXR”. The federal model has been introduced in recent years, which try to produce the results by updating the global fitness value and differential equations proved its popularity by working on the adjacent databases in the distributed systems. The generator always tries to maximize the falsification, and the Discriminator always tries to minimize the misclassification to reduce the losses. In this paper, the model is implemented as a client-server by collecting random samples of different patients and using RESNET-51, it diagnosis the health condition of the patient.

Soumya Ranjan Nayak[7] performed a comparative study on various deep learning techniques using pre-trained models to discuss the best method to predict Covid-19. In this paper, the author compared 8 pre-trained models to classify whether a person suffers from Covid-19 or not based on the Chest X-ray. The comparative study is performed using the ImageNet dataset. The paper proposed the best pre-trained model based on the hyperparameters, which are well fine-tuned. The author illustrated that ResNet-50 and ResNet-34 are the best models based on accuracy.

Emtiaz Hussain[8] proposed a novel approach known as “CoroDet” for detecting Covid-19 based on the Chest X-ray images. In this paper, the author considered 4 classes of images to classify, namely normal, pneumonia-viral, pneumonia-bacterial and covid-19. The author generated the dataset by combining data from various sources and performed a data labelling task. Later, the author designed 22-Layer neural networks using three different activation functions and compared the performance of the model based on 2-classifier, 3- classifier and 4-classifier. The proposed model consists of 5 main components: The convolutional layer, the most important layer, which extracts the main features with the help of filters. The major goal of this layer is to extract important features like edges, colour and shape and also to perform crucial operations like blur, sharpening and detections. The pooling layer reduces the size of an image to achieve the computations fast. Flatten layer to transform the matrix values into a single column. The dense layer, a fully connected layer, helps in deciding which class the features belongs to most. An activation function helps the model classify the network images. In this paper, the author has chosen three different activation functions at different layers. The model has achieved better accuracy when compared to another state of the art methods.

Stefanos Karakanis[9] proposed a lightweight model for both binary and multi-classification. The model for binary classification has not implemented transferred learning. The model has implemented pre-trained model ResNet8 with Adam, and cross-entropy optimization functions and training process is performed with the help of Conditional GANS, which are familiar for creating new synthetic images. The model evaluation of neural networks is based on the loss function, which evaluates the difference between the actual labels and predicted labels. The task of the optimizer is to choose the loss function which gives maximum accuracy with the least loss. Among the optimizers available, the proposed model uses Adam optimizer, which calculates the average of momentum exponentially. The major advantage of Adam is it uses bias correction when the process becomes slow. The conditional GANs generate the new images based on the conditions of the receiving information. The GANs suffer from the problem of controlling the incoming information. To solve this issue, Conditional GANs uses vector information of features. Finally, the model is evaluated with the help of the Grad-CAM heat map. The proposed model has achieved 98.3% accuracy.

III. MATERIALS AND METHODS

In the proposed research, the model worked on the dataset collected and published by the NLM, USA. The dataset contains 326 healthy images, 330 pneumonia and 400 COVID patient details. A few samples of CXR images are illustrated in figure 2.

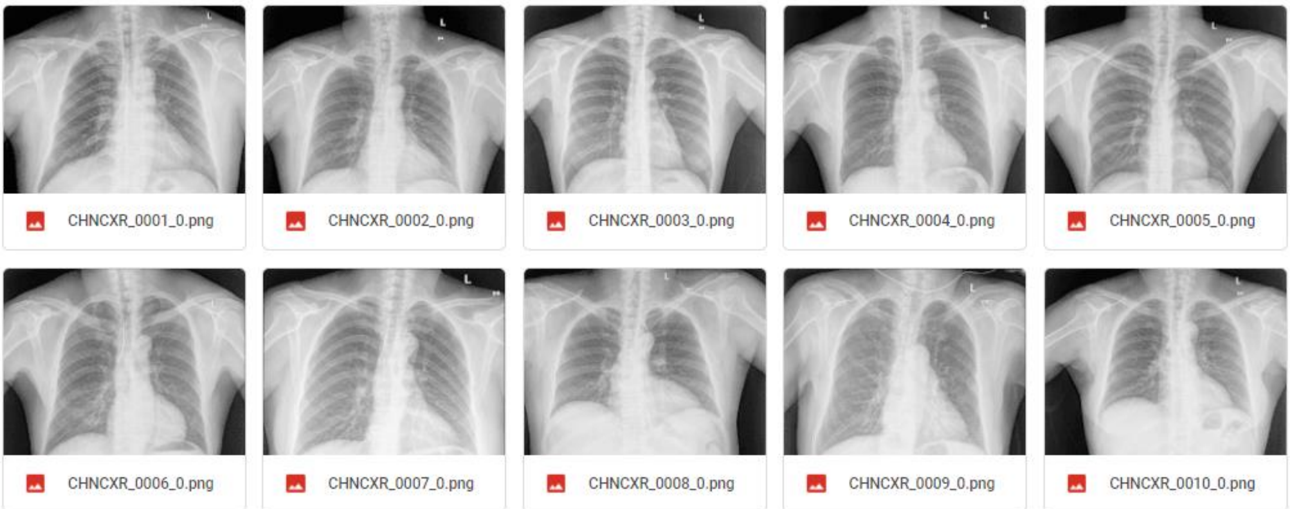
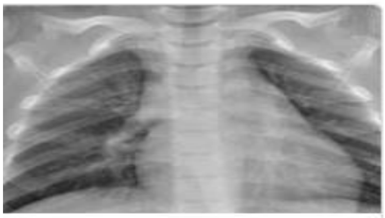
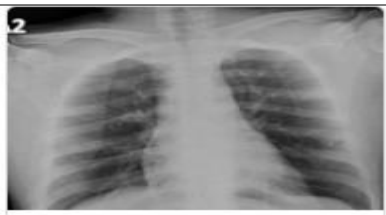


Fig. 2 Sample Images of CXR collected by NLM, in association with various hospitals

The dataset also provides the clinical recordings of the images in the textual format where the annotated data consists of the age, gender and abnormality of the patient. Few sample recordings are illustrated in table 1.

Table 1. Clinical Findings of the Patients

S.No	Image	Details of the Patient
1		Gender: Male Age: 45 yrs Abnormality: normal
2		Gender: Female Age: 29 yrs Abnormality: Pneumonia
3		Gender: Female Age: 68 yrs Abnormality: COVID

IV. PROPOSED METHODOLOGY

To design an efficient deep learning neural network, these small amounts of samples are not sufficient; because of this, it results in either high bias or high variance. To solve this issue, the model aims to generate synthetic data using UNET integrated GAN's in which augmented images are created for masked ones by translating and reconstructing the entire image. The entire process is illustrated in below sections:

A. Data Visualization

In machine learning, it is easy to understand the data relation between the features, but in deep learning, the relationship between the complex features cannot be presented with the linear elements but learning the features from training data is very important in learning applications. So, to analyze the data, the proposed research using traditional approach plotted images as shown in figure 3

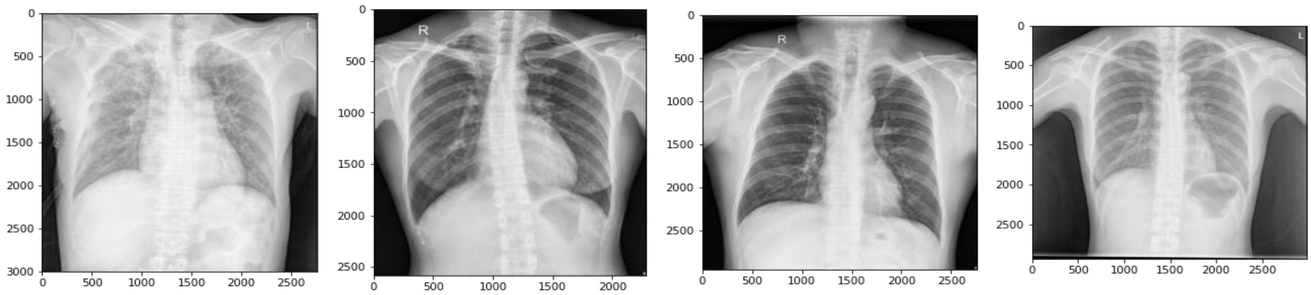


Fig. 3 Analysis of the Original Chest X-ray Images

a) Cyclic GAN: The main intention to implement cycle GAN in the proposed research is it translates one image to another by performing pixel by pixel operation. In the CGAN process, a simple image is represented as a map vector and using the encoder, it performs an embedded *operation and produces it as an output of a generator, which is* known as "Translated Styles". The Discriminator gets two inputs one is "Translated Styles" and "Targeted Styles", and it constructs a vector map representing fake as "0" and real as "1". The latent space is utilized to perform the vector operations. In this research, more consideration is given for loss incurred due to consistency because of the cycle formation, i.e., translation of one image to another and reconstructing the original image from the generated image. The consistency loss can be computed as shown in equation 1.

$$\mathbf{Loss}(\mathbf{Image}) = \mathbf{G}\left(\mathbf{G_{New}}(\mathbf{Image})\right) - \quad (\mathbf{1})$$

It also reduces the cost and complexity because it doesn't need any pairing between the images. To construct a CGAN, the model needs two generators, and the conversion process is illustrated simply in figure 4.

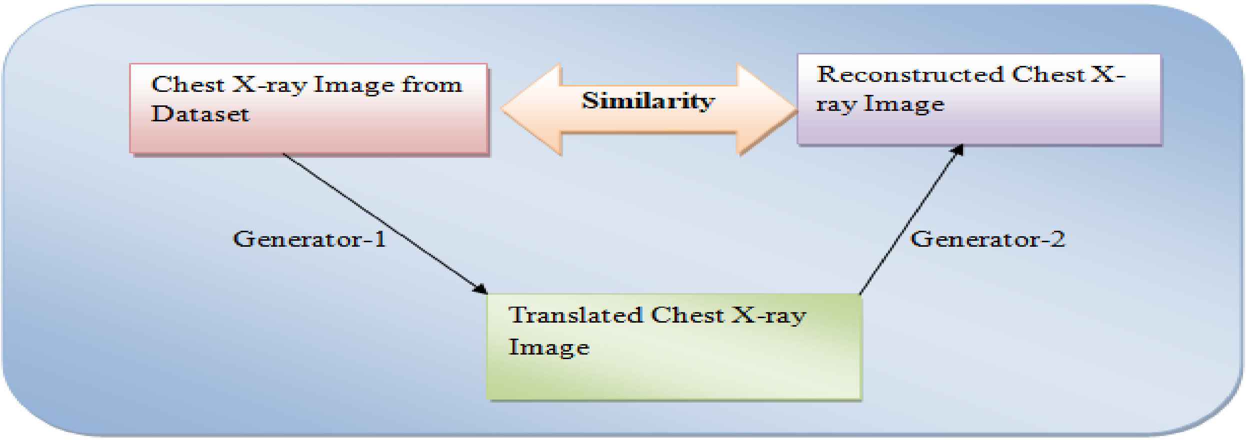


Fig. 4 Generators Application on Chest X-ray Images

The flowchart for the discriminator design is illustrated in figure 5 with instance normalization technique

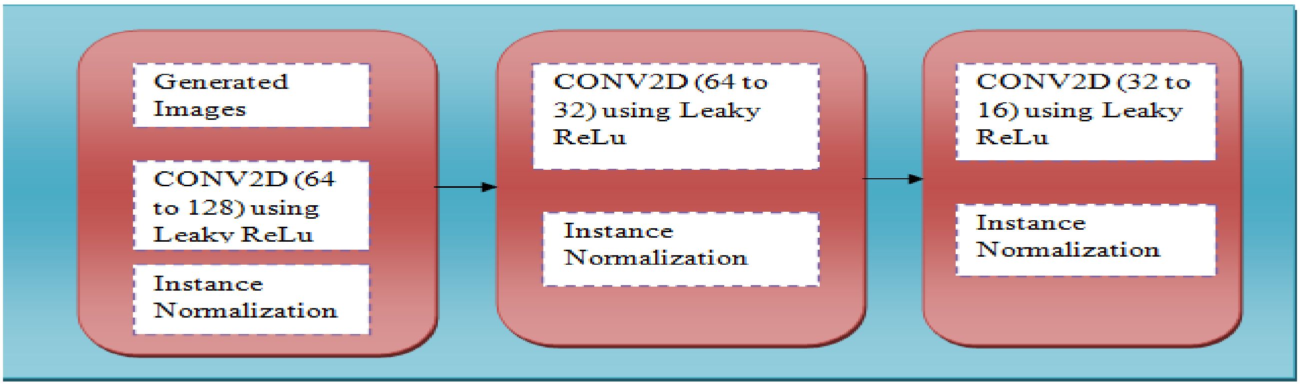


Fig. 5 Discriminator Architecture

b) UNET Segmentation: The main reason for combining UNET with CGAN's is to produce semantic images, and UNETS are popular for semantic segmentation. This semantic segmentation process assigns a class label to each image pixel and tries to find the predicted class labels in deep. The UNET model consists of an encoder designed using the covenant layers combined with pooling layers to extract the factors that impact each component of the layer and a decoder designed using the transposed CNN to identify the localized parameters, which are arranged in fully connected mode. The segmented images are shown in figure 6.

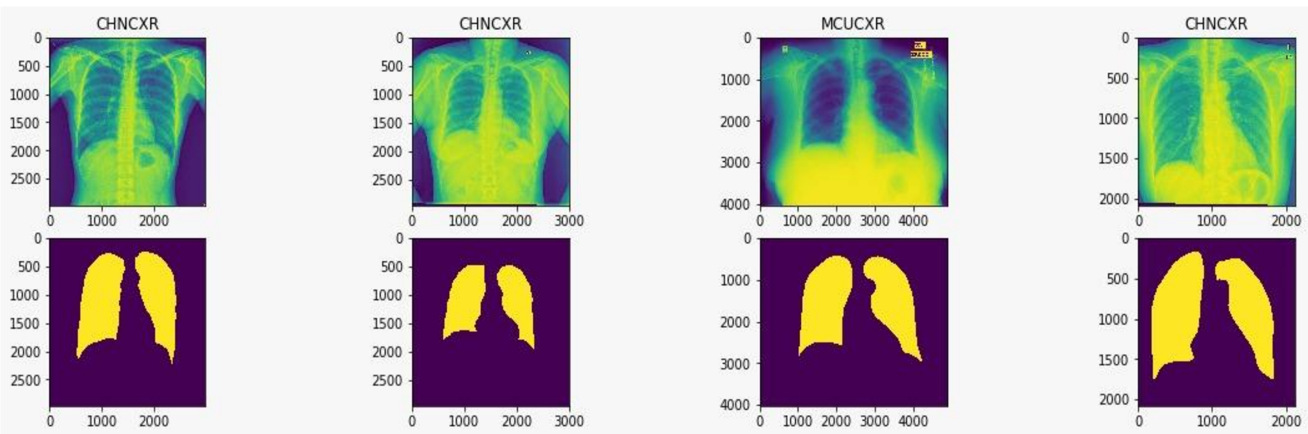


Fig. 6 Segmentation on Chest X-ray Images after performing the UNET operation

The model generates a dense feature map with the same dimensions as the input RGB image (height and width). There are as many channels in the output feature map as there are classes for each pixel to predict from. The upsampling process trains and gets the details of the learning parameters of the model by repeating some blocks, namely CONV2D, MAXPOOL, and DROPOUT. It decreases the resolution of the image by designing the contrasting path to get the details of the labels associated with the pixel, and the expanding path tries to optimize the location parameters. The UNET architecture in the proposed research is designed as shown in figure 7.

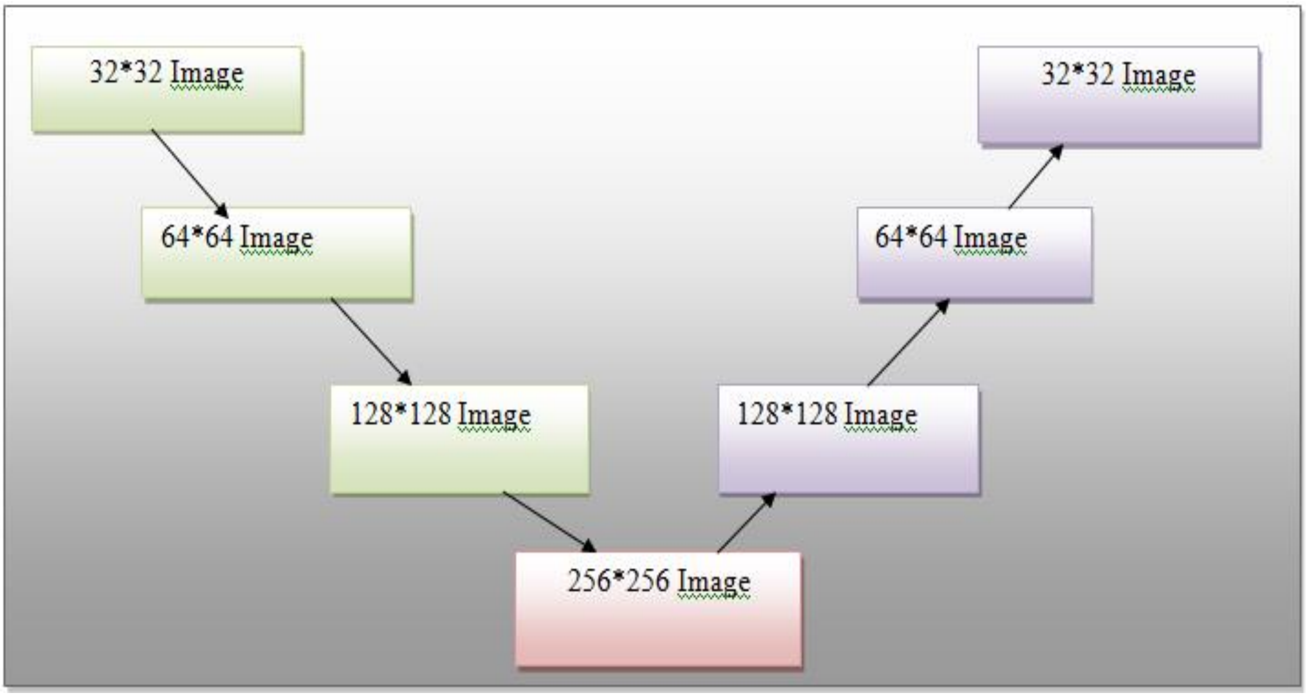


Fig. 7 Proposed UNET Architecture for Semantic Segmentation

V. RESULTS AND DISCUSSION

This section discusses the various output produced during the training and GAN designing phase. Figure 8 displays the Generator and Discriminator during the contraction phase of UNET Segmentation. The translation and reconstruction architecture sizes are represented clearly

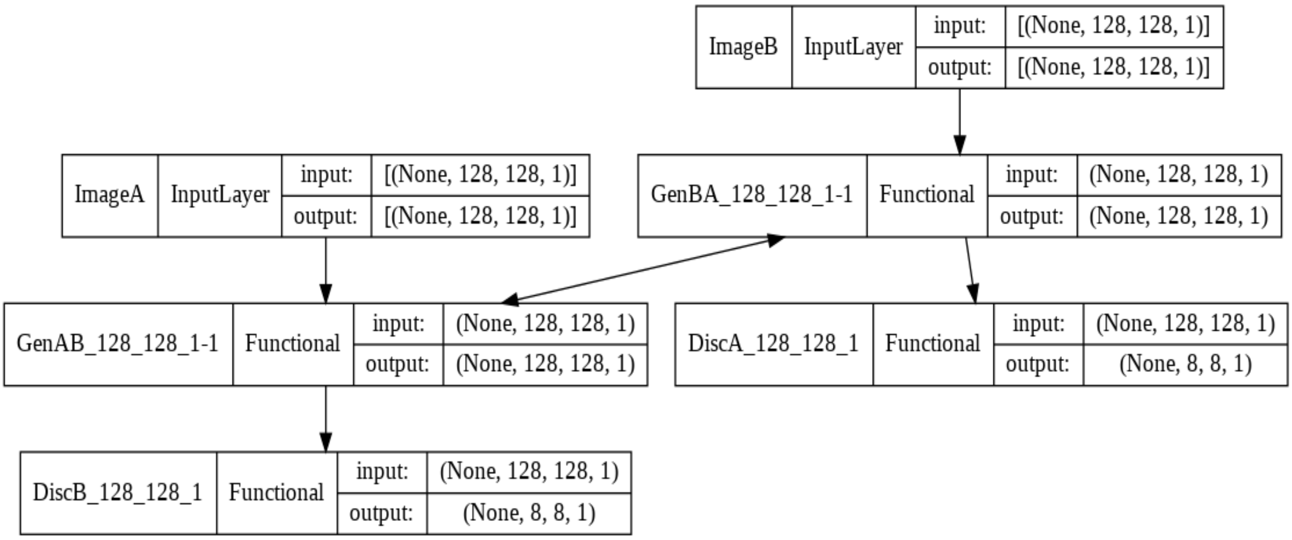


Fig. 8 Generator and Discriminator Functional Creators for both Images

Figure 9 represents the overall architecture designed during the discrimination phase using the instance normalization concept.

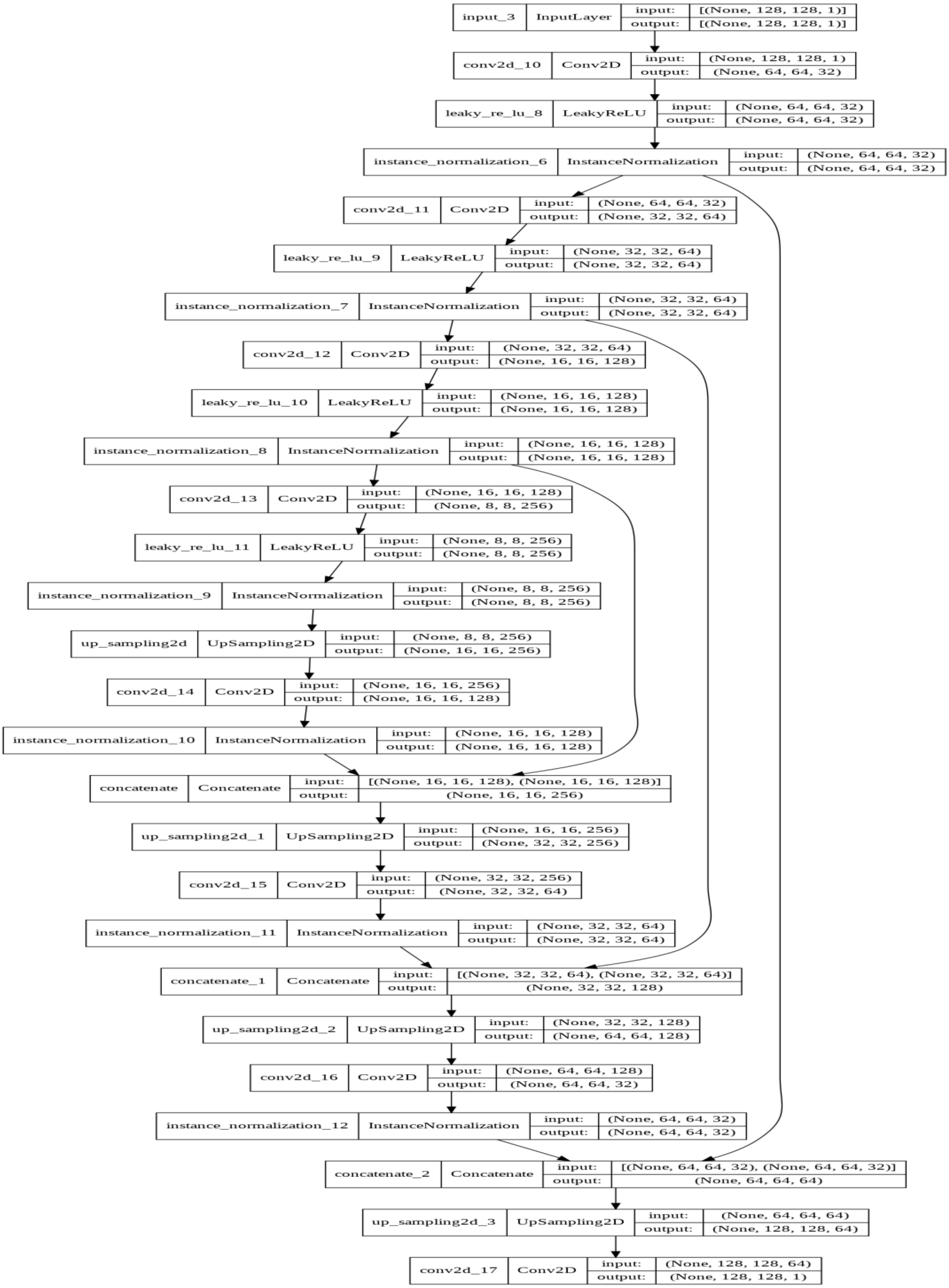


Fig. 9 Number of Trainable Parameters at each layer at Discrimination Phase

Figure 10 discuss the epochs iterating process to make the system get familiarity with the distinct features involved in the image. Less number of epochs may miss learning about the important and high context features, and more number of epochs may slow down the entire system, and the efficiency of the epochs are associated with the learning rate.

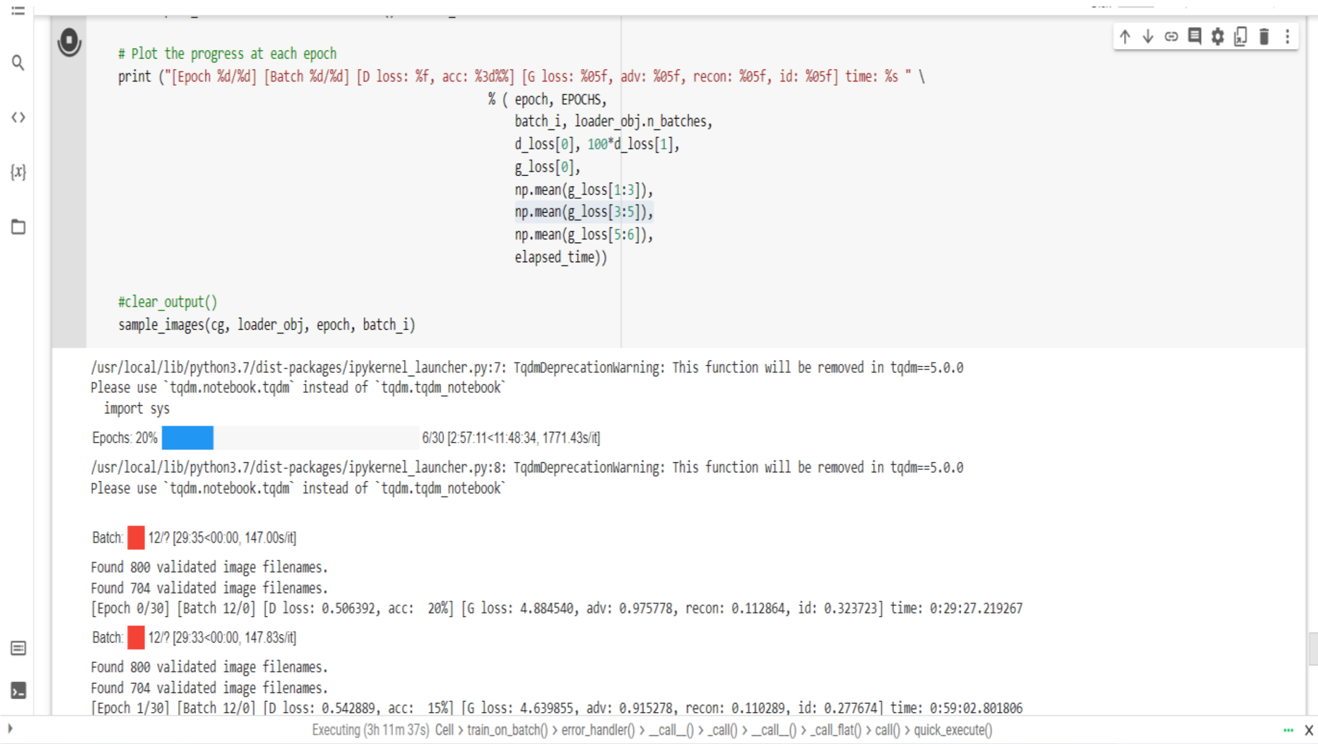


Fig. 10 Training during the Initial Stages of epochs

In Figure11, the proposed model original A represents the domain A of Generator and Original B represents domain B of Generator. Using double generators in the proposed research, the model produces a cycle by reconstructing the original image.

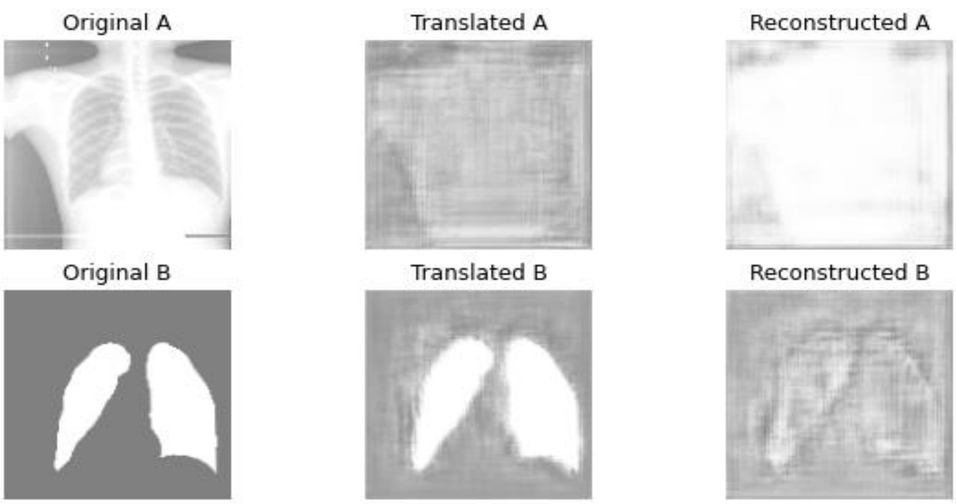


Fig. 11 Translated and Reconstructed Images for Original Images

The final output produced by the HCU GAN is represented in figure 12. These images are passed as input for the Discriminator to classify the images as real and fake.

Fig. 12 Augmented Images creation of Chest X-rays using GAN's

In [10], Khalid EL ASNAOUI et al. experimented with different deep learning architectures. It has applied the min-max normalization technique to process the pixel values of the image and CLAHE to improve the pixel intensity of low-quality scanned images. After that, it has studied pre-trained models like VGG-16,19, DenseNet, and others. The performance of these models is illustrated in table 2.

Table 2. Comparative Study of Different Pre-Trained Models

S. No	Model Name	Number of epochs	Accuracy
1	CNN	300	84.18
2	VGG16	300	86.26
3	VGG19	300	85.25
4	Inspection_V3	300	94.59
5	ResNet-50	300	96.60
6	UNET+CGAN	300	97.19

VI. CONCLUSION

The conditions are uncontrollable in a real-time context, and the system may encounter issues such as illness dataset labelling, addressing the data imbalance problem, and overfitting issues due to the reduced size of diagnostic hints in the real-time environment. The system also suffers from overfitting because selecting and applying all of the potential geometric transformations necessitates a large number of computations. So, by leveraging pre-trained models and developing a solid policy schema, such as Hybrid Cyclic UNET GAN, the auto augmentation process can effectively enhance accuracy. Because the Hybrid Cyclic UNET GAN modules create additional training data, the system is able to handle the problem of overfitting that happens as a result of simple picture processing. From table2, it is a clear observation that when GAN's are combined with pre-trained models, it has got less accuracy, but the proposed model instead of trained models it combined with semantic segmentation component known as "UNET" and obtained the best accuracy of 97.19, which approximately 1.5% better than the base model.

REFERENCES

[1] Albahli S., Efficient GAN-based Chest Radiographs (CXR) augmentation to diagnose coronavirus disease pneumonia. International journal of medical sciences, 17(10) (2020) 1439–1448. https://doi.org/10.7150/ijms.46684

[2] Loey, M., Smarandache, F., & M. Khalifa, N. E., Within the Lack of Chest COVID-19 X-ray Dataset: A Novel Detection Model Based on GAN and Deep Transfer Learning. In Symmetry 12(4) (2020) 651. https://doi.org/10.3390/sym12040651

[3] Shah, P. M., Ullah, H., Ullah, R., Shah, D., Wang, Y., Islam, S. ul, Gani, A., & Rodrigues, J. J. P. C., DC-GAN-based synthetic X-ray images augmentation for increasing the performance of EfficientNet for COVID -19 detection. In Expert Systems. Wile (2021)y. https://doi.org/10.1111/exsy.12823

[4] Srivastav, D., Bajpai, A., & Srivastava, P., Improved Classification for Pneumonia Detection using Transfer Learning with GAN based Synthetic Image Augmentation. 11th International Conference on Cloud Computing, Data Science & Engineering (Confluence) (2021). doi:10.1109/confluence51648.2021

[5] 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:2004.01184 [Cs, Eess], (2020). arXiv.org, http://arxiv.org/abs/2004.01184.

[6] Zhang, L., Shen, B., Barnawi, A. et al. FedDPGAN: Federated Differentially Private Generative Adversarial Networks Framework for the Detection of COVID-19 Pneumonia. Inf Syst Front 23 (2021) 1403–1415. https://doi.org/10.1007/s10796-021-10144-6

[7] Soumya Ranjan Nayak, Deepak Ranjan Nayak, Utkarsh Sinha, Vaibhav Arora, Ram Bilas Pachori, Application of deep learning techniques for detection of COVID-19 cases using chest X-ray images: A comprehensive study, Biomedical Signal Processing and

Control, 64, (2021) 102365, ISSN 1746-8094, https://doi.org/10.1016/j.bspc.2020.102365.

[8] Emtiaz Hussain, Mahmudul Hasan, Md Anisur Rahman, Ickjai Lee, Tasmi Tamanna, Mohammad Zavid Parvez, CoroDet: A deep learning-based classification for COVID-19 detection using chest X-ray images, Chaos, Solitons & Fractals, (2020) 110495, ISSN 0960-0779, https://doi.org/10.1016/j.chaos.2020.110495.

[9] Stefanos Karakanis, Georgios Leontidis, Lightweight deep learning models for detecting COVID-19 from chest X-ray images, Computers in Biology and Medicine, 130 (2021) 104181, ISSN 0010-4825, https://doi.org/10.1016/j.compbiomed.2020.104181.

[10] Asnaoui, Khalid El, et al. "Automated Methods for Detection and Classification Pneumonia Based on X-Ray Images Using Deep Learning." ArXiv:2003.14363 [Cs, Eess], (2020). arXiv.org, http://arxiv.org/abs/2003.14363.