Article

Generating Synthetic Images for Healthcare with Novel Deep Pix2Pix GAN

Abeer Aljohani * and Nawaf Alharbe ![ORCID]

Computer Science Department, Applied College, Taibah University, Madinah 46537, Saudi Arabia
* Correspondence: aahjohani@taibahu.edu.sa

Abstract: Due to recent developments in deep learning and artificial intelligence, the healthcare industry is currently going through a significant upheaval. Despite a considerable advance in medical imaging and diagnostics, the healthcare industry still has a lot of unresolved problems and unexplored applications. The transmission of a huge number of medical images in particular is a difficult and time-consuming problem. In addition, obtaining new medical images is too expensive. To tackle these issues, we propose deep pix2pix generative adversarial networks (GAN) for generating synthetic medical images. For the comparison, we implemented CycleGAN, Pix2Pix GAN and Deep Pix2Pix GAN. The result has shown that our proposed approach can generate a new synthetic medical image from a different image with more accuracy than that of the other models. To provide a robust model, we trained and evaluated our models on a widely used brain image dataset, the IXI Dataset.

Keywords: GAN; CycleGAN; Pix2Pix GAN; medical images; image synthesis; image transmission

1. Introduction

Modern machine learning (ML) is extensively applied in the field of medicine and has demonstrated incredible results year by year. One of the conventional methods is the variational autoencoder (VAE), which was initially proposed for a variational inference problem in deep covert Gaussian algorithm [1,2]. Recent deep learning framework, Generative Adversarial Networks (GANs) [3] got a lot of attention due to its ability to create synthetic medical images that are very close to real ones. One of the challenges of medical image collection is the infrequency of medical image datasets. It is difficult to find the dataset that deals with medical images when compared to other domains. It is usually hard to find because of the proprietary values, the data being personally identifiable information, IP, the monetary value associated with it and more. There is also a challenge in terms of lesser devices being in the field such as MRIs (Magnetic resonance imagings). Moreover, medical images play an essential role that human radiologists diagnosing patients. This is obviously limited by speed, tiredness and experience. Therefore, it takes years and great financial cost to teach a radiologist, to become a qualified radiologist. In addition, MRI scans take a long time and taking a different view of the patient's organ requires more time. Instead of obtaining long clinical scanning and a costly process, GAN helps to create synthesic images that show other views of the scanning organ. Therefore, the opportunity of artificially created medical images is significantly engaging. However, there are still complex unsolved issues in creating synthetic high-quality medical images for current medical research. Recent medical image studies have progressed through more reliable and modern techniques, able to reliably methodically acquire the proper features of the different acquisition devices. When medical experts' knowledge reached anatomical and physiological information, it was a great assistance to produce realistic images. This allows us to confirm medical image inspection techniques for medical analysis, therapy planning, and other medical applications. For example, [4] proposed the GAN-based framework, which has been shown to produce high-resolution, high-fidelity images in an unsupervised

manner in the field of medical image processing. Those generated synthetic of T1-weighted brain MRI images are comparable with real images. [5] represented in their work generating high-resolution skin lesion images which are absolutely the same as real images by applying progressive growing of GAN. One more work [6] which used unsupervised learning with Deep Convolutional-Generative Adversarial Networks to generate lung nodule samples realistically. The GAN model represented that images of lung cancer nodules are completely the same as the original images. According to [7] their method of deep convolution GAN generates restricted output image quality. However, deep convolution GAN has been applied to create high-quality image examples of lungs nodule and liver lesions. The lung image data set synthetically creates nodules touching the lung border in [8]. This method allows making better pathological lung segmentation. Moreover, segmentation of brain MR by using conditional GAN generators allows to create augmentation of training samples for brain tumour segmentation [9], and trained CycleGAN model to accurate geometric infringement in expansion MR [10]. As we introduced various GAN models can greatly help pathology, and improve the quality of medical service

The rest of the paper is organized as follows. Section 2 represents related works. Section 3 explains the models which we used. Section 4 reports the experimental results and the discussion. Finally, Section 5 concludes our work.

2. Related Works

In recent years, machine learning have been and deep learning widely applied for medical image processing and have shown incredible achievements. Medical image processing aims to extract abnormal information about a patient's medical condition. Medical images are achieved by several medical technologies, which are Magnetic Resonance Imaging (MRI), Computed Tomography (CT), Positron Emission Tomography (PET) and Ultrasound (US). The achieved images are processed using deep learning techniques to extract essential information about the disease that appears on the image.

One of the research fields in medical image processing is generating synthesized images based on GAN. The framework architecture includes two networks: one that generates fake images and the second one differentiates the original and synthetic images among them. GANs have attracted a lot of achieved attention in medical image analysis systems and various GAN models have recently focused on generating high-quality synthetic images. One of the purposes of the GAN framework is to generate images of one mode from another mode [11], and GAN used image inpainting [12].

Recently, the GAN framework has been applied to several medical imaging tasks. Most of the research has been accomplished by the GAN, which is an image-to-image framework to generate image translation. Costa et al. [13] propose an approach that learns to generate eye fundus images through the data. The authors matched real eye fundus images with their proper vessel tree and trained them to train retinal vessel segmentation images. Then, the model has trained the interpretation from the vessel tree to a generated retinal image. The authors, Dai et al. [14], purpose Structure Correcting Adversarial Network (SCAN) to segment lung parts and the heart in scanned images and the heart from chest X-ray images. Xue and his colleague [15] represented two GAN models as a segmentor and critic and tested their methods using datasets from the MICCAI BRATS brain tumor segmentation. Nie et al. [16], propose a method that uses data-driven to address where a fully convolutional network to generate computed tomography (CT) given a magnetic resonance (MR) image, a patch-based GAN was applied for translation. Ben-Cohen et al. [17] also used conditional GAN and fully convolution networks (FCN) to generate CT image to a PET image that best parts of liver lesions. One of the research areas has been focused on the GAN-based framework for image inpainting. The authors, Schlegl and his colleague [18], represented the AnoGAN model in their work, which used a deep convolutional generative adversarial network to learn a manifold of normal anatomical variability and was trained with good patches of the retinal area to use the data allocation of good tissue.

In this study, we proposed that we generate synthetic images from one view to another view by using deep GAN, and this means it is enough to have one view. It is not necessary to make an MRI of the other side by using expensive equipment and taking a staff member's time as well. We found that if we have one view; we make another view with the same accuracy as the first (PSSNR error of X). This also allows us to have more training data.

3. Methods

3.1. GAN Neural Network

In 2014, Ian Goodfellow and his team put forth the GAN model [3]. This model comprises two neural networks that are simultaneously trained to produce artificial data that is realistically dependent on the input. A generator that simulates a transfer function makes up the initial network. Input noise is used to generate data, which as nearly as possible resembles the goal data distribution. A discriminator in the GAN model's second network must be able to tell genuine data from generated data.

An ideal total loss for GAN is 0.5, and it is determined by the expression (1) below:

$$min_G max_D L(G, D) = P_{xr}[log D(x)] + P_{ng} log([1 - D(G(x, n))]) \tag{1}$$

where G and D are loss functions of the generator and discriminator, respectively, x represents the input image, n represents noise, P_{xr} represents the likelihood that the image are real, and P_{ng} represents the likelihood that the image are created. While the discriminator must identify discrepancies, increasing error, the generator aims to reduce error between both the real and created images [19].

3.2. CycleGAN

Making a new type of image from an existing one is a crucial aspect of image synthesic. Patients infrequently have CT and MR at the same time because paired pictures, for instance, are challenging to collect. Manpower is also needed to label other types of data. For translation from unpaired image to image, CycleGAN [20] is suggested with the aim of converting a picture from one sector into a consistent image from another sector. This is a versatile and cutting-edge framework that is extensively applied across numerous sectors. CycleGAN features two discriminators and two generators, unlike regular GAN. The first generator creates reliable domain B images from reliable domain A images, while the second generator does the opposite. CycleGAN is used to synthesize the picture from sector A to sector B, and then sector A is generated to complete the loop. The discriminator's job is to make sure that every image that CycleGAN creates is as close to the original as possible. In other words, by training two mutual GAN networks, CycleGAN completes the mutual translation of pictures across several domains. CycleGAN has relatively few data needs, is more task-adaptive, and does not demand paired sector A and sector B images throughout the training phase [21].

3.3. Pix2Pix GAN

Pix2Pix is a generator and a discriminator system that is built on the Conditional Generative Adversarial Network (CGAN) [22,23]. The generator discovers a mapping between the source picture x and random noisy image z to the target image y, i.e., $x, z \rightarrow y$. Discriminator makes a distinction between actual and fake $y|x$. The following equation represents Pix2Pix's objective function:

$$L_{Pix2Pix}(G, D) = E_{x,yPdata(x,y)}[log D(x, y)] + E_{xPdata(x),zPz(z)}[log(1 - D(x, G(x, z)))] \tag{2}$$

It is vital to note that Pix2Pix and the conventional GAN differ significantly from one another [24]. For instance, the Pix2Pix discriminator employs the PatchGAN structure to separate the genuine image from the fake image when given a pair of images as input rather than a single image as shown in Figure 1.
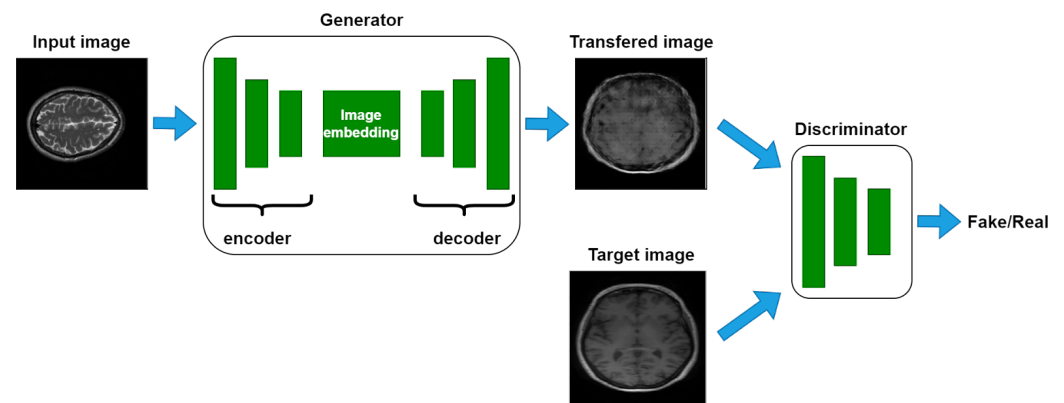
Figure 1. Overall flowchart of image-to-image transmission with Pix2Pix GAN.

In order to increase the performance of the pix2pix model for generating synthetic images from one view to another view, we developed the generator, which is modeled based on the architecture of U-Net [25], in this study. The network's conformation, which resembles the letter "U", inspired its name. An ascending path called the decoder and a descending path called the encoder make up the architecture. The encoder is a straightforward CNN that converts the input image into a feature vector by performing convolutions and pooling operations. Normal convolutions and transposed convolutions, commonly referred to as "up-convolutions", are carried out by the decoder. So, the size of the feature vector grows steadily until it is bigger than the size of the input image. The up-sampled feature map in the appropriate up-convolution layer in the decoder is concatenated with the feature information extracted in each convolution layer of the encoder during the up-sampling process. As a result, U-Net generates a segmentation map that is more accurate. We increased, respectively, the number of layers of the encoder and the decoder. The overall structure of the suggested generator architecture is illustrated in Figure 2. The hyperparameters of the suggested model are manually selected. The kernel dimensions in the convolutional layer of the encoder and decoder are the same 4×4. The stride size is 2×2 and the padding is the same. The rest of the hyperparameters of the proposed generator such as number of kernels in each layer and activation functions are can be found in Figure 2. The pseudo code of the proposed model is shown in Algorithm 1.

Algorithm 1 Pseudo code of training of Pix2Pix GAN. k is the number of network updates per training iteration

---

1: **for** *iteration* =number of training... **do**
2:     **for** k steps... **do**
3:         Sample minibatch of m samples $y^1, ..., y^m$ from data distribution $p_{data}(y)$.
4:         Sample minibatch of m noise samples $z^1, ..., z^m$ from noise prior $p_g(z)$.
5:         Sample minibatch of m samples $x^1, ..., x^m$ from data generating distribution $p_{data}(x)$.
6:         Update the discriminator by ascending its stochastic gradient: $\nabla_{\theta d} \frac{1}{m} \sum_{i=1}^{m} [log D(x^i, y^i)] + log(1 - D(x^i, G(x^i, z^i)))]$
7:     **end for**
8:     Sample minibatch of m noise samples $z^1, ..., z^m$ from noise prior $p_g(z)$.
9:     Update the generator by descending its stochastic gradient: $\nabla_{\theta g} \frac{1}{m} \sum_{i=1}^{m} log(1 - D(x^i, G(x^i, z^i)))$
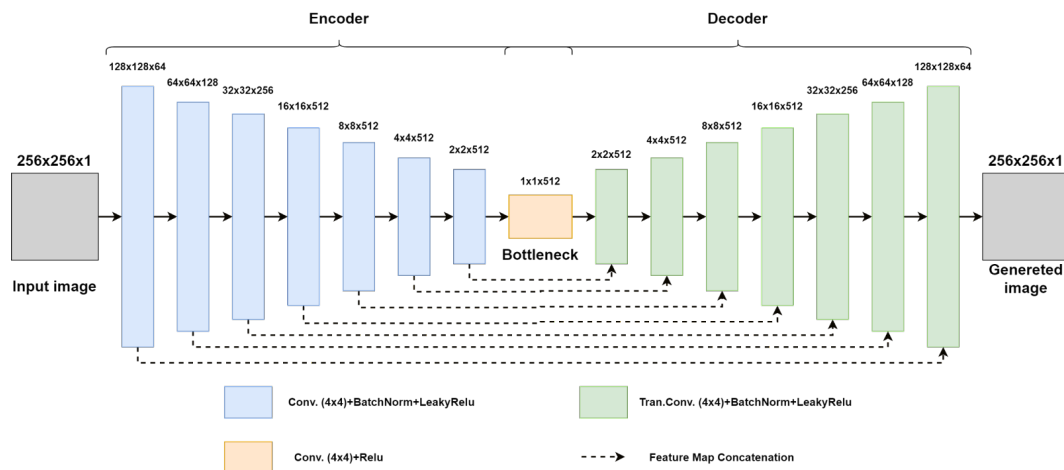10: **end for**

Figure 2. Structure of the proposed generator.

4. Experimental Results and Discussion

We carried out comprehensive experiments on the Brain slices images (IXI Dataset) [26] dataset, using a total of 2000 images (1000 images of them are used as an input domain and the rest of them is used as a target domain) for the training dataset and 300 images for the testing dataset, to confirm the efficacy and accuracy of the proposed deep pix2pix model. The outcomes of the suggested method have been compared to the results of Cycle GAN and the original pix2pix GAN models.

Training parameters: We used TensorFlow open-source framework (v2.10.0) and python programming language (version: 3.7) to build the architecture of the suggested model and to implement it. When it comes to a hardware system, the model was performed on a Windows 10 operation system and an AMD Ryzen 9 5900X 12-Core CPU running at 3.7 GHz, a graphics processing unit (Model: NVIDIA GeForce GTX 3060 with 24 GB), and CUDA for accelerated training were used in the experiment. There are 200 experimental training iterations. Adam optimization function was chosen as an optimizer. Batch size was set up at 64 and the learning rate is 0.0001.

Dataset: To generate good quality synthetic brain images, we picked up IXI dataset which 600 MR brain images were collected from a normal, healthy human brain by three different hospitals in London [26]. The MR brain image dataset includes T1, T2 and PD-weighted images, MRA images, and diffusion-weighted images (15 directions). The data were gathered for the IXI—Information extraction from Images (EPSRC GR/S21533/02) project. The T1 dataset has 430 participants and each participant's folder includes 50 vertical scanned brain images. Like T2, T2 also contained the same number of participants and their images but in a horizontal view.

Details about network implementation and structure: All of the models used in the study have 256 × 256 reshaped pictures. Train and test data separately load, respectively, 481 and 96. Then, we create the generator and discriminator models by calling the Tensorflow library. In addition, we define training staff and single training step as a step function. The deep pix2pix model by default employs a generator resembling PatchGAN, U-net and loss of Vanilla GAN.

Evaluation metrics: Two main indices, including Peak signal to noise ratio (PSNR) expressed in Equation (3) and structural similarity index measure (SSIM) expressed in Equation (4), are used to evaluate the new generated images in order to more objectively represent the image color rendering quality of various models. These two indices are frequently employed in the image processing evaluation measures [27]. *PSNR* is an objective metric to assess how well a color image was produced. The following is the calculating formula:

$$PSNR = 10log10(2n - 1)2MSE \tag{3}$$

$$MSE = 1HWi = 1Hj = 1WX(i,j) - Y(i,j)2 \tag{4}$$

where H and W stand for the image's height and width, (i,j) for each pixel, and n for the number of bits per pixel, respectively, and X and Y stand for two images.

The $SSIM$ index is utilized for further comparison because the $PSNR$ index, which has certain flaws as well, cannot accurately describe the uniformity of the visual effect and the quality of the images. A statistic to assess how similar two photos are is called $SSIM$. The efficiency and accuracy of this approach are shown by contrasting the model's output with the original color image. The following is the calculating formula:

$$SSIM = (2xx + c1)(2xy + c2)(x2y2 + c1)(x2y2 + c2) \tag{5}$$

where x and y stand for the mean value of the real image and the generated fake image, respectively. $x2$ and $y2$, respectively, demonstrate the difference of the real image and the generated fake image, xy indicates the covariance of the real image and the generated face image, $c1 = (k1, L)2$ and $c2 = (k2, L)2$ are constants that maintain stability, and L is the dynamic range of pixel value, $k1 = 0.01, k2 = 0.03$. For quantitative analysis, three models, including cyclegan, pix2pix and proposed deep pix2pix, attempt to generate synthetic images from one view to another view. All models, including Cycle GAN, Pix2Pix GAN and the proposed model, are trained 50 times, and an average value is used for analysis. The performance evaluation metrics are shown in Table 1. According to Table 1, the proposed deep pix2pix model performs very well in pixel loss, indicating that the deep generator of the proposed model can effectively decrease the interference of leaving interconnections on the decoder network.

Table 1. The results of the three models in Brain slices images dataset using MAE, PSNR and SSIM as standards.

Method	MAE	PSNR	SSIM
Cycle GAN	0.059	59.56	0.0344
Pix2Pix GAN	0.057	60.51	0.354
Deep Pix2Pix GAN	**0.038**	**61.32**	**0.391**

The three models, including cyclegan, pix2pix and proposed deep pix2pix, try to generate synthetic images from one view to another view for qualitative analysis. As shown in Figure 3, the proposed deep pix2pix model shows the most accurate result compared to the rest of the models used in the experiment, while the cyclegan model performs the worst qualitative analysis.
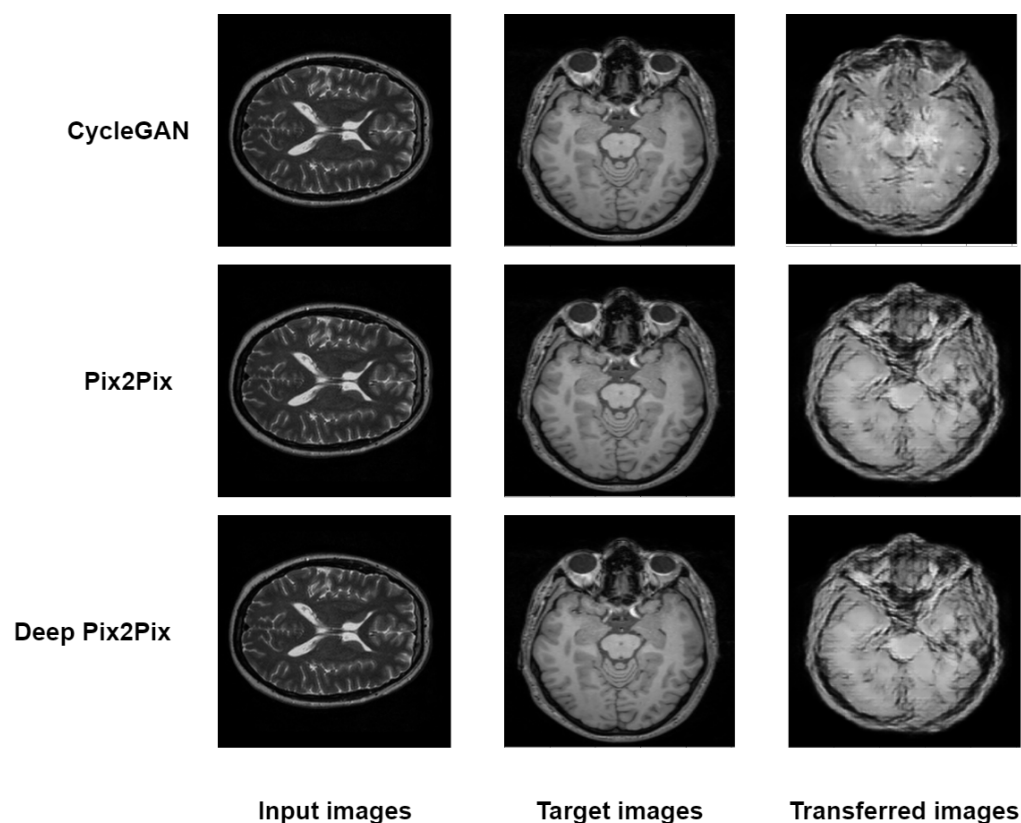
Figure 3. Testing images of CycleGan, Pix2Pix and Deep Pix2Pix.

5. Conclusions

In this work, a robust deep pix2pix (DPP) approach is proposed. The robustness of the proposed approach is evaluated by training an image translation model on the MRA |X| dataset. The results showed an increase in the quality of translated images. Comparisons are made with the proposed approach and previous state of the art methods, namely, cycleGAN and pix2pix. The comparison has shown that the result of the proposed model is better than that of cycleGAN and pix2pix. Our results suggest that making deeper, more powerful GAN architectures improves the performance of the medical image translation model. The hyperparameters of the deep pix2pix model are manually selected in this study. To increase the overall performance, particle swarm optimization (PSO) will be adapted to find the best proper hyperparameters for the model.

Author Contributions: Conceptualization, N.A.; methodology, N.A.; software, N.A.; resources, N.A.; supervision, A.A.; project administration, A.A. All authors have read and agreed to the published version of the manuscript.

Conflicts of Interest: The authors declare no conflict of interest.

References

1.　Kingma, D.P.; Welling, M. Auto-encoding variational bayes. *arXiv* **2013**, arXiv:1312.6114.
2.　Jimenez Rezende, D.; Mohamed, S.; Wierstra, D. Stochastic backpropagation and approximate inference in deep generative models. In Proceedings of the 31st International Conference on Machine Learning, Beijing, China, 21–26 June 2014; Voulme 32, pp. 1278–1286.
3.　Goodfellow, I.; Pouget-Abadie, J.; Mirza, M.; Xu, B.; Warde-Farley, D.; Ozair, S.; Courville, A.; Bengio, Y. Generative adversarial nets. In Proceedings of the 2014 Advances in Neural Information Processing Systems, Montreal, QC, Canada, 8–13 December 2014.

4. Bermudez, J.D.; Happ, P.N.; Feitosa, R.Q.; Oliveira, D.A. Synthesis of multispectral optical images from SAR/optical multitemporal data using conditional generative adversarial networks. *IEEE Geosci. Remote. Sens. Lett.* **2019**, *8*, 1220–1224. [CrossRef]

5. Baur, C.; Albarqouni, S.; Navab, N. Generating highly realistic images of skin lesions with GANs. In *OR 2.0 Context-Aware Operating Theaters, Computer Assisted Robotic Endoscopy, Clinical Image-Based Procedures, and Skin Image Analysis*; Springer: Cham, Switzerland, 2018; pp. 260–267.

6. Chuquicusma, M.J.M.; Hussein, S.; Burt, J.; Bagci, U. How to fool radiologists with generative adversarial networks? A visual turing test for lung cancer diagnosis. In Proceedings of the 2018 IEEE 15th International Symposium on Biomedical Imaging (ISBI 2018), Washington, DC, USA, 4–7 April 2018; pp. 240–244.

7. 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* **2018**, *321*, 321–331. [CrossRef]

8. Jin, D.; Xu, Z.; Tang, Y.; Harrison, A.P.; Mollura, D.J. CT-realistic lung nodule simulation from 3D conditional generative adversarial networks for robust lung segmentation. In Proceedings of the International Conference on Medical Image Computing and Computer-Assisted Intervention, Granada, Spain, 16–20 September 2018; pp. 732–740.

9. Mok, T.C.W.; Chung, A.C.S. Learning data augmentation for brain tumor segmen-tation with coarse-to-fine generative adversarial networks. *arXiv* **2018**, arXiv:1805.11291.

10. Gu, X.; Knutsson, H.; Nilsson, M.; Eklund, A. Generating diffusion MRI scalar maps from T1 weighted images using generative adversarial networks. In *Scandinavian Conference on Image Analysis*; Springer: Norrköping, Sweden, 2019; pp. 489–498.

11. Radford, A.; Metz, L.; Chintala, S. Unsupervised representation learning with deep convolutional generative adversarial networks. *arXiv* **2015**, arXiv:1511.06434.

12. Yeh, R.A.; Chen, C.; Yian, L.T.; Schwing, A.G.; Hasegawa-Johnson M.; Do, M.N. Semantic image inpainting with deep generative models. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Honolulu, HI, USA, 21–26 July 2017; pp. 5485–5493.

13. Costa, P.; Adrian, G.; Maria, I.M.; Michael, D.A.; Meindert, N.; Ana, M.M.; Aurélio, C. Towards adversarial retinal image synthesis. *arXiv* **2017**, arXiv:1701.08974.

14. Dai, W.; Dong, N.; Wang, Z.; Liang, X.; Zhang, H.; Xing, E.P. Structure correcting adversarial network for organ segmentation in chest x-rays. In *Deep Learning in Medical Image Analysis and Multimodal Learning for Clinical Decision Support*; Springer: Cham, Switzerland, 2018; pp. 263–273.

15. Xue, Y.; Xu, T.; Zhang, H.; Long, LR.; Huang, X. SegAN: Adversarial network with multi-scale L1 loss for medical image segmentation. *Neuroinformatics* **2018**, *16*, 383–392. [CrossRef] [PubMed]

16. Nie, D.; Trullo, R.; Lian, J.; Petitjean, C.; Ruan, S.; Wang, Q.; Shen, D. *Medical Image Synthesis with Context-Aware Generative Adversarial Networks*; Springer: Berlin, Germany, 2017; pp. 417–425.

17. Ben-Cohen, A.; Klang, E.; Raskin, S.P.; Amitai, M.M.; Greenspan, H. *Virtual PET Images from CT Data Using Deep Convolutional Networks: Initial Results*; Springer: Berlin, Germany, 2017; pp. 49–57.

18. Schlegl, T.; Seeb, P.; Waldstein, S.M.; Schmidt-Erfurth, U.; Langs, G. Unsupervised anomaly detection with generative adversarial networks to guide marker discovery. In *Advances in Neural Information Processing Systems*; Springer: Berlin, Germany, 2014; pp. 146–157.

19. Rocca, J. Understanding generative adversarial networks (gans). *Medium* **2019**, *20*. Available online: https://towardsdatascience.com/understanding-generative-adversarial-networks-gans-cd6e4651a29 (accessed on 21 October 2022).

20. Zhu, J.Y.; Park, T.; Isola, P.; Efros, A.A. Unpaired image-to-image translation using cycle-consistent adversarial networks. In Proceedings of the IEEE International Conference on Computer Vision, Venice, Italy, 22–29 October 2017; pp. 2223–2232.

21. Wu, H.; Jiang, X.; Jia, F. UC-GAN for MR to CT image synthesis. In *Workshop on Artificial Intelligence in Radiation Therapy*; Springer: Berlin, Germany, 2019; pp. 146–153.

22. Mirza, M.; Osindero, S. Conditional generative adversarial nets. *arXiv* **2014**, arXiv:1411.1784.

23. Odena, A.; Olah, C.; Shlens, J. Conditional image synthesis with auxiliary classifier gans. In Proceedings of the International Conference on Machine Learning, Sydney, NSW, Australia , 6–11 August 2017; pp. 2642–2651.

24. Wang, X.; Yan, H.; Huo, C.; Yu, J.; Pant, C. Enhancing Pix2Pix for remote sensing image classification. In Proceedings of the 2018 24th International Conference on Pattern Recognition (ICPR), Beijing, China, 20–24 August 2018; pp. 2332–2336.

25. Popescu, D.; Deaconu, M.; Ichim, L.; Stamatescu, G. Retinal blood vessel segmentation using pix2pix gan. In Proceedings of the 2021 29th Mediterranean Conference on Control and Automation (MED), Virtually, 22–25 June 2021; pp. 1173–1178.

26. Ixi Dataset. Available online: https://brain-development.org/ixi-dataset/ (accessed on 21 October 2022).

27. Yao, S.; Tan, J.; Chen, Y.; Gu, Y. A weighted feature transfer gan for medical image synthesis. *Mach. Vis. Appl.* **2021**, *32*, 1–11. [CrossRef]