

Colorectal cancer prediction via histopathology segmentation using DC-GAN and VAE-GAN

R. Sujatha¹, Mahalakshmi K² and Mohamed Sirajudeen Yoosuf^{3*}

^{1,2} School of Information Technology & Engineering, Vellore Institute of Technology, Vellore, Tamil Nadu, India

^{3*} School of Computer Science and Engineering (SCOPE), VIT-AP University, Andhra Pradesh, India

Abstract

Colorectal cancer ranks as the third most common form of cancer in the United States. The Centres of Disease Control and Prevention report that males and individuals assigned male at birth (AMAB) have a slightly higher incidence of colon cancer than females and those assigned female at birth (AFAB) Black humans are more likely than other ethnic groups or races to develop colon cancer. Early detection of suspicious tissues can improve a person's life for 3-4 years. In this project, we use the EBHI-seg dataset. This study explores a technique called Generative Adversarial Networks (GAN) that can be utilized for data augmentation colorectal cancer histopathology Image Segmentation. Specifically, we compare the effectiveness of two GAN models, namely the deep convolutional GAN (DC-GAN) and the Variational autoencoder GAN (VAE-GAN), in generating realistic synthetic images for training a neural network model for cancer prediction. Our findings suggest that DC-GAN outperforms VAE-GAN in generating high-quality synthetic images and improving the neural network model. These results highlight the possibility of GAN-based data augmentation to enhance machine learning models' performance in medical image analysis tasks. The result shows DC-GAN outperformed VAE-GAN.

Keywords: Generative Adversarial Network, Variational Autoencoder GAN, Colorectal Cancer, Medical image

Received on 05 December 2023, accepted on 07 March 2024, published on 12 March 2024

Copyright © 2024 R. Sujatha *et al.*, licensed to EAI. This is an open access article distributed under the terms of the [CC BY-NC-SA 4.0](#), which permits copying, redistributing, remixing, transformation, and building upon the material in any medium so long as the original work is properly cited.

doi: 10.4108/eetpht.10.5395

*Corresponding author. Email: msyoosuf.research@gmail.com

1. Introduction

Colorectal cancer, often known as CRC, can affect both the colon and the rectum. The second most prevalent disease) in women is colorectal cancer (CRC) and in males, colorectal cancer is the third most prevalent type of cancer, is one of the most common kinds of gastrointestinal cancer. The incidence of CRC is expected to increase globally by 80% by 2035, despite previously existing variations such as geo-graphic distribution, age, and gender inequalities. The main reason for the rising incidence of CRC is changing in lifestyle, notably dietary patterns. The majority (70–80%) of CRCs are sporadic, although about one-third are hereditary. The term CRC refers to a broad range of carcinoma subtypes with various morphological traits and molecular alterations. Recently, cancer detection and treatment have benefited from the application of machine learning [21] [22] approaches for medical imaging, such as

Generative Adversarial Networks (GANs). Data augmentation can boost the effectiveness of machine learning algorithms for analyzing medical images. In this article, we investigate how to segment colorectal cancer histology using generative adversarial networks (GANs) for data augmentation. We evaluate how well two GAN models—DC-GAN and VAE-GAN—perform when it comes to creating convincing synthetic images for use in training a model based on neural networks that can be developed to predict the occurrence of cancer. Our research intends to increase the precision of cancer prediction models using advanced techniques that can improve the precision of detecting and treating colorectal cancer.

Accurate segmentation of histopathology images is crucial for the early detection and effective treatment of cancer. However, the limited availability of annotated data can hinder the performance of these models. Data augmentation using GANs has shown promise in

generating synthetic images that can improve these models. The study aims to compare the effectiveness of two GAN models, DC-GAN and VAE-GAN, in generating synthetic images in order to train a neural network model for cancer prediction. The findings of this study can help to improve the medical image analysis models and ultimately contribute to early detection and colorectal cancer treatment. Colorectal Cancer is one of the world's major causes of cancer death, and early detection is critical for effective treatment and improved outcomes. Despite advances in diagnostic technologies, medical professionals' manual examination of medical imaging data can still be time-consuming and prone to human error, resulting in missed diagnoses or false positives. Create a machine learning model that accurately diagnoses and classifies colorectal cancer using a set of medical imaging data, such as CT scans, colonoscopy images, or MRI scans. To make accurate predictions, the model should take into account both the visual appearance and structural information of the images. The model should also be capable of identifying and highlighting regions of the images that are indicative of CRC, as well as providing a quantitative measure of the degree of malignancy for each region.

Scope

- Create a GAN-based deep learning model to diagnose colorectal cancer from medical image datasets such as EBHI-seg
- To improve accuracy and robustness, train the GAN model on a large, annotated dataset of medical images.

Objective

- To improve the accuracy and efficiency of the colorectal cancer diagnostic process by developing a deep learning model Using GANs.
- To provide medical professionals with a tool to aid the detection, and diagnosis of colorectal cancer and to shorten the time required for manual im-age examination.
- One of the primary objectives is to decrease the number of false positives and missed diagnoses in the colorectal cancer diagnosis procedure.
- To advance the medical domain imaging analysis by demonstrating the utility of GANs in cancer detection.

2. Literature Review

[1] Used the CNN architecture to test the classification findings. Then, using artificial liver lesions instead of the traditional data augmentation methods, examined the impact of data augmentation. Put the two techniques for creating artificial lesions into practice. They discovered that the Deep Convolutional GAN (DCGAN) technique performed better in testing. [2] Demonstrate that such a

SegAN framework outperforms the latest U-net segmentation method in terms of effectiveness and stability for the segmentation task. Using MICCAI BRATS brain tumour seg-mentation challenge data, they evaluated our SegAN approach. Numerous experimental findings show how effective the suggested SegAN with multi-scale loss is. [4-5] Assess Cycle GAN's potential for enhancing data in CT scan segmentation tasks. Using a large image collection, A Cycle GAN was trained to convert CT images with contrast to those without contrast. Then, utilising these created non-contrast images, they added to our training using the trained Cycle GAN. The researchers evaluated the segmentation performance of U-nets on two separate datasets- the initial contrast CT dataset utilized for segmentation creation, and a second dataset that consisted solely of non-contrast CTs, after training the U-Nets on the original dataset and a combined dataset containing both the original data and the generated non-contrast images. The in-distribution dataset and the out-of-distribution dataset are the names given to these two separate datasets.[9] This study uses Deep Convolution GANs (DCGANs) to address the issue of a small, labelled dataset. A visualisation programme called ImageJ was employed to examine how closely the actual and artificial images resembled one another. With the aid of medical professionals, a visual Turing test was carried out to verify the proposed model.

[10] For the purpose of segmenting the tumour in breast ultrasound pictures, they recommend employing a technique depending on a Generative Adversarial Network (GAN). The discriminator is a CNN classifier, while the segmentation and generation modules are implemented using Residual-Dilated-Attention-Gate-NET (RDAU-NET). The total Accuracy, PR-AUC, ROC-AUC, and F1-score achieved were, respectively, 0.98, 0.95, 0.89, and 0.88 in comparison to the majority of conventional deep net models. The outcomes also indicate how the WGAN-RDA-UNET model can be used to address the flaws in RDA U-Net, Convolutional neural networks, and other models. [11] Proposes employing a semi-supervised multichannel-based generative adversarial network (MGAN) to rate Diabetic Retinopathy. The suggested semi-supervised MGAN can recognise the subtle damage feature utilising high-resolution fundus pictures without confining by reducing the reliance on labelled data. Experimental findings on the open Messidor data set demonstrate the effectiveness of the suggested model in grading DR. [15] proposed a GAN based on NU-Net. The primary architecture for this approach is cGAN. The NU-Net is the method's generating network, and the FCN is its discriminator. Utilising the MM-WHS 2017 challenge dataset, our experiment about automatic complete heart seg-mentation outperformed most Deep learning techniques with a Dice score of 0.899. [17] To create fake medical images that correspond to masks from a single training image, we describe the SinGAN-Seg synthetic data production process. The re-searchers demonstrated that models trained on synthetic data can outperform those trained on real data, provided that both datasets comprise a

substantial amount of training data. This was achieved by training UNet++ on a combination of real and synthetic data generated through the SinGAN-Seg pipeline. However, when training datasets lack a significant amount of data, synthetic data produced by the SinGAN-Seg pipeline enhances segmentation model performance.

3. Materials and Methods

3.1 Data Distribution

This study used a dataset with 2228 colorectal enteroscopy biopsy histopathological hematoxylin and eosin images (EBHI-seg) divided into six classes: "Adenocarcinoma," "High-grade IN," "Normal," and "Lowgrade." IN," "Polyp," and "Serrated adenoma." The dataset can be obtained at <https://figshare.com/articles/dataset/EBHI-SEG/21540159/1>.

3.2 Data Pre-Processing & Augmentation

At this point, the dataset's images have all been scaled to a fixed size of 96x96. The data augmentation is to provide more representative and diverse data to improve the performance of machine learning models. The model can learn to be more durable and generalise to new, unknown data by producing modifications of the original data.

3.3 DCGAN

DCGANs, or "Deep Convolutional Generative Adversarial Networks," are used to produce images. Convolutional neural networks (CNNs) are used as both the generator and discriminator networks in the conventional Generative Adversarial Network (GAN) architecture, which is modified to create DCGANs [12].

The generator network generates images that resemble the real images in the training dataset using input as a random noise vector is used. The discriminator network attempts to distinguish between real and fraudulent images as input. With the aim of improving both the discriminator's accuracy in identifying false and the generator's capability to create realistic images, in a competitive approach, the two networks are trained concurrently. DCGANs have a number of architectural characteristics that aid in the creation of high-quality images. For instance, the discriminator network downsamples the image while the generator network upsamples the input noise vector using transposed convolutional layers. Additionally aiding in the stabilisation of the training process and raising the caliber of the generated images are the usage of batch normalisation and the LeakyReLU activation function [13].

A wide range of image types, including faces, natural settings, and even artwork, have been produced using DCGANs.

3.4 VAE-GAN

The term "Variational Autoencoder Generative Adversarial Network," or simply "VAE-GAN," refers to a specific type of deep learning model that, in order to produce images, combines the benefits of both Variational Autoencoders (VAE) and Generative Adversarial Networks (GAN).

A particular class of generative model known as a VAE produces new samples by learning a low-dimensional representation of the data input. GANs, on the opposite hand, there is a category of adversarial model that employs a discriminator network to separate real data from fake data and a generator network to make the fake data. By employing the VAE to develop a latent representation of the data, which the GAN subsequently uses to produce realistic images, VAE-GAN combines the benefits of both models [8].

In VAE-GAN, while the generator network uses a random noise vector as input, the discriminator attempts to distinguish between real and fake images. The input image is mapped by the VAE encoder network to a latent representation, which is subsequently utilised to recreate the original image. The VAE decoder network creates a reconstructed image using the latent representation as input. The VAE encoder and decoder networks are trained using a reconstruction loss and a KL-divergence loss, while the generator network and the VAE decoder network are trained together using the adversarial loss from the discriminator network [8].

VAE-GAN has been found to produce high-quality photos with crisper features and fewer artifacts. It has been used to create realistic faces and settings from nature, among other image-generation jobs.

4. Implementations

The Keras and Tensor Flow libraries are used to implement both the VAE and DCGAN models in Python.

4.1 DCGAN

In DCGAN, the generator network utilises deconvolutional layers to create new images that resemble the training dataset using a random noise vector as input. A CNN-based classifier called the discriminator network uses realistic images from the training dataset or created images from the generator network to try to determine whether

they are real or fake. The dataset is loaded. Figure 1 illustrates the entire flow.

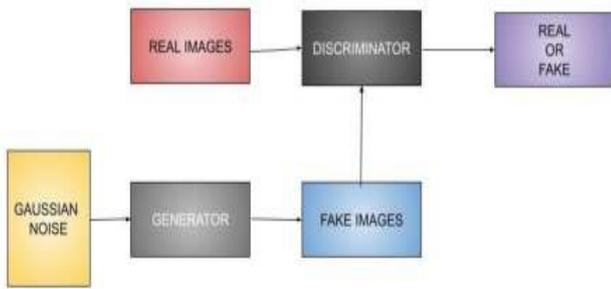


Figure 1. DCGAN Architecture

Generator: The generator begins with a dense layer that receives an input of a seed (random noise). Following that, a normal layer with ReLU as activation and a Batch Normalization layer is present. The noise vector is fed into the model, which is then repeatedly upsampled until it reaches the size of the images. Three deconvolutional layers are used for this. All with a kernel size of 5x5 and utilise bias option set to false and an ordinary layer that uses activation. A batch normalisation layer and an initial deconvolutional layer with a 1x1 stride size following ReLU. A batch Normalization layer, a normal layer utilising activation ReLU, and a second deconvolutional layer with a stride size of 2x2 are all added after the second deconvolutional layer [18].

Discriminator: An image classifier built on the CNN serves as the discriminator. It employs 3 convolutional layers with 64 and 128 filters, respectively. It employs a stride size of 2x2 and a kernel size of 5x5. A regular layer with LeakyReLU as the activation and a dropout layer was incorporated into the model architecture with a 0.2 dropout rate following each convolutional layer [18].

Loss function: Due to the fact that they are trained on two different networks, the discriminator and generator optimizers are separated.

Discriminator loss: A technique is used to gauge the discriminator's effectiveness in differentiating between real and fake images. It evaluates the discriminator's predictions against a variety of other ones using real-world images as its input. The discriminator's predictions on fake images are handled in the same way, although in this case, a comparison is made using an array of zeros [19].

Generator loss: The effectiveness of the generator's deception of the discriminator is measured by a method that is put into practice. The discriminator classification on the resulting images is compared to an array of ones to accomplish this. Developing and producing images to train the network, a training loop is established. A random seed is first utilised as input to create an image using a technique that simultaneously trains the generator and discriminator. After each iteration of the loop, the model calculates the loss and adjusts the gradients before the discriminator distinguishes between the real and fake images. After each iteration of the loop, the images are created and saved during the training loop [19].

4.2 VAE-GAN

By first creating samples in the latent space of the VAE using the generator portion of the GAN and then mapping those samples to the data space using the decoder portion of the VAE, the VAE-GAN integrates these two concepts. The reconstruction error between the original and reconstructed data as well as the adversarial loss between the real and fake samples are both minimised during the training of the VAE-GAN [6]. Flow is depicted in Figure 2. First, the datasets are loaded.

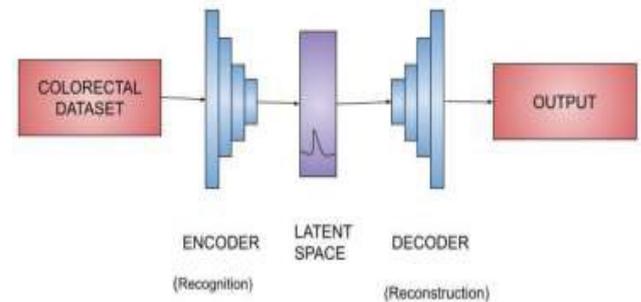


Figure 2. VAE-GAN Architecture

Network Architecture: An input layer is utilised for the encoder network and accepts the images as input. Two convolutional layers are added after that, the first of which employs 64 filters and the second of which is doubled to 128 filters. Both have a 3-kernel size and a 2x2 stride size. They are followed by a layer that flattens, and the output size of the subsequent fully connected dense layer is twice the dimension of the latent space. The encoder's structure is somewhat mirrored in the decoder. It consists of a dense fully connected layer whose output size corresponds to that of images and the activation ReLU, followed by an input

layer whose output size corresponds to that of the decoder network.

Three deconvolutional layers are then applied, each with a kernel that is the same size as the convolutional layers in the encoder network and uses the same activation function for the first two deconvolutional layers. The deconvolutional layers decrease from 128 filters at stride size 2 to 64 filters at stride size 2 to either 1 or 3 filters at stride size 1 for color images [14] [20].

Loss function: The reconstruction loss and the KL divergence loss are the two terms that make up the loss function in VAE. The model's ability to reconstruct the input data is measured by the reconstruction loss, and the latent variables are encouraged to follow a normal distribution by the KL divergence loss. The reconstruction loss was calculated using mean squared error (MSE). The KL divergence loss encourages the latent variables to follow a normal distribution. It is described as the discrepancy between the latent variable distribution and a typical normal distribution. The sum of the reconstruction loss and the KL divergence loss is the total loss function for VAE [7]. Utilising optimization methods like stochastic gradient descent and backpropagation, the objective of training VAE is to reduce this total loss function.

Training and Generating: The dataset is iterated through and then given to the encoder network for training. The mean and log variance parameters are taken from the estimated posterior distribution during each iteration. When selecting samples from the approximate posterior distribution the reparametrization approach is used. The samples are given to the decoder as the final training step to acquire the logits from the generating distribution. The images are generated and saved by selecting a collection of latent vectors from the Gaussian prior distribution. To produce predictions of the specified distribution, the generator converts the latent sample to logits of the observation. The generated images are these predictions [3].

5. Results and Discussions

The Python programming language is used in this work to train and evaluate the model using Google Colab Notebook. The keras and torch libraries are utilized for VAE-GAN and DC-GAN. Figures 3 and 4 show the sample images used in DCGAN and fake images generated in DCGAN. Figure 5 and 6 shows the sample images used in VAE-GAN and produces fake images via reconstruction and average VAE-GAN process.

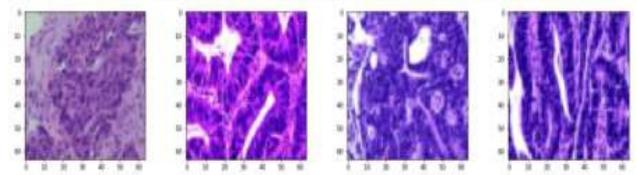


Figure 3. Sample images – DCGAN

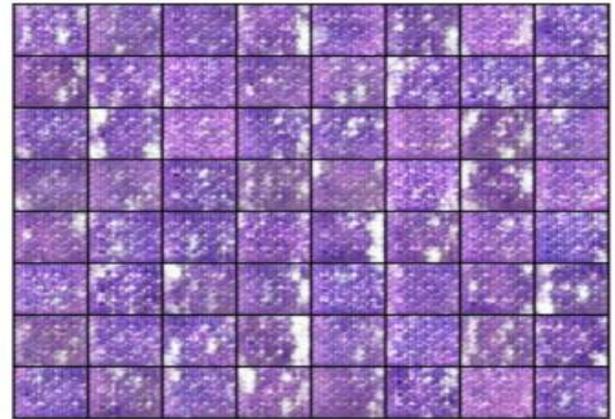


Figure 4. Fake images – DCGAN

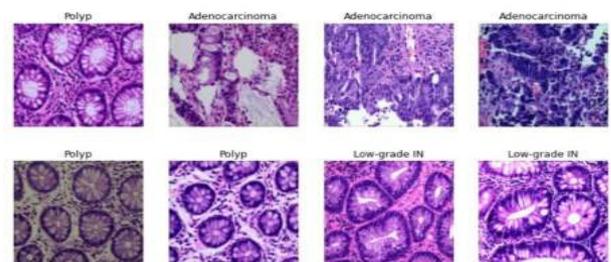


Figure 5. Sample images - VAE-GAN

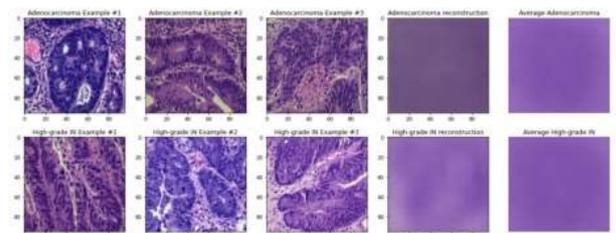


Fig. 6. Reconstruction and Average images of VAE-GAN

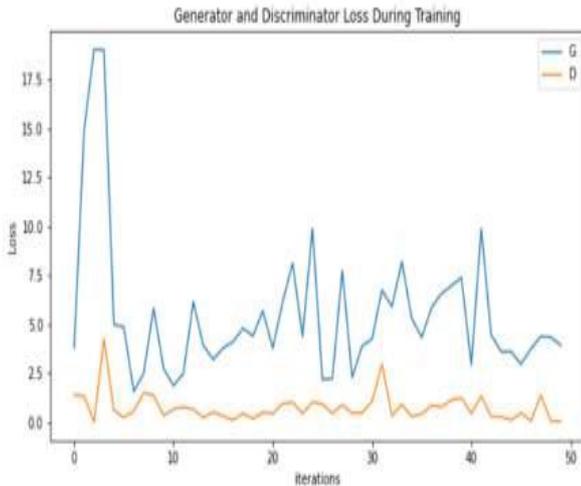


Figure. 7. Generator and Discriminator loss during training

The loss functions of the generator and discriminator are used to assess each network's performance during training. Typically, the discriminator loss is a measurement of how well it can tell the difference between generated and real data, whereas the generator's loss is a measure of how well the generated data can deceive the discriminator. Due to their random initialization and lack of data knowledge, the generator and discriminator both have significant loss values at the start of training. The loss of the generator reduces during training as it learns to produce more realistic data that can deceive the discriminator. On the other hand, because the discriminator improves at differentiating between fake and real data, it experiences a decrease in loss. Figure 7 graphically represents the generator and discriminator loss during training. To assess the effectiveness of the prediction model, the loss measure is used. Figure 8 shows the Loss D and G in the case of DCGAN and VAE-GAN in generating high-quality synthetic images and improving the neural network model. Based on this DCGAN outperforms by showing minimal discriminator loss.

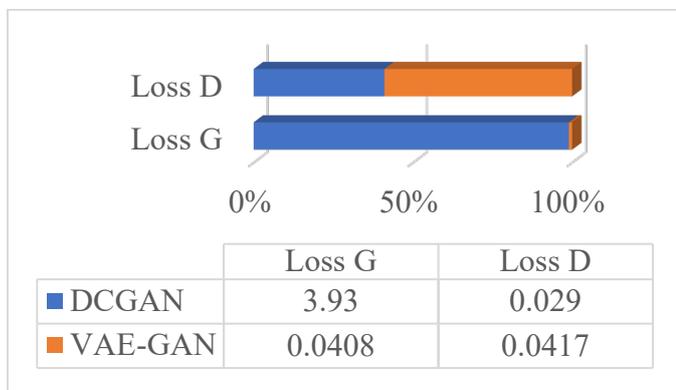


Figure. 8. DCGAN vs VAE-GAN – Loss Comparison

6. Conclusion and Future Work

In summary, our study demonstrated the effectiveness of using GAN for data augmentation in medical image analysis, specifically improving the neural network model for colorectal cancer prediction using histopathology images. The study compared the performance of two GAN models, DC-GAN and VAE-GAN, and found that DC-GAN outperformed VAE-GAN in generating realistic synthetic images and improving the neural network model. To delve deeper into data augmentation for medical image analysis tasks, future studies could investigate alternative GAN architectures and hyper parameters. Additionally, the accuracy of the model could be increased by researching transfer learning techniques and multi-task learning and enabling the prediction of other aspects of cancer diagnosis and prognosis. Furthermore, the development of standardized evaluation metrics and benchmarks for medical image analysis tasks could facilitate the comparison and validation of different models and techniques. Finally, the application of these techniques to larger and more diverse datasets could enhance their effectiveness and impact on the early detection and treatment of colorectal cancer.

References

- [1] Frid-Adar, M., Diamant, I., Klang, E., Amitai, M., Goldberger, J., & Greenspan, H, "GAN-based synthetic medical image augmentation for increased CNN performance in liver lesion classification" *Neurocomputing*, Vol. 321,321-33 (2018)
- [2] Xue, Y., Xu, T., Zhang, H., Long, L. R., & Huang, X. "Segan: Adversarial net-work with multi-scale l 1 loss for medical image segmentation", *Neuroinformatics*, 16, 383-392 (2018).
- [3] Way, G. P., & Greene, C. S, "Extracting a biologically relevant latent space from cancer transcriptomes with variational autoencoders". in *pacific symposium on biocomputing 2018: Proceedings of the Pacific Symposium (pp. 80-91)*, 2018
- [4] Sandfort, V., Yan, K., Pickhardt, P. J., & Summers, R. M., "Data augmentation using generative adversarial networks (Cycle GAN) to improve generalizability in CT segmentation tasks", *Scientific reports*, 9, (2019).
- [5] Liu, X., Guo, S., Zhang, H., He, K., Mu, S., Guo, Y., & Li, X. (2019). Accurate colo-rectal tumor segmentation for CT scans based on the label assignment generative ad-versarial network. *Medical physics*, 46(8), 3532-3542.s
- [6] Asperti, A. (2019). About generative aspects of variational autoencoders. In *Machine Learning, Optimization, and Data Science: 5th International Conference, LOD 2019, Siena, Italy, September 10–13, 2019, Proceedings 5* (pp. 71-82). Springer International Publishing.
- [7] Odaibo, S. (2019). Tutorial: Deriving the standard variational autoencoder (vae) loss function. arXiv preprint arXiv:1907.08956.
- [8] Kingma, D. P., & Welling, M. (2019). An introduction to variational autoencoders. *Foundations and Trends in Machine Learning*, 12(4), 307-392.

- [9] Desai, S. D., Giraddi, S., Verma, N., Gupta, P., & Ramya, S. (2020, September). Breast cancer detection using GAN for limited labeled dataset. In 2020 12th International Conference on Computational Intelligence and Communication Networks (CICN) (pp. 34-39). IEEE.
- [10] Negi, A., Raj, A. N. J., Nersisson, R., Zhuang, Z., & Murugappan, M. (2020). RDA-UNETWGAN: an accurate breast ultrasound lesion segmentation using wasserstein generative adversarial networks. *Arabian Journal for Science and Engineering*, 45, 6399-6410.
- [11] Wang, S., Wang, X., Hu, Y., Shen, Y., Yang, Z., Gan, M., & Lei, B. (2020). Diabetic retinopathy diagnosis using multichannel generative adversarial network with semisupervision. *IEEE Transactions on Automation Science and Engineering*, 18(2), 574-585.
- [12] Kora Venu, S., & Ravula, S. (2020). Evaluation of deep convolutional generative adversarial networks for data augmentation of chest x-ray images. *Future Internet*, 13(1), 8.
- [13] Bushra, S. N., & Shobana, G. (2020, December). A Survey on Deep Convolutional Generative Adversarial Neural Network (DCGAN) for Detection of Covid-19 using Chest X-ray/CT-Scan. In 2020 3rd International Conference on Intelligent Sustainable Systems (ICISS) (pp. 702-708). IEEE.
- [14] Bushra, S. N., & Shobana, G. (2020, December). A Survey on Deep Convolutional Generative Adversarial Neural Network (DCGAN) for Detection of Covid-19 using Chest X-ray/CT-Scan. In 2020 3rd International Conference on Intelligent Sustainable Systems (ICISS) (pp. 702-708). IEEE.
- [15] Lin, C. C., Hung, Y., Feris, R., & He, L. (2020). Video instance segmentation tracking with a modified vae architecture. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition* (pp. 13147-13157).
- [16] Lou, Z., Le, K., & Tian, X. (2021, June). Nu-net based gan: Using nested u-structure for whole heart auto segmentation. In 2021 IEEE International Conference on Artificial Intelligence and Computer Applications (ICAICA) (pp. 601-604). IEEE.
- [17] Xi, Y., & Xu, P. (2021). Global colorectal cancer burden in 2020 and projections to 2040. *Translational oncology*, 14(10), 101174.
- [18] Thambawita, V., Salehi, P., Sheshkal, S. A., Hicks, S. A., Hammer, H. L., Parasa, S., ... & Riegler, M. A. (2022). SinGAN-Seg: Synthetic training data generation for medical image segmentation. *PloS one*, 17(5), e0267976.
- [19] Liu, B., Lv, J., Fan, X., Luo, J., & Zou, T. (2022). Application of an Improved DCGAN for Image Generation. *Mobile Information Systems*, 2022.
- [20] Raju, M. S. N., & Rao, B. S. (2022). Colorectal Cancer Disease Classification and Seg-mentation Using A Novel Deep Learning Approach. *International Journal of Intelligent Engineering Systems*, 15(4), 227-236.
- [21] Rafique, R., Islam, S. R., & Kazi, J. U. (2021). Machine learning in the prediction of cancer therapy. *Computational and Structural Biotechnology Journal*, 19, 4003-4017.
- [22] Li, Y., Wu, X., Yang, P., Jiang, G., & Luo, Y. (2022). Machine Learning for Lung Cancer Diagnosis, Treatment, and Prognosis. *Genomics, Proteomics & Bioinformatics*.