# Image-to-Image Translation Based on CycleGAN: From CT to MRI

Chenjie Ni
*School of Artificial Intelligence, Southeast University, Nanjing, China*

Keywords: Computed Tomography, Magnetic Resonance Imaging, Image translation, CycleGAN.

Abstract: Computed Tomography (CT) and Magnetic Resonance Imaging (MRI) have equal importance in routine examinations. However, in some cases, one certain type may not be available due to limitations in condition. Therefore, it is necessary to establish a connection between CT and MRI images. With the idea of image-to-image translation, this study proposes using the Cycle-Consistent Generative Adversarial Networks (CycleGAN) structure to build a mapping between these two kinds of medical images. Through the combination of Resnet Generator as well as Patch Generative Adversarial Networks (PatchGAN) Discriminator, the CycleGAN model is trained bidirectionally to achieve cyclic translation. Both qualitative and quantitative evaluations are implemented to highlight the model's effectiveness in transforming CT or MRI images from either direction to the other. In addition, the CycleGAN model excels particularly in cycle consistency, meaning a realistic recovery of the transformed images. Therefore, this study presents a powerful way for achieving mutual conversion between CT and MRI images, which is especially meaningful to diagnosis with limited information. In addition, this research also suggests the potential of image-to-image translation in medical image processing. Future research directions can be set upon this study to further improve the clarity of images and reduce noise so that the generated results can be truly used for clinical diagnosis.

## 1 INTRODUCTION

CT and MRI are two basic ways of getting information about the diseased region during diagnosis (Kidwell and Amie 2006). Yet each of these two methods has its advantages and limitations. For CT, the advantages lie in its short examination time, low cost, and wider application range (Angela and Müller 2011). However, CT has radiation and is not suitable for pregnant women and children. Meanwhile, the contrast resolution of CT is relatively low. Concerning MRI, it is non-invasive to the human body, with diverse parameters and the freedom to choose the orientation for imaging (Beek and Eric 2008). But it also brings drawbacks such as long scanning time, large noise, and expensive equipment. In addition, due to the strong magnetic field during operation, it cannot be used for patients with ferromagnetic substances in their bodies. Considering the equal importance of these two methods, it is necessary to establish a connection between CT and MRI images to provide more information for constrained diagnosis.

Many studies have proposed meaningful methods to build this link or create new images based on existing information. For example, Han Xiao attempted to reconstruct CT images from MRI by using a deep convolutional neural network (Han 2017). Toda Ryo et al. attempted to use semi-conditional Information Generative Adversarial Networks (InfoGAN) to synthesize CT images of certain types of lung cancer (Toda et al 2021). Alrashedy, Halima Hamid N. et al. proposed Brain Generative Adversarial Networks (BrainGAN), combining Generative Adversarial Networks (GAN) architectures with Convolutional Neural Network (CNN) models to generate MRI images (Alrashedy et al 2022). Kwon Gihyun et al. used auto-encoding generative adversarial networks to generate 3D brain MRI images (Kwon et al 2019). However, the above-mentioned studies as well as most of the existing methods can only achieve unidirectional image synthesis like synthesizing MRI images with CT images. This deficiency has put some constraints on doctors to get full information on the patients. Yet in recent years, the task of image-to-image translation has been broadly discussed, bringing some new ideas for connecting CT and MRI images (Isola et al 2017). Using a training set of aligned image pairs, image-to-image translation aims to learn the mapping between an input image and an output image. While there have already been a lot of existing applications of image



Ni, C.
Image-to-Image Translation Based on CycleGAN: From CT to MRI.
DOI: 10.5220/0012799400003885

In *Proceedings of the 1st International Conference on Data Analysis and Machine Learning (DAML 2023)*, pages 229-233
ISBN: 978-989-758-705-4

translation (e.g. Chen et al 2021), little attention was paid to the field of medicine. The idea of image translation is very suitable for constructing a bidirectional change pathway between CT and MRI.

Given the facts above, the main objective of this study is to enable a free switch between CT and MRI images. Specifically, the ct2mri dataset is preprocessed first, including the partition of the training set and test set, as well as resizing images. Second, CycleGAN structure is introduced to achieve this translation process (Zhu et al 2017). CycleGAN is a powerful model that can learn to translate images between different styles without paired examples. This independence of paired images is especially helpful to the connection of CT and MRI because their images always vary greatly in properties. The process of CycleGAN can be concluded as training one pair of generator and discriminator for each direction. For valid image translation, constraints on loss are added to ensure consistent content with different styles. Through pairs of generator and discriminator in CycleGAN, features of the images are extracted and reorganized to construct mappings between two domains. Thus, a direct connection between images of different domains is learned, allowing the model to convert any related images into each other. The experimental results demonstrate a satisfying performance in the bidirectional translation of CT and MRI images. This kind of translation model can help doctors quickly and effectively obtain the necessary information when conditions are limited, such as when one of the medical images is unavailable due to patient reasons.

# 2 METHODOLOGY

## 2.1 Dataset Description and Preprocessing

The dataset used in this study is sourced from Kaggle called CT and MRI brain scans (CT and MRI brain scans 2020). It contains a total number of 4974 images of the results of CT and MRI brain scans. The size of these images is not uniform, for the training process of CycleGAN is unpaired, which means it is not affected by whether the image size corresponds or not. The images have been pre-adjusted to make sure that the results of brain scans are in the center and take up approximately even space in every image.

The goal of the experiment is to learn a map between CT and MRI images in the dataset. To be loaded for use in a CycleGAN implementation for image-to-image translation, all of the CT and MRI brain images are organized into a directory structure and labeled as A and B respectively, with 2486 CT images for A and 2488 MRI images for B. Also, for model evaluation, a training set and a testing set are created from each of the parts with a ratio of 70% to 30%. Fig. 1 displays a typical pair of CT and MRI images from this collection.

(a) CT image in dataset (b) MRI image in dataset

Figure 1: An illustration of a CT and MRI image from the dataset of CT and MRI brain scans (Picture credit: Original).

## 2.2 Proposed Approach

The core issue of this proposed approach for CT and MRI image translation lies in constructing a complete structure of CycleGAN. This involves choosing proper network structure for both generator and discriminator in each direction, as well as a powerful loss function to drive the entire training process. For the Discriminator, it is chosen to have a PatchGAN structure with a patch size of 70x70; For the Generator, several Resnet Blocks are utilized to build the whole network. With regard to the loss function, GAN loss and cycle consistent loss are combined to ensure better performance. Fig. 2 illustrates the structure of the system.

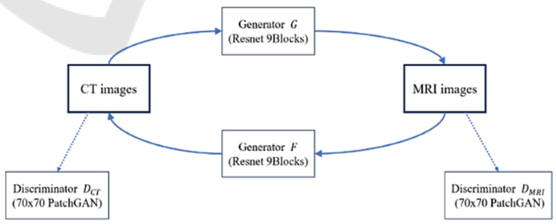


Figure 2: Composition of the model (Picture credit: Original).

### 2.2.1 ResNet Generator

ResNet is a well-known convolutional neural network with efficient performance regarding vanishing or exploding gradient problems. Resnet Block takes a step further. It draws on the core ideas of Resnet, which is "skip connection", and generalizes into a universal neural network layer structure to have two convolutional layers and a skip connection. This method largely prevents the degradation of deep neural networks. The generators in this study are designed to be mainly made up of 9 Resnet Blocks, with reflection padding inside convolutional layers to preserve edge information of images. In addition, downsampling is implemented prior to the input of Resnet Blocks to reduce subsequent computational complexity, together with upsampling after Resnet Blocks to recover the image size. In the last convolutional layer, 64 generator filters with size 7x7 are created to contain the generated information. All of the operations above contribute to achieving better results of feature extraction, pushing the generator to generate more realistic images as well as enhancing robustness. Fig. 3 shows the basic sequence of the generator structure.

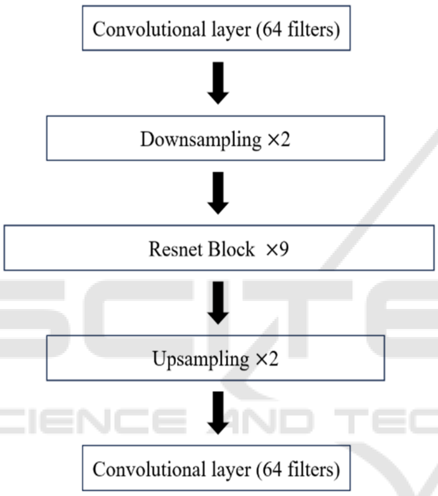


Figure 3: Sequential layer structure of the Resnet generator (Picture credit: Original).

### 2.2.2 PatchGAN Discriminator

Initially, PatchGAN is proposed to solve ambiguous generation in L2 or L1 loss cases. Instead of dealing with the whole image at one time, PatchGAN focuses on local image patches step by step and penalizes structure at the scale of patches. Convolutationally scanning across the image, this discriminator aims to decide whether each patch is fake or not, and finally collect all responses to provide the ultimate output. This kind of patch-based structure has fewer parameters than a discriminator dealing with full images, which can greatly accelerate the training process. Besides, discriminator with PatchGAN structure can be applied to arbitrarily-sized images process, providing great convenience for this study. According to the suggestions in the original paper, the patch size in this study is set to 70x70 to get an optimal performance. In addition, for network architecture, the PatchGAN discriminator is constructed through 3 main convolutional layers with an increasing number of filters and converges to one output channel by performing convolution processing again in the end to get the predicted results.

### 2.2.3 Loss function

It is critical to choose the right loss function in the training of deep learning models, especially in generational ones. As for this image-to-image translation task, the full objective loss function mainly consists of two terms: The first is adversarial losses. Here an improved version of vanilla GAN losses proposed in Zhu et al's study is implemented. It is called LSGAN loss:

$$l_{LS}(G_1, D_Y) = \mathbb{E}_{y\sim p_{data}(y)}[(D_Y(y) - 1)^2] + \mathbb{E}_{x\sim p_{data}(x)}\left[D_Y(G_1(x))^2\right] \quad (1)$$

The above formula illustrates the form of LSGAN loss, where $G$ represents the generator mapping from $X$ to $Y$, while $D_Y$ denotes the discriminator on domain $Y$. LSGAN loss substitutes a least square loss for the original negative log-likelihood, which brings a more stable training as well as better performance. For the opposite direction, there is also a similar function $l_{LS}(G_2, D_X)$.

The second part is defined as cycle consistency loss:

$$l_{cyc}(G_1, G_2) = \mathbb{E}_{x\sim p_{data}(x)}\left[\|G_2(G_1(x)) - x\|_1\right] + \mathbb{E}_{y\sim p_{data}(y)}\left[\|G_1(G_2(x)) - y\|_1\right] \quad (2)$$

where $G$ and $F$ represent two generators. The cycle consistency loss guarantees that the cycle of image translation is able to bring the input back to the original image as similarly as possible. Then the full objective is established through a combination of the following forms:

$$l(G_1, G_2, D_X, D_Y) = l_{LS}(G_1, D_Y) + l_{LS}(G_2, D_X) + \lambda l_{cyc}(G_1, G_2) \quad (3)$$

where $\lambda$ controls the relative weight of two different types of loss. This parameter was determined through hyperparameter tuning to ensure optimal performance. To prevent overfitting, instance normalization is implemented. In comparison to traditional batch normalization, instance normalization performs better in image translation because, for this type of task, each pixel of the input sample is crucial to the training process.

### 2.3 Implementation Details

In the training process of the suggested model, several important aspects are highlighted. Firstly, Adam is chosen to be the optimizer of all generators and discriminators in CycleGAN because of its satisfying performance concerning gradient descent in high-dimensional spaces. Speaking of hyperparameters, in the first 50 training epochs, the learning rate is fixed at 0.0002, and in the subsequent 50 training epochs, it decreases linearly to zero. This can make sure that the model learns more at the beginning, and keeps the parameters almost unchanged near the end to reduce the probability of overfitting. The momentum term of Adam is set to be 0.5. Limited by equipment RTX3060, the batch size during training is constrained to 2, and the model trains for a total of 100 epochs.

## 3 RESULTS AND DISCUSSION

As a generative model, evaluation of the performance usually focuses on observing the generation results through the test set on the trained model. Specifically, the results of this study will be discussed through the method of visualization as well as generation accuracy. For testing and evaluation, 744 unpaired CT and MRI images are prepared to give translation. Here only the translation results of the model from CT images to MRI images and back to CT will be shown. It is because, for CycleGAN, the results of image translation from both two directions (which is CT-MRI-CT and MRI-CT-MRI) should be equivalent in performance.

### 3.1 Visualization Analysis

Some typical test outputs are selected to be demonstrated in Fig. 4 below. From left to right, the generated MRI image, original MRI image, restored CT image, and original CT image are sequentially displayed in columns.

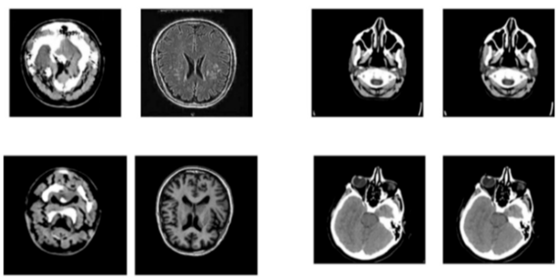

Figure 4: Typical outputs of the constructed CycleGAN (Picture credit: Original).

It can be intuitively seen from Fig. 4 that the CycleGAN model constructed in this study effectively maps the given CT images into MRI ones, with necessary details as well as correct contour. Thanks to the delicate structure of the Resnet Generator, the CycleGAN model has such a strong feature extraction ability that it can rebuild most of the detailed information of the real images. Besides, PatchGAN Discriminator enhanced the refinement of the generator as well by serving as an adversarial part, forcing the generator to pay more attention to details. Though defects can be observed such as there is still residual information from the original image, it is caused by the nature of CycleGAN, which tends to preserve the content. Nevertheless, the CycleGAN model still establishes a valid connection between CT and MRI images from a visual perspective.

At the same time, the model almost perfectly recovers the transformed images back into the original ones. This means that the CycleGAN model in this study has a strong cycle consistency, which should be attributed to the powerful constraint of cycle consistency loss in the loss function on the generation of image content. In addition, the results imply that the parameter $\lambda$ is not obtained too morbidly to cause failures in image generation, proving a success in hyperparameter tuning.

### 3.2 Generation Accuracy

In this work, the structural similarity index measure (SSIM) is utilized to assess the trained model's generation accuracy. The SSIM metric extracts three key features from an image: brightness, contrast, and structure, which are used to measure the similarity between two given images. Implementing this metric through the outputs of the test set, the model gets an average score of 0.4038 on the generated MRI images and 0.9642 on the recovery of the translated images. SSIM metric provides a quantified summary of the performance of the CycleGAN model. Combined with the visualization results, it can be concluded that the CycleGAN model has no problems in generating most of the image details, but still faces challenges in terms of image brightness and clarity, which is caused by CycleGAN's property of keeping the original structure information as is discussed before. This observation raises the necessity for some structural alteration on the CycleGAN model to eliminate excess information.

# 4 CONCLUSION

This article introduces an approach employing a CycleGAN architecture to decipher the intricate mapping relationship between CT and MRI images, with a curated dataset of brain scans serving as the primary data source. The model exhibits remarkable performance in feature analysis and extraction, leveraging the Resnet and PatchGAN architectures for its generator and discriminator components. This choice empowers the model to excel in capturing salient features and fostering discriminative capabilities.

An extensive series of experiments has been meticulously conducted to evaluate the proposed methodology, employing a range of qualitative and quantitative metrics. The results garnered from these experiments on the CT and MRI brain scan dataset are highly promising. The CycleGAN model successfully forges a meaningful connection between CT and MRI images, preserving intricate details and structural integrity. Moreover, the model demonstrates robust cycle consistency, affirmed through both visual inspection and the SSIM.

The model's remarkable image generation capabilities can be attributed to ResNet's ability to retain vital input information and PatchGAN's effectiveness in scrutinizing generated images at the patch level. It is important to acknowledge that future research endeavors will be primarily dedicated to refining the model's architecture to address any identified limitations. Additionally, the exploration of a diverse range of models for enhancing performance in the domain of image translation will remain a focal point in upcoming research pursuits. This commitment to continuous improvement underscores the model's potential contributions to the field of medical imaging.

# REFERENCES

S. Kidwell, W. Amie, "Imaging of the brain and cerebral vasculature in patients with suspected stroke advantages and disadvantages of CT and MRI," Current neurology and neuroscience reports, vol 6, 2006, pp. 9-16.

C. Angela, P. Müller, "Introduction to computed tomography," Kgs. Lyngby: DTU Mechanical Engineering, 2011

J. R. Beek, A. Eric, "Hoffman Functional imaging: CT and MRI Clinics in chest medicine, vol. 29, 2008, pp. 195-216

X. Han, "MR - based synthetic CT generation using a deep convolutional neural network method Medical physics, vol. 44, 2017, pp. 1408-1419

R. Toda et al, "Synthetic CT image generation of shape-controlled lung cancer using semi-conditional InfoGAN and its applicability for type classification," International Journal of Computer Assisted Radiology and Surgery, vol. 16, 2021, pp. 241-251

H. Alrashedy, N. Hamid et al, "BrainGAN: brain MRI image generation and classification framework using GAN architectures and CNN models," Sensors, vol. 22, 2022, pp. 4297

G. Kwon, H. Chihye, D. Kim, "Generation of 3D brain MRI using auto-encoding generative adversarial networks," International Conference on Medical Image Computing and Computer-Assisted Intervention Cham, Springer, 2019

P. Isola et al, "Image-to-image translation with conditional adversarial networks," Proceedings of the IEEE conference on computer vision and pattern recognition, 2017

Z. Chen et al. "Semantic segmentation for partially occluded apple trees based on deep learning," Computers and Electronics in Agriculture, vol. 181, 2021, p. 105952

J. Y. Zhu et al, "Unpaired image-to-image translation using cycle-consistent adversarial networks," Proceedings of the IEEE international conference on computer vision, 2017

CT and MRI brain scans https://www.kaggle.com/datasets/darren2020/ct-to-mri-cgan