Synthesizing MRI Images Using Generative AI and Image Style Transfer GAN

PL.Chithra[1], Dhivya.S.D[2]

{ chitrasp2001@yahoo.com[1], dhivyadharmalingam@gmail.com[2]}

Department of Computer Science, University of Madras, India[1], Department of Computer Science, University of Madras, India[2]

Abstract. This work proposes to realistically synthesize the missing MRI sequence using generative artificial intelligence (GAI). GAI framework learns the patterns, data distribution, and structure of the input MRI images by using training data with those MRI images and then generates new data with similar data distribution as input MRI images. In this proposed work, Deep Convolution Generative Adversarial Network (DCGAN) and Pix2PixGAN generate a synthesized image using a real input image and random noise drawn from the dataset distribution. Its output is approximated to the original image using Image Style Transfer GAN, which gives an impressive outcome with an accuracy of 95.6%. Accuracy and loss are assessed using the performance metrices, for content and style images and the data distribution of real and fake images of DCGAN and Pix2Pix GAN. To treat brain tumor patients, it is vital to examine all four modalities, such as T1, T2, T1w, and Flair, which are often missing due to image artifacts and time constraints. These missing brain tumor modalities lead to difficulties while treating patients, creating class imbalance and reducing performance and accuracy while training the dataset for classification and segmentation in deep learning and machine learning.

Keywords: Generative Artificial Intelligence, DCGAN, Pix2PixGAN, Image Style Transfer GAN, modalities, class imbalance.

1 Introduction

While computer Tomography (CT) is better at imaging bones, magnetic resonance imaging (MRI) offers better visualization of soft tissue, cartilage, ligaments, and organ comparison. CT distinguishes between fat, water, muscle, and other soft tissue. The benefit of an MRI experiment is that there's less radiation, in contrast to X-rays, CT scans, and Positron Emission Tests (PET), so it is able to be used greater adequately on individuals who are greater liable to the consequences of radiation. Researching more about cancer after they find it through examination is necessary, to locate the exact location and nature of the tumor. To plan cancer treatments, such as surgery or radiation therapy, and analyze how well the treatment works. In detecting and planning treatment for brain tumor T1, T1w, T2, and FLAIR images, these MRI

IACIDS 2023, November 23-25, Lavasa, India

DOI 10.4108/eai.23-11-2023.2343219

images are sometimes missed and visually unappealing in clinical practices due to time limits or image artifacts such as patient motion, etc. This reduces the chances of detecting tumor and planning treatment for tumor, and most algorithms require these four sequences. To overcome these missing sequences, the proposed work realistically synthesizes the missing MRI sequence using generative artificial intelligence (GAI) Also, missing some modalities leads to class imbalance, which reduces performance and accuracy while training the dataset.

The synthetic image generation process goes through two processes: a learning process and a generation process, the generator gets samples from a normal distribution and learns features and data distribution of MRI images during the training process. Using a random noise vector as input, sequential upsampling procedures are used to generate an MRI image by learned features, data distribution during the training process is known as the generation process. The discriminator attempts to distinguish between MRI images generated by the generator and those with original input MRI images. The rest of the work based on the DCGAN which consists of generator and discriminator blocks and Style Transfer GAN, followed by Pix2Pix GAN 's discriminator and generator block, and finally Style Transfer GAN. This work is organized as follows Section 2 explains the background on GANs, Section 3 the methodology, Section 4 MRI dataset and preprocessing, Section 5 results and performance matrices followed by Section 6 presents the conclusion and future work.

2 Literature Survey

Generative image models have been extensively researched and are classified into parametric and nonparametric types. Non-parametric models are frequently employed in texture creation [1], super-resolution [2], and in-painting [3] by matching patches of images from a database of existing images. Parametric models for image generation have been extensively studied and experimented on MNIST digits or texture creation [4]. A variational sampling technique for image generation has had some success [5], although the samples are frequently fuzzy. This method employs an iterative forward diffusion process to generate images [6].G. M. Conte et al., work, used the publicly available, data set and trained two Generative Adversarial models to generate missing MRI sequences in one for generating T1 sequences and another for FLAIR sequences using Mean Squared Error (MSE) and Structural Similarity Index (SSE) [7].Changhee Han. et al., worked with Convolutional Neural Network-based brain tumor detection, which is challenging via conventional GAN difficulties arise due to unstable GAN training with high resolution. Progressive growing GAN(PGGAN)--based data augmentation (DA) method has shown better performance, combined with classical DA, in tumor detection and other medical imaging tasks. [8]. Martin Arjovsky et al. Proposed a method improving the stability of learning, eliminating mode collapse, and providing learning curves used for debugging and hyperparameters, the deep connections to different distances between distributions [9]. The main objective of the proposed work focuses on combining two different variants of GAN namely DCGAN and Pix2Pix, its outcomes are approximated with Style transfer GAN to achieve more resembling images similar to real or input images. In DCGAN 4 dense layers including input and output layers are used to reduce internal complexity and to overcome the vanishing gradient problem often raised in GAN. To normalize input value 0 to 1 batch normalization is used in each layer, and transposed convolution is used to retain the important features in input images, and LeakyRelu is used instead of Relu to scale the output from each layer -1 to 1. Pix2pix GAN, to generate synthetic MRI images of brain tumors batch normalization used in the UNet architecture to normalize mean and variance of input passed in each layer. Style transfer is used to maintain approximately the same texture as the original input image and it is a technique to enhance the image resemblance along with the above-mentioned features included in this model making it to obtain an accuracy of 95.6 %.

3 Methodology

The architecture of the proposed method consists of DCGAN and Pix2pix GAN with Style Transfer GAN as shown in Figure 1.

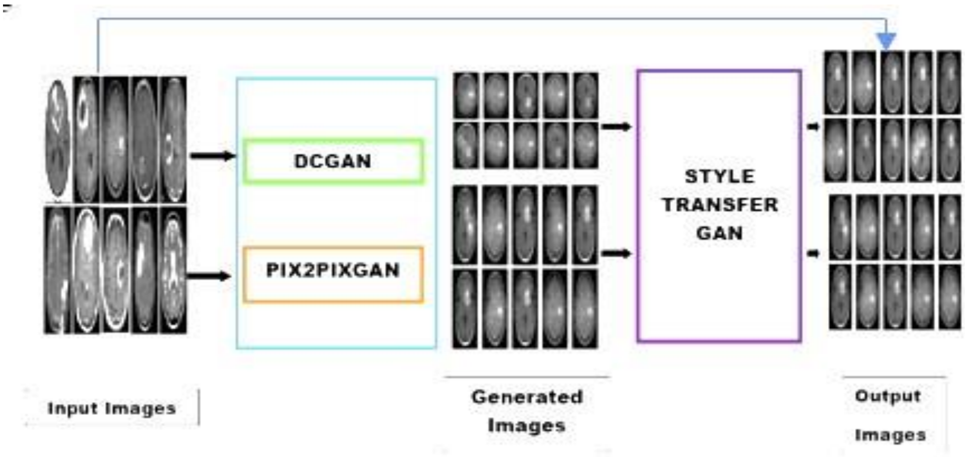


Figure. 1. Architecture of the Proposed Work

3.1 DCGAN with Style Transfers

The DCGAN consists of two blocks, discriminator block D and generator block G. The data distribution is confiscated by the generative model. The chance of the sample being chosen from the training MRI dataset to the generative model is determined by the discriminative model. An adversarial procedure is used to train the two models simultaneously. This architecture follows a game theory strategy that resembles a minimax two-player [10]. The output MRI image from DCGAN is fed into Style Transfer GAN as a style image and the original input MRI image is fed as a content image into Style Transfer GAN to obtain the final output MRI Image.

3.1.1 DCGAN Generator Block

In the generator blocks G, the first layer generates uniform noise distribution Z as input is fed to the dense net layer, in contradiction to the process of upsampling noise distribution using learned features while training the MRI image dataset. The noise vector is taken as an input MRI image and upsampled using transpose convolution to an input MRI image of size 128, as shown in Figure 2. To retain the learned features of an image and to regularize the flow of input to each unit so that it has a mean value of 0 and variance as 1, this makes gradient flow in deeper models and helps to solve training MRI image dataset issues such as sample oscillation and model vulnerability. The deep generators starts to learn and stops the generator from combining all of the samples as one point. The LeakyReLu [12] activation is used in all layers except the output layer, and the Tanh activation scales the output between −1 and 1. The 2 x 2 fractional-(strided convolutions) Generator, allows the network to learn (spatial downsampling) on their own. Remove related hidden layers to avoid model instability and stabilize the convergence speed. Backpropagation and an optimizer Adam with a constant learning rate of 0.01 are used to update the generator weights. The generator loss is calculated using binary cross entropy. The generator network of a DCGAN consists of 4 hidden layers, including the input layer, 1 hidden layer, and 1 output layer. Transposed convolution is performed in hidden layers to upsample the images, which are followed by batch normalization and LeakyReLU activation functions.

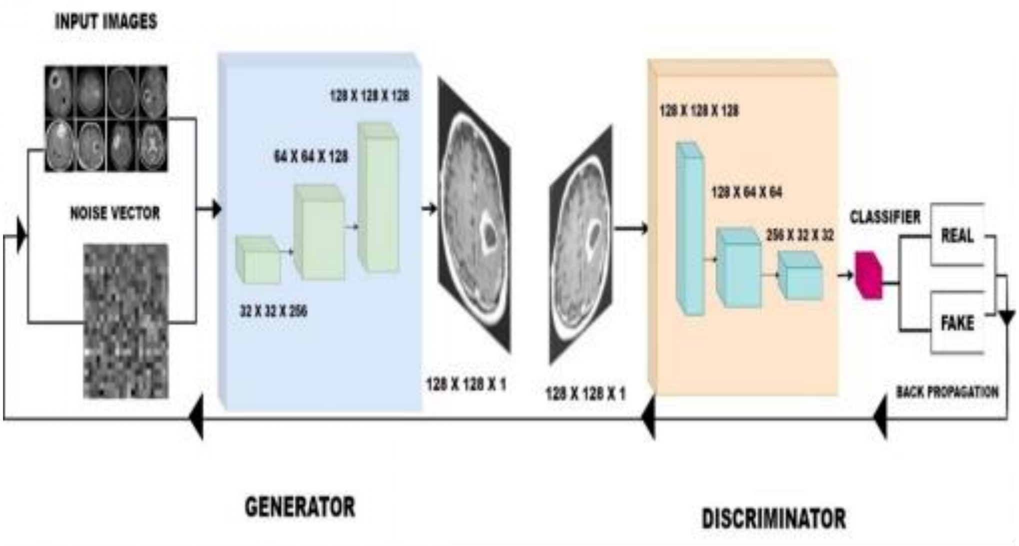


Figure. 2. Architecture of DCGAN

3.1.2 DCGAN Discriminator Block

The discriminator block consists of 4 layers, the input layer as the first layer followed by two layers and the output layer. Each layer has batch normalization, except the first layer. LeakyReLU activation used for all layers except output layer, and Sigmoid is used for the output layer, as shown in Figure 2. The input channel has a stride of 1, and all hidden layers have a stride value of 2 and a padding value of 1 so that the output image sizes will be half the input images. As image sizes increase in deeper layers, the number of channels increases by twice. Backpropagation and an optimization phase are used to update the discriminator's weights. Batch normalization is used to regularize the discriminator's learning rate. LeakyReLU activation function is used gradients to avoid backward flow of the layer. The last layer classifies the output.

3.1.3 Style Transfer GAN for DCGAN

The style Transfer GAN works with the principle of texture transfer. The pre-trained model of VGG consists of 16 convolutions and five pooling layers. The network scales the weights to normalize the values , the mean activation of each convolutional filter across images and positions is one. The texture transfer creates a texture from a input MRI image by limiting texture synthesis to preserve the target MRI image's content. This is accomplished by matching the deep feature's Gram matrix statistics through optimization. The Style Transfer GAN uses the VGG 19 network, which comprises rectified linear activation functions and no normalizing or pooling over feature maps. Re-scaling is possible without affecting its output. For image synthesis, the max pooling operation is replaced by average pooling, yielding a slightly more appealing outcome. Assume X and Y are the content and style images in the domain where no pairings exist between them. To characterize the tumor image domain X, a brain tumor dataset is used to obtain features, regions of interest, tones, and textures. $\{x_a\}$ a =1. . . . , N, $x_a \in X$. To train the style block output of DCGAN images, $\{y_b\}$ b = 1,..., M, $y_j \in Y$, $Y_\beta \subset Y$ refers to a sub-domain of Y that contains tumor images of β. When $y_b \in Y\beta$, we denote it as $y\beta_b$. The style GAN trains two separate neural networks, G and F, one to transfer input images and another to replace DCGAN output images with original input MRI images. This is done by using diverse training, Generator G can generalize its learning to the entire input image dataset domain and can transfer regions of interest to arbitrary regions at run time. DCGAN and Pix2PixGAN output, $y_\beta \in Y\beta$, and input images, $x \in X$, the transfer network $G: X \times Y_\beta \rightarrow Y_\beta$ extracts the tumor region from y_β and applies it to x to maintain its identity. Our result G (x, y_β), highlighted in Figure 3, should belong to the domain Yβ. Given the same photo y_β, the final output network $F: Y \rightarrow X$ learns to identify the region of interest in the identity of y_β

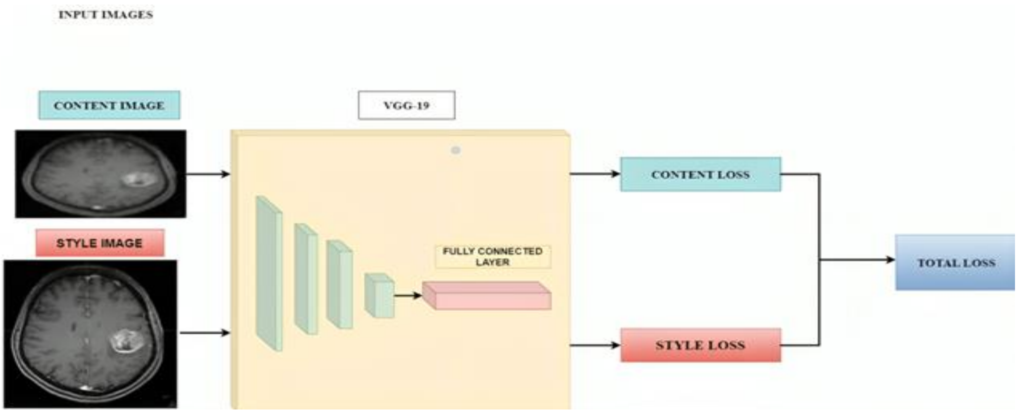


Figure 3. Architecture of Style Transfer GAN

The G and F are unbalanced functions. G takes a pair of images as input to transfer the style from one style to the other. F learns to reproduce ROI.. Total Loss of Style Transfer GAN is calculated by optimizing a min-max.

3.2 Pix2Pix GAN with Style Transfer GAN:

Pix2Pix is a conditional GAN (CGAN) that uses conditional distribution as an instruction to distribute the output data. It pertains to class labels with the condition. As a result, during dataset training, the UNet architecture receives the images with their actual labels. The Pix2Pix GAN uses real MRI image data, noise, along with condition labels to generate images. The Pix2Pix consists of two blocks, namely the generator UNet and the discriminator patch GAN, as shown in Figure 4.

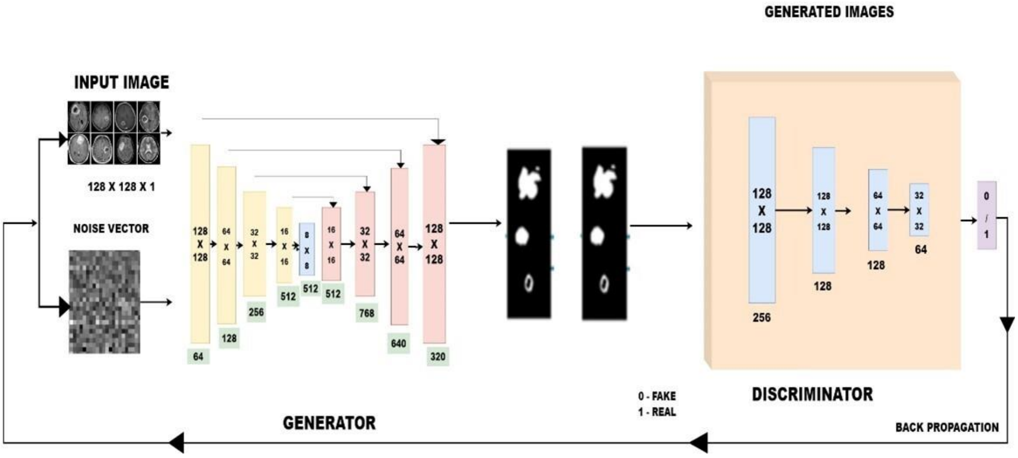


Figure. 4. Architecture of a Pix2Pix GAN

3.2.1 Generator Block of Pix2Pix GAN

The generator G learns a mapping from real data distribution X and noise distribution Z and Y are labeled outputs or target data from the generator G: {X, Z}->Y. The generator attention UNet consists of two blocks, namely the encoding and decoding. Convolution layers make up the encoding block, also called the contractive path. The proposed method consists of 4 encoding units of convolution layer, leaky Relu, and batch normalization. They extract spatial information from the input image while downsampling data. Then they pass the data to the subsequent layer until it reaches the middle, or the bottleneck, from which the expansive path originates. The decoder is also known as the expansive path; the transpose convolution is performed in upsampling, and concatenation, the information and a skip connection are used to retain the information lost during downsampling

3.2.2 Discriminator Block of Pix2Pix GAN

The discriminator D learns representation from label data Y and real data distribution and distinguishes real or fake D=Y | X. The discriminator uses Patch GAN, it splits an image into actual and fake segments; it classifies a patch of (n*n) in the image as real and fake. As a result, more restrictions are imposed, and high-frequency details are visualized. This patching method operates more quickly than classifying the entire image. The discriminator follows patch pairings: an input image and a generated image, as well as an input image and a target image, merge the input pairs together. The discriminator loss consists of the sum of two losses, one between an array of ones and an actual image and the other between an array of zeros and a synthetic image. The sigmoid cross-entropy loss is used to calculate the loss.

3.2.3 Style Transfer GAN for Pix2Pix GAN

Pix2Pix GAN is an advancement of GAN. It efficiently handles the transfer of image styles between data sets and the model to extract MRI image features more effectively. The pre-trained model of VGG consists of 16 convolutions and five pooling layers. The fact that the data set is not paired and there is also an edge blur edge. The Gaussian smoothing technique is used to retain the edge information more clearly as shown in Figure 3.

4 Dataset and Preprocessing

The dataset contains 920 tumor images, without tumor 300, for a total of 1220 images from the public brain tumor datasets of 2021, 2000, and 2019. The preprocessing techniques included resizing the images and scale between [-1, 1]. The MRI image data is given as a NumPy array so that pixel values are 8-bit unsigned integer values in the range [0, 255].

5 Result and Performance Metrices

The overall output of Image style Transfer GAN data distribution of real images and fake images is based on content loss and style loss is plotted as the graph is shown in Figure 5. The parameters required to train and test the DCGAN, Pix2Pix GAN, and progressive GAN are tabulated in TABLE 1

Table 1: Parameter configuration of DCGAN, PIX2PIX GAN and Style Transfer GAN

Parameters	DCGAN	Pix2PixGAN	Progressive GAN
Mital batch size	120	80	100
No. of Epochs	3750	2000	2500
Discriminator Learning rate	0.0001	0.0001	0.0001
Generator Learning rate	0.0002	0.0002	0.0002
Optimizer	Adam	Adam	Adam
Discriminator Loss	4.2999	3.8712	7.988
Generator Loss	8.2126	6.5682	9.1280
Accuracy without Style Transfer	87.5%	89.02%	83%
Accuracy Style Transfer	95.6%	93.7%	92.8%

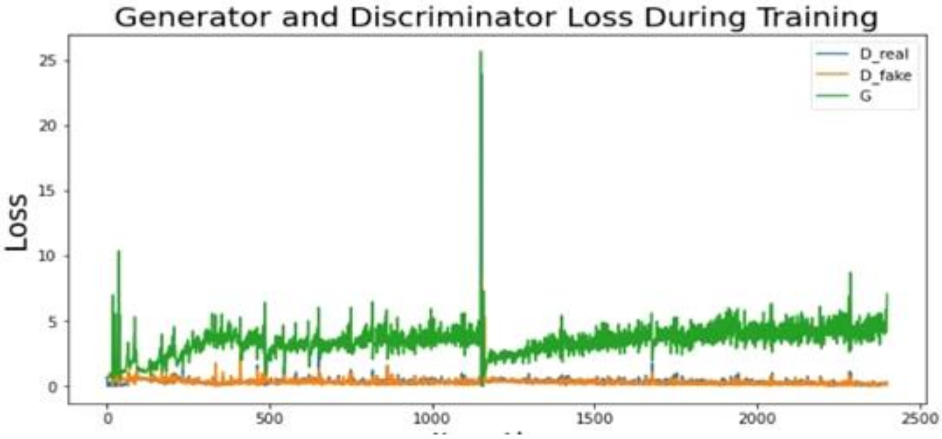
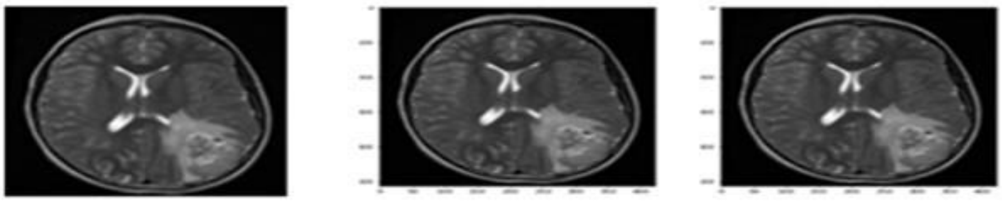


Figure. 5. Distribution of real and fake images of Style GAN

The input MRI image, the ground truth of DCGAN with Style Transferred GAN image, and with Style Transferred GAN are shown in Figure 6.



(a) (b) (c)

Figure 6**.** (a) Input Image (b) DCGAN with Style Transferred GAN (c) Pix2pix with Style Transferred GAN

The accuracy of DCGAN, PIX2PIX GAN, and progressive GAN has been compared with and without style transfer GAN. Is shown in Figure 7.

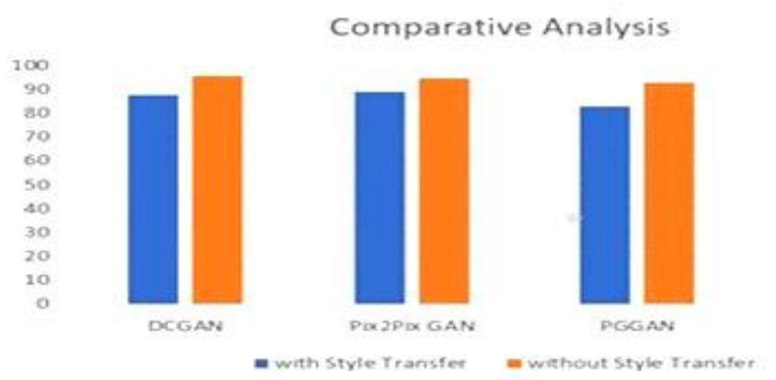


Figure 7. Comparative analysis of DCGAN, PIX2PIX GAN, and Progressive GAN with and without Style Transfer GAN

6 Conclusion and Future Work

The proposed work for synthesizing images using the GAN framework DCGAN and Pix2Pix GAN outcomes are passed to Style Transfer GAN received outcomes are remarkable and it closely resembles the real input images with an accuracy of 95.6 % and its performance is evaluated using data distribution of the real image and fake images and content image and style images. Future work includes overcoming model instability caused due to increasing layers, overcoming gradient degradation while training dense layers, reducing noise at times region of interest, and increasing accuracy.

References

[1] Efros, Alexei, Leung, Thomas K, et al. Texture synthesis by non-parametric sampling. In Computer Vision, 1999. The Proceedings of the Seventh IEEE International Conference on, volume 2, pp. 1033–1038. IEEE, 1999

[2] Freeman, William T, Jones, Thouis R, and Pasztor, Egon C. Example-based super-resolution. Computer Graphics and Applications, IEEE, 22(2):56–65, 2002

[3] Hays, James, and Efros, Alexei A. Scene completion using millions of photographs. ACM Transactions on Graphics (TOG)

[4] Portilla, Javier and Simoncelli, Eero P. A parametric texture model based on joint statistics of complex wavelet coefficients. International Journal of Computer Vision, 40(1):49–70, 2000.

[5] Kingma, Diederik P and Ba, Jimmy Lei. Adam: A method for stochastic optimization. arXiv preprint arXiv:1412.6980,26(3)

[6] Soki-Dickstein, Jascha, Weiss, Eric A, Maheswaranathan,Niru, and Ganguli, Surya. Deep unsupervised learning using nonequilibrium thermodynamics. arXiv

[7] G. M. Conte, A. D. Weston, D. C. Vogelsang, K. A. Philbrick, J. C. Cai, M. Barbera, F. Sanvito, D. H. Lachance, R. B. Jenkins, W. O. Tobin, “Generative adversarial networks to synthesize missing t1

and flair MRI sequences for use in a multisequence brain tumor segmentation model," Radiology, vol. 299, no. 2, pp. 313–323, 2021.
[8] Changhee Han, Leonardo Rundo, Ryosuke Araki, Yujiro Furukawa, Giancarlo Mauri, Hideki Nakayama, and Hideaki Hayashi," Infinite Brain MR Images: PGGAN-based Data Augmentation for Tumor Detection"
[9] Arjovsky, M., Chintala, S. & Bottou, L. Wasserstein generative adversarial networks. In: International conference on machine learning, pp. 214–223 (PMLR, 2017).
[10] Goodfellow, I. J., et al. Generative adversarial networks. arXiv preprint arXiv:1406.2661 (2014).
[11] ssIoffe, Sergey and Szegedy, Christian. Batch normalization: Accelerating deep network training by reducing internal covariate shift. arXiv preprint arXiv:1502.03167, 2015