RETINAL IMAGE SYNTHESIS FOR DIABETIC RETINOPATHY ASSESSMENT USING DCGAN AND VAE MODELS

[1]Y.SRAVANI DEVI, [2]S.PHANI KUMAR

[1]Research Scholar, GITAM Deemed to be University
[2]HOD-CSE, GITAM Deemed to be University
E-mail: [1]y.sravanidevi@gnits.ac.in, [2]psingams@gitam.edu

Abstract - Amongst the most significant considerations in classification of the image is the data amounts particularly in the medical images. Although the main challenge in the healthcare sector is ‘attaining the datasets. In this, we display the images of the synthesized retinal fundus by preparing a VAE i.e., Variational Autoencoder & another model known as the DCGAN, adversative model on almost 3662 images of retina which have been captured from a dataset known as the APTOS- Blindness dataset. The finding of this method is in creating the images of retina without the usage of vessel segmentation that is previously used. This enables the system to become independent. The models which are acquired are the synthesizers of the image that are proficient in producing resized images of retina of any amount from a basic regular distribution. Moreover, a lot of images than this have been utilized for the purpose of training than any other task in literature. The assessment or appraisal of a synthetic image is done by giving an output to a CNN model & the average squared error was counted between the average 2-Dimensional hologram of images that were real and synthetic as well. Later, by analyzing the latent space and average loss of the images. The achieved outcomes out of the analysis inferred that the general images have less extent of loss in DCGAN as opposed to Variational Auto Encoders.

Keywords - Diabetic Retinopathy, Data Augmentation, Generative Adversarial Network, DCGAN, Variational Auto Encoder.

I. INTRODUCTION

Diabetic retinopathy (DR) is a very common disease-is responsible for the loss of vision among diabetic people. ophthalmologists or eye specialists usually detect, & check DR severity based on number of similar lesions and types. In accordance with the international convention, the intensity of Diabetic Retinopathy can be classified in 5 levels: normal, mild, moderate, severe non-proliferative DR (NPDR) and PDR. The lesions comprise of exudates, soft exudates, haemorrhages, microaneurysms, laser marks, proliferate membranes, etc. It is very laborious & uneasy for even for the eye specialists to detect DR, so automated DR grading models have started to be considered over the past few years. Numerous past works, choose deep models to carry out DR grading and acquire significant advancement over other techniques. Training an efficient deep CNN model generally needs a huge amount of varied and balanced data. Despite that, the data distribution of DR over various grades are immensely disbalanced as aberrant images of fundus only make up a little portion. For instance, in the biggest public DR dataset, EyePACS , images of the levels of DR 3 and 4 are 2.35% and 2.16% of the whole, individually, while regular images of level 0 are 73.67%. Accepting such disbalanced data will make the model less responsive to various samples with greater severity levels of the DR and will result in overfitting. No wonder that basic techniques of data augmentation [11] like flipping & random cropping & rotation can alleviate the issue, the poor variety of the samples from those levels are still responsible for restricting the performance of model. Therefore, in this research paper, we suggest an image generation DCGAN model which creates more multifarious images of DR with various grading levels and utilize these generated images to aid in the training of a grading model.

A. (GANs) known as Generative Adverbial Networks.

In this deep learning process 2 neural networks are utilized. The first is the generator network and the second is discriminator network. Each one of them is a deep neural network. Basically, in the Generator network, an arbitrary noise is taken as an input to create data that are considered as samples as practicable as possible to an original dataset & also a discriminator network that differentiates between the actual or original data and the generated data as depicted in figure 1.

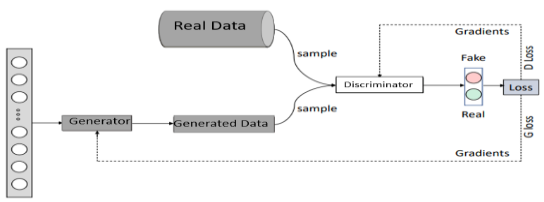


Fig 1: Architecture of GAN

GANs [9] are basically structures that are prepared to develop realistic designs or actual objects that are hard to discern from the existing real and actual objects i.e. GANs take the training data distribution and create examples that are new from the very same distribution. The GANs are multipurpose models that comprises of 2 different neural network models: a generator and a discriminator. The main objective of the generator network is to create reasonable fake examples. On the contrary, the objective of the discriminator network is to make a distinction between the real example created by the training data & the fake one which is produced from the generator network. Instinctively, a person can see a 'generator network' as forger which falsifies examples to make it look as actual or realistic as possible & the discriminator network as an examiner which tries to distinguish the real and the fake one. During the training, the generator network gets ameliorated at creating artificial examples. In the similar manner, the discriminator network prepared to become an improved investigator or examiner which appropriately distinguishes between the real and the fake samples i.e., it acquires to model the likeliness of an example and identifies it as fake or real. The likeliness or the chances of the model of being real or forged from the discriminator is the one that aids the generator network to create improved samples over the course of time. The balance of the game is where the production of realistic samples that are fake by the help of generator which look identical to the actual samples acquired from the training data. At the same time, the discriminator is left speculating at a fifty percent of the likeliness or the probability that discerns whether the example is fake or real.

II. RELATED WORK

The method or procedure of a deep convolutional denoising autoencoder is founded on the complete changing multi-norm loss function. the minimization feature along with the batch normalization methods have been presented to restore the fundus. It is applied for the restoration of the fundus and to low down the level of noise. On the top of that, the network speed is basically fast for the loud images aa opposed to few other models. It is basically done by fine adjustment of the tunes of the network which requires the dropout tools.

A DR produced generative adversarial network (DR-GAN) [1] to create fundus images that have a high definition can be changed with lesion information and arbitrary grading. The conditioning of Retina generator takes place on the lesion and constructional masks. Moreover, as vectors that are adaptive grading which are modelled from the latent grading spaces that can be used to manage the created grading intensity.

The efficacy of the multitask learning which is related to the problems associated with regression. In materials science, the conducted experiments take place on one ionic conductivity dataset and 7 benchmark datasets. The inference of the experiments that have been conducted indicate that an improvement of performance in generalization of differently variable linear regression examples which take place in the multi-task learning.

With the usage of DCGAN [2], development of retinal images take place which no longer require the procedure of vessel segmentation. Thus, the new procedure makes it totally unreliable. The models or the examples that have been acquired have the potential of synthesizing trimmed retinal images of any amount from a regular normal distribution. Moreover, many images were utilized in the process of training than any other persisting model. Another method in the synthesis of retinal image used as a system trained on the vessel networks & their equivalent images of retinal fundus. To put in another way, a transformation has been learnt between the retinal fundus and the vessel trees. The prominent drawback of their approach is the reliability of an autonomous algorithm to separate the vessels.

On the foundation of Transfer learning on AlexNet & GoogLeNet models, this methodology is being used an image of DR classification model [6] was explained. The usage of this model is for grading the DR level. Moreover, the EyePacs dataset is utilis ed to get a training set which comprises of images that are 35,126 in number & also the test set 53,576. We observe increased percentage by 90 in the achievement of sensitivity and specificity by the suggested DR interpretable classifier & enables it to identify more intense Diabetic Retinopathy cases. This disease seems to be a continuous disease that needs immediate detection as it is significant to stop the growth of this disease.

Likewise, a regular screening is important in the protection of the eyes of the patient. Thus, the generation of proficient & dependable structure of computer assisted diagnosis of DR as CAD system. The diagnosis of DR detected by the identification of unusual structures in the fundus especially Exudates, Haemorrhages, bright and dark lesions, cotton wool spots, Microaneurysms. Therefore, it is essential to do the segmentation of these parts or elements in a very accurate manner to have a better identification, detection, and localization. The significant methods [12] that are used for detecting the major clinical elements of the DR are 'supervised and unsupervised learning techniques. Disease offering automated screening based on the images of retina.

III. PROPOSED VAE AND DCGAN METHODS

A. Variational Auto Encoders:

1.1. It comprises of 2 neural networks: the first is the encoder also known as the approximate inference network which is responsible for mapping a training sample to a latent or a hidden space & the decoder network which plans or maps from the latent space to an artificial sample. In this task, the latent space is centred isotropic multivariate Gaussian & the encoder & decoder which are completely linked or associated neural networks which consist of an individual hidden layer. Moreover, in the phase of learning also called

Training phase, the encoder acquires the latent variables z from the input data & the decoder extracts those sorts of variables to produce a sample. After that, during the phase of generation, VAE pulls samples from the latent space that rush through the decoder to ultimately acquire an artificial sample, also known as the synthetic sample. The framework of VAE can be depicted from the Figure 2(a)

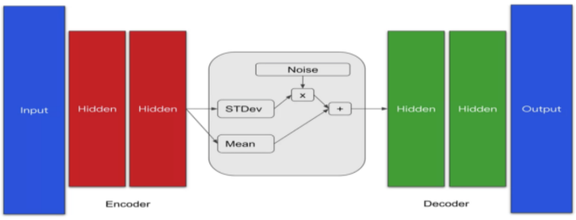


Fig.2. Architecture of Variational Auto Encoder

B. Deep Convolutional Generative Adversarial Networks:

GAN's or GAN are the deep neural net framework which is composed of the two nets. One is known as the generator and the other (the adversary) is known as the discriminator. A grade of CNN also known as the Deep Convolutional Generative Adversarial Networks (DCGAN) that are supported on a particular strategy. The main advancement on the very first GAN is this framework which produces improved or enhanced quality images & more stability during the stage of training. Following the instructions to create the generator & the discriminator as explained in the research paper by Radford et al., we applied & trained them on the retinal images that have been cropped with the usage of the generator cost functions & the original discriminator. On the similar lines, as in the VAE technique, artificial image generation with the usage of DCGAN [10] majorly comprises of the 2 phases: One is the learning phase and another one as the generation phase. In the first phase, the generator basically pulls out the examples from an N-dimension regular distribution that rush through the generator to acquire an artificial sample & the discriminator effort to make a distinction between the images taken from the generator & the training set images. We can also look at the picture of DGCAN framework or architecture in the Figure 4(b)

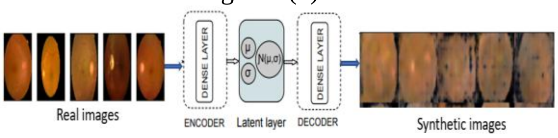


Fig.3. a) Schema of VAE architecture

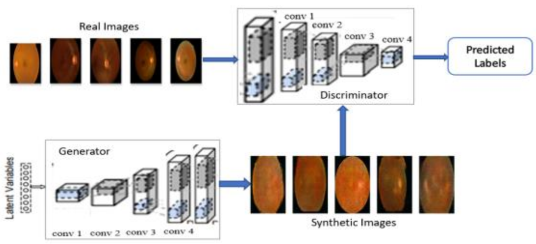


Fig.3b) Schema of DCGAN architecture

The framework has a few developments on the current GAN's. One of the changes are the substitution of all the pooling layers along with the strides convolutions in the fractional strides convolutional in the engine , the usage of Batch normalization in each one of the generator & the discriminator, the complete substitution of completely associated or linked hidden layers with the mean pooling in the end, the usage of LeakyReLU stimulation in the engine or the generator for all of the complete layers excluding the output & the usage for output & use of LeakyReLU stimulation pf every layer in the discriminator. No wonder that the research works have successfully been enhanced in the adversarial models, but the major demanding task is basically the model training. To attain the solution of this issue, we have tracked the suggestion advised in [2] to acquire stability while training tge DCGAN. Suggestions for example the normalizing the input images between -1 & -1, ADAM optimizer used for the engine or the generator with the utilization of a Gaussian distribution for the mini batches which consists of original images that have been utilized for the training of the models and also the latent space.

IV. RESULTS AND DISCUSSIONS

A. DATASET:

An amount of 3662 images from the APTOS Blindness. The testing images almost 1928 in number and the Train images almost 3662 in number have been utilised for the training of the models utilized in this work. The complete dataset comprises of almost 18590 fundus images which are separated into almost 3662 training, 1928 validation & testing image which are almost 13000 in number by the arrangers of Kaggle competition.

In this task, all of the complete experiments have been conducted, a deep learning library consisting of an open-source NVIDIA Titan Xp GPU & Keras were uitlized.

B. Loss Functions [9]:

As we are aware that in the normal GAN, the model of DCGAN imitates a sort of a competition in which the Generator tries to create original or natural images developed by the generator. Increasing the misclassification error of the discriminator while developing more real like images & attempting to trick the discriminator is one of the major objectives of the DCGAN model. This is also known as a 2-players minimax game and can be explained here-

$$Ex\,[\log(D(x)] + Ez\,[\log(1 - D(G(z)))]$$

Where Ex refers to the expected value over all actual data instances, $D(x)$ which is basically the discriminator's estimate of the possibility that the actual data instance x is real, $G(z)$ also known as the generator's output when an actual noise has been given z, $D(G(z))$ is the discriminator's evaluation of the possibility that a take instance is genuine. The expected value (Ez) over all the arbitrary inputs to a generator.

Thus, the arrangement is basically qualified to minimize $\log(1 - D(G(z)))$ & maximize $\log(D(x))$. Now, basically the training of the VAE & DCGAN architectures on the re-scaled & the pre-processed images of retina from the APTOS Blindness dataset with no usage of any data augmentation. For every image size, testing of a scope of N-dimensional latent spaces from almost 32 To 100 latent variables. Every latent variable. Every latent space was examined to control all the systems that don't remember the training database & meanwhile it produces reasonable images of the retina. To do this, Estimation of intermediate latent representation points is done. To train the model of VAE, we conducted a lot of tests & concluded that greater results have been acquires when we utilized a 512-dimension latent space & also 1008 * 1008 spatial resolution. With a little batch size of 64, and running for almost fifty epochs, we acquired the artificial images as depicted in the Figure 4.

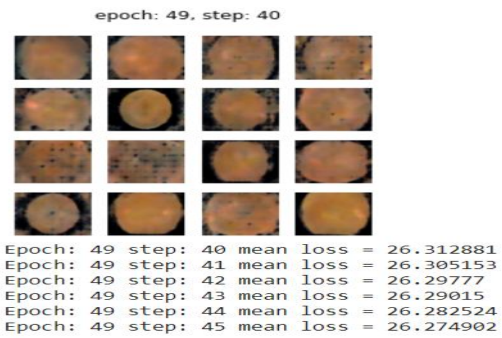


Fig.4a)

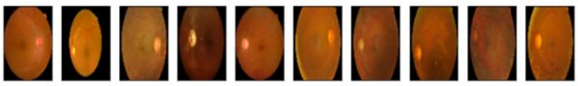


Fig.4b)

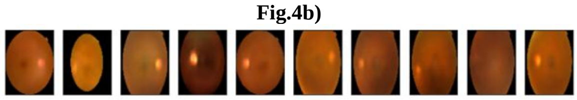


Fig.4c)

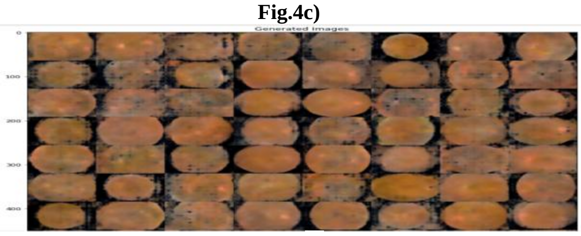


Fig. 4d)

Fig.4. Samples of images produced using the VAE architecture: a). Produced images for each epoch with mean and loss. b) & c). plotting the sample and predicted images. d) plotting the generated images in a grid.

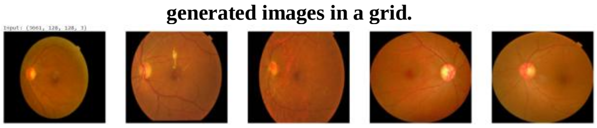


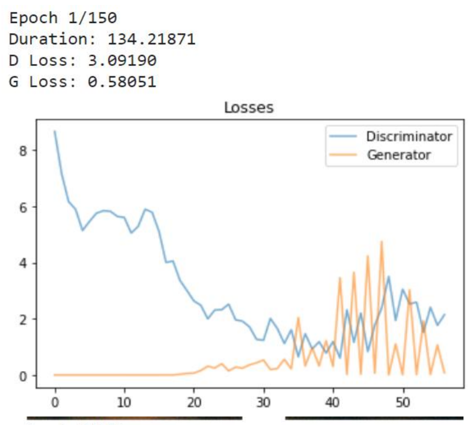
Fig.5: Examples of synthetic images generated by the DCGAN architecture

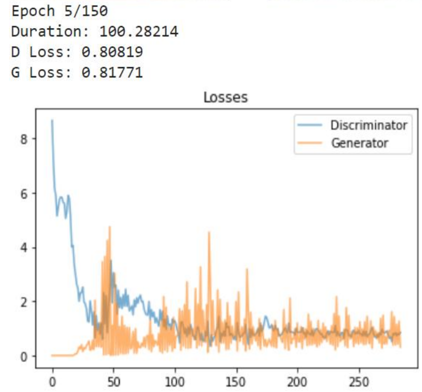
Even though the texture of the synthetic images acquired from VAE is like the realistic images, they are hazy, high loss and don't consist of the expected properties in a fundus image. With the reference to the DCGAN architecture, we have been able to find that realistic images were obtained when using an image size of 128×128 pixels, a small batch size of 64 and 32 epochs. Examples of them are shown in Fig. 5.

The major merit of using this architecture is basically the sharpness of these synthetic images that those which have been created using the VAE technique. For Confirming the discriminator images and the generator, we can count the loss for both. VAE has suffered more loss as compared to the DCGAN. Therefore, the estimation or the evaluation of the images solely created by the DCGAN was continued. The acquired results are expressed in the Table-1 & the figure as well.

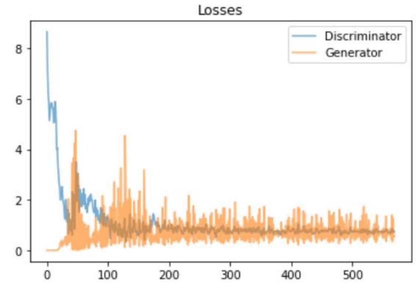
S. No	Epoch	Discriminator Loss	Generator Loss
1	1	3.09	0.58
2	5	0.80	0.81
3	10	0.74	0.75
4	15	0.728	0.726
5	20	0.72	0.77
6	25	0.70	0.71
7	30	0.68	0.77
8	32	0.68	0.77

Table. I Discriminator and generator loss of DCGAN at random epochs

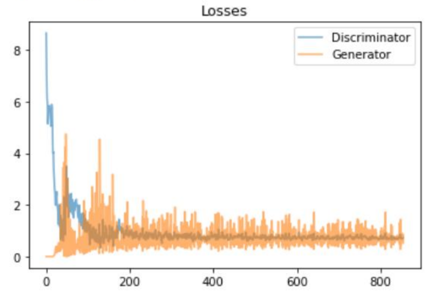




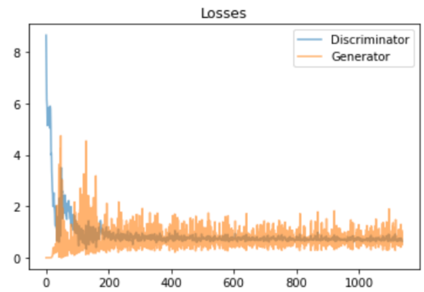
```
Epoch 10/150
Duration: 100.52345
D Loss: 0.74571
G Loss: 0.75054
```



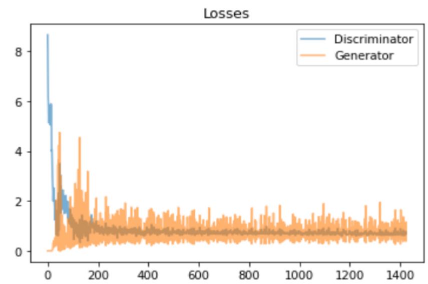

```
Epoch 15/150
Duration: 100.61579
D Loss: 0.72872
G Loss: 0.72687
```



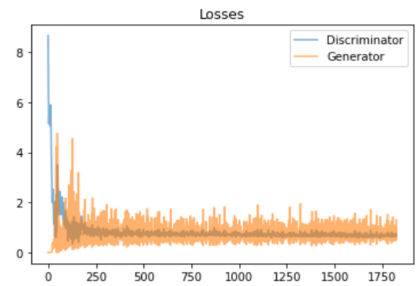

```
Epoch 20/150
Duration: 100.44941
D Loss: 0.72267
G Loss: 0.77153
```




```
Epoch 25/150
Duration: 100.29597
D Loss: 0.70593
G Loss: 0.77279
```




```
Epoch 32/150
Duration: 100.57005
D Loss: 0.68908
G Loss: 0.77840
```




```
Epoch 30/150
Duration: 100.40553
D Loss: 0.68087
G Loss: 0.77577
```

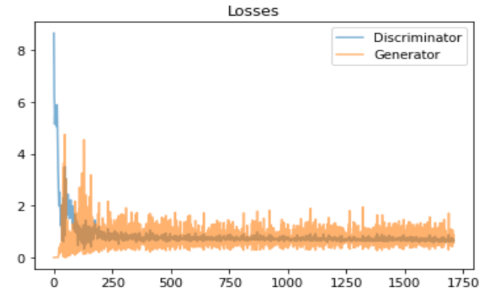


Fig.6. losses of discriminator and generator for random epochs represented in Table I.

C. Observations

Tabulated the Discriminator and Generator losses of random epochs. For the discriminator, true sample label is marked as 1, and generator sample label is marked as 0, nevertheless of the quality of the image sample generated. So, $D(x)$ should near to 1 and $D(G(z))$ should near to 0. In our experiment we save just the images that $D(G(z))$>=0.76 which is close to 1.

V. CONCLUSION

In this paper, the 2 generative suggested models which are founded based on VAE & DCGAN framework have been trained on the retinal images that have been resized using the APTOS Blindness 2019 dataset. On the contrary, based on vessel masks, the previous methods have been utilized for the training of the system, the suggested models shown here need not have the vessel masks to create the images. Moreover, with the usage of DCGAN, credible as well as cropped images of retina with no loss were produced & estimated by the medical experts. Inferences after assessing have evidently shown that this system is an acceptable solution and right approach towards a model which is proficient in creating labelled images of the retina.

REFERENCE

[1] "DR-GAN: Conditional Generative Adversarial Network for Fine-Grained Lesion Synthesis on Diabetic Retinopathy Images", Yi Zhou, Member, IEEE, Boyang Wang, Xiaodong He, Shanshan Cui and Ling Shao, Senior Member, IEEE, BIOMEDICAL AND HEALTH INFORMATICS, Nov-11, 2020.

[2] "Retinal Image Synthesis for Glaucoma Assessment using DCGAN and VAE Models", Andres Diaz-Pinto, Adrian Colomer, Valery Naranjo, Sandra Morales, Yanwu Xu, and Alejandro F Frangi, 19th International Conference on Intelligent Data Engineering and Automated Learning: IDEAL 2018, 21-23 Nov 2018, Madrid, Spain. Springer Verlag, pp. 224-232. ISBN 9783030034924.

[3] "DVAE: Deep Variational Auto-Encoders for Denoising Retinal Fundus Image", Biswajit Biswas, Swarup Kr Ghosh and Anupam Ghosh, Hybrid Machine Intelligence for Medical Image Analysis, Studies in Computational Intelligence 841, Springer Nature Singapore Pte Ltd. 2020.

[4] "Auto-encoder-based generative models for data augmentation on regression problems", Hiroshi ohno,

Springer-Verlag GmbH Germany, part of Springer Nature 2019.

[5] Costa, P., Galdran, A., Meyer, M.I., Niemeijer, M., Abr`amoff, M., Mendona, A.M., Campilho, A.: End-to-end Adversarial Retinal Image Synthesis. IEEE Transactions on Medical Imaging PP (99) (2017) 1–1.

[6] Sivaji Dutta, Bonthala CS Manideep, Syed Muzamil Basha, Ronnie D. Caytiles, N. Ch. S. N. Iyengar. Classification of Diabetic Retinopathy Images by Using Deep Learning Models. International Journal of Grid and Distributed Computing Vol. 11, No. 1 (2018), pp.89-106 http://dx.doi.org/10.14257/ijgdc.2018.11.1.09.

[7] Orlando JI, Prokofyeva E, del Fresno M, Blaschko MB. An ensemble deep learning based approach for red lesion detection in fundus images. Comput Methods Progr Biomed 2018;153:115–27.

[8] S. Chintala, E. Denton, M. Arjovsky, and M. Mathieu, "How to Train a GAN? Tips and tricks to make GANs work," https://github.com/soumith/ ganhacks, 2016.

[9] Dewi, C.; Chen, R.-C.; Liu, Y.-T.; Yu, H. Various Generative Adversarial Networks Model for Synthetic Prohibitory Sign Image Generation. Appl. Sci. 2021, 11, 2913. https://doi.org/10.3390/app11072913.

[10] Kora Venu, S.; Ravula, S. Evaluation of Deep Convolutional Generative Adversarial Networks for Data Augmentation of Chest X-ray Images. Future Internet 2021, 13, 8. https://doi.org/10.3390/fi13010008

[11] Shorten, C.; Khoshgoftaar, T.M. A survey on image data augmentation for deep learning. J. Big Data 2019, 6, 60. [CrossRef]

[12] Yi, X.; Walia, E.; Babyn, P. Generative adversarial network in medical imaging: A review. Med. Image Anal. 2019, 58, 101552. [CrossRef] [PubMed]

★★★