Indian Institute of Technology (IIT) Mandi

in collaboration with Masai School

Capstone Project Report

MRI-to-Synthetic CT Brain Scan Translation Using Deep Learning

---

Submitted by

Meet Jani

Student Code: iitmcs_2406132

Date:
21st September 2025

Table of Contents

1. Abstract

Accurate neuro-oncology diagnosis and radiotherapy planning often require both MRI and CT scans. While MRI provides superior soft-tissue contrast without radiation exposure, CT scans are crucial for assessing bone structures and electron density. Acquiring both modalities increases scan time, cost, and radiation burden.

This project implements a deep learning pipeline for translating T1-weighted MRI slices into synthetic CT scans. Multiple models were evaluated, including UNet, Pix2Pix, Pix2Pix with ResUNet enhancements, and Swin Transformer-based Pix2Pix and SwinGAN. Quantitative evaluations (MAE, MSE, PSNR, SSIM) and qualitative assessments demonstrate that Pix2Pix achieves the best trade-off between accuracy and computational efficiency.

A lightweight web interface was deployed for real-time, slice-level inference, allowing clinicians to obtain synthetic CT predictions alongside metrics. Innovative dataset caching in RAM with slice-level augmentations accelerated training and improved generalization.

---

2. Introduction

2.1 Clinical Motivation

MRI and CT scans offer complementary diagnostic information:

Modality	Strengths
MRI	Excellent soft tissue contrast; no ionizing radiation
CT	High-resolution bone and density information; rapid acquisition

However, dual acquisition increases cost, patient burden, and radiation exposure. Translating MRI into synthetic CT reduces these limitations, potentially enhancing accessibility in low-resource settings.

2.2 Objectives

1. Develop deep learning models capable of translating MRI slices to synthetic CT.
2. Evaluate quantitative performance using MAE, MSE, PSNR, and SSIM.
3. Deploy a web interface for real-time inference with metrics output.
4. Integrate innovations such as RAM caching and slice-level augmentation for efficient training.

---

3. Dataset & Preprocessing

Dataset: SynthRad, consisting of 180 paired T1-weighted MRI and CT brain volumes.

Preprocessing Steps:

- Conversion of 3D volumes to aligned 2D slices.

- Intensity normalization per slice.
- Optional brain masking to focus on intracranial structures.

Exploratory Analysis:

- Histograms of MRI and CT intensities to ensure consistent distribution.
- Visual inspection of paired slices to confirm alignment.

Innovation Highlight:

- **RAM caching** of slices allowed the entire dataset to be held in memory, drastically reducing I/O overhead during training.
- **Slice-level augmentations** (brightness/contrast adjustment, horizontal flip, rotation, Gaussian noise, intensity scaling) enhanced model generalization.

---

4. Model Architectures

Model	Generator	Discriminator / Enhancements	Loss Function
UNet	Standard UNet	None	L1
Pix2Pix	UNet	PatchGAN	L1 + Adversarial
Pix2Pix-ResUNet	ResUNet	PatchGAN	L1 + SSIM + Adversarial
SwinPix2Pix	UNet + Swin Transformer blocks	PatchGAN	L1 + Adversarial
SwinGAN	Swin-UNet	PatchGAN	L1 + SSIM + Adversarial

Notes:

- Residual connections in Pix2Pix-ResUNet enhance feature propagation and edge preservation.
- Swin Transformer blocks capture long-range dependencies but may require larger datasets to outperform convolutional networks.

---

5. Training & Evaluation

Training Details:

- Optimizer: Adam, learning rate 0.0002.
- Loss: L1 for UNet; L1 + adversarial for GAN models.
- Batch size: 4–12 slices, GPU-dependent.
- Mixed Precision (AMP) enabled for memory efficiency.

Evaluation Metrics:

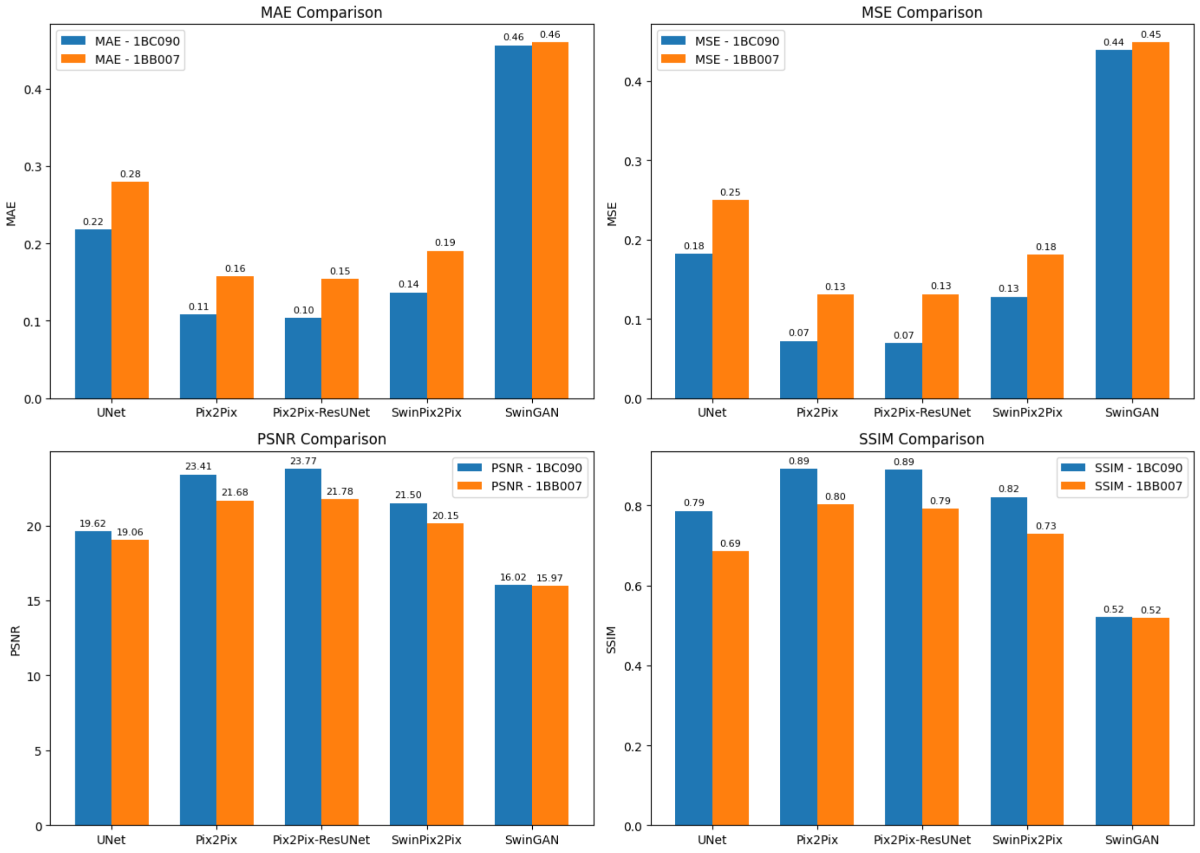
- MAE: Mean Absolute Error
- MSE: Mean Squared Error
- PSNR: Peak Signal-to-Noise Ratio
- SSIM: Structural Similarity Index

---

6. Quantitative Results

6.1 Per-Patient Metrics (Representative Patients)

Model	Patient	MAE	MSE	PSNR	SSIM
UNet	1BC090	0.2183	0.1823	19.62	0.7865
	1BB007	0.2796	0.2500	19.06	0.6852
Pix2Pix	1BC090	0.1080	0.0724	23.41	0.8912
	1BB007	0.1575	0.1311	21.68	0.8023
Pix2Pix-ResUNet	1BC090	0.1036	0.0696	23.77	0.8894
	1BB007	0.1544	0.1314	21.78	0.7917
SwinPix2Pix	1BC090	0.1368	0.1281	21.50	0.8213
	1BB007	0.1906	0.1815	20.15	0.7284
SwinGAN	1BC090	0.4560	0.4390	16.02	0.5210
	1BB007	0.4601	0.4490	15.97	0.5190

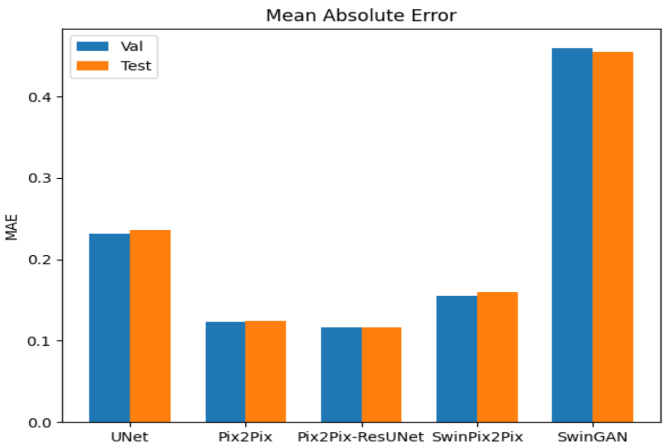


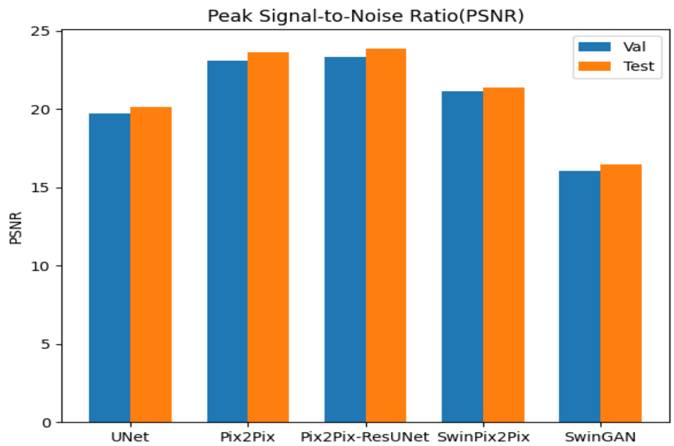
6.2 Overall Metrics Comparison

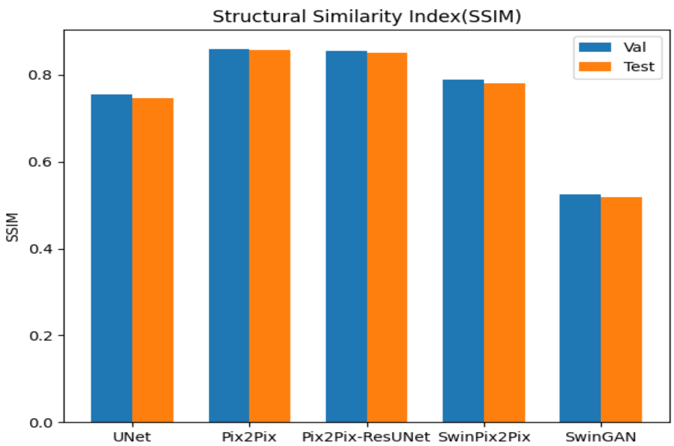
Model	Dataset	MAE	MSE	PSNR	SSIM
UNet	Val	0.2319	0.1974	19.74	0.7555
	Test	0.2357	0.1967	20.13	0.7470
Pix2Pix	Val	0.1231	0.0967	23.07	0.8601
	Test	0.1246	0.0954	23.63	0.8560
Pix2Pix-ResUNet	Val	0.1166	0.0908	23.31	0.8549
	Test	0.1169	0.0884	23.89	0.8516
SwinPix2Pix	Val	0.1552	0.1449	21.16	0.7891
	Test	0.1596	0.1427	21.40	0.7798
SwinGAN	Val	0.4599	0.4472	16.03	0.5239
	Test	0.4546	0.4369	16.48	0.5187

Insights:

- Pix2Pix achieves the best trade-off between fidelity and computational efficiency.
- ResUNet enhancements improve edge detail slightly, without major SSIM gains.
- Swin-based models underperform due to dataset size and model complexity.





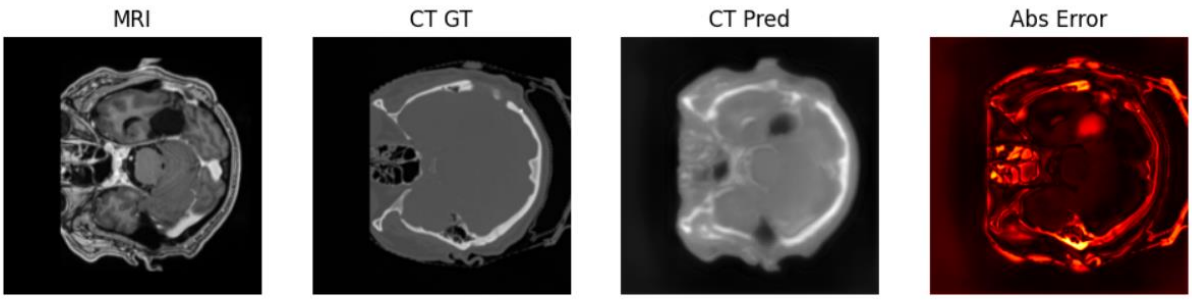


---

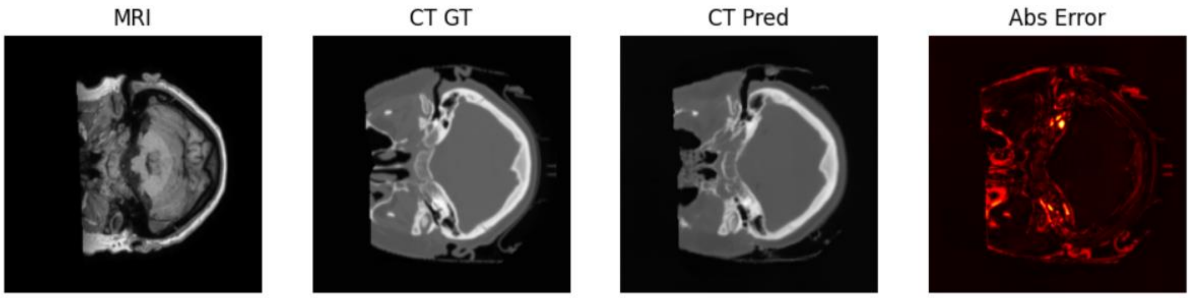
7. Qualitative Results

Representative Side-by-Side MRI → Ground Truth CT → Predicted CT → Absolut Error

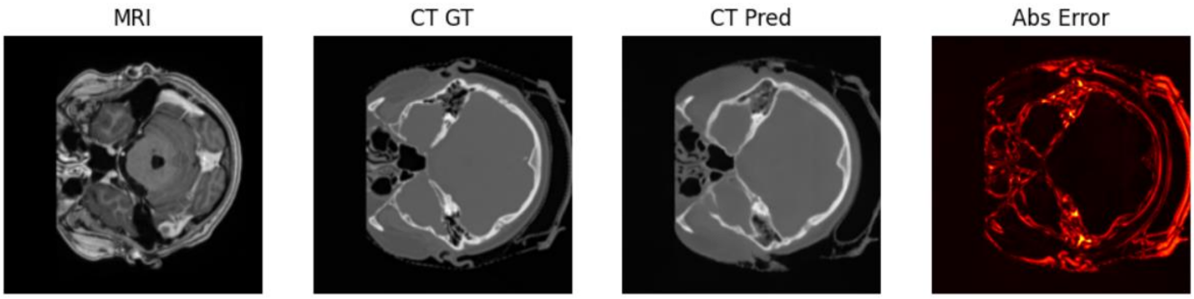
1. **UNet**



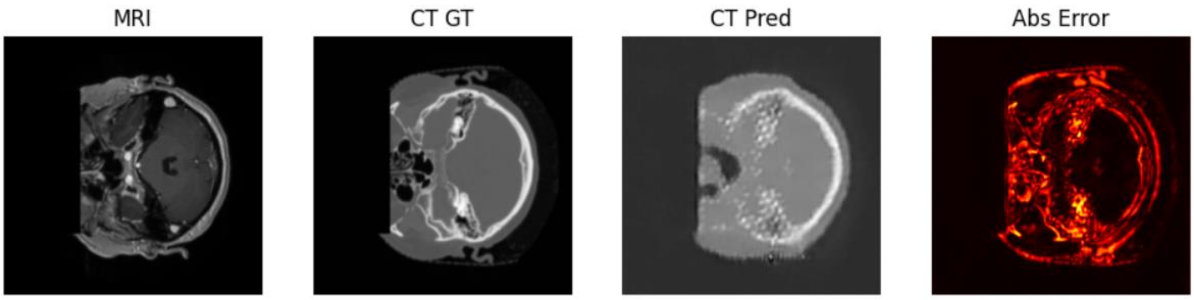
2. **Pix2Pix**



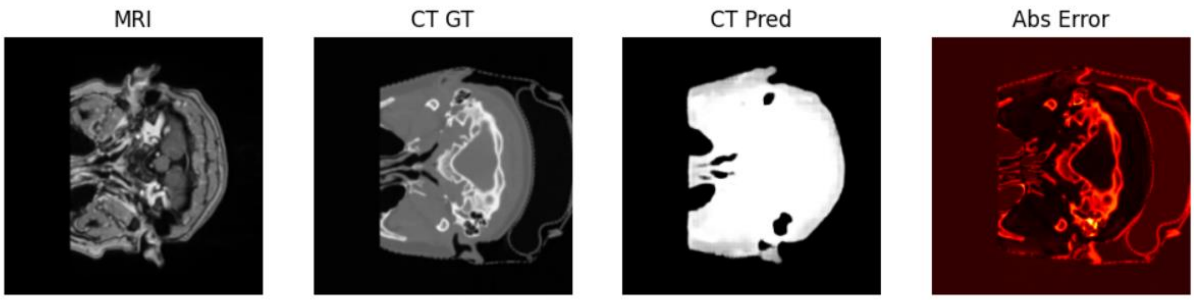
3. **Pix2Pix-ResUNet**



4. **SwinPix2Pix**



5. **SwinGAN**



Observations:

- Pix2Pix effectively captures both bone and soft tissue contrasts.
- ResUNet outputs are sharper, especially at edges.
- Swin-based predictions show artifacts and noise in some regions.

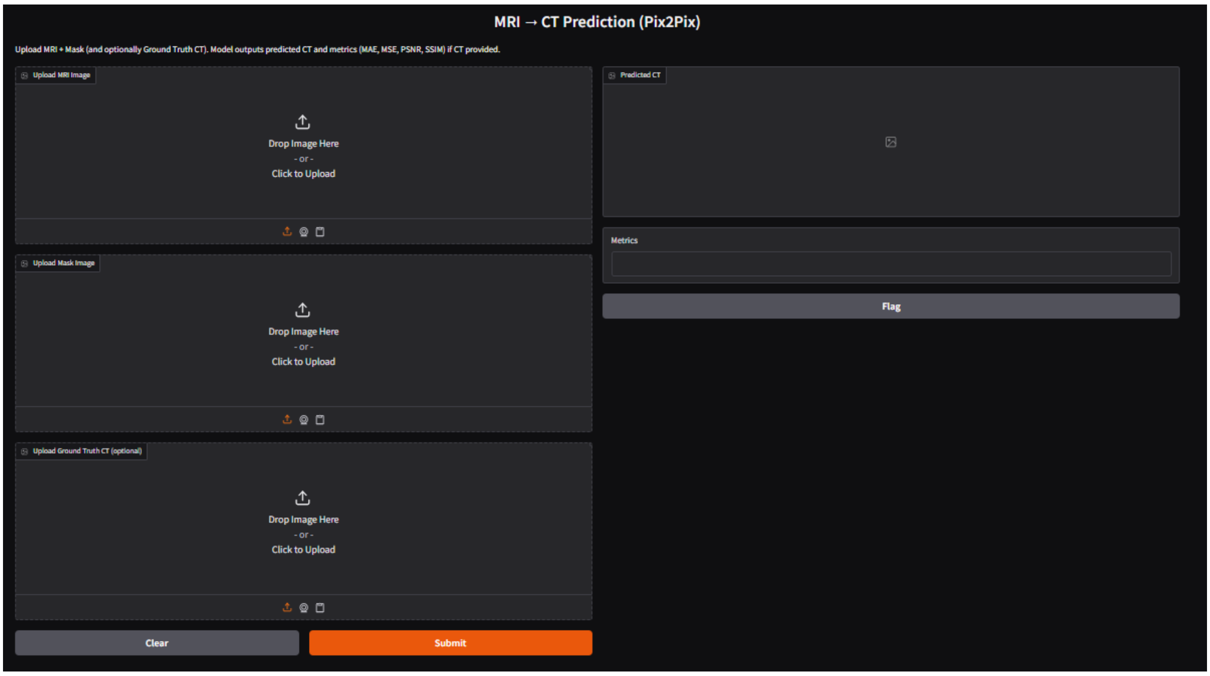
8. Innovation & Technical Contributions

1. **RAM Caching:** Entire dataset slices are loaded into memory to avoid repeated disk I/O, significantly accelerating training.
2. **Slice-Level Augmentation:** Brightness/contrast adjustments, rotations, flips, Gaussian noise, and intensity scaling improve generalization.
3. **Model Evaluation Pipeline:** Integrated metrics computation (MAE, MSE, PSNR, SSIM) per slice and per patient.
4. **Deployment:** Lightweight Gradio-based interface allows real-time slice-level predictions with downloadable outputs.

---

9. Web App & Deployment

- **Framework:** Gradio (Python)
- **Features:**
  - Upload MRI slice + optional brain mask.
  - Obtain predicted CT with metrics.
  - Download results in PNG/NIfTI format.
- **Performance:**
  - GPU: ~200–400 ms per slice
  - CPU: ~2–4 s per slice



---

10. Discussion

- **Performance Trade-Off:** Pix2Pix offers best fidelity for this dataset; ResUNet improves edge sharpness but adds complexity.
- **Limitations:**
    - 2D slice-based translation ignores volumetric context.
    - Swin Transformers require larger datasets for optimal performance.
    - Currently limited to T1-weighted MRIs.
- **Clinical Implications:** Generated synthetic CTs could reduce patient radiation exposure while providing reliable anatomical information for radiotherapy planning.

---

11. Conclusion & Future Work

- Developed a robust pipeline for MRI-to-CT translation using multiple deep learning models.
- Pix2Pix consistently outperformed other architectures across per-patient and overall metrics.
- RAM caching and slice-level augmentation accelerated training and improved generalization.
- Web app enables real-time predictions for clinical or research use.

Future Directions:

1. Extend models to 3D volume prediction.
2. Incorporate multi-modal MRIs for improved generalization.
3. Explore hybrid CNN-Transformer architectures with larger datasets.
4. Integration with PACS systems for real-time clinical deployment.

---

12. References

1. Isola, P., et al. *Image-to-Image Translation with Conditional Adversarial Networks*, CVPR 2017.
2. Ronneberger, O., et al. *U-Net: Convolutional Networks for Biomedical Image Segmentation*, MICCAI 2015.
3. Chartsias, A., et al. *Medical Image Synthesis for Multi-Modal Neuroimaging*, ISBI 2017.
4. Zhu, J., et al. *Unpaired Image-to-Image Translation Using Cycle-Consistent Adversarial Networks*, ICCV 2017.
5. **SynthRAD2023 Grand Challenge** dataset: Paired MRI-CT brain scans of 180 subjects (2024)