UNIVERSITY OF COPENHAGEN

MRI to Synthesis-CT Generation Using Pix2Pix Framework

Reza Karimzadeh, Bulat Ibragimov

Department of Computer Science

Contact Information:
Universitetsparken 5
2100 København Ø

Email: reza.karimzadeh@di.ku.dk

Abstract

This study explores the synthesis of CT images from MR scans, a cost-effective approach in medical imaging. It employs a Generative Adversarial Network (GAN) with a SwinUnet backbone to generate synthetic CT images from MR data, reducing radiation exposure and imaging costs. Using a patch-based Pix2Pix model, fine-grained CT details are synthesized, enhancing clinical procedures and treatment planning. The study utilizes the SynthRad 2023 dataset for training and evaluation. Results demonstrate promising performance in terms of Mean Absolute Error (MAE), Peak Signal-to-Noise Ratio (PSNR), and Structural Similarity Index (SSIM), suggesting the potential of this technique in medical imaging applications.

Introduction

The synthesis of CT images from MR scans is an innovative and cost-effective approach in medical imaging. It involves generating synthetic CT images using advanced deep learning models trained on paired MR and CT datasets. The aim is to reduce radiation exposure and imaging costs for patients by eliminating the need for separate CT scans. During testing, the trained model takes an MR scan as input and produces a synthetic CT image that closely resembles a real CT scan. This technology has significant potential to streamline clinical procedures, improve diagnostic accuracy, and enhance treatment planning in various medical applications, particularly in radiation therapy planning. Ongoing research and advancements in deep learning and medical imaging hold promising prospects for the widespread adoption of this technique, offering non-invasive and patient-friendly imaging solutions in the future [3].

Main Objectives

1. Create a Generative Adversarial Network (GAN) model to generate synthetic CT images from MR images for data augmentation in medical imaging.
2. Use the SwinUnet transformer architecture as the core component of the GAN model to improve spatial feature extraction.
3. Implement a 3D patch-based Pix2Pix model to ensure accurate synthesis of fine-grained CT details from MR images.

Materials and Methods

Overall, this study aimed to explore the effectiveness of a patch-based Pix2Pix [2] models with SwinUnet [1] backbone in synthesizing realistic CT images from MR data. The models were trained and evaluated using the SynthRad 2023 challenge dataset, and the obtained results provide valuable insights into their performance and potential clinical applications.

Dataset Description and Preprocessing

The dataset used in this study was obtained from the synthRad 2023 challenge, comprising 360 paired registered MR and CT images from three different medical centers for brain and pelvis (180 3D images per task). Each pair consists of an MR image and its corresponding ground truth CT image. Prior to training, data preprocessing was carried out to ensure consistency and optimize model performance. Both MR and CT images underwent normalization to scale their intensity values within the range of -1 to 1. Additionally, data augmentation techniques, such as rotation and affine transformations, were applied to augment the dataset and improve the model's generalization. We split the data randomly to 80 and 20 percent for sub-training and sub-validation.

$$PI_{CT} = \frac{I_{CT} + 1024}{4024} \times 2 - 1 \qquad PI_{MR} = \frac{I_{MR}}{Max(I_{MR})} \times 2 - 1 \tag{1}$$

Where PI is prepossessed image and I is the original image.

Model Architectures

Two models with same architecture were employed for MR to CT translation for brain and pelvis images. The model was based on the Pix2Pix framework, employing SwinUnet as the backbone. This model operated in a patch-based 3D manner, where patches of size 64x64x64 were used for both training and inference. For inference, an aggregation technique was employed, involving 50 percent overlapping patches. The loss function used in both models is a weighted sum of Mean Absolute Error (MAE) with an emphasis on valuable regions and the Structural Similarity Index (SSIM) loss (eq. 2). This combination of losses encourages the preservation of important anatomical details while ensuring overall image fidelity.

$$L = \frac{1}{N}\Sigma|W(n)| \times |G(n) - P(n)| + \alpha \times SSIM(P,G) + \lambda \times GAN_{Loss} \tag{2}$$

$$W = \frac{(1+G)}{2}$$

Where G, P and N are ground-truth,predicted CT and the number of voxels, respectively. W, α and λ are weights for MAE, SSIM and GAN cost functions.
The models were trained with random patches generated using a sub-training set in 1000 epochs with AdamW optimizer with a constant learning rate of 1e-5. The best model was selected based on the sub-validation set during the training.

Results

After training, the performance of both models was evaluated quantitatively using metrics such as Mean Absolute Error (MAE), Peak Signal-to-Noise Ratio (PSNR) and SSIM to assess the quality of the synthesized CT images. The results are presented in table 1. Figure 1, demonstrates an axial view of a CT generation for brain images.

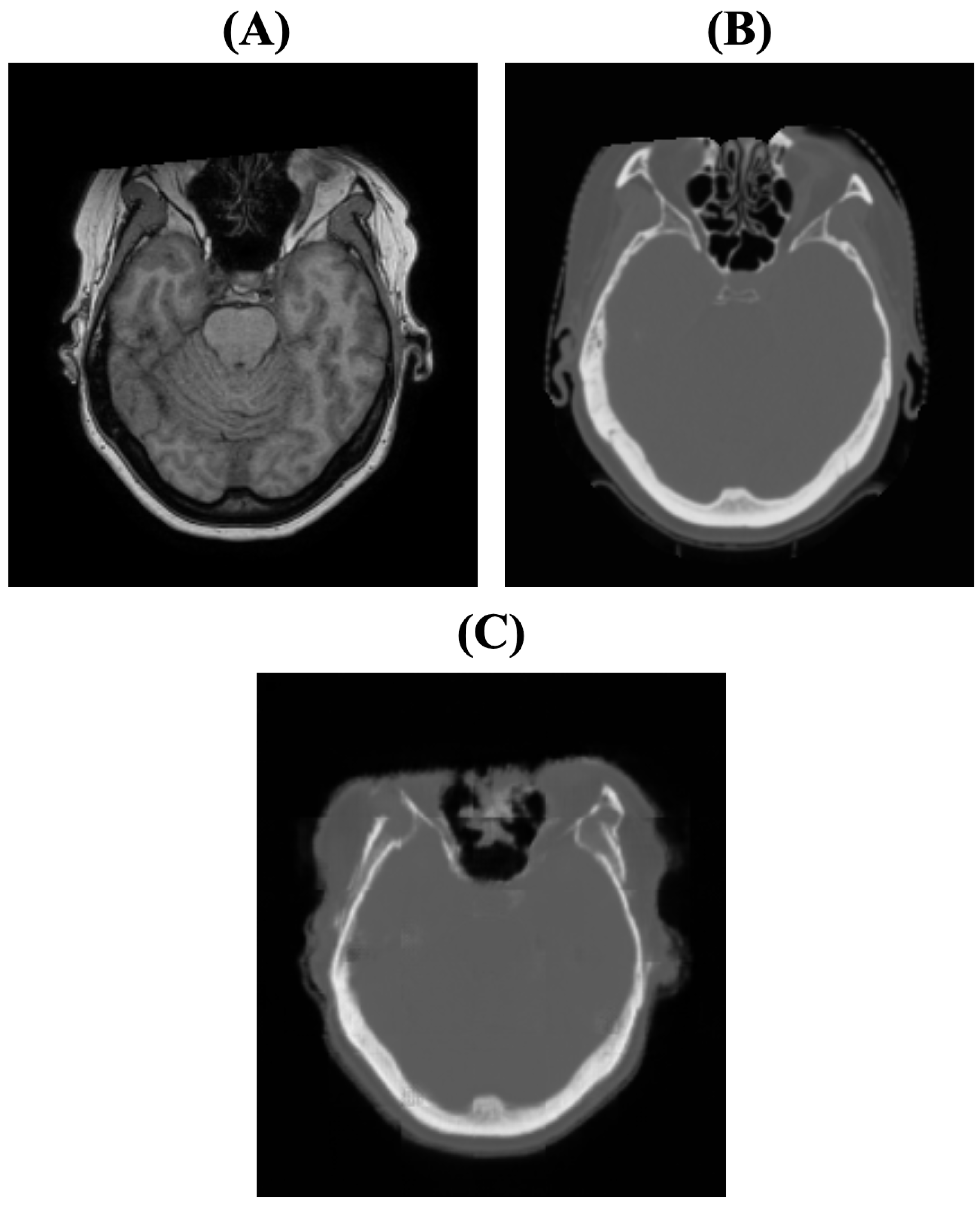

Figure 1: An axial cross section of MRI (A), proportional CT as ground-truth (B), and sCT generated by Pix2Pix framework (C)

Model/Metric	MAE	PSNR	SSIM
Pix2Pix	113.38 ± 20.34	24.70 ± 1.42	0.76 ± 0.03

Table 1: MAE, PSNR and SSIM metrics for Pix2Pix model

Conclusions

In this research, we have developed and evaluated a novel approach for synthesizing CT images from MR scans using a GAN-based framework with a SwinUnet backbone. The results showcase the model's ability to generate synthetic CT images that closely resemble ground truth CT scans, as evidenced by low MAE values, high PSNR scores, and favorable SSIM metrics. This approach holds promise for reducing patient radiation exposure and imaging costs, making it a valuable tool for enhancing medical imaging practices and treatment planning. Future work may involve further refinement of the model and its integration into clinical workflows for improved patient care.

References

[1] Ali Hatamizadeh, Vishwesh Nath, Yucheng Tang, Dong Yang, Holger R Roth, and Daguang Xu. Swin unetr: Swin transformers for semantic segmentation of brain tumors in mri images. In *International MICCAI Brainlesion Workshop*, pages 272–284. Springer, 2021.

[2] Phillip Isola, Jun-Yan Zhu, Tinghui Zhou, and Alexei A Efros. Image-to-image translation with conditional adversarial networks. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 1125–1134, 2017.

[3] Tonghe Wang, Yang Lei, Yabo Fu, Jacob F Wynne, Walter J Curran, Tian Liu, and Xiaofeng Yang. A review on medical imaging synthesis using deep learning and its clinical applications. *Journal of applied clinical medical physics*, 22(1):11–36, 2021.