# Capstone Project Report

**Name :** Vishal Kumar
**Student code :** iitmcs_2406329

## Project Title : MRI-to-Synthetic CT Brain Scan Translation Using Deep Learning

## Introduction

In modern clinical workflows, particularly for radiotherapy planning of brain tumors, Magnetic Resonance Imaging (MRI) and Computed Tomography (CT) scans serve complementary roles. MRI offers superior soft-tissue contrast for accurate tumor delineation, while CT is essential for radiation dose calculation. However, the necessity of acquiring both scans imposes significant logistical burdens, including increased costs, extended patient scan times, and exposure to ionizing radiation from the CT scan. This project addresses these challenges by developing a deep learning framework to generate synthetic CT (sCT) images directly from T-weighted MRI scans, enabling a more efficient "MRI-only" clinical pathway.

Utilizing the SynthRad dataset of 180 preprocessed and aligned 3D paired brain scans, this study aims to optimize the Pix2Pix architecture for high-fidelity, 2D slice-level sCT generation. While Pix2Pix provides a strong baseline, its performance is highly dependent on its training configuration. Therefore, this project systematically evaluates two key modifications against an 81-epoch baseline to enhance image quality and training efficiency:

1. Altering the **Generator-to-Discriminator (G:D) training ratio**.
2. Replacing the standard L1 reconstruction loss with a more robust **Charbonnier loss function**.

The primary objective is to identify an optimal configuration that yields superior sCT image quality in significantly fewer training epochs than the standard approach.

## Methodology and Experiments

A series of experiments were conducted using a paired MRI-CT dataset to identify an optimal training configuration for the Pix2Pix framework. Performance was evaluated using **Peak Signal-to-Noise Ratio (PSNR)**, **Structural Similarity Index (SSIM)**, and **Mean Absolute Error (MAE)**. The experimental setups are detailed below, progressing from a baseline to exploratory tests and culminating in a two-phase fine-tuning strategy.

| Exp. ID | Description | Loss Function | Epsilon (ε) | G:D Ratio | Epochs |
|---|---|---|---|---|---|
| E1 | Baseline Pix2Pix | L1 Loss | N/A | 1:1 | 81 |
| E2 | G:D Ratio Test | L1 Loss | N/A | 2:1 | 10 |
| E3 | Two-Phase (L1 → Charb.) | L1 → Charbonnier | Phase 2: 1e-4 | Phase 1: 3:1 Phase 2: 2:1 | 45(10+35) |
| E4 | Charbonnier Test | Charbonnier | 1.00E-04 | 2:1 | 55 |
| E5 | Charbonnier Test | Charbonnier | 1.00E-06 | 1:1 | 20 |
| E6 | Two-Phase Fine-Tuning | Charbonnier | Phase 1: 1e-6 Phase 2: 1e-4 | Phase 1: 1:1 Phase 2: 2:1 | 78 (40+38) |

## Quantitative Results

The performance of each experimental configuration on the test set is summarized in the table below. The **"Two-Phase Fine-Tuning"** experiment emerged as the clear top performer, achieving the best overall results across all key metrics. This optimized model surpassed the strong performance of the fully-trained 81-epoch baseline model. Notably, the baseline model saved at 54 epochs based on validation loss also showed very strong results, confirming the effectiveness of the standard approach when properly validated.

| Experiment Description | # Total Epochs | G:D Strategy | Loss Strategy | # PSNR (↑) | # SSIM (↑) | # MAE in HU (↓) |
|---|---|---|---|---|---|---|
| Two-Phase Fine-Tuning | 78 | 1:1 → 2:1 | Charb. → Charb. | 35.74 | 0.962 | 11.49 |
| Baseline (Final) | 81 | 1:1 | L1 | 34.75 | 0.954 | 13.22 |
| Baseline (Best Val) | 54 | 1:1 | L1 | 34.38 | 0.951 | 13.6 |
| Charbonnier (Initial Test) | 20 | 2:1 | Charbonnier | 33.81 | 0.946 | 14.68 |
| Two-Phase (L1 → Charb.) | 45 | 3:1 → 2:1 | L1 → Charb. | 33.35 | 0.94 | 16.05 |
| G:D Ratio Test | 10 | 3:1 | L1 | 33.21 | 0.942 | 16.45 |
| Charbonnier (Standard) | 40 | 1:1 | Charbonnier | 32.75 | 0.937 | 16.31 |
| G:D Ratio Test | 10 | 2:1 | L1 | 32.35 | 0.93 | 17.56 |

## Analysis and Interpretation

The experimental results reveal a clear and logical path to optimizing the Pix2Pix framework for this task. The analysis progressed from establishing a strong performance benchmark with the baseline model to identifying a novel, state-of-the-art training strategy that ultimately surpassed it.

### 1. Baseline Performance and Early-Stage Insights

The fully-trained baseline model (**E1**) established a robust performance benchmark, achieving a final **PSNR of 34.75** after 81 epochs. This confirmed the viability of the standard Pix2Pix architecture but also highlighted the significant training time required. The initial exploratory experiments provided a key insight: increasing the G:D ratio to 3:1 accelerated initial convergence, achieving a respectable PSNR of 33.21 in just 10 epochs. This suggested that modifying the training dynamic could be a powerful tool for improving efficiency.

### 2. The Critical Impact of the Charbonnier Loss

The introduction of the Charbonnier loss function proved to be the most critical modification for efficient learning. The initial 20-epoch experiment using this loss with $\varepsilon$=1e-4 (**E4**) yielded a remarkable **PSNR of 33.81**. This result approached the performance of the 54-epoch validation-saved baseline in less than half the time, strongly indicating that the Charbonnier loss provides a more effective and stable learning signal than the standard L1 loss.

### 3. State-of-the-Art Results Through Two-Phase Fine-Tuning

The final experiment (**E6**) combined the insights from all previous tests into a novel two-phase training strategy. By first establishing a strong foundation with a stable 1:1 ratio and Charbonnier loss ($\varepsilon$=1e-6) for 40 epochs, the model was then **fine-tuned** for an additional 38 epochs using an accelerated 2:1 G:D ratio and a smoother loss ($\varepsilon$=1e-4).

This fine-tuned model became the undisputed top performer, achieving a final **PSNR of 35.74** and an **SSIM of 0.962**. This result not only surpasses the 81-epoch baseline in fewer total epochs but also demonstrates the power of a curriculum-based training approach: establishing a solid foundation and then aggressively optimizing to achieve state-of-the-art results that a standard, monolithic training process could not reach.

## Conclusion

This study successfully developed and validated a novel two-phase training methodology that enhances both the performance and efficiency of the Pix2Pix framework for MRI-to-sCT translation.

Our experiments confirmed that a standard Pix2Pix model with L1 loss serves as a strong performance benchmark. However, our definitive conclusion is that this performance can be significantly surpassed. The introduction of the **Charbonnier loss** was the critical factor, enabling the model to achieve high-fidelity results with far greater efficiency than the standard L1 loss.

The optimal configuration was a **two-phase fine-tuning approach**, which achieved a new state-of-the-art result (**PSNR of 35.74**), surpassing the fully-trained baseline in fewer total epochs. This demonstrates that a sophisticated training strategy, which combines a stable initial training phase with an aggressive fine-tuning period, is the most effective path to maximum quality and efficiency.

The success of this methodology warrants future work focused on:

1. **Clinical Validation:** Using the generated sCTs for radiotherapy dose calculation.
2. **Model Optimization:** Applying techniques like quantization for deployment in clinical software.
3. **Generalization:** Testing this training strategy on other medical image translation challenges.