Proceedings of the 2023 International Conference on Software Engineering and Machine Learning
DOI: 10.54254/2755-2721/8/20230124

Inter-translation between CT and MRI brain scans based on cycle consistent adversarial networks

Haoran Li

The department of computer science, Sichuan university, Chengdu, 610207, China

lihaoran@stu.scu.edu.cn

Abstract. Since Computed Tomography (CT) and Magnetic Resonance Imaging (MRI) are often used together in diagnosing brain diseases, it is rather time-consuming and expensive for patients to obtain two scans simultaneously. In this paper, the author proposed a new idea for CT-MRI inter-translation, using dataset of human brain scans to achieve unpaired image-to-image translation between CT scans and MRI scans. More specifically, the author proposed using Cycle Consistent Adversarial Networks (CycleGAN) to realize the idea of style transfer between CT and MRI. Briefly, this paper had trained two sets of generator and discriminator to form a "cycle". This model can retain the main characteristics of a scan while transferring the scan's style. To keep this translation process "cycle-consistent", the $Loss_{cyc}$ is used to keep the main content. Furthermore, the $Loss_{GAN}$ ensures that the generated images exhibit a close stylistic resemblance to the target domain and the $Loss_{identity}$ guarantees that the hue of the generated images is retained during the translation process. In addition, the author has added weights to the $Loss_{cyc}$ and the $Loss_{identity}$ to help the model perform better. Finally, this paper visualize the training process by presenting the original images and generated images together. Experimental results indicates that CycleGAN achieves a competitive performance can hopefully be a good auxiliary in brain diseases diagnosis.

Keywords: cycle consistent adversarial networks, CT scans, MRI scans, Res Net.

1. Introduction

Computed Tomography (CT) and Magnetic Resonance Imaging (MRI) are two essential ways used in brain scans to identify strokes, tumors, or other issues in brains. A CT scan is a diagnostic imaging process that produces images of the inside of the body using computer technology and X-rays [1]. In addition, MRI aims to create precise images of virtually every internal body structure. MRI scanners combine radiation and a powerful magnet to make images of the human body. No radiation is released during an MRI exam, in contrast to X-rays [2].

Today 700 million people worldwide suffer from brain diseases [3]. CT and MRI are often used together in brain scans, because they have different effects in diagnosing different brain diseases. CT mainly aims at haemorrhagic brain diseases and its imaging speed is faster than MRI. MRI is mainly used to check the infarction in the brain and the image quality is very high.

Consider the following facts: 1) CT and MRI are often used together for the diagnosis of brain diseases; 2) Some patients are already very weak and are not suitable for excessive X-ray radiation caused by CT; 3) The price of MRI is relatively more expensive than CT, and MRI imaging speed is slower. Therefore, unpaired image-to-image translation between CT and MRI can not only save a lot of time and money for the patients, but also prevent patients from absorbing excessive X-ray radiation, which can be considered in this case.

In 2014, Ian Goodfellow et al. proposed Generative Adversarial Networks (GAN), and GAN launched a deep learning revolution [4]. GAN bears a lot of advantages, and it has various application scenarios. First, the most powerful feature of GAN is that it has an unsupervised learning process, and thus GAN does not need labeling the data. In addition, GAN can augment images, generate images of high quality, create images from text, transform images from one domain to another, etc. Cycle-Consistent Adversarial Networks (CycleGAN) which is proposed in 2017 by Jun-Yan Zhu et al. is a derivative of GAN [5]. CyleGAN is widely used to achieve unpaired image-to-image translation. This translation process is also called style transfer, because this process will transfer the style of an image while retaining the image’s main content.

Nowadays, deep learning methods are broadly used in the diagnosis of brain diseases. The previous researches mainly focus on the identification and classification based on feature extraction techniques [6]. However, image translation is rarely used despite its huge potential. This study hopes to achieve image-to-image translation between CT and MRI, and thus reduce the cost of time and money spending on brain scans and prevent the patients from absorbing excessive X-ray radiation.

This paper uses CycleGAN to achieve the translation between CT and MRI. CT and MRI brain scans dataset which is from Kaggle is used in this paper [7]. In addition, the dataset is divided into trainA, trainB, testA and testB, with domain A being CT scans and domain B being MRI scans. There are totally 4972 images in the dataset. After the training process, the generated images are very close to the pictures in the other domain. Using the method in this paper, CT and MRI can be translated to the other one. The translated images can possibly be a good auxiliary in brain diseases diagnosis.

2. Methodology

2.1. Dataset description and preprocessing

Dataset used in this paper is CT and MRI brain scans which is from Kaggle [7]. There are 4972 images in total in the dataset. In addition, the dataset folders are divided into trainA, trainB, testA and testB, and domain A are CT scans (Figure 1) and domain B are MRI scans (Figure 2). In trainA and testA, the size of a single image is 512×512 and the color mode of the image is gray. In trainB and testB, the size of a single image is not uniform, and the color mode of the image is RGB.

The degree of X-ray absorption by organs and tissues is depicted in CT images using various gray scales. Hence, white shadows denote high absorption regions, that is, regions of high density, such as bones, while black shadows denote low absorption regions, that is, regions of low density, such as the lungs. This is similar to the black-and-white images exhibited in X-ray imaging. However, CT images have a better density resolution when compared to X-ray images.

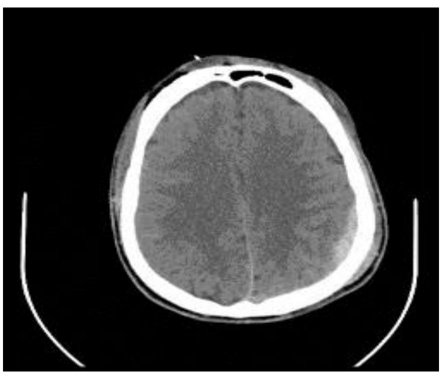

Figure 1. The sample data of CT scans in the collected dataset.

Magnetic field, which is caused by the MRI equipment, will arrange the atoms in brain similarly. Then, the atoms will be moved from their original positions due to the radiation from MRI machine. When the radiation ends, those atoms in brain will move back to their initial places, generating signals. The generated signals will be translated into an image, showing the bodily region under study by a computer. The image can then be displayed in grayscale on a monitor [2]. CT scans are less transparent than MRI scans.

Figure 2. The sample data of MRI scans in the collected dataset.

Data preprocessing is also carried out to help generalize the model. First, this study resized the image height to 1.12 times the original and use bicubic to get higher resolution. Next, the RandomCrop as data augmentation was utilized and the random horizontal flip was carried out. Last, converting the image to a tensor and then normalize the RGB channels with expectation being 0.5 and variance being 0.5.

2.2. CycleGAN

Since the proposal of the Generative Adversarial Networks (GAN), GAN has been widely used in areas such as image translation, image generation, image super-resolution, object detection, video generation, etc. In general, the structure of the GAN is mainly based on the Convolutional Neural Network which has been employed in many computer vision related tasks [8-10]. What followed are all kinds of GAN derivatives, among which CycleGAN shows a wealth of application scenarios and amazing generative capability.

The primary application of CycleGAN is "style transfer." To be more precise, CycleGAN can alter an image's aesthetic while keeping its essential elements. The pix2pix model existed before CycleGAN, but it required paired training data, which is challenging to come by in practice. CycleGAN can use more training data because it doesn't need paired training data. In order to ensure that CycleGAN maintains the original characteristics of an image while converting the style of it, that is, to be "cycle consistent", two sets of generators (G, F) and discriminators (D_X, D_Y) and cycle consistency loss are used to achieve F(G(x)) ≈ x and G(F(y)) ≈ y. In addition, identity loss is used to keep the hue of the picture during the translation process.

First, preparations before training are needed. Generators and discriminators will be initialized or loaded. In addition, the optimizer, learning rate, batch size and normalization will be set. Next, the training process will start. After convolution, domain transferring and deconvolution, generators will generate the corresponding image in the other domain. Discriminators will judge whether the input image is real or fake using convolutional layers (Figure 3). Lastly, loss will be computed and used to update the models.

DOI: 10.54254/2755-2721/8/20230124

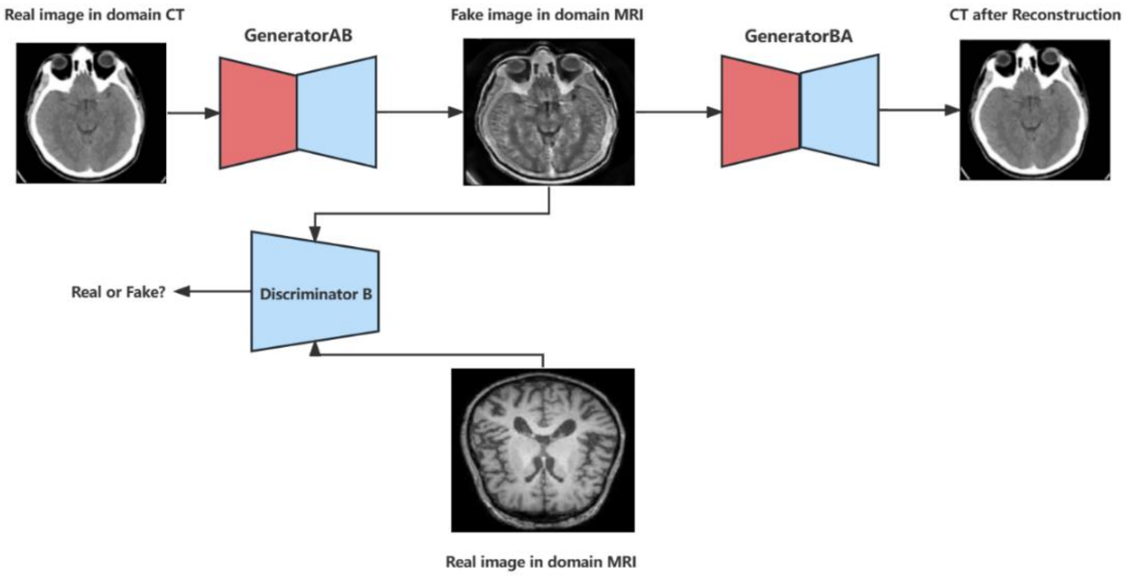

Figure 3. The structure of the model.

2.3. Implementation details

Pytorch is used as the framework of this work. There are 200 epochs for training process. The learning rate is 0.0002 and learning rate will start to decay proportionally at 100 epochs to help achieve better results. Adam is chosen as the optimizer and betas are 0.5 and 0.999. The training batch size is set to 1 to make the model focuses on one picture at a time. For the loss function, the more details about the formulates can be found as follows.

$$\boldsymbol{Loss_{GAN}} = \frac{1}{2}\{\ (D_Y(G(x)) - 1)^2 + \ (D_X(F(y)) - 1)^2\ \} \tag{1}$$

$$\boldsymbol{Loss_{Cyc}} = E_{x\sim P_{data}(x)}[||F(G(x)) - x\ ||_1] + \ E_{y\sim P_{data}(Y)}[||G(F(y)) - y\ ||_1] \tag{2}$$

$$\boldsymbol{Loss_{identity}} = E_{x\sim P_{data}(x)}[||F(x) - x\ ||_1] + \ E_{y\sim P_{data}(Y)}[||G(y) - y\ ||_1] \tag{3}$$

$$\text{The generator loss: } Loss = Loss_{GAN} + 4 * Loss_{cyc} + 8 * Loss_{identity} \tag{4}$$

$$\text{The discriminator loss: } Loss = \frac{1}{2}\{(D(y) - 1)^2 + \left(D(G(x))\right)^2\} \tag{5}$$

3. Result and discussion

Following the training phase, two proficient generators exhibit the capacity to perform translations between CT and MRI scans. The generated images, as depicted in Figure 4, Figure 5, Figure 6 and Figure 7, demonstrate a discernible convergence towards the corresponding domain over the course of the training.

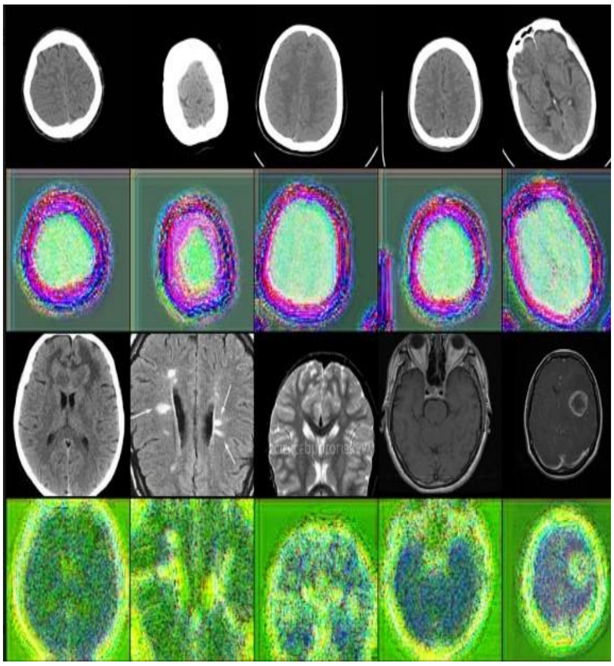

Figure 4. Top to bottom for CT, CT to MRI, MRI and MRI to CT, iters 0.

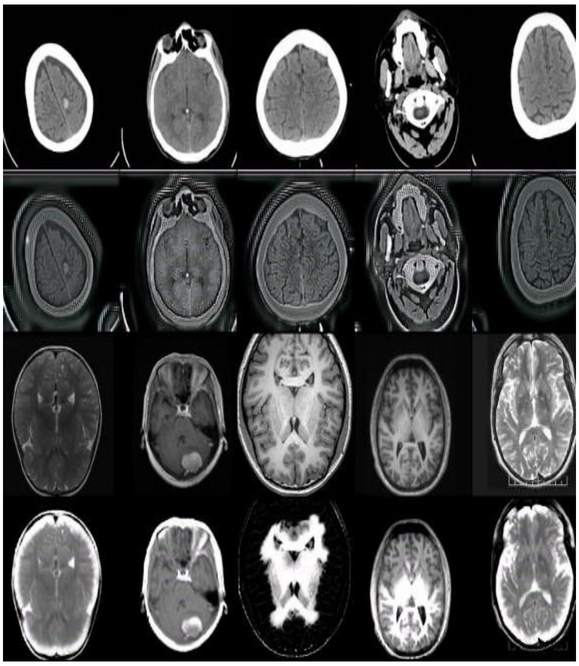

Figure 5. Top to bottom for CT, CT to MRI, MRI and MRI to CT, iters 5000.

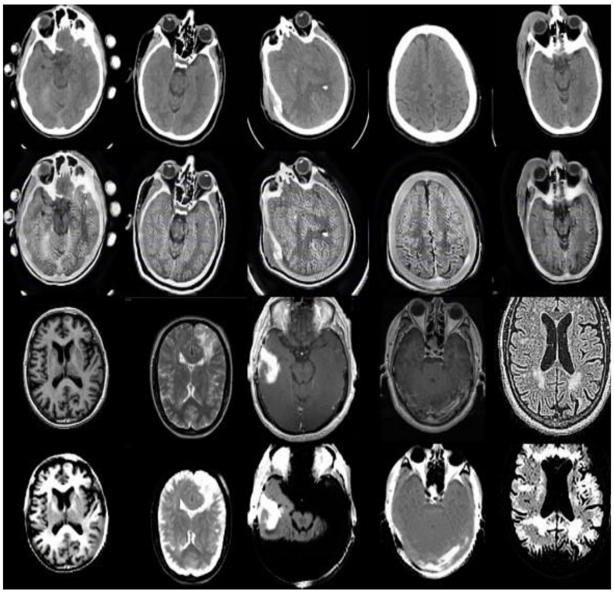

Figure 6. Top to bottom for CT, CT to MRI, MRI and MRI to CT, iters 10000.

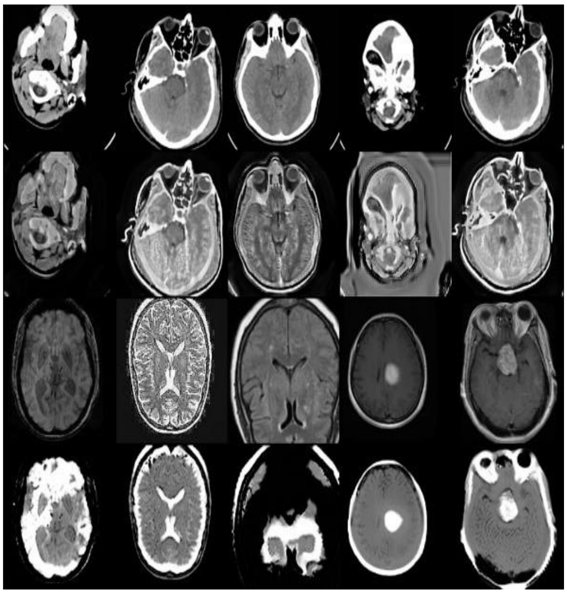

Figure 7. Top to bottom for CT, CT to MRI, MRI and MRI to CT, iters 15000.

Also, after just 10 epochs, the loss of generators and the loss of discriminators are very close to converging, like it is depicted in Figures 8, Figure 9 and Figure 10.

DOI: 10.54254/2755-2721/8/20230124

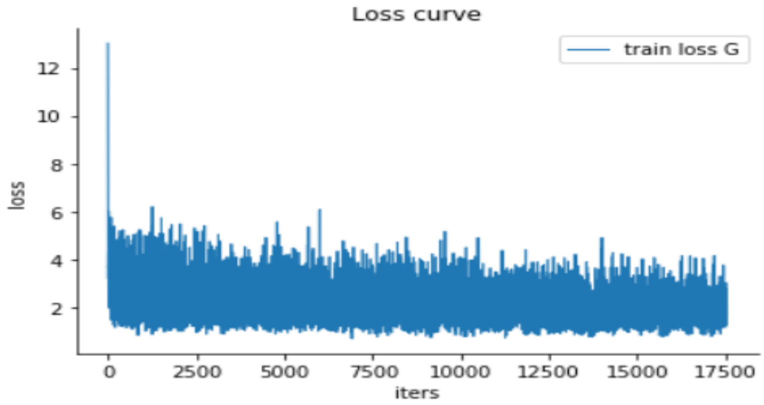

Figure 8. Loss curve of generators.

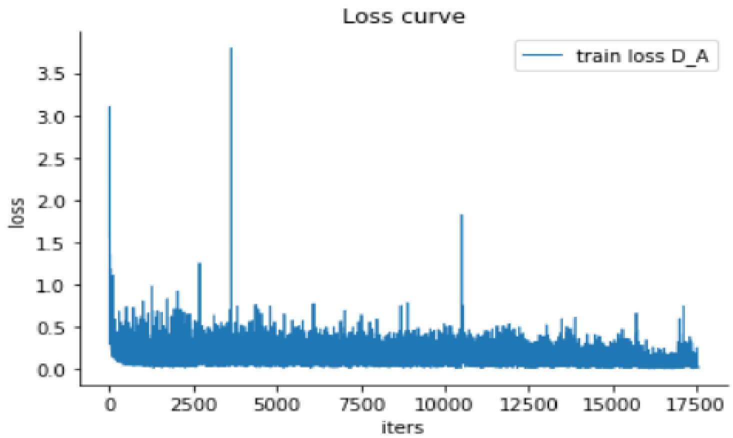

Figure 9. Loss curve of discriminator_A.

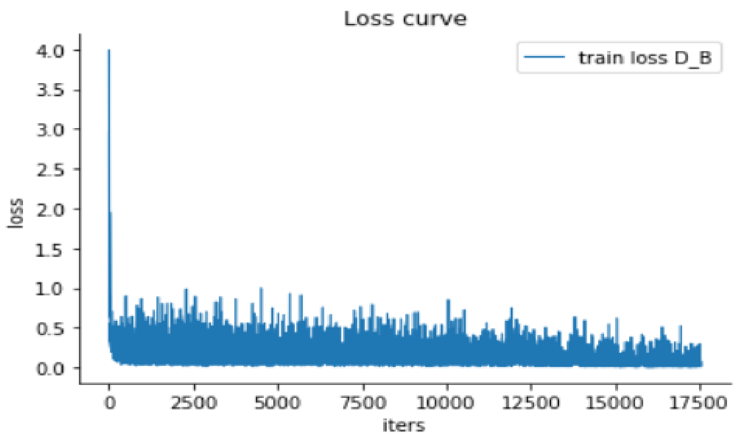

Figure 10. Loss curve of discriminator_B.

Figure 4, Figure 5, Figure 6 and Figure 7 illustrates that the quality of the generated images demonstrates a marked improvement over the course of the training epochs. The $Loss_{GAN}$ ensures that the generated images exhibit a close stylistic resemblance to the target domain. Meanwhile, the $Loss_{cyc}$ guarantees that the main content of the generated images remains consistent with the source domain. Lastly, the $Loss_{identity}$ ensures that the hue of the generated images is retained during the translation process.

Few generated images may not be too “real” due to the missing features of the original scans or the limited training epochs. Also, the $lambda_{cyc}$ and $lambda_{identity}$ can be changed to help the generators perform better. Lastly, Resnet34 is used in the generators, using Resnet50 may result in better performance but higher cost.

4. Conclusion

In this paper, the author proposed unpaired image to image translation between CT scans and MRI scans by using Cycle Consistent Adversarial Networks and dataset of CT and MRI brain scans of human. The author has trained two sets of generator and discriminator with adjusted $Loss_{GAN}$, $Loss_{cyc}$ and $Loss_{identity}$ to keep the translation process “cycle consistent”. Some experiments were carried out to evaluate the proposed method. The results of experiment demonstrated that the models trained with the method in this paper are comparable to or even better than the centralized baseline under various hyperparameter settings. In the future, the author plans to improve the proposed method to let the translated images be a good auxiliary in brain diseases diagnosis and in turn saves plenty of time and money for the patients.

References

[1] JOHNS HPKINS 2023 Computed Tomography (CT) scan https://www.hopkinsmedicine.org/health/treatment-tests-and therapies/computed-tomography-ct-scan
[2] JOHNS HPKINS 2023 Magnetic Resonance Imaging (MRI) https://www.hopkinsmedicine.org/health/treatment-tests-and-therapies/magnetic-resonance-imaging-mri
[3] LundBeck 2023 Understanding brain diseases https://www.lundbeck.com/cn/cn/patients/understanding-brain-diseases
[4] Goodfellow I Pouget-Abadie J Mirza M et al. 2020 Generative adversarial networks Communications of the ACM 63(11) 139-144
[5] Zhu J Y and Park T Isola P et al. 2017 Unpaired image-to-image translation using cycle-consistent adversarial networks Proceedings of the IEEE international conference on computer vision 2223-2232
[6] Khan P et al. 2021 Machine Learning and Deep Learning Approaches for Brain Disease Diagnosis: Principles and Recent Advances IEEE Access vol. 9 pp. 37622-37655 doi: 10.1109/ACCESS.2021.3062484
[7] Kaggle 2020 CT to MRI cgan https://www.kaggle.com/datasets/darren2020/ct-to-mri-cgan
[8] Qiu Y Yang Y Lin Z et al. 2020 Improved denoising autoencoder for maritime image denoising and semantic segmentation of USV China Communication 17(3) 46-57
[9] Yuan F Zhang Z Fang Z 2023 An effective CNN and Transformer complementary network for medical image segmentation Pattern Recognition 136 109228
[10] Zhu F Wang S Li D et al. 2023 Similarity attention-based CNN for robust 3D medical image registration Biomedical Signal Processing and Control 81 104403