International Journal of Current Research and Applied Studies, (IJCRAS)

ISSN: 2583-6781

available at https://ijcras.com/

Volume 2 Issue 5 Sep-Oct 2023

Page 16-29

DIABETIC RETINOPATHY IMAGE SYNTHESIS USING DCGAN AND VAE MODELS

K.Shayam Kishore[1] and Y Sravani Devi[2]

[1]Student, ME, IARE, Hyderabad, INDIA
18951A0395@iare.ac.in

[2]Research Scholar, CSE GITAM Deemed to be university
Hyderabad, India
y.sravanidevi@gnits.ac.in

ABSTRACT

The amount of data, particularly in medical imaging, is one of the most important factors in classifying the image. Nevertheless, obtaining the datasets is the biggest obstacle in the healthcare industry. In this, we prepare a VAE (variational autoencoder) and another model known as the DCGAN (deep convolutional generative adversarial networks), on almost 3662 retinal images that have been captured from a dataset known as the APTOS- Blindness dataset, to display the images of the synthesized retinal fundus. The advantage of this method is that retinal pictures can be produced without the preceding vessel segmentation technique. As a result, the system can become autonomous. The models that are acquired are the image synthesizers that are adept at synthesizing resized retinal images of any amount from a fundamentally regular distribution. Furthermore, more images than this have been used in literature for training purposes than for any other endeavor. Giving an output to a CNN model allows for the evaluation or appraisal of a synthetic image, and the average squared error between the average 2-Dimensional hologram of actual and synthetic images was also calculated. by examining the average loss and latent space of the images later. The analysis's successful results suggested that DCGAN, as opposed to Variational Auto Encoders, has less loss in general images.

Keywords: Data Augmentation, DC-GAN, Variational Auto Encoder (VAE), Diabetic Retinopathy, Generative Adversarial Networks, CNN.

1. INTRODUCTION

The general illness known as diabetic retinopathy (DR) is what causes diabetics to lose their eyesight or perhaps become blind. Human eye specialists typically identify and categorize this disease's severity depending on the kind and wide range of connected lesions. The severity of DR can be classified into 5 stages, according to the international consensus [1], [2]: normal, mild, moderate, severe non-proliferative diabetic retinopathy (NPDR), and PDR. Hard exudates, soft exudates, hemorrhages, microaneurysms, laser marks, proliferating membranes, etc. are some of the related lesions. Even our numerous ophthalmologists find it difficult and time-consuming to diagnose diabetic retinopathy, hence automated grading models of DR [3], [4], [5], [6] have begun to be taken into consideration during the past few years. Numerous previous studies [7], [8], [9], and [10] support deep models, allowing the deployment of DR classification to significantly outperform a number of other types of approaches. For several vision tasks, such as visual categorization [11], object recognition [12], semantic segmentation [13], [14], and image synthesis [15], deep convolutional neural networks (CNNs) have dramatically outperformed manual feature extraction and regular machine learning methods. The data distribution of DR over various other grades is exceedingly uneven since atypical images of the fundus only make up a small percentage of the data needed to train an effective or powerful deep CNN model. For instance, the photos of DR levels 3 and 4 only accounts for 2.35% and 2.16% of the whole DR, respectively, in the largest publicly available DR dataset, EyePACS [16], but the images that are typical for level 0 account for 73.67%. Accepting this kind of oversized data makes the model less sensitive or responsive to many examples with higher DR severity levels, which leads to overfitting. Even while the fundamental techniques for data augmentation, including as rotation, random cropping, and flipping, can resolve the problem, the model's performance is still constrained by the insufficient diversity of samples at those levels.

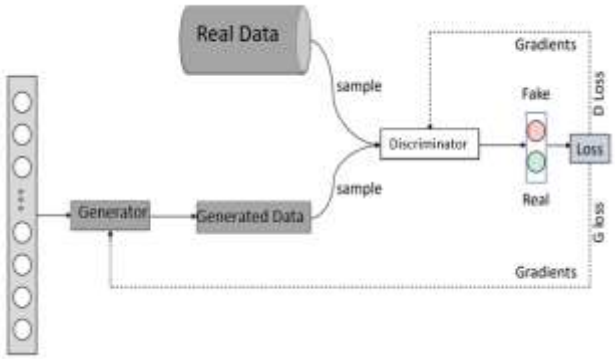


Fig 1: Architecture of GAN

GANs (Generative Adversarial Networks) [17] are widely used for a variety of picture producing jobs. The structure of GAN typically consists of a discriminative model D and generative model G competing against one another in a min-max game, which has led to a significant advancement in the synthesis of photorealistic images. In particular, a neural network is used to generate realistic data, and another is in charge of distinguishing between genuine and synthetic data. The deconvolutional layer, which is in charge of scaling

up the process to transform low-resolution images into higher-resolution ones, is used by DCGAN to extend GAN. By integrating a one-hot vector with the random noise vector, CGAN [18] seeks to define the requirements for the generator. On top of that, CycleGAN [19] is used to do unpaired image-to-image translation from a source domain to a target domain. Another method for training class-conditional pictures and scaling the size of the batch and some model parameters is BigGAN [20], which combines a range of excellent current methodologies. One and the other, which are images with great quality and resolution, are routinely produced as a result. We were able to construct a retina generator that generates realistic, better resolution images because to these well-developed taught GAN frameworks.

The development of retinal fundus image synthesizers is the primary topic of this paper. In contrast to earlier studies, this unique approach utilised more images throughout the training phase and did not require vessel masks. Using images from the APTOS Blindness dataset, we trained the Deep Convolutional Generative Adversarial Networks (DCGAN) [7] and the Variational Autoencoder (VAE) [6]. Then, using these models, we created artificial retinal samples that we ultimately evaluated using the loss function.

2. RELATED WORK

A. GANs in the Synthesis of Medical Image:

The absence of extensive and varied annotated databases could be overcome by synthesizing medical images using GANs. Numerous strategies have been put forth for a number of medical imaging domains, including computed tomography (CT) [22], [23], [24], magnetic resonance imaging (MRI) [25], [26], [27], & chest X-rays [28]. For instance, CT imaging's radiation exposure raises the danger of cancer. A series of 3D fully convolutional networks were used to illustrate the synthesis of CT images from MRI data in [22]. A reconstruction according to the pixels, loss, and an image gradient loss were selected for producing as well as adversarial learning. Mahapatra et al. used conditional GAN and Bayesian neural networks for the synthesis.

B. GANs in Synthesis of Retinal Image:

Recently, some researchers have also used GANs to produce retinal fundus images. The first U-Net framework was proposed by Costa et al. [31] for the transfer of vessel segmentation masks to fundus images using a standard GAN architecture. However, the examples that have been made have block flaws and lack manageable grading information. Tub-SGAN [32] was proposed in order to expand the style transfer to the generator and so broaden the diversity of synthesized samples. Although this is largely prospering and displaying a positive outcome, it is impossible to manufacture the DR-associated lesions and specifics of the physiological retina. Recently, Niu et al. [33] tried to create pictures of the fundus using pathological descriptors and vascular segmentation.

The retina's unhealthily generated images have been discussed by a number of other academics and researchers. A GAN-based method is suggested by Pujitha and Sivaswamy for the creation of images containing hemorrhages. For the model training phase of this technique, a vascular tree and lesion binary annotations are required. By giving the model the vascular tree and lesion (hemorrhage) masks, a fresh image of the retina can be produced. Despite this, the results were deemed unsatisfactory since the reproduction of

retinal pictures was deemed to be insufficient. These technologies have significant drawbacks despite producing images of the retina that are generally reasonable. They require training data with pixel-level annotations, for instance, but they are exceedingly expensive to get. Additionally, the model's capacity for creation is determined by the training set variability. Additionally, the high computational requirements frequently compromise the qualities of the created images, particularly by providing low resolution images that are insufficient for retinography, which has greater resolutions. Similarly, as was already indicated, these activities frequently rely on the retinal vasculature for both the creation of the image and the evaluation of its plausibility. However, the produced images do not depict vascular networks that have been given medical approval. Additionally, as far as we can tell, it has not yet been possible to create medical images using a certain DR grade and a specific type of lesion. The most recent studies suggested that it was possible to create images with varying numbers of hemorrhages, but the ability to do so wasn't further investigated.

With lesion information and arbitrary grading, a DR-produced generative adversarial network (DR-GAN) [34] can produce fundus images with great quality. On the lesion and construction masks, the retina generator is condition. Additionally, to control the generated grading intensity, adaptive grading vectors that are modeled from latent grading spaces can be used.

The effectiveness of multitask learning in relation to the issues with regression. In materials science, one ionic conductivity dataset and seven benchmark datasets are used in the research. According to the conclusions drawn from the trials, multi-task learning leads to an improvement in performance when generalizing examples of linear regression with multiple variables.

The creation of retinal pictures without the need for vessel segmentation is made possible by the use of DCGAN [35]. Therefore, the new method completely undermines its reliability. The collected models or instances have the capacity to synthesize trimmed retinal pictures of any amount from a regular normal distribution. In addition, more photos were used in the training process than with any previous persistent model.

Another technique for creating retinal images uses a system that has been trained on vascular networks and the comparable retinal fundus images. In other words, a transition between the retinal fundus and the vascular trees has been discovered. The main flaw in their strategy is how well-reliable an autonomous algorithm is at separating.

C. Variational Auto Encoders for image synthesis:

The VAE framework was first introduced by Kingma and Welling [36] in 2014, and it has since been hailed as one of the greatest contributions to generative modeling or representation learning in general. The VAE method provided a cutting-edge method for integrating deep learning with probabilistic models. The main distinction between VAEs and traditional autoencoders is that they learn latent variables through repeated divisions, which has proven to be a very useful trait when using generative modeling functions. A distribution rather than the individual numbers could be returned thanks to the cleverly designed VAE encoding. To put it more plainly, the encoder creates a pair of vectors, one of means () and the other of standard deviations ().

As a result, instead of studying a deterministic mapping like typical autoencoders do, the VAE seeks to learn the divisions of latent variables based on the mean values and their variations. Based on the samples of and values, it is hypothesized that the latent dimensional space exists, and Figure 2 depicts the general structure of the VAE framework. We advise reading Kingma and Welling's tutorial [37] for a comprehensive description of the VAE approach, even if it is outside the purview of this work.

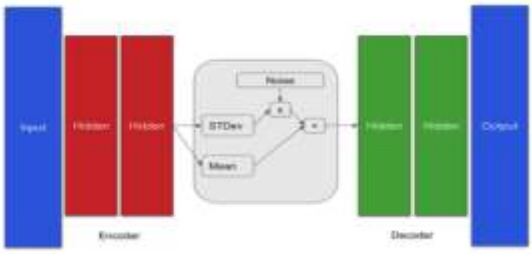


Fig.2 Architecture of Variational Auto Encoder

The VAE technique has gradually gained acceptance across a range of generative modeling challenges since its inception. For instance, text creation using an RNN-based VAE architecture was implemented. Similar to this, a recent study [38] produced a blended or hybrid architecture of recurrent neural networks (RNN) and convolutional neural networks (CNN) for the purpose of text production, while other research papers [39,40,41] investigated the potential of VAE for the development of natural images. It is also important to note that the generative adversarial network (GAN) [42] is another well-known generative modeling technique, however the current study is not primarily concerned with it.

3. PROPOSED VAE AND DCGAN METHODS

A. Variational Auto Encoders:

1.1. It consists of two neural networks: the encoder network, also known as the approximate inference network, is in charge of mapping a training sample to the latent or hidden space, while the decoder network plans or maps from the latent space to an artificial sample. In this challenge, the encoder and decoder are fully connected neural networks with one hidden layer each, and the latent space is an isotropic multivariate Gaussian that is centered. During the learning phase, sometimes referred to as the training phase, the encoder extracts the latent variables z from the input data, and the decoder eliminates those types of variables to produce a sample. Then, during the generation stage, VAE extracts samples from the latent space. The framework of VAE can be depicted from the Figure 2(a)

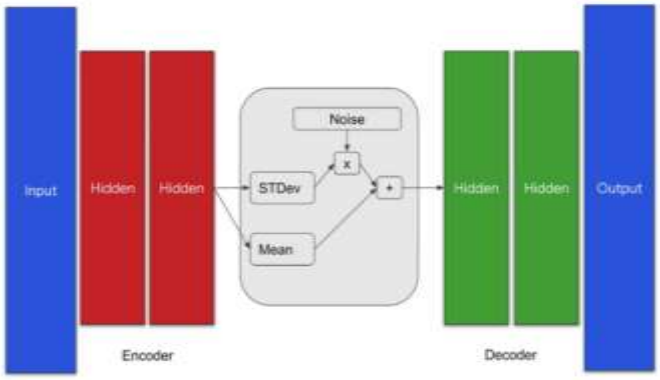


Fig.2. Architecture of Variational Auto Encoder

B. Deep Convolutional Generative Adversarial Networks:

The deep neural net framework known as GANs, or GANs, is made up of two nets. The opposing party is referred to as the discriminator, and the former is known as the generator. a CNN grade also referred to as Deep Convolutional Generative Adversarial Networks (DCGAN) that is based on a specific tactic. This framework, which generates images of increased or enhanced quality and more stability during the training phase, is the primary improvement over the very first GAN. We implemented and trained the generator & discriminator on the cropped retinal pictures using the generator cost functions & the original discriminator, as per the directions provided in the study publication by Radford et al. Similar to the VAE technique, DCGAN-based artificial picture generation [35] primarily consists of two phases: one is the learning phase and the other is the generating phase. In the first stage, the discriminator works to distinguish between the images acquired from the generator and the training set images while the generator essentially picks out samples from an N-dimension regular distribution that rush through the generator to acquire a fake sample. Figure 4(b)'s depiction of the DGCAN framework or architecture is another option.

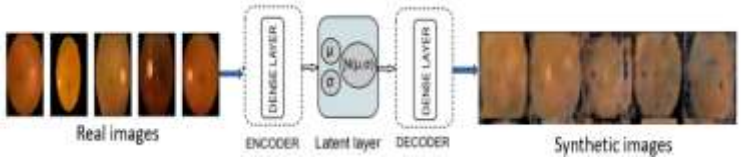


Fig.3. a) Schema of VAE architecture

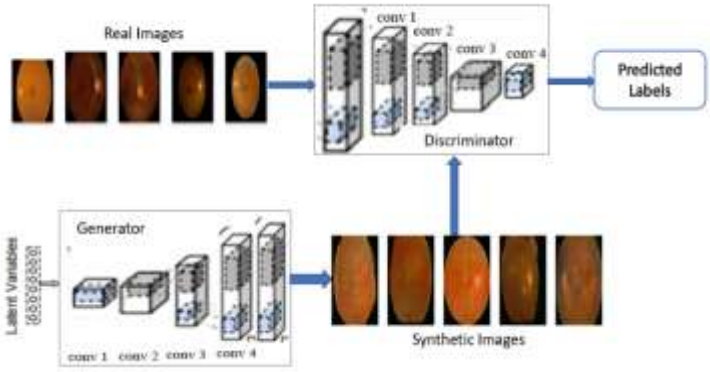


Fig.3b) Schema of DCGAN architecture

A few improvements over the current GANs are made in the framework. One of the changes includes the use of Batch normalization in each generator and discriminator, the complete replacement of completely associated or linked hidden layers with the mean pooling in the end, and the use of Leaky Re LU stimulation in the engine or the generator for all of the complete layers excluding the output and the u. It is no surprise that the research has effectively improved adversarial models, but the most difficult challenge is really model training. In order to find a solution to this problem, we have followed the advice given in [35] to gain stability when training the DCGAN. Examples include normalizing the input images between -1 and -1, using the ADAM optimizer for the engine or generator with a Gaussian distribution for the mini batches, which are made up of the original images used to train the models and the latent space, and more.

4. RESULTS AND DISCUSSIONS

A. DATASET:

3662 photos total from the APTOS Blindness. The training of the models used in this work was done using the testing photos, which were almost 1928 in number, and the train images, which were almost 3662 in number. The organizers of the Kaggle competition divided the over 18590 fundus photos in the whole dataset into roughly 3662 training, 1928 validation, and nearly 13000 testing images.

An open-source NVIDIA Titan Xp GPU and Keras were used in this task's deep learning library to conduct all the necessary experiments.

B. Loss Functions [9]:

We are aware that the DCGAN model imitates a contest in which the generator tries to create distinctive or organic visuals. The DCGAN model's primary objective is to trick the discriminator by producing more realistic images while simultaneously increasing the discriminator's misclassification error. This can also be described as a two-player minimax game here:

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

where G(z) is the output of the generator after an actual noise has been applied to z, D(G(z)) is the discriminator's judgment of the likelihood that a take instance is genuine, and D(x) is the discriminator's estimate of the likelihood that the actual data instance x is real. Ex is the expected value for all occurrences of the actual data. a generator's predicted value (Ez) for each random input. As a result, the configuration is basically qualified to maximize log(D(x) and minimize log (1 - D(G(z))).

Now, basically, the training of the VAE and DCGAN architectures using only the rescaled and pre-processed retinal pictures from the APTOS Blindness dataset. Testing of a range of N-dimensional latent spaces from roughly 32 to 100 latent variables for each image size. each latent factor. To regulate all the systems that don't remember the training database, every latent space was looked at. In the meantime, it produces believable retinal images. Estimation of intermediate latent representation points is carried out to achieve this. We performed numerous tests to train the VAE model and came to the conclusion that using a 512-dimension

latent space and 1008 * 1008 spatial resolution produced the best results. We obtained the artificial images shown in Figure 4 by running for roughly fifty epochs with a small batch size of 64.

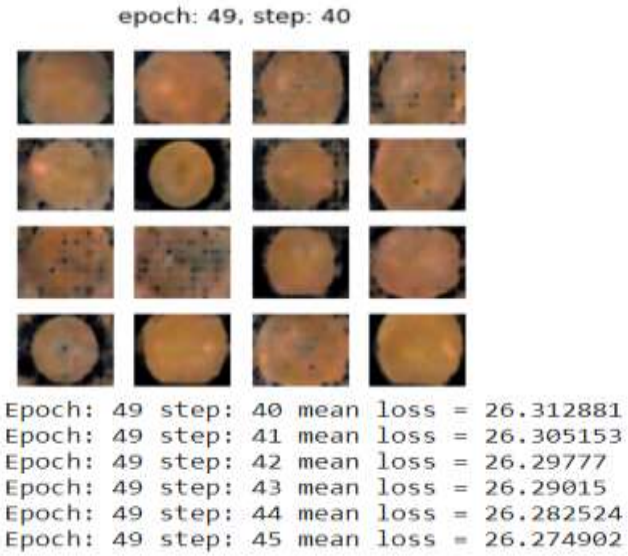


Fig.4a)

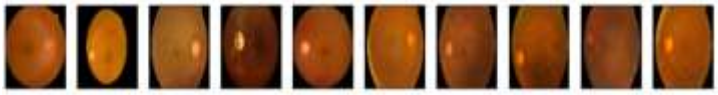


Fig.4b)

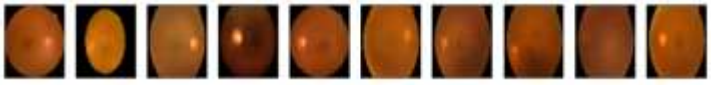


Fig.4c)

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.

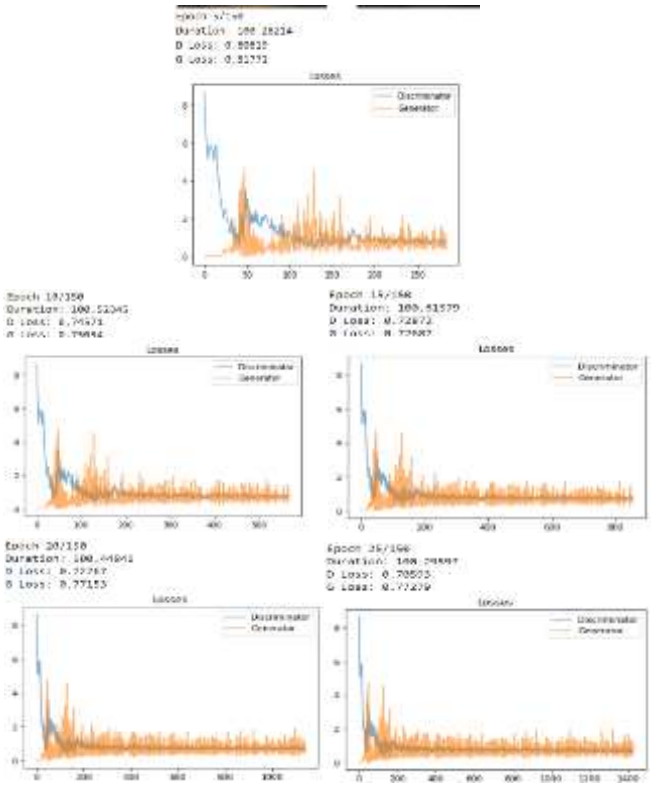
Fig.5: Examples of synthetic images generated by the DCGAN architecture

Although the synthetic images obtained using VAE have a texture similar to realistic photos, they lack the typical characteristics of a fundus image and are fuzzy and high loss. We were able to determine, with the use of the DCGAN architecture, that realistic images were produced with an image size of 128128 pixels, a

modest batch size of 64, and 32 epochs. In Fig. 5, examples of these are presented. Basically, the clarity of these synthetic images compared to those produced using the VAE technique is the main benefit of employing this architecture. We can calculate the loss for both the generator and the discriminator pictures in order to confirm them. When compared to the DCGAN, VAE has incurred more loss. As a result, the estimation or evaluation of the images produced only by the DCGAN was kept going. The obtained results are shown in Table 1 and in the figure as well.

TABLE I.

S.No	**Discriminator and generator loss of DCGAN at random epochs**		
	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



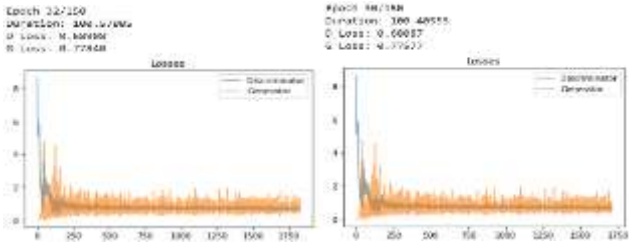


Fig.6. losses of discriminator and generator for random epochs represented in Table I.

C. Observations:

The random epochs' Discriminator and Generator losses were tabulated. Regardless of how well the image sample was generated, the genuine sample label is marked as 1 and the generator sample label is recorded as 0. D(x) should therefore be close to 1, and D(G(z) should be close to 0. In our experiment, we only save the photographs when D(G(z)>=0.76, or nearly 1, in the image.

D. Evaluation of DCGAN & VAE:

Frechet Inception Distance (FID) can be used to analyze the images produced by GAN and auto encoders. Compared to Inception Score (IS), FID assesses image similarity to actual images more precisely. The Inception Score (IS) measures the quality of the images based on how effectively the top-performing image classification system InceptionV3 recognizes a set of synthetic photos as one of 1000 recognized things. For each synthetic image, the scores consider the confidence in the conditional class predictions as well as the integral of the marginal probability of the predicted classes (diversity). How phony images compare to real ones is not considered by the conceptualization Score. The FID score was developed with the intention of evaluating synthetic images utilizing information from a collection of pictures.

```
47.png
FID score for DCGAN for image set 47.png.png:  30.690502651115935

277.png
FID score for DCGAN for image set 277.png.png:  29.9437015228026

46.png
FID score for DCGAN for image set 46.png.png:  27.899642932693496

245.png
FID score for DCGAN for image set 245.png.png:  22.99150092901897

263.png
FID score for DCGAN for image set 263.png.png:  21.1447459901583

178.png
FID score for DCGAN for image set 178.png.png:  19.737705663710074

107.png
FID score for DCGAN for image set 107.png.png:  17.99402296802066

40.png
FID score for DCGAN for image set 40.png.png:  14.961998408523982

63.png
FID score for DCGAN for image set 63.png.png:  11.014112833375293
```

Fig.7 FID scores of DC-GAN generated images.

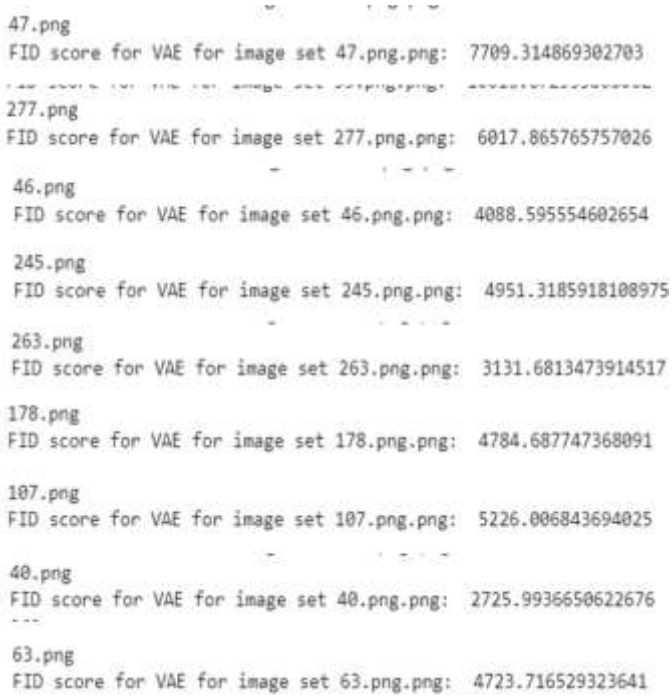


Fig.8 FID scores of VAE generated images.

Based on the FID scores of both DCGAN & VAE generated images, DCGAN generated images have low values and VAE generated images values are very high. So, for retina images DCGAN showing the better results than VAE.

5. CONCLUSIONS

The retinal pictures from the APTOS Blindness 2019 dataset were used to train the two generative proposed models in this paper, which are based on the VAE and DCGAN frameworks. On the other hand, whereas the prior methods have been used to train the system using vessel masks, the suggested models illustrated here do not require vessel masks in order to produce the images. Additionally, by using DCGAN, retinal pictures that were cropped without losing quality were obtained. Conclusions following evaluation have clearly demonstrated that this method is a workable option and the correct strategy for developing a model capable of producing annotated images of the retina.

REFERENCES

1. V. Gulshan, L. Peng, M. Coram, M. C. Stumpe, D. Wu, A. Narayanaswamy, S. Venugopalan, K. Widner, T. Madams, J. Cuadros et al., “Development and validation of a deep learning algorithm for detection of diabetic retinopathy in retinal fundus photographs,” Jama, vol. 316, no. 22, pp. 2402–2410, 2016.

2. "International clinical diabetic retinopathy disease severity scale," American Academy of Ophthalmology, 2012.
3. L. Seoud, J. Chelbi, and F. Cheriet, "Automatic grading of diabetic retinopathy on a public database," in MICCAI. Springer, 2015.
4. H. Pratt, F. Coenen, D. M. Broadbent, S. P. Harding, and Y. Zheng, "Convolutional neural networks for diabetic retinopathy," Procedia Computer Science, vol. 90, pp. 200–205, 2016.
5. Y. Yang, T. Li, W. Li, H. Wu, W. Fan, and W. Zhang, "Lesion detection and grading of diabetic retinopathy via two-stages deep convolutional neural networks," in MICCAI. Springer, 2017, pp. 533–540.
6. X. He, Y. Zhou, B. Wang, S. Cui, and L. Shao, "Dme-net: Diabetic macular edema grading by auxiliary task learning," in MICCAI. Springer, 2019, pp. 788–796.
7. Z. Wang, Y. Yin, J. Shi, W. Fang, H. Li, and X. Wang, "Zoom-innet: deep mining lesions for diabetic retinopathy detection," in MICCAI. Springer, 2017, pp. 267–275.
8. "A Scoping review of diabetic retinopathy detection techniques using deep learning: taxonomy, methods, and recent developments", Y. Sravani Devi, S. Phani Kumar, High Technology Letters, ISSN NO : 1006-6748, Volume 26, Issue 11, Nov-2020.
9. Z. Lin, R. Guo, Y. Wang, B. Wu, T. Chen, W. Wang, D. Z. Chen, and J. Wu, "A framework for identifying diabetic retinopathy based on antinoise detection and attention-based fusion," in MICCAI. Springer, 2018, pp. 74–82.
10. Y. Zhou, X. He, L. Huang, L. Liu, F. Zhu, S. Cui, and L. Shao, "Collaborative learning of semi-supervised segmentation and classification for medical images," in CVPR, 2019.
11. K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in CVPR, 2016, pp. 770–778.
12. L. Liu, W. Ouyang, X. Wang, P. Fieguth, J. Chen, X. Liu, and M. Pietikainen, "Deep learning for generic object detection: a survey," ¨ International Journal of Computer Vision, vol. 128, no. 2, pp. 261–318, 2020.
13. A. Garcia-Garcia, S. Orts-Escolano, S. Oprea, V. Villena-Martinez, P. Martinez-Gonzalez, and J. Garcia-Rodriguez, "A survey on deep learning techniques for image and video semantic segmentation," Applied Soft Computing, vol. 70, pp. 41–65, 2018.
14. D.-P. Fan, T. Zhou, G.-P. Ji, Y. Zhou, G. Chen, H. Fu, J. Shen, and L. Shao, "Inf-net: Automatic covid-19 lung infection segmentation from ct images," IEEE Transactions on Medical Imaging, 2020.
15. X. Yi, E. Walia, and P. Babyn, "Generative adversarial network in medical imaging: A review," Medical Image Analysis, vol. 58, p. 101552, 2019.
16. "Kaggle diabetic retinopathy detection competition," https://www. kaggle.com/c/diabetic-retinopathy-detection.
17. I. Goodfellow, J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley, S. Ozair, A. Courville, and Y. Bengio, "Generative adversarial nets," in NeurIPS, 2014, pp. 2672–2680.
18. A. Radford, L. Metz, and S. Chintala, "Unsupervised representation learning with deep convolutional generativeadversarial networks," in ICLR, 2016.
19. M. Mirza and S. Osindero, "Conditional generative adversarial nets," ArXiv, vol. abs/1411.1784, 2014.

20. J.-Y. Zhu, T. Park, P. Isola, and A. A. Efros, “Unpaired image-to-image translation using cycle-consistent adversarial networkss,” in ICCV, 2017.
21. A. Brock, J. Donahue, and K. Simonyan, “Large scale GAN training for high fidelity natural image synthesis,” in ICLR, 2019.
22. S. Kazeminia, C. Baur, A. Kuijper, B. Ginneken, N. Navab, S. Albarqouni, and A. Mukhopadhyay, “Gans for medical image analysis,” Artificial Intelligence in Medicine, vol. 109, p. 101938, 2020.
23. D. Nie, R. Trullo, C. Petitjean, S. Ruan, and D. Shen, “Medical image synthesis with context-aware generative adversarial networks,” MICCAI, vol. 10435, pp. 417–425, 2016.
24. Y. Ruan, D. Li, H. Marshall, T. Miao, T. Cossetto, I. Chan, O. Daher, F. Accorsi, A. Goela, and S. Li, “Mb-fsgan: Joint segmentation and quantification of kidney tumor on ct by the multi-branch feature sharing generative adversarial network,” Medical Image Analysis, p. 101721, 2020.
25. C. Xu, L. Xu, P. Ohorodnyk, M. Roth, B. Chen, and S. Li, “Contrast agent-free synthesis and segmentation of ischemic heart disease images using progressive sequential causal gans,” Medical Image Analysis, p. 101668, 2020.
26. J. Jiang, Y.-C. Hu, N. Tyagi, P. Zhang, A. Rimner, G. S. Mageras, J. O. Deasy, and H. Veeraraghavan, “Tumor-aware, adversarial domain adaptation from ct to mri for lung cancer segmentation,” MICCAI, vol. 11071, pp. 777–785, 2018.
27. T. Zhou, H. Fu, G. Chen, J. Shen, and L. Shao, “Hi-net: hybrid-fusion network for multi-modal mr image synthesis,” IEEE Transactions on Medical Imaging, 2020.
28. J. Zhao, D. Li, Z. Kassam, J. Howey, J. Chong, B. Chen, and S. Li, “Tripartite-gan: Synthesizing liver contrast-enhanced mri to improve tumor detection,” Medical Image Analysis, p. 101667, 2020.
29. D. Mahapatra, B. Bozorgtabar, J.-P. Thiran, and M. Reyes, “Efficient active learning for image classification and segmentation using a sample selection and conditional generative adversarial network,” MICCAI, pp. 580–588, 2018.
30. W. Wei, E. Poirion, B. Bodini, S. Durrleman, N. Ayache, B. Stankoff, and O. Colliot, “Learning myelin content in multiple sclerosis from multimodal mri through adversarial training,” MICCAI, pp. 514–522, 2018.
31. Maayan Frid-Adar, Eya lKlang, Michal Amitai, Jacob Goldberger, and Hayit Greenspan, Synthetic data augmentation using GAN for improved liver lesion classification. 2018 IEEE 15th International Symposium on Biomedical Imaging ,pages 289–293, (2018)
32. C. P, G. A, M. M I, N. M, A. M, M. A M, and C. A, “End-toend adversarial retinal image synthesis,” IEEE Transactions on Medical Imaging, vol. 37, no. 3, pp. 781–791, 2018.
33. H. Zhao, H. Li, S. Maurer-Stroh, and L. Cheng, “Synthesizing retinal and neuronal images with generative adversarial nets,” Medical Image Analysis, vol. 49, pp. 14–26, 2018.
34. 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.
35. “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.

36. Y. Niu, L. Gu, F. Lu, F. Lv, Z. Wang, I. Sato, Z. Zhang, Y. Xiao, X. Dai, and T. Cheng, "Pathological evidence exploration in deep retinal image diagnosis," in AAAI, 2019.
37. Kingma, D.P.; Welling, M. Auto-encoding variational bayes. In Proceedings of the 2nd International Conference on Learning Representations (ICLR), Banff, AB, Canada, 14–16 April 2014.34
38. Kingma, D.P.; Welling, M. An introduction to variational autoencoders. arXiv 2019, arXiv:1906.02691. Available online: https: //arxiv.org/abs/1906.02691 (accessed on 2 May 2021). [CrossRef]35
39. Bowman, S.R.; Vilnis, L.; Vinyals, O.; Dai, A.M.; Jozefowicz, R.; Bengio, S. Generating sentences from a continuous space. arXiv 2015, arXiv:1511.06349. Available online: https://arxiv.org/abs/1511.06349 (accessed on 2 May 2021).36
40. Semeniuta, S.; Severyn, A.; Barth, E. A hybrid convolutional variational autoencoder for text generation. arXiv 2017, arXiv:1702.02390. Available online: https://arxiv.org/abs/1702.02390 (accessed on 2 May 2021). 37
41. "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.
42. "Auto-encoder-based generative models for data augmentation on regression problems", Hiroshi ohno, Springer-Verlag GmbH Germany, part of Springer Nature 2019.