Paper Seminar Report

Cross Modality Medical Image Synthesis and Registration through Machine Learning

600.456/656 Computer Integrated Surgery II

Ping-Cheng Ku (pku1@jh.edu)

Apr 2021

1 Project summary

Magnetic Resonance Imaging (MRI) is a commonly used imaging technology for the diagnosis of the hip osteonecrosis, which is a disease resulting in the death of bone cells. Core decompression is a commonly used surgical method for the removal of the ostenecrotic tissue (dead bone) from the femoral head.

In order to remove necrotic tissue in the femoral head, surgeons rely on preoperative MR scans for tool trajectory planning and take intraoperative X-ray shots to monitor the procedure. However, there are currently no simple method that could easily translate the tool trajectories annotated on the MR scans to the intraoperative X-ray shots. Thus, the motivation of this project is to propose machine learning based algorithms that registers the paths between the two image modalities.

Figure 1: Illustration of the annotated drilling path in both MRI (left) and X-ray (right) during core decompression surgery [1]

2 Paper Selection

The paper – **Cross-modality image synthesis from unpaired data using CycleGAN: Effects of gradient consistency loss and training data size** – is selected for the relevance to the project. This project focuses on cycleGAN training in order to generated synthesized CT and MR images, while in our project the gneration of synthesized CT from MR is a crucial step.

2.1 Introduction and Background

The authors first introduces the two main imaging modalities used in orthopedic procedures: Computed tomography (CT) and magnetic resonance imaging (MRI). The authors explained that CT images excels at identifying bone boundaries while MRI images work better identifying muscle structures. The author then describes the need of MR-to-CT synthesis in order to delineate bone structures directly in MRI to reduce patients' radiation exposure while taking CT shots.

The authors then introduces image synthesis techniques available, including patch based learning [2] and deep networks like convolutional neural networks (CNN) and generative adversarial networks (GAN). Conventional approaches often requires paired images as training dataset, which limited the applications as medical images are harder to acquire. The introduction of cycleGAN [3], however, does not required paired images, and the authors

decided to perform MR-to-CT synthesis base on the cycleGAN network. The authors made some modifications to the network by introducing gradient consistency (GC) loss, which encourages edge alignment between images in the two different modalities.

2.2 Method

The authors collected two sets of dataset: a MRI dataset consisting of 302 unlabeled volumes and CT dataset consisting of 613 unlabeled, and 20 labeled volumes which are associated with manual segmentation labels of 19 muscles around hip and thigh, pelvis, femur and sacrum bone as shown in Fig.2.

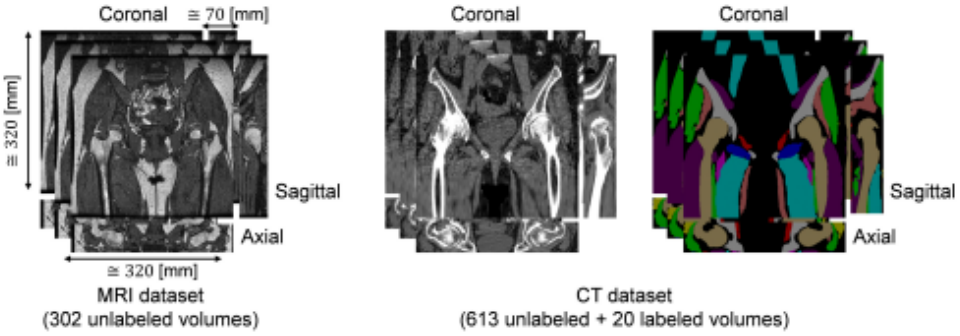


Figure 2: Training datasets used in [4]

The cycleGAN workflow the authors adopted is illustrated in Fig.3 below. The main structure of the proposed network by the authors remains the same as the cycleGAN network introduced by Zhu et al. [3], which includes two generator models (G_{CT}, G_{MR}) and two discriminator models (D_{CT}, D_{MR}).

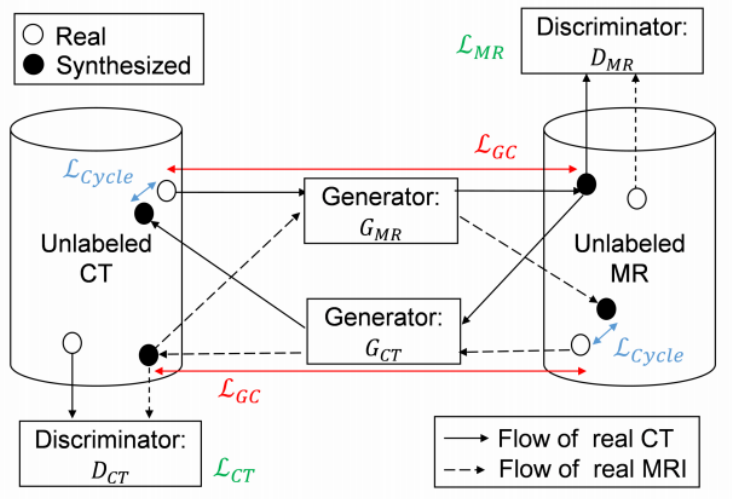


Figure 3: Workflow of the proposed method in[4]

The loss functions corresponding to the model includes the generative adversarial loss ($\mathcal{L}_{CT}$, $\mathcal{L}_{MR}$), cycle-consistency loss ($\mathcal{L}_{Cycle}$), and gradient consistency loss ($\mathcal{L}_{GC}$). Both generative adversarial loss and cycle-consistency loss were introduced by Zhu et al. in the original cycleGAN paper and are not modified in this work. The authors added the gradient consistency (GC) loss to the proposed method, which aims to preserve the boundaries of the synthesized images due to the significant intensity difference between the two imaging modalities.

The calculation of the gradient consistency was given by the authors as follow:

$$GC(A,B) = \frac{1}{2}\{NCC(\nabla_x A, \nabla_x B) + NCC(\nabla_y A, \nabla_y B)\}$$

$$\text{where, } NCC(A,B) = \frac{\sum_{(i,j)}(A-\bar{A})(B-\bar{B})}{\sqrt{\sum_{(i,j)}(A-\bar{A})^2}\sqrt{\sum_{(i,j)}(B-\bar{B})^2}}$$

The corresponding gradient consistency loss is then calculated through the following formula:

$$\mathcal{L}_{GC} = \frac{1}{2}\{\sum_{x \in I_{CT}}(1 - GC(x, G_{MR}(x))) + \sum_{y \in I_{MR}}(1 - GC(y, G_{CT}(y)))\}$$

The overall loss function of the entire cycleGAN network is updated correspondingly by adding the GC loss with a weighting constant (λ_{GC}) as follow:

$$\mathcal{L}_{total} = \mathcal{L}_{CT} + \mathcal{L}_{MR} + \lambda_{Cycle}\mathcal{L}_{Cycle} + \lambda_{GC}\mathcal{L}_{GC}$$

2.3 Evaluation

The quantitative evaluation of this work is achieved by comparing the results by adjusting: (1) the number of training data, and (2) the implementation of gradient consistency loss.

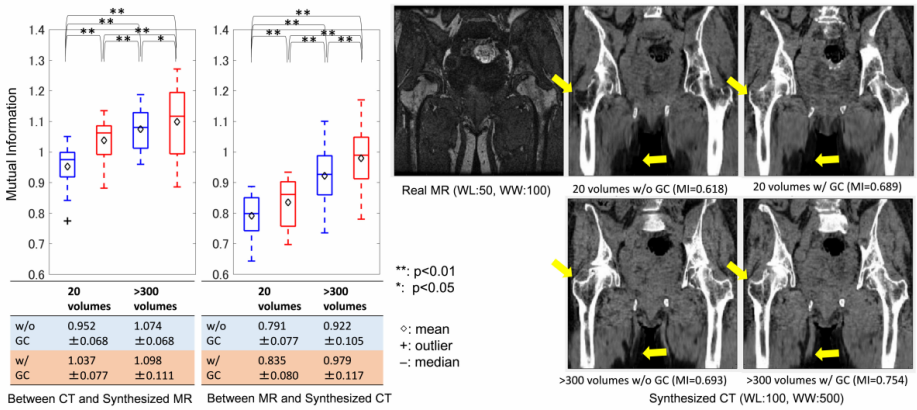
To evaluate the proposed network, the authors trained the network with 20 MR and 20 CT volumes and with 302 MR and 613 CT volumes separately, both with and without the implementation of GC loss. The first experiment used three sets of paired MR and CT volumes of the same patient as the test data, and image similarity metrics including mean absolute error (MAE) and peak-signal-to-noise ratio (PSNR) are used for result evaluation. In the second experiment, unpaired 10 MR and 20 CT volumes were tested and mutual information (MI) between synthesized CT and original MR was used for evaluation since the ground truth of the original MR is no longer available.

The result of the first experiment where paired MR and CT volumes are used as test sets is shown below:

		20 volumes		>300 volumes	
		w/o GC	/w GC	w/o GC	/w GC
MAE	Patient #1	30.121	30.276	26.899	26.388
	Patient #2	26.927	26.911	22.319	21.593
	Patient #3	33.651	32.155	29.630	28.643
	Average ± SD	30.233 ± 2.177	29.781 ± 1.777	26.283 ± 1.367	25.541 ± 1.129
PSNR	Patient #1	14.797	14.742	15.643	15.848
	Patient #2	15.734	15.628	17.255	17.598
	Patient #3	14.510	14.820	15.674	15.950
	Average ± SD	15.014 ± 0.330	15.063 ± 0.380	16.190 ± 0.273	16.465 ± 0.296

Figure 4: First experiment result (paired MR and CT test set)[4]

The result of the second experiment where unpaired MR and CT volumes are used as test sets is shown below:



2.4 Conclusion

The authors mentioned that both experiments conducted showed that the increased number of volumes in the training datasets significantly improves the overall network performance. The introduction of gradient consistency loss slightly improves the performance in both experiments as well. By inspecting the synthesized CT images, the authors claims that with GC loss implemented, the shape of the femoral head and the adductor muscles would be more likely preserved.

3 Critical Assessment

3.1 Pros

The paper provides thorough training methods, and loss function implementations. The use of mathematical notations are align with most generative adversarial network papers and image similarity papers, allowing readers to easily go through the equations. The authors also came up with good evaluation methods by adapting to available data. With unpaired

MR-CT volumes, mutual information (MI) metric may be the best evaluation method, while with the few paired volumes, MAE and PSNR were used for more accurate analysis. The few labeled CTs were also used for quantitative evaluation on segmentation.

3.2 Cons

In the evaluation section of the paper, the effect of the implementation of gradient consistency loss was tested. In both experiments, the introduction of this additional loss function does make slight improvements in both MI and PSNR values. However, there are no additional tests that modifies the weight of the loss (λ_{GC}). The weight is set to 0.3 throughout the experiment, which is significantly lower than the cycle-consistency loss (which is set to 3) and it may be interesting to see the effects of increasing this value, or perhaps set is as an additional variable and conduct another set of experiments.

Both the datasets used and their code for the implementation are not publicly available. Releasing the dataset might not be possible due to patient privacy issues and is understandable. However, open-sourcing the code implementation would really help researchers reproduce their outcome or perhaps utilize their results as benchmark for evaluation.

3.3 Relevance

This paper provides good introduction of cycleGAN implementation on hip regions and provides clear mathematical approaches for loss function calculation, which are very useful for my project as I am also currently using a similar cycleGAN setup.

The authors' adaptability to provide different evaluation methods base on the dataset they possess is also a great ability to learn, and I will be utilizing some of their methods during the evaluation step of my project.

References

[1] *Osteonecrosis of the Hip - OrthoInfo - AAOS*. URL: https://orthoinfo.aaos.org/en/diseases--conditions/osteonecrosis-of-the-hip.

[2] A. Torrado-Carvajal et al. "Fast patch-based pseudo-ct synthesis from T1-WEIGHTED Mr images FOR pet/mr Attenuation correction in brain studies". In: *Journal of Nuclear Medicine* 57.1 (2015), pp. 136–143. DOI: 10.2967/jnumed.115.156299.

[3] Jun-Yan Zhu et al. "Unpaired Image-to-Image Translation Using Cycle-Consistent Adversarial Networks". In: *2017 IEEE International Conference on Computer Vision (ICCV)* (2017). DOI: 10.1109/iccv.2017.244.

[4] Yoshito Otake et al. "Robust 3D–2D image registration: application to spine interventions and vertebral labeling in the presence of anatomical deformation". In: *Physics in Medicine and Biology* 58.23 (2013), pp. 8535–8553. DOI: 10.1088/0031-9155/58/23/8535.