T1 – T2 WEIGHTED SYNTHETIC IMAGE GENERATION USING CYCLE GAN

PROJECT REPORT

Wahid Alam
Department of Biomedical Engineering
The University of Iowa
iowa city, IA 52242
mohammadwahidul-alam@uiowa.edu

Subin Erattakulangara
Department of Biomedical Engineering Engineering
The University of Iowa
iowa city, IA 52242
subin-erattakulangara@uiowa.edu

August 23, 2022

ABSTRACT

MRI is a non-invasive imaging technology with plethora of possibilities in disease detection, diagnosis, and treatment monitoring. The most common MRI sequences are T1-weighted and T2-weighted scans both of these contrast weighted techniques have its own advantage and limitations. From machine learning point of view T2 weighted images have better contrast on tissue regions. But T2 contrast weighted method is highly susceptible to motion artifacts due to its longer readout time. In this project we are trying to solve that issue by synthetically generating T2 weighted images from T1 weighted images. We have used two types of datasets to understand the performance of the network. The network was able to perform well on static MRI images. One of the drawbacks found was its bias towards special positions on training set.

Keywords Cycle GAN · T1 · T2

1 Introduction

Magnetic resonance imaging or MRI is a non-invasive imaging technology that produces three dimensional detailed anatomical images. It is often used for disease detection, diagnosis, and treatment monitoring. It is based on sophisticated technology that excites and detects the change in the direction of the rotational axis of protons found in the water that makes up living tissues. The method is well suited for imaging non-bony part tissues or soft tissues. Compared to Computed tomography MRI relatively safe since It does not generate any ionizing radiation like x-rays. The brain, spinal cord, and nerves, as well as muscles, ligaments, and tendons are seen much more clearly with MRI than with regular x-rays and CT; for this reason, MRI is often used to image knee and shoulder injuries. In the brain, MRI can differentiate between white matter and grey matter and can also be used to diagnose aneurysms and tumors. Because MRI does not use x-rays or other radiation, it is the imaging modality of choice when frequent imaging is required for diagnosis or therapy, especially in the brain. However, MRI is more expensive than x-ray imaging or CT scanning.

The most common MRI sequences are T1-weighted and T2-weighted scans. T1-weighted images are produced by using short TE(echo time) and TR(Reputation time). The TE and TR can be considered as a variable to switch between T1 and T2 images. The contrast and brightness of the image are predominately determined by T1 properties of tissue. Conversely, T2-weighted images are produced by using longer TE and TR times. In these images, the contrast and brightness are predominately determined by the T2 properties of tissue. Using these T1 and T2 protocols we can highlight certain tissue properties which are specific for diagnosis.

Machine learning has been aggressively applied to the field of magnetic resonance imaging in multiple ways. People have found its application ranging from data acquisition to analysis of MRI data. Machine learning has been in the forefront of helping clinicians to diagnose multiple diseases using MRI data. From MRI we can acquire data in multiple forms like static images, Dynamic (2D +time), 3D data and 4D data (3D + time). The ability of MRI to generate such kinds of data is helpful for machine learning since it's a data driven modality. Even though there is huge possibility to generate varies types of data MRI have some limitations and trade-offs on acquiring these. These trade-offs related to T1 and T2 protocols introduce certain constraints for collecting data from the hardware. We are unable to go beyond these physical constraints and generate data. But If we can go beyond these constraints there is a huge benefit for machine learning algorithms. In this project we are trying to address this problem by generating synthetic data which we cannon practically generate using MRI machines due to hardware limitations.

2 Problem definition

As we know there are certain advantages for T2 and T1 weighted imaging protocols, But In some cases It's hard to generate those specific protocol data. For example, if we are trying to take a dynamic MRI scan; T1 is the default protocol used since it can capture data at a higher frame rate compared to T2. From the machine learning standpoint T2 data have higher advantage over T1 since the images acquired using T2 protocol have a greater number of features embedded in the soft tissues. Since T2 protocol is a slow process, in practical it is not possible to acquire data of a moving object.
Due to the abundance of features on T2 weighted data machine learning algorithms will have higher performance on T2 compared to T1 weighted images. Due to the problem mentioned above we have a practical limitation on generating sufficient data to train a network for classification/segmentation on T2 weighted images. Since there is a possibility to generate dynamic data / video from T1 weighted imaging in an easy and fast manner, in this project we are trying to solve that limitation by synthetically generating T2 weighted images from T1 weighted data [1]. We are also trying the possibility of generating a dynamic T2 weighted sequence which is practically impossible due to the hardware constraints[2].

3 Dataset

The problem statement requires T1 and T2 images from an MRI dataset. So, we have choses MICCAI BraTS dataset. The dataset consists of T2, T1 and Flair dataset. All the data is acquired using clinically acquired 3T multimodal MRI scans.

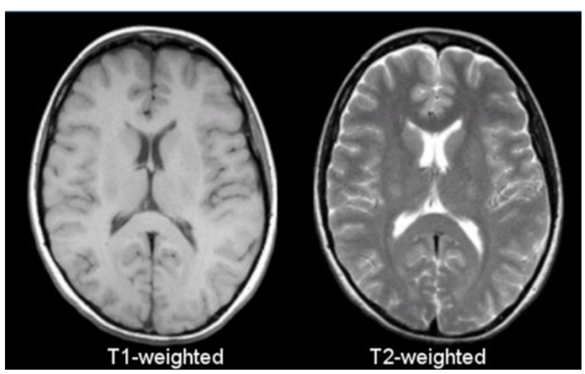

Figure 1: T1-T2 transformation pair.

The dataset is in a acquired as 3D data with multiple slices. For the current problem we won't be needing all the slices in the 3D data. So, the dataset processing pipeline will select few images in a range and will be saved separately to a folder.
Brats Dataset - `https://www.med.upenn.edu/sbia/brats2018/data.html`

The second dataset we have used here is airway dataset. The collection contains images from in house T1 dataset creating using University of Iowa research MRI scanner and T2 images from open-source dataset [3]. The in-house dataset is a dynamic speech dataset which frames where subject is speaking. All our in-house data were acquired on a GE 3Tesla scanner with a fast GRE sequence (midsagittal plane; spatial resolution: $2.7mm^2$; $\sim 6frames/second$, FOV: $20 \times 20cm^2$).

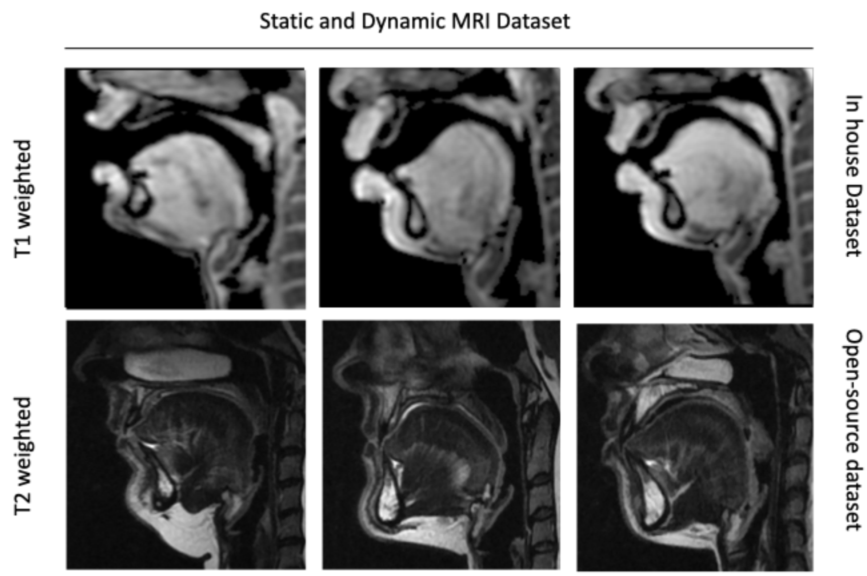

Figure 2: shows some examples from airway MRI dataset we used.

4 Methods

In this section, we introduce the proposed use of cycle GAN for creating synthetic T2 weighted contrast images from T1 Contrast MRI images. Our method contains use of open-source Brats brain dataset and In-house Upper airway MRI dataset for training and testing purposes. The model architecture is comprised of two generator models: one generator (Generator-A) for generating images for the first domain (Domain-A) and the second generator (Generator-B) for generating images for the second domain (Domain-B).

4.1 Preprocessing

We have considered two set of preprocessing for both the data. First the Brats contain volumetric MRI data in T1, T2 and FLAIR contrast. For this problem we are only considering T1 & T2 contrast. The data is arranged in a format where T1 & T2 contrast data is stored in folder with subject ID. To separate the data into different folders we implemented a loop to read and separate T1 & T2 volumetric files. Later, we read the data separately and slice the volumetric data in axial direction. Since the number of slices are constant amount all the subjects, we have selected 20 slices from one subject. After slicing we have 1400 T1 images and 2500 T2 images. This data is divided into train and test sets. Both T1 and T2 test set have 45 images and rest of the images are used for training.

4.2 Architecture

The model architecture consists of two generators (G1 & G2) where G1 generate image for T2 weighted and G2 generating T1 weighted images. The mapping of generators is given by this equation $G_1 : T_2 \rightarrow T_1$, $G_2 : T_1 \rightarrow T_2$. These generator models create the image translation where G1 takes and image from T1 domain and G2 takes an input from T2 domain. Each of these generators have individual discriminators (D1 & D2). The discriminator D1 takes an image from G1 and predicts if it's real or fake. The second model takes images from D2 and tries to generate the same prediction. Both generators and Discriminators are trained on an adversarial sum process like normal GAN models. In this process the generator learns to better fake images and discriminator to better understand the differences. Together the model find equilibrium in process.

Additional to this the generator models are regularized by adding generated images as input to the next generator model and compared the output to original images. Passing images though both the generators are called a cycle. Together

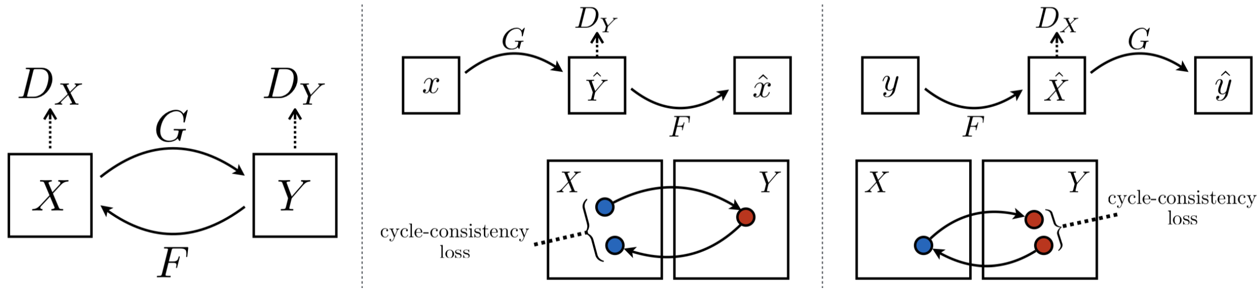

Figure 3: (a) Our model contains two mapping functions $G : X \rightarrow Y$ and $F : Y \rightarrow X$, and associated adversarial discriminators D_Y and D_X. D_Y encourages G to translate X into outputs indistinguishable from domain Y, and vice versa for D_X and F. To further regularize the mappings, we introduce two *cycle consistency losses* that capture the intuition that if we translate from one domain to the other and back again we should arrive at where we started; (b) forward cycle consistency loss: $x \rightarrow G(x) \rightarrow F(G(x)) \approx x$, and (c) backward cycle-consistency loss: $y \rightarrow F(y) \rightarrow G(F(y)) \approx y$.

each generator models are trained to do better reproduce the original image refereed as cycle consistency. Figure 3 shows the relationship between the discriminator and generators in our cycle GAN architecture.

$$\mathcal{L}_{cyc}(G,F) = \mathbb{E}_{x\sim\mathcal{P}_{data}(x)}[||F(G(x)) - x||_1] + \mathbb{E}_{y\sim\mathcal{P}_{data}(y)}[||G(F(y)) - y||_1] \tag{1}$$

$$\mathcal{L}_{GAN}(G, D_Y, X, Y) = \mathbb{E}_{y\sim\mathcal{P}_{data}(y)}[\log D_Y(y)] + \mathbb{E}_{x\sim\mathcal{P}_{data}(x)}[\log(1 - D_Y(G(x)))] \tag{2}$$

One of the specialties of this network the use of adversarial and cycle consistency loss. Adversarial loss are applied to both mapping functions. For mapping function gx xg the objective is expressed as $Eq(1)$. where G tries to generate images $G(x)$ that look similar to images from domain Y , while D_Y aims to distinguish between translated samples $G(x)$ and real samples y. G aims to minimize this objective against an adversary D that tries to maximize it. Similar way we have cycle consistency loss which is a way to incentivize the combination of forward consistency and backward consistency $Eq(2)$.

4.3 Training

The network is built upon Keras which is a higher-level API of TensorFlow. The code is modularized to into multiple folders for ease of customization. The training process is same for both Brats dataset and inhouse airway images. Since the training process for cycle GAN is computationally intensive, we have limited the number of epochs based on the size of the dataset. For Brats Brain MRI images which have a relatively large dataset (~ 4000 images) is trained for 100 epochs, in house upper airway MRI dataset which is 40 images strong is trained for 400 epochs. Both the training is conducted on argon high performance computing cluster.

5 Results and Discussion

We have performed the experiments on Dynamic & Static MRI images. The experiment involved two type of input data and some variations in number of epochs and dataset size. This type of variations was introduced to study the characteristics of the GAN network on MRI data. Table 1 provides us the experimental setup we used, and the results based on visual observation. From this table we can summarize that the quality of output images has a very good correlation with the number of epochs. Increasing the dataset size by huge number and average epoch number weren't able to generate good set of results compared to very small dataset size and higher number of epochs.

Table 1: Experimental setup parameters and observation results

Data	Number of images	Number of epochs	Visual Analysis Result
Airway	67	10	Low
Brain MRI	67	10	Low
Airway	67	100	Low
Brain MRI	67	100	Average
Airway	67	400	Good
Brain MRI	67	400	Better
Brain MRI	4000	100	Average

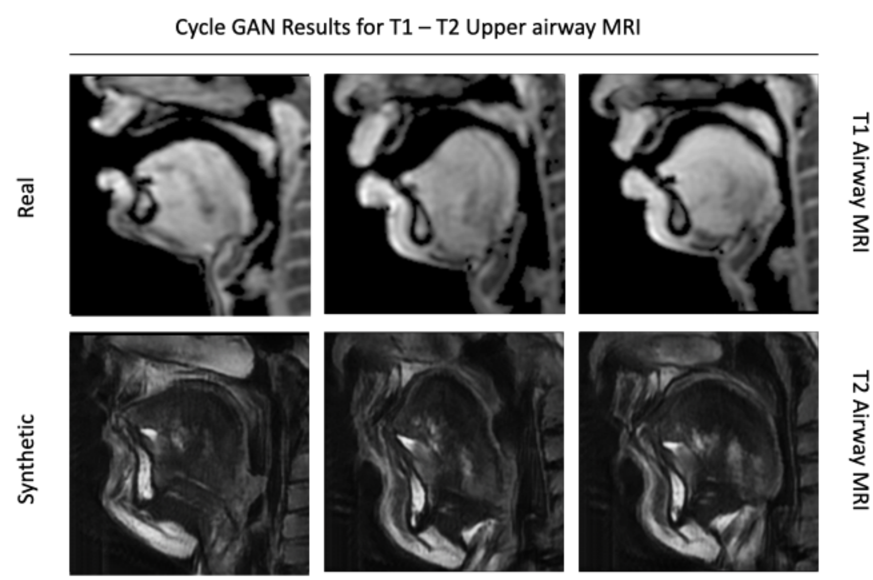

Figure 4: represents the output from T2 & T1 airway images. You can observe here that the cycle GAN network has produced a good result which look comparable to the actual T2 airway MRI images. In this result we would like you to focus on a specific articulator called velum. All the output has velum at resting posture but none of the T1 images uses to test the network had velum in resting posture. The issue we think happening here is that the cycle GAN network is getting biased towards the special position of velum in T2 airway MRI training set. All the images in training set had T2 airway velum in the resting posture.

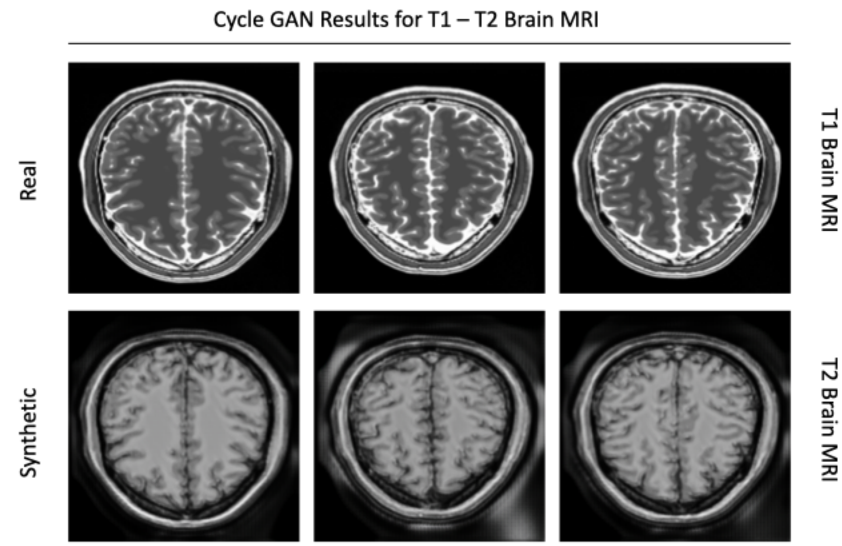

Figure 5: shows us the cycle GAN output and input images used for Brain MRI. Compared to the airway dataset Brain MRI dataset is a static MRI dataset. On visual observation we can appreciate that the network is able to create a good result on the static brain MRI dataset.

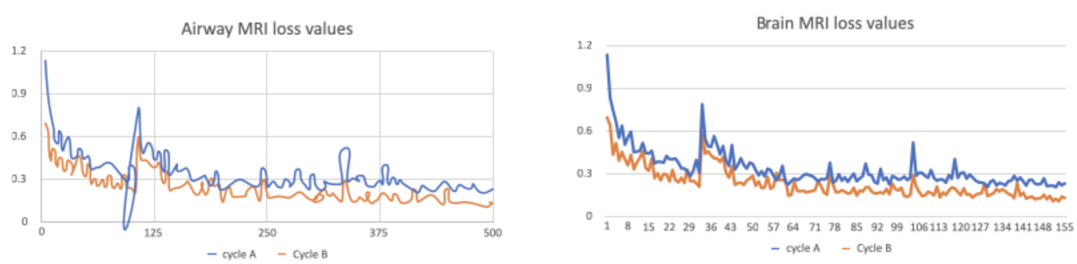

Figure 6: represents the relationship between loss values and the number of epochs for both the dataset. We have only chosen few values from the entire list of loss values. Here we can appreciate the decrease in loss values with increase in the number of epochs.

6 Conclusion

The research we conducted here was not just to create a synthetic T2 data from T1 data but also to understand the characteristics of network on medical images like relationship with input data and other parameters. In here we have used two types of data i.e., Static & Dynamic. We have observed that the network is performing excellent of static images and average on dynamic MRI data. The reason for this observation may be the presence of moving elements in the images. We also experimented the dataset size & number of epochs. Based on results obtained we summarize that the network has better performance on increasing the number of epochs compared to increase in dataset size. One of the reasons may be affecting this observation is that Brain MRI images are fairly look similar for different subjects compared to natural images that are used. This lack of variation in dataset may be the cause of no significant improvement in network performance on increasing dataset size. One of the drawbacks we found on using airway images is that the network tent to preserve the special position of velum on the training images.

References

[1] Bauer, D. F., Russ, T., Waldkirch, B. I., Segars, W. P., Schad, L. R., Zöllner, F. G., & Golla, A. K. (2020). Generation of Multimodal Ground Truth Datasets for Abdominal Medical Image Registration Using CycleGAN. arXiv preprint arXiv:2012.01582.

[2] Hiasa, Y., Otake, Y., Takao, M., Matsuoka, T., Takashima, K., Carass, A., ... & Sato, Y. (2018, September). Cross-modality image synthesis from unpaired data using CycleGAN. In International workshop on simulation and synthesis in medical imaging (pp. 31-41). Springer, Cham.

[3] Lim, Yongwan, Asterios Toutios, Yannick Bliesener, Ye Tian, Sajan Goud Lingala, Colin Vaz, Tanner Sorensen et al. "A multispeaker dataset of raw and reconstructed speech production real-time MRI video and 3D volumetric images." arXiv preprint arXiv:2102.07896 (2021).