

VNetLS: Supervised 3D VNet-Based MRI Enhancement with Loss Optimization

by

**Syeda Sunehra Sakhawat (223014171)
Nusrat Jannat Tithy (223014116)
Romana Sultana (223014170)
Kazi Ifat Jahan Zim (223014017)
Meherun Nessa Eva (223014219)**

*Capstone project report (CSE 4098C) submitted in partial fulfillment of the
requirements for the degree of*

Bachelor of Science in Computer Science and Engineering

Under the supervision of

Nafees Mansoor, PhD



**DEPARTMENT OF COMPUTER SCIENCE AND ENGINEERING
UNIVERSITY OF LIBERAL ARTS BANGLADESH**

SPRING 2026

© *Syeda Sunehra Sakawat, Nusrat Jannat Tithy, Romana Sultana, Kazi Ifat Jahan Zim
and Meherun Nessa Eva*
All rights reserved

DECLARATION

Project Title VNetLS: Supervised 3D VNet-Based MRI Enhancement with Loss Optimization
Authors Syeda Sunehra Sakhawat, Nusrat Jannat Tithy, Romana Sultana, Kazi Ifat Jahan Zim and Meherun Nessa Eva
Student IDs 223014171, 223014116, 223014170, 223014017 and 223014219
Supervisor Nafees Mansoor, PhD

We declare that this capstone project report entitled *VNetLS: Supervised 3D VNet-Based MRI Enhancement with Loss Optimization* is the result of our own work except as cited in the references. The capstone project report has not been accepted for any degree and is not concurrently submitted in candidature of any other degree.

Syeda Sunehra Sakhawat
223014171

Department of Computer Science and Engineering
University of Liberal Arts Bangladesh

Nusrat Jannat Tithy
223014116

Department of Computer Science and Engineering
University of Liberal Arts Bangladesh

Romana Sultana
223014170

Department of Computer Science and Engineering
University of Liberal Arts Bangladesh

Kazi Ifat Jahan Zim
223014017

Department of Computer Science and Engineering
University of Liberal Arts Bangladesh

Meherun Nessa Eva
223014219

Department of Computer Science and Engineering
University of Liberal Arts Bangladesh

Date: June 08, 2026



Department of Computer Science and Engineering
University of Liberal Arts Bangladesh
Mohammadpur, Dhaka - 1207

CERTIFICATE

This is to certify that the capstone project report entitled **VNetLS: Supervised 3D VNet-Based MRI Enhancement with Loss Optimization**, submitted by **Syeda Sunehra Sakhawat** (Student ID: 223014171), **Nusrat Jannat Tithy** (Student ID: 223014116), **Romana Sultana** (Student ID: 223014170), **Kazi Ifat Jahan Zim** (Student ID: 223014017) and **Meherun Nessa Eva** (Student ID: 223014219) are undergraduate students of the **Department of Computer Science and Engineering** has been examined. Upon recommendation by the examination committee, we hereby accord our approval of it as the presented work and submitted report fulfill the requirements for its acceptance in partial fulfillment for the degree of *Bachelor of Science in Computer Science and Engineering*.

Nafees Mansoor, PhD
Associate Professor

Department of Computer Science and Engineering
University of Liberal Arts Bangladesh

Foysal Ahmed
Lecturer

Department of Computer Science and Engineering
University of Liberal Arts Bangladesh

Atanu Shuvam Roy

Lecturer

Department of Computer Science and Engineering
University of Liberal Arts Bangladesh

Prof. Muhammad Golam Kibria, PhD, SMIEE

Professor and Head

Department of Computer Science and Engineering
University of Liberal Arts Bangladesh

Place: Dhaka

Date: June 08, 2026

ACKNOWLEDGEMENTS

We would like to express our deep and sincere gratitude to our research supervisor, *Nafees Mansoor, PhD*, for giving us the opportunity to conduct research and providing invaluable guidance throughout this work. His dynamism, vision, sincerity and motivation have deeply inspired us. He has taught us the methodology to carry out the work and to present the works as clearly as possible. It was a great privilege and honor to work and study under his guidance.

We are greatly indebted to our honorable teachers of the Department of Computer Science and Engineering at the University of Liberal Arts Bangladesh who taught us during the course of our study. Without any doubt, their teaching and guidance have completely transformed us to the persons that we are today.

We are extremely thankful to our parents for their unconditional love, endless prayers and caring, and immense sacrifices for educating and preparing us for our future. We would like to say thanks to our friends and relatives for their kind support and care.

Finally, we would like to thank all the people who have supported us to complete the project work directly or indirectly.

Syeda Sunehra Sakawat, Nusrat Jannat Tithy, Romana Sultana, Kazi Ifat Jahan Zim and Meherun Nessa Eva

University of Liberal Arts Bangladesh

Date: June 08, 2026

Dedicated to
*my beloved parents for their
love, care and support*
– Syeda Sunehra Sakhawat

To my beloved parents
*for their unconditional love
and prayer*
– Nusrat Jannat Tithy

To my family
for their endless support
– Romana Sultana

To my beloved parents
the source of my inspiration
– Kazi Ifat Jahan Zim

To my Family
for their immense support
– Meherun Nessa Eva

ABSTRACT

Magnetic Resonance Imaging (MRI) is essential for medical diagnosis, but image quality depends heavily on scanner strength. High-field MRI systems such as 7T provide superior resolution and contrast; however, their high cost and limited availability restrict widespread clinical use in under developed areas. This study proposes a deep learning-based framework, VNetLS, to enhance low-field 3T MRI images into high-quality 7T-like outputs. The method employs a structured pre-processing pipeline, including skull stripping, bias field correction, registration, and intensity normalization, to ensure data consistency. A 3D VNet encoder-decoder architecture is used to effectively capture volumetric anatomical features. The key contribution is a multi-component loss function that combines Mean Absolute Error (MAE), Structural Similarity Index Measure (SSIM), and gradient loss, improving image sharpness, structural preservation, and edge detail. In addition, several fine-tuning strategies were explored to further optimize model performance, where replacing the optimizer with AdamW demonstrated the most significant improvement in output quality. Experimental results show that the proposed model outperforms baseline approaches, achieving higher PSNR and SSIM values, indicating improved image quality and anatomical fidelity. This framework offers a cost-effective, software-based solution for MRI enhancement, particularly beneficial for resource-limited healthcare settings. Future work includes expanding paired datasets and incorporating multi-modal MRI inputs to further improve robustness and performance.

Keywords: 3D Convolutional Neural Network, Image Super-Resolution, Medical Image Processing, Multi-Component Loss Function, Volumetric Imaging, Image Reconstruction

Contents

1	Introduction	1
1.1	Problem statement	3
1.2	Aims, objectives and Motivation	4
1.2.1	Aims	4
1.2.2	Objective	4
1.2.3	Motivation	5
1.3	Project Specification	5
1.3.1	Functional Specifications	5
1.3.2	Performance Specifications	6
1.3.3	Data and Training Specifications	6
1.3.4	Evaluation and Validation Specifications	6
1.3.5	System Constraints	7
2	Literature Review	8
2.1	Literature Review Table	12
3	Methodology	18
3.1	Dataset collection	18
3.2	Data Preprocessing	20
3.2.1	Skull Stripping	21
3.2.2	Bias Field Correction (N4 Correction)	22
3.2.3	Registration	23
3.2.4	Resampling	25
3.2.5	Intensity Normalization	25
3.3	Model Selection	26
3.4	Model Optimization and Training	26

3.4.1	Epoch Tuning	27
3.4.2	Loss Function Design	27
3.4.3	3D Training Strategy	27
3.4.4	Optimizer Selection	28
3.5	Testing	28
3.6	Evaluation	28
3.6.1	Quantitative Evaluation	29
3.6.2	Qualitative Evaluation	29
4	Design and Development	30
4.1	Data Preparation Module	30
4.2	Data Loading and Padding Module	32
4.3	Dataset Splitting	32
4.3.1	VNet-Based Enhancement Module	32
4.3.2	Loss Function Design	33
4.4	Training and Optimization Module	34
4.5	Evaluation Module	34
4.6	Website Implementation	35
5	Result analysis	37
5.1	Model Selection and Performance Analysis	37
5.2	Quantitative Results	38
5.2.1	Result Analysis	39
5.3	Qualitative Analysis	44
5.3.1	Performance Comparison	48
6	Plans and Deliverables	49
6.1	Budget	49
6.1.1	Project Timeline	50
7	Conclusions	51
7.1	Social, Legal, Ethical, and Environmental Issues	51
7.1.1	Social Impact	51
7.1.2	Ethical and Legal Considerations	51
7.1.3	Environmental Issues	52

7.2	Brief Summary	53
7.3	Future works	53
References		55

List of Figures

3.1	Methodology	19
3.2	Skull Stripping of 3T MRI images	21
3.3	Skull Stripping of 7T MRI images	21
3.4	Bias Field Correction	22
3.5	Bias Field Correction Graph	22
3.6	Registration of Pairs	24
3.7	Transformed and Resampled	25
4.1	Architecture of VnetLS	31
4.2	VnetLS pipeline	31
4.3	User Interface	35
4.4	Showing The 3T Input & 7T Output	36
5.1	Vnet Mask vs Vnet Without Mask	40
5.2	VnetLS Mask vs VnetLS Without Mask	41
5.3	Vnet vs VnetLS	41
5.4	Vnet vs Vnet Fine Tune	42
5.5	VNetLS vs VNetLS Fine Tune	43
5.6	Vnet vs VNetLS with AdamW	44
5.7	Predicted 7T Comparison with Target in validation	45
5.8	Predicted 7T Comparison with Target testing	46
5.9	Predicted 7T Comparison with Target testing	47
6.1	Total Budget	49
6.2	Project Timeline	50

List of Tables

2.1	Literature Review Analysis Table	12
2.2	Result Matrix Analysis Matrices across Different Studies	17
5.1	Impact of skull stripping on model performance	38
5.2	Comparison of model performance on different fine tuning experiments	39

Chapter 1

Introduction

Magnetic Resonance Imaging or MRI is an important tool in modern medical diagnosis. It can produce clear images of internal body structures. It is especially useful for viewing soft tissues. MRI is widely used for brain analysis. It helps in tumor detection and disease monitoring. One major advantage of MRI is that it does not use ionizing radiation which makes it safer for patients [Shamaei et al. \(2025a\)](#). However, MRI images are not always perfect. Image quality can be affected by noise. Patient movement is another common issue. Hardware limitations also reduce quality. Fast scanning methods can further degrade the image.

These problems are more noticeable in low-field and accelerated MRI systems. In such systems, the signal strength is weak. Spatial resolution is also lower. As a result, the images show poor contrast. Fine anatomical details are often lost [Zhang et al. \(2025a\)](#). High-field MRI scanners such as 3 Tesla and 7 Tesla systems provide much better image quality. They offer higher signal-to-noise ratio. Anatomical structures appear clearer. But these systems are very expensive. They need advanced infrastructure and trained professionals. Because of this they are not widely available. This is especially true in developing countries like Bangladesh. Most hospitals rely on low or mid-field MRI scanners. These scanners often produce blurry images with low contrast. This makes clinical interpretation more difficult [Islam et al. \(2025a\)](#).

Although high-field MRI systems provide significant advantages, their availability is largely limited to technologically advanced healthcare environments.

Due to high installation cost, maintenance complexity, and infrastructure requirements, in many developing and underdeveloped regions, including countries like Bangladesh, such systems are not widely accessible. As a result, healthcare providers often rely on lower-field MRI machines, where subtle anatomical details may not be clearly visible, potentially affecting accurate diagnosis and treatment planning. Replacing existing systems with high-field MRI scanners is economically challenging for most institutions. Therefore, there is a growing need for computational approaches that can enhance MRI image quality without requiring hardware upgrades. Similar concerns regarding accessibility and cost limitations of high-field MRI systems have been discussed in prior studies [Z. Wang et al. \(2025a\)](#).

To overcome these limitations, MRI image enhancement has become an effective solution. It is software-based approach that improves image quality without repeating scans and avoids costly hardware upgrades [Cui et al. \(2024a\)](#). Traditional enhancement methods include filtering techniques, histogram-based contrast adjustment, and interpolation-based super-resolution methods. These techniques can reduce noise and improve contrast, but they often fail to preserve important anatomical details, which limits their reliability for diagnosis [Z. Wang et al. \(2025a\)](#).

Recently, deep learning has shown strong potential in MRI image enhancement. These methods learn directly from data. Convolutional Neural Networks and Generative Adversarial Networks are widely used for denoising, resolution enhancement, and contrast improvement. They can model complex non-linear patterns in MRI data [Shamaei et al. \(2025a\)](#). Unpaired learning methods such as CycleGAN have further expanded this field by enabling enhancement without paired training data. Low-field MRI images can be transformed into high-field-like images using this approaches [Jiang et al. \(2025a\)](#). Newer models such as diffusion-based methods and transformers have shown improved performance by preserving fine details and maintaining structural consistency.

In this project, a supervised deep learning framework is introduced to translate 3T MRI images to approximated 7T-like quality, given paired datasets. A 3D VNet-based encoder-decoder architecture is then employed to learn volumetric spatial characteristics and reconstruct a high-quality MRI image. Additionally, a

multi-component loss optimization strategy is integrated to enhance structural preservation and edge details for high visual quality and anatomical consistency of standard contrast-enhance images. The method we propose aims to be a solution that is cost effective and practical for improving the quality of MRI images specifically in limited resource healthcare areas.

1.1 Problem statement

MRI or Magnetic Resonance Imaging is where we get to look into the inside of the body. It is very useful for soft tissues. Doctors use it to detect disease of the brain, spine, joints, and other organs. High-field MRI machines like 7 Tesla make very clear images than low tesla images. They show small details and have good signal quality. But these machines cost a lot. They need special rooms and trained staff. In developed countries, these high-field MRI systems are available and widely used for advanced diagnosis. But in countries like Bangladesh, it is very rare to see this kind of machine used for its high cost and lack of infrastructure. It keeps many hospitals from being able to utilise them especially in Bangladesh. Many hospitals are running standard, low- or mid-field MRI machines like 3 Tesla. Those MRI images are not always very good and clear to diagnose. They can be noisy and have low contrast. Fine details may look blurry. This makes it harder for doctors to find problems.

This project's goal is to remove noise and improve contrast for clear and detailed MRI images. Important structures in the body must stay visible. This method does not need extra scans. Hospitals can use the machines they already have. Instead of replacing expensive MRI hardware, this project focuses on a software-based solution using deep learning to enhance image quality. Better images help doctors find tumors or injuries more clearly. It also helps with planning treatment. The study focuses on simple, affordable ways to improve MRI images. The idea is to make MRI more useful in everyday hospitals. It uses computer techniques to make images look closer to high-field quality. This can save time, cost, and effort. Patients get faster and more accurate diagnoses.

1.2 Aims, objectives and Motivation

1.2.1 Aims

The project aims to upgrade low-field MRI images using a software based approach. Images from low-field MRI scanners often suffer from noise. Contrast is also weak in many cases. These problems make medical analysis harder for doctors. The main goal is to reduce noise and improve contrast. At the same time important anatomical structures must remain clear. The proposed approach enhances image quality without modifying MRI hardware. No need to do any extra scanning.

1.2.2 Objective

The objectives of this project are as follows:

1. To learn the basic and deep knowledge on MRI enhancement using deep learning and analyze the relationship between MRI field strength and its impact on image quality and diagnostic performance.
2. To preprocess MRI datasets to ensure data consistency using techniques such as skull stripping, bias field correction, and intensity normalization.
3. To design and implement a 3D VNet-based deep learning model for MRI image enhancement.
4. To incorporate different types of loss functions in order to improve structural preservation and edge detail in the enhanced images.
5. To enhance the 3T MRI images to make them visually appear more like a 7T MRI image.
6. To evaluate the performance of the proposed model via quantitative metrics which are Peak Signal-to-Noise Ratio (PSNR) and Structural Similarity Index
7. To compare the enhanced MRI outputs with original images to estimate improvements in visual quality and structural accuracy.

1.2.3 Motivation

In many hospitals of our country high-quality MRI machines are not easy to get. Advanced high-field MRI systems are mostly limited to developed regions, making them less accessible in countries like ours. Because of this doctors have to use MRI images which are not very clear. Many times the images contain noise and low contrast, so small details are hard to see. When the image is not clear, diagnosis becomes difficult and it may affect proper treatment.

There are traditional image enhancement techniques already, but they do not always work well. Some methods only improve brightness or contrast but important body structure can be lost. In medical images this is a big problem because doctors need original information, not distorted images. So these methods are not enough in real clinical use. For this, we need a low cost and simple solution. Without replacing the high quality machine or adding extra hardware, enhancing MRI image quality is much more efficient and important. Therefore, a software-based deep learning approach is considered as a practical alternative. Using a deep learning approach, this study tries to enhance MRI images in a better way. Clearer images can help doctors understand patient conditions more easily and make better medical decisions.

1.3 Project Specification

The design specifications show the number of criteria for evaluating the performance and successfulness of the MRI image enhancement system. These specifications may evolve during development as part of an iterative refinement process.

1.3.1 Functional Specifications

- The system shall accept 3T MRI images as input
- The system shall generate enhanced 7T-like MRI images as output
- The model shall support both T1-weighted and T2-weighted MRI modalities

- The system shall preserve anatomical structures without introducing artificial features
- The system shall process 3D MRI volumes using a VNet-based architecture

1.3.2 Performance Specifications

- The system shall achieve a minimum PSNR improvement over input images
- The model shall achieve an SSIM value greater on test data
- The inference time per image shall be within acceptable computational limits
- The model shall maintain consistent performance across unseen test samples

1.3.3 Data and Training Specifications

- The model shall be trained using paired 3T–7T MRI datasets
- The dataset shall be divided into training, validation, and testing sets
- The model shall be trained using optimized hyperparameters within available GPU constraints
- Data preprocessing shall include skull stripping, bias correction, registration, and normalization

1.3.4 Evaluation and Validation Specifications

- Model performance shall be evaluated using PSNR and SSIM
- Visual comparison shall be used to assess structural preservation
- The model shall be validated on unseen data to ensure generalization

1.3.5 System Constraints

- The system shall be implemented on Google colab using Python and deep learning frameworks (e.g., PyTorch)
- Training shall be performed under GPU memory limitations
- The system shall use efficient memory handling techniques (e.g., patch-based training)
- The model complexity shall be limited to ensure feasibility within available computational resources

Chapter 2

Literature Review

Magnetic Resonance Imaging (MRI) enhancement has become a critical research area due to the inherent trade-off between image quality and accessibility. High-field MRI systems such as 7T provide superior spatial resolution as well as tissue contrast but their high cost and limited availability prohibits mass clinical use, leading to increasing interest in computational methods for enhancing lower-field MRI data [Sarracanie & Salameh \(2020\)](#). Recent advances in deep learning have significantly improved MRI reconstruction, denoising, and super-resolution, enabling software-based solutions to address these limitations [Z. Chen et al. \(2023\)](#).

Early deep learning approaches relied on convolutional neural networks (CNNs) and U-Net architectures for MRI super-resolution and enhancement. For example, [Kalluvila \(2023\)](#) demonstrated that U-Net-based models can effectively recover high-frequency details from low-resolution MRI, achieving improved PSNR values, though limitations remain in capturing long-range dependencies. Similarly, [Roy et al. \(2024\)](#) showed that CNN-based super-resolution models such as SRResNet outperform simpler architectures by improving structural fidelity and perceptual quality.

More advanced architectures have extended these approaches into volumetric processing. [Cui et al. \(2024b\)](#) proposed a 3D V-Net-based framework for synthesizing 7T MRI from 3T scans, demonstrating strong performance

improvements in both PSNR and SSIM. Likewise, [Acs & Zhuang \(2025\)](#) introduced a 3D FR-U-Net with multi-scale fusion and residual learning, achieving improved anatomical consistency in synthesized MRI volumes. These studies highlight the importance of 3D architectures in preserving volumetric anatomical structures.

Generative Adversarial Networks (GANs) have been widely adopted for MRI enhancement due to their ability to produce realistic textures. CycleGAN-based methods enable translation between MRI domains without paired datasets, making them practical for real-world scenarios. Studies such as the 3T-to-7T translation framework (2024) and [Kawahara & Nagata \(2021\)](#) demonstrate that GANs can enhance contrast and structural detail while maintaining anatomical consistency through cycle-consistency loss.

Further advancements include conditional GANs for contrast enhancement [Solak et al. \(2025\)](#) and LoHiResGAN for ultra-low-field MRI enhancement [Islam et al. \(2023\)](#), which integrate adversarial, L1, and SSIM losses to improve both realism and structural fidelity. Despite their success, GAN-based methods often face challenges such as potential artifact generation and reliance on high-quality training data.

More recently, diffusion-based models have emerged as powerful alternatives for MRI synthesis. [Jiang et al. \(2025b\)](#) introduced a frequency-aware diffusion model that incorporates wavelet decomposition to better preserve high-frequency details, achieving superior PSNR and SSIM compared to GAN-based approaches. Similarly, [J. Wang et al. \(2023\)](#) leveraged latent diffusion models for 3D MRI super-resolution, demonstrating high reconstruction fidelity but at the cost of significant computational complexity. These methods show improved realism and structural preservation, although their high computational requirements and reliance on large datasets remain limitations.

To address the scarcity of paired MRI datasets, several studies have explored unpaired and self-supervised learning techniques. [Zhang et al. \(2025b\)](#) proposed a Pretext Task Adversarial (PTA) framework that combines contrastive

learning and adversarial training to synthesize high-field MRI from unpaired data, achieving strong perceptual quality improvements. Similarly, [Islam et al. \(2025b\)](#) introduced an INR-based style transfer method that enhances ultra-low-field MRI using unpaired high-field references, balancing perceptual enhancement with structural consistency. These approaches reduce dependency on paired datasets and improve scalability, though maintaining strict anatomical fidelity remains a challenge.

Recent research has shifted toward multi-stage and hybrid frameworks that integrate multiple tasks within a single pipeline. [Shamaei et al. \(2025b\)](#) proposed a longitudinal MRI reconstruction framework combining reconstruction, registration, and transformer-based enhancement, achieving improved structural fidelity by leveraging prior scans. Similarly, [Sun et al. \(2025\)](#) introduced a foundation model (BME-X) that integrates denoising, super-resolution, and harmonization with tissue-aware learning, demonstrating strong generalization across datasets and downstream tasks.

Hybrid approaches combining physics-based modeling with deep learning have also shown promise. [Z. Wang et al. \(2025b\)](#) demonstrated that physics-informed synthetic data generation improves model generalizability across different acquisition settings, addressing limitations of data scarcity and domain shifts.

MRI enhancement is also closely tied to denoising and preprocessing strategies. [Xue et al. \(2025\)](#) proposed SNRAware training, incorporating SNR information and g-factor maps to improve denoising performance across diverse datasets. [Jeme V & Jerome S \(2024a\)](#) demonstrated that combining advanced denoising, contrast enhancement, and attention-based CNNs significantly improves segmentation performance.

Attention mechanisms have further improved enhancement quality by enabling models to focus on relevant anatomical features. [B. M. Li et al. \(2022\)](#) showed that attention-based super-resolution models outperform traditional

methods in clinical MRI data, achieving high PSNR and SSIM values.

Despite significant advancements, several challenges remain in MRI enhancement research. Many models rely on limited datasets or single-center data, which restricts generalizability across scanners and patient populations [Y. Li et al. \(2024\)](#); [Islam et al. \(2023\)](#). Additionally, maintaining anatomical fidelity while enhancing image quality remains a critical concern, particularly in unpaired and generative models.

Another important challenge is computational complexity, especially in diffusion-based and multi-stage frameworks, which may limit real-time clinical applicability. Furthermore, evaluation metrics for example PSNR and SSIM are widely used, and do not always reflect clinical relevance, highlighting the need for task-based and radiologist-driven validation.

2.1 Literature Review Table

Table 2.1: Literature Review Analysis Table

Paper	Year	Algorithm/Model	Result	Limitation
Shamaei et al. (2025b)	2025	Transformer, Unet, E2EVarNet	SSIM: 0.9457 PSNR: 34.73 dB NMSE:0.0051	Small Dataset, Healthy subjects only, focused solely on T1-weighted brain images, Two experts evaluated six subjects
Zhang et al. (2025b)	2025	CycleGAN-like unpaired translation model with pretext-task supervision.	FID:16.892 IS:1.933 MS-SSIM ↓:0.324	Generalization challenges Increased training complexity
Islam et al. (2025b)	2025	Baseline (64mT MRI), N2N, SIREN, ULF MRI	PIQE: 50.71 CNR: 5.01 MLC: 0.9874 MSLC: 0.2090	2D network, Unpaired setting
Acs & Zhuang (2025)	2025	Semi-supervised 3D multi-scale fusion residual U-Net	SSIM: 0.7368 PSNR: 23.254 NMSE: 0.3616	small-volume paired datasets lack of validation on pathological cases
Z. Wang et al. (2025b)	2025	PISF framework (Physics-Informed Synthetic data learning framework)	Improves PSNR by 8% over the plain version	Performance depends on realism of synthetic data. May not fully capture complex real-world scanner and patient variability. Not validated through large-scale prospective clinical studies

Paper	Year	Algorithm/Model	Result	Limitation
Xue et al. (2025)	2025	SNRAware training scheme	PSNR = 54.90 dB SSIM: 0.70	Requires noise calibration data (noise scans) for SNR unit scaling, which many public MRI datasets lack. Generalization tested on a limited set of sequences and anatomies; broader evaluation needed.
Solak et al. (2025)	2024	CycleGAN (Cycle-consistent GAN)	MSE = 0.0041 SSIM = 0.7 (0.6–0.8) PSNR = 25.65 (22.23–27.22)dB RMSE = 2.1 (1.8–2.3)	The method relies on GAN-generated synthetic contrast images, which may introduce hallucinated or non-clinical features.
Jiang et al. (2025b)	2025	FADM (Frequency-Aware Diffusion Model)	PSNR = 31.897 $\pm$ 1.342 SSIM = 0.951 $\pm$ 0.008 FID = 33.304	Computational load is higher than simple GAN models
Y. Li et al. (2024)	2024	VMamba and Transformer	BRISQUE $min \approx 3.1084$ PIQE $mean \approx 13.122$	Not validated in real-time MRI Limited generalization to variable MRI protocols Needs evaluation by radiologists Training bias (dataset-dependent)

Paper	Year	Algorithm/Model	Result	Limitation
Islam et al. (2023)	2024	CycleGan with ResNet generators (3, 7, 9 residual blocks) and PatchGAN/PixelGAN discriminators	best model is resnetgen patchGAN SSIM =83.80, psnr = 26.25, NMSE= 0.0088, NMAE= 0.0630	Limited to T1-weighted axial MRI and small dataset (10 subjects); does not handle other orientations or T2-weighted images; no explicit lesion modeling; performance depends on skull stripping and may not generalize to other datasets or clinical scenarios
Islam et al. (2025b)	2024	An anisotropic U-Net variant designed for the Image Quality Transfer (IQT) inverse problem.	$PSNR \approx 26.84$ $SSIM \approx 0.796$	does not correct global image artefacts or distortions for being patch-based. Model training relies on simulated low-field data which may limit generalisability to unseen scanner/protocol variations.
Cui et al. (2024b)	2024	Multiple custom 3D V-Net variants	TBI dataset, SSIM 0.914 ± 0.016 PSNR 25.60 ± 0.77 , WATNet (SSIM = 0.881 ± 0.008 , PSNR = 23.60 ± 0.37); perceptual V-Net (SSIM ≈ 0.908 , PSNR ≈ 25.51) V-Net-GAN (SSIM ≈ 0.899 , PSNR ≈ 24.33)	Needs more validation including formal assessment of hallucinations and uncertainty; does not yet handle multiple MR contrasts jointly

Paper	Year	Algorithm/Model	Result	Limitation
Jeme V & Jerome S (2024a)	2024	An Attention Module-based Convolutional Neural Network (AM-CNN) for skull stripping.	Mean Dice <i>coefficient</i> ≈ 0.998 <i>Sensitivity</i> ≈ 0.9975 <i>Specificity</i> ≈ 0.9985	Validated on a limited dataset, so generalisation to varied scanners or patient populations may be uncertain. Pre-processing and enhancement steps add complexity and may be sensitive to parameter settings or noise types.
Kalluvila (2023)	2023	3D Brain Latent Diffusion Model (LDM) with two SR variants:a) InverseSR(LDM) b)InverseSR (Decoder) Uses a pretrained 3D Latent Diffusion Model	Slice 4mm: SSIM=0.803 ± 0.030 , PSNR=29.64 ± 1.64 Slice 8mm: SSIM= 0.754 ± 0.038 , PSNR= 27.92 ± 1.60	Requires large computational resources due to long DDIM sampling chains; optimization is slow because each gradient step needs a full diffusion pass; reconstruction quality limited by LDM generator capacity and output variability.
J. Wang et al. (2023)	2022	Critic-guided deep attention super-resolution framework based on U-Net + Attention Gates + a separate Critic network	$MAE = 3.78310^{-3}$ $NMSE = 4.3210^{-10}$ PSNR = 35.39 SSIM = 0.9852.	Small paired dataset- only 61 patients, paired LR/HR MRIs Critic overfitting hence a high dropout rate of 0.5 is needed. Warp registration underperforms — Warping consistently underperformed s.

Paper	Year	Algorithm/Model	Result	Limitation
B. M. Li et al. (2022)	2022	Modified 2D U-Net (19 convolutional layers; dropout; tanh output activation; upsampling instead of deconvolution)	T1 $\rightarrow$ T2: PSNR = 29.45 ± 1.72 , SSIM=0.932, MAE=0.0124, MSE=0.0012	1. The model is 2D due to limited computational resources. 2. Synthetic FLAIR images show lower quality — loss of details, blurring, imperfections 3. Slight underestimation of intensities across translations
Kawahara & Nagata (2021)	2021	CGAN with a U-Net-style generator and discriminator (GAN framework)	T2 from T1: rMAE: 0.15 rMSE: 0.07 rRMSE: 0.24 PSNR: 58.53 T1 from T2: rMAE: 0.17 rMSE: 0.10 rRMSE: 0.31 PSNR: 56.90 MI: 1.15	Only healthy anatomy, 2D slices, limited scanner/-generalization, possible synthesis errors

Table 2.2: Result Matrix Analysis Matrices across Different Studies

Literature	PSNR	SSIM	NMSE	MSE / RMSE	Dice	Sensitivity	Specificity	FID / IS
Shamaei et al. (2025b)	✓	✓	✓					
Zhang et al. (2025b)		✓						✓
Islam et al. (2025b)								
Diniz et al. (2025)					✓			
Sun et al. (2025)	✓	✓	✓					
Z. Wang et al. (2025b)	✓							
Xue et al. (2025)	✓	✓						
Jeme V & Jerome S (2024a)	✓	✓		✓				
Jiang et al. (2025b)	✓	✓						✓
Y. Li et al. (2024)								
Jeme V & Jerome S (2024b)	✓	✓			✓	✓	✓	
Islam et al. (2023)	✓	✓	✓					
Islam et al. (2025b)	✓	✓						
Asiri et al. (2024)					✓	✓	✓	
Cui et al. (2024b)	✓	✓						
Rasheed et al. (2023)					✓	✓		
Jeme V & Jerome S (2024a)	✓							
Kalluvila (2023)	✓	✓						
J. Wang et al. (2023)	✓	✓	✓					
B. M. Li et al. (2022)	✓	✓		✓				
Kawahara & Nagata (2021)	✓			✓				

Chapter 3

Methodology

To develop the enhancement model a Supervised deep learning frame-work is proposed to convert low-field (3 Tesla) MRI images into high-field (7 Tesla) quality using a paired dataset. The model handles 3D MRI volumetric data for synthesizing high quality MRI images. The methodology consists of dataset selection, requirement analysis, preprocessing, model selection, model optimization, training, and evaluation. The workflow of this project is shown with a UML diagram.

3.1 Dataset collection

The dataset used in the study is published in Scientific Data. [X. Chen et al. \(2023\)](#) was created to support research on MRI enhancement, image synthesis, and cross-field strength translation, particularly for developing and evaluating deep learning models that aim to generate high-field MRI images from lower-field scans. The dataset consists of brain MRI scans from 10 healthy adult subjects, where each subject was scanned at both 3 Tesla (3T) and 7 Tesla (7T) magnetic field strengths. For every participant, two commonly used structural MRI contrasts were acquired: T1-weighted (T1w) and T2-weighted (T2w) images. As a result, the dataset provides us with four image volumes for every patient. They are 3T-T1w, 3T-T2w, 7T-T1w, and 7T-T2w . With these we get a fully paired dataset in both contrast type and field strength. 3T MRI images were acquired using Siemens Magnetom Prisma scanner and

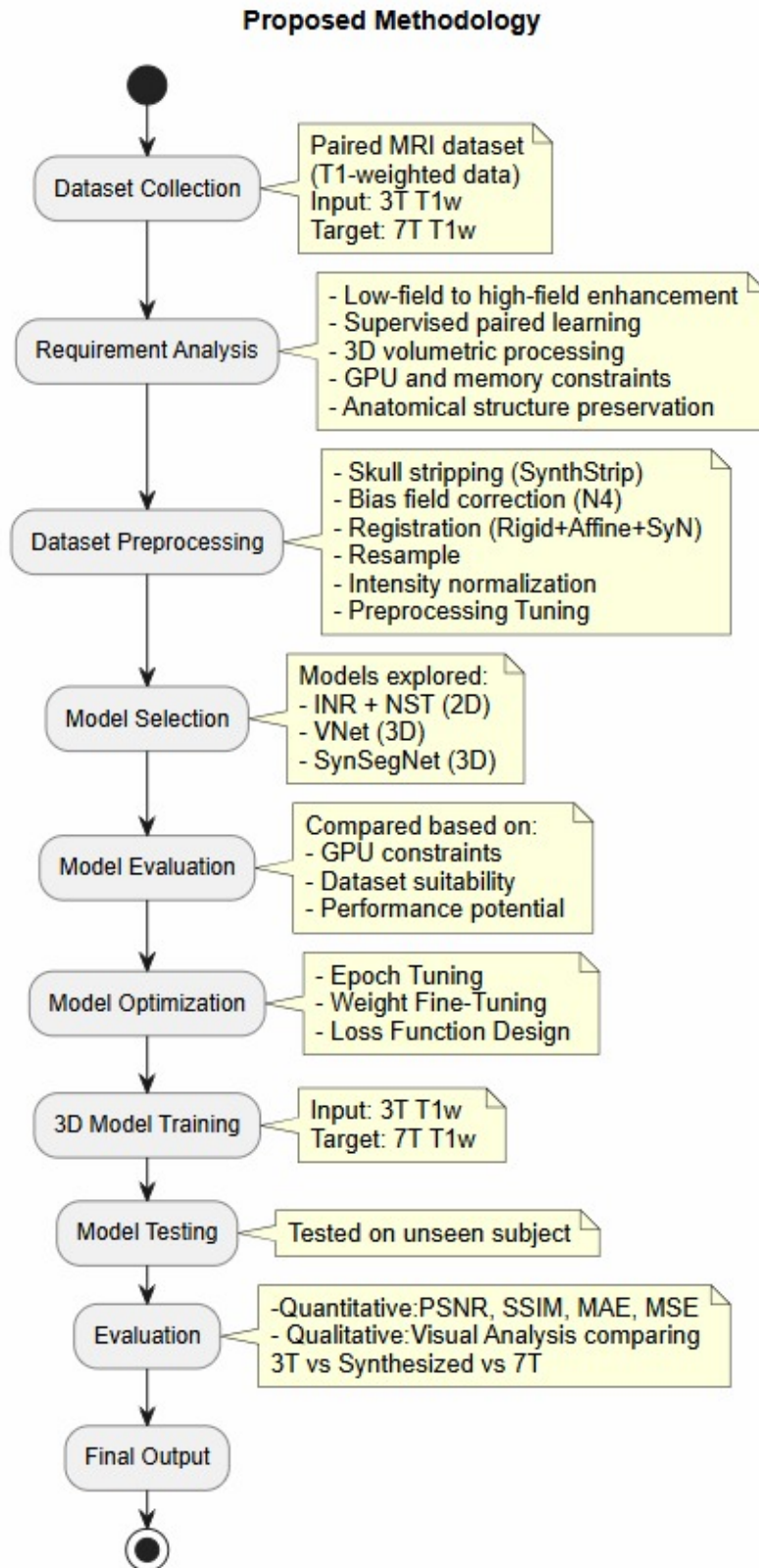


Figure 3.1: Methodology

7T MRI images were acquired using a Siemens Magnetom Terra Scanner. After acquisition, the 3D volumetric images were converted from DICOM to NIfTI format and were organized in the Brain Imaging Data Structure (BIDS) standard. This standardized organization improves data accessibility and reproducibility. Only Facial features were removed from all images protecting participant privacy and they also performed linear registration in all modalities for each subject to be spatially aligned so that it enables voxel-wise comparison between 3T and 7T scans. Image quality of this dataset was assessed using MRIQC with signal-to-noise ratio (SNR) and contrast-to-noise ratio (CNR) metrics. The reported quality measures confirm that the scans show good tissue contrast and structural clarity. Overall, this dataset serves as a valuable resource for developing and validating MRI synthesis methods which very closely aligns with our project goal of translating 3T MRI into high-quality 7T images.

3.2 Data Preprocessing

The paired dataset of Magnetic Resonance Imaging (MRI) data contain non-brain tissues, intensity inhomogeneity, inter-session misalignment, and inconsistent spatial dimensions. These issues can negatively affect downstream tasks. Therefore, a preprocessing pipeline was applied to ensure data consistency, improve image quality, and enhance model strength. The preprocessing steps applied in this project include:

- Skull stripping
- Bias field correction
- Registration (Rigid +Affine+SyN)
- Resample
- Intensity normalization

3.2.1 Skull Stripping

Skull stripping was performed using SynthStrip, a deep learning-based skull stripping method developed by FreeSurfer. SynthStrip is modality-agnostic and suitable for both T1-weighted and T2-weighted MRI scans. Raw MRI scans contain non-brain tissues such as the skull, scalp, fat, and neck regions. These irrelevant structures can introduce noise into learning-based models and reduce enhancement accuracy. Removing non-brain tissue ensures that subsequent processing focuses exclusively on brain anatomy.

Output:

- Brain-extracted image
- Binary brain mask

Visualization: Figure X shows representative T1-weighted and T2-weighted images before and after skull stripping. The results demonstrate complete removal of non-brain tissues while preserving cortical and subcortical brain structures. (Reference: Figure X – Before vs After Skull Stripping)

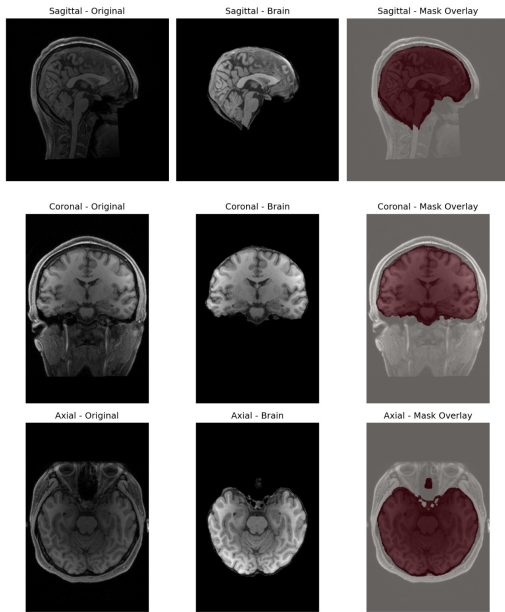


Figure 3.2: Skull Stripping of 3T MRI images

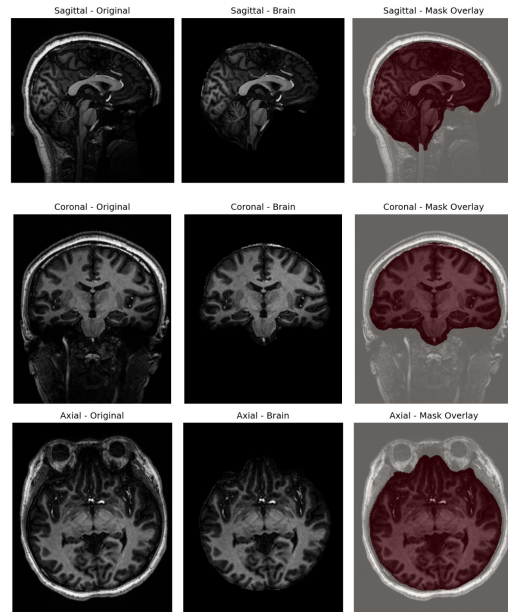


Figure 3.3: Skull Stripping of 7T MRI images

3.2.2 Bias Field Correction (N4 Correction)

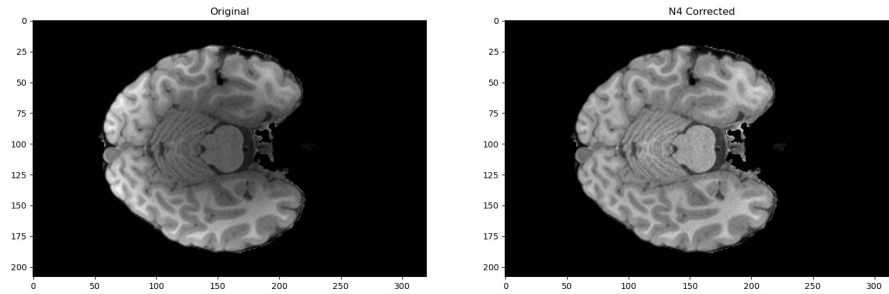


Figure 3.4: Bias Field Correction

MRI intensity values may vary smoothly across the image due to scanner-related artifacts rather than actual tissue differences. Hence, Bias field correction was applied using the N4ITK algorithm, which corrects low-frequency intensity non-uniformities caused by magnetic field inhomogeneities to remove the slow-varying b1 magnetic field effects that could distort the images and facilitate further brain extraction and scan alignment.

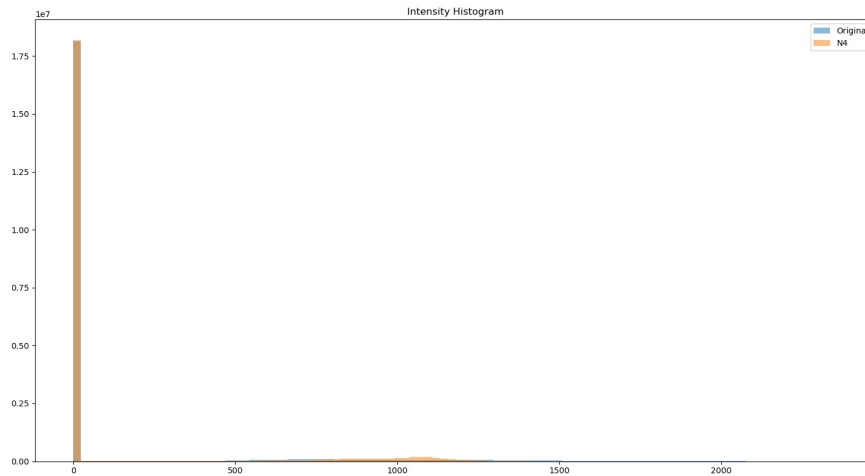


Figure 3.5: Bias Field Correction Graph

3.2.3 Registration

Although the dataset is paired, perfect alignment is not guaranteed due to differences in scanning sessions and equipment. Inter-session MRI scans differ in head orientation and position. Registration ensures Anatomical correspondence across subjects, reliable voxel-wise comparison. Therefore, a three-stage registration pipeline was applied :

- Rigid Registration: The first stage of registration was rigid transformation, which corrects global spatial differences. This includes Translation (shifting along x , y , z axes) and Rotation meaning orientation correction. Rigid registration makes sure overall brain structures alignment without altering shape or size. This step is essential for establishing an initial coarse alignment while preserving anatomical integrity.
- Affine Registration: After rigid alignment, affine registration is applied to refine the alignment by adding transformation parameters. After translation and rotation, affine transformation adds Scaling (size adjustment) and Shearing (skew correction) which corrects differences in voxel spacing and geometric distortions between the 3T and 7T MRI scans.
- Symmetric Normalization (SyN): Even after affine alignment, small local mismatches were seen due to anatomical variability and scanner-induced distortions. To address this, Symmetric Normalization (SyN), a non-linear deformable registration method was applied. SyN performs voxel-level alignment and ensures precise matching of fine anatomical structures. Also, Alignment of cortical folds and subcortical regions and Correction of local distortions are corrected by SyN.

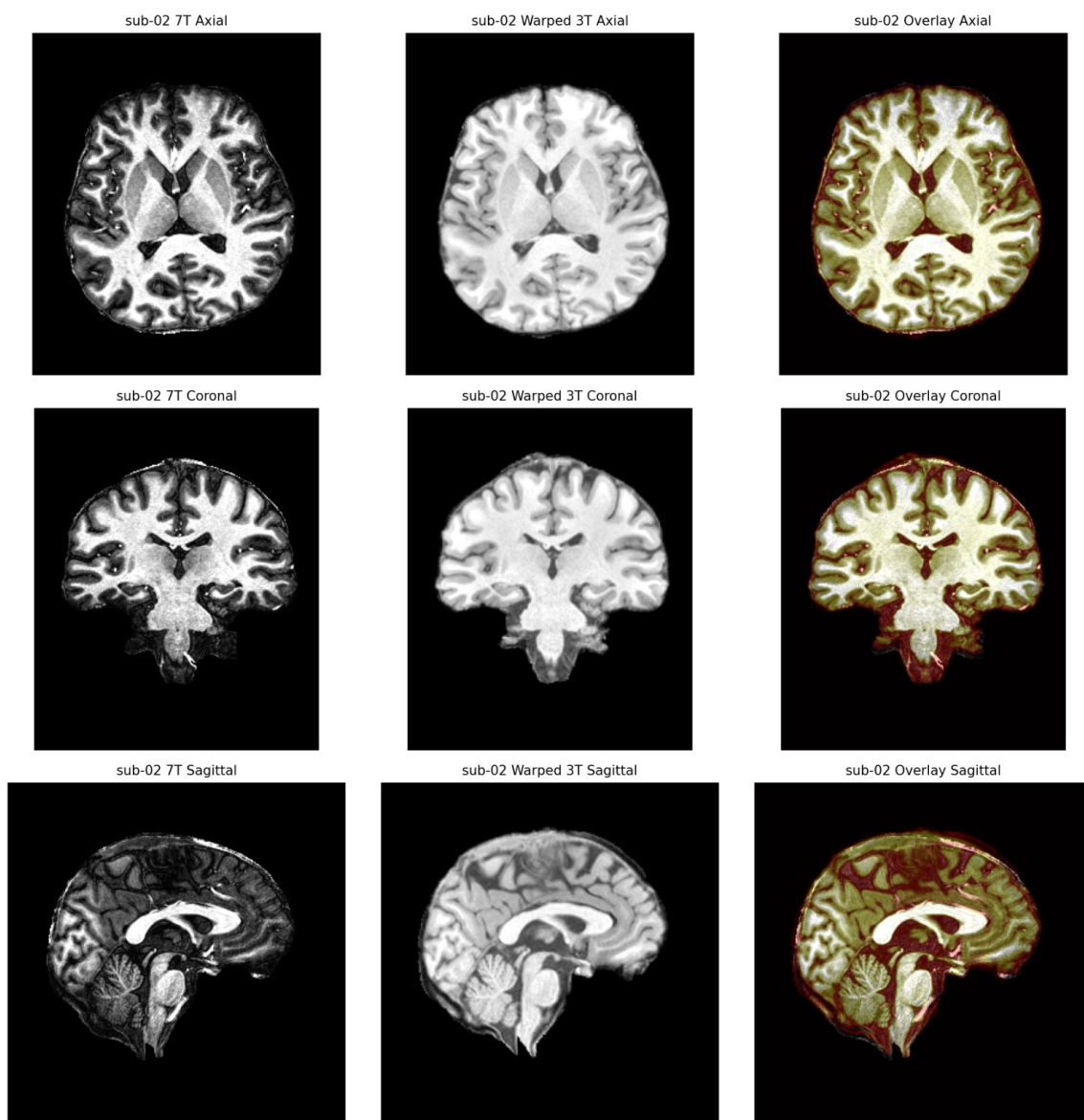


Figure 3.6: Registration of Pairs

3.2.4 Resampling

After the registration process, the MRI volumes still differ in spatial resolution and voxel spacing due to different acquisition protocols and scanner configurations. For pairwise training, it requires standardized input dimensions. For this, all MRI volumes were resampled to a uniform spatial resolution and voxel spacing. Resampling interpolates the image data to a fixed grid

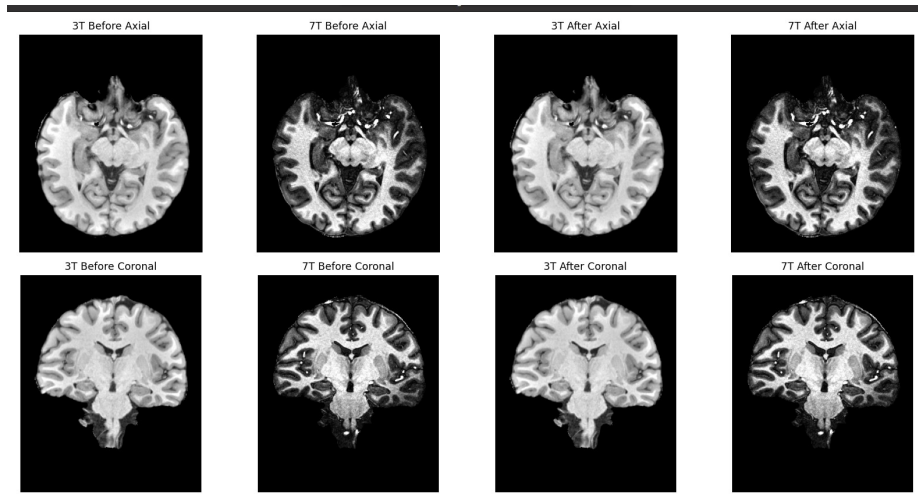


Figure 3.7: Transformed and Resampled

3.2.5 Intensity Normalization

Intensity normalization reduces inter-subject and inter-session intensity variability, stabilizes model training and Improves convergence and generalization. That's why we applied Z-score normalization to all non-zero brain voxels.

$$I_{norm} = \frac{I - \mu}{\sigma}$$

where μ and σ represent the mean and standard deviation of brain voxel intensities. The intensity values of the MRI scan sequences were in the range $[0, 1]$.

Preprocessing Optimization Multiple preprocessing configurations were tested

and refined to achieve Better alignment, intensity consistency and Minimal information loss This tuning step was crucial for achieving near ideal input data for the model.

3.3 Model Selection

As we have quite a few project constraints such as GPU limitations, dataset size, and task complexity, three different model approaches were explored. They include:

- Implicit Neural Representation (INR) with Neural Style Transfer (NST) (2D)
- VNet-based 3D model
- SynSegNet (3D)

Each model was evaluated depending on feasibility, computational cost, and expected performance for the MRI enhancement task. The INR and NST approach was a 2D model and could not fully capture volumetric spatial dependencies. On the other hand, SynSegNet model handles 3D MRI scans but has higher computational demands and complexity relative to available resources. In contrast, the VNet-based model showed an effective balance between performance and efficiency. It is specifically designed for volumetric medical imaging and is capable of capturing spatial context across 3D MRI volumes. Therefore, the VNet architecture was selected as the primary model for this study for its suitability for the task, computational feasibility, and potential for high-quality MRI enhancement.

3.4 Model Optimization and Training

After the selection of VNet architecture, a series of systematic experiments were conducted to optimize the model's performance. The goal that was considered here was to ensure a stable training, improve convergence, and enhance the quality of the generated 7T MRI images . Multiple aspects of the training process like epoch selection, weight optimization, and loss function implementation were carefully analyzed and refined.

3.4.1 Epoch Tuning

The dataset used in this project has only 10 subjects: 3D MRI scans. The referenced paper (paper citation) used 300 epoch for their bigger dataset. As our dataset is much smaller and subject to overfitting, Epoch tuning was one of our main consideration. The number of training epochs plays a crucial role in model performance. Training for too few epochs can result in underfitting, where the model fails to learn meaningful patterns, while excessive training can lead to overfitting, where the model memorizes the training data and performs poorly on unseen data. For this reason, multiple training runs were conducted with varying epoch values. The model performance was continuously monitored using evaluation metrics, and the optimal number of epochs was selected based on the similarity between ground truth and predicted image.

3.4.2 Loss Function Design

After that, with the goal of enhancing the models performance, a composite loss function was designed by combining multiple loss functions. The baseline loss function was L1 loss to measure the pixel-wise absolute difference between the predicted image and the ground truth. Alone L1 loss alone does not adequately capture structural and perceptual loss of the image. Therefore, additional loss functions were added. The Structural Similarity Index (SSIM) loss was included to improve perceptual quality to focus on structural information of image like texture and contrast. Also, a gradient loss was added to preserve edge information and fine anatomical details. . The combination of these loss functions helps the model to generate sharper images with improved structural consistency compared to using a single loss function.

3.4.3 3D Training Strategy

The model was trained using full 3D MRI volumes so that it can capture spatial data across all slices rather than training slice wise to learn each slice independently. This approach helps the network to learn volumetric features and better understanding of anatomical continuity within the brain. 8 subjects were trained on the selected model using preprocessed MRI volume with T1w volumes to enhance MRI image quality by mapping lower-quality inputs to higher-quality outputs using 3D

volumetric learning. This training strategy leverages the advantages of 3D MRI scans.

3.4.4 Optimizer Selection

During training, different optimization techniques were explored to improve model performance and ensure stable learning. Initially, the model was trained using the Adam optimizer. Later, during the fine-tuning stage, the optimization strategy was changed to the AdamW optimizer.

The optimizer was configured with: Weight decay = 1×10^{-6}

AdamW is an improved version of the Adam optimizer that combines adaptive learning with weight regularization. It automatically adjusts the learning process for different parameters, helping the model learn important image features more effectively. AdamW also applies weight decay separately during training, which helps reduce overfitting and improves performance on unseen MRI data. Switching from Adam to AdamW during fine-tuning helped the model train more smoothly and improved the quality of the reconstructed images. It also improved structural consistency and preserved finer anatomical details in the generated MRI images. Therefore, AdamW with weight decay regularization was selected for the final stage of the proposed MRI enhancement framework.

3.5 Testing

Model performance is tested on one unseen subject to assess its generalization and assess capability. This subject is not included while we did the training to ensure an unbiased evaluation of the enhancement performance on new data. By using completely unseen data, the testing phase provides a realistic assessment of how well the model can generalize beyond the training dataset

3.6 Evaluation

The generated MRI images are evaluated using both quantitative and qualitative methods.

3.6.1 Quantitative Evaluation

For Quantitative evaluation , standard image quality metrics were employed the metrics include Peak Signal-to-Noise Ratio (PSNR), Structural Similarity Index (SSIM), Mean Absolute Error (MAE), and Mean Squared Error (MSE). These metrics measure the similarity between the synthesized 7T images and the ground truth images. PSNR evaluates reconstruction quality, SSIM measures structural similarity. Then MAE and MSE calculate pixel-wise differences between input and target.

3.6.2 Qualitative Evaluation

Besides evaluation performance with metrics, , qualitative evaluation was performed through visual inspection of the generated images. Only depending on the metrics value does not ensure the predicted images correctness . So, this step is essential in medical imaging. The synthesized images were compared with both the original 3T inputs and the ground truth 7T images. Key aspects such as structural preservation, tissue clarity, contrast enhancement, and overall visual quality were carefully analyzed. This comparison helps verify whether the model successfully enhances image quality while maintaining anatomical integrity.

Chapter 4

Design and Development

The proposed model, VnetLS is designed as a supervised 3D MRI enhancement framework that translates low-field 3T T1-weighted MRI into high-field 7T-like T1-weighted MRI. It uses paired 3T and 7T MRI volumes, where the 3T image acts as the input and the corresponding 7T image acts as the ground truth target. A brain mask is also used during training and evaluation to ensure that the model focuses only on brain regions and ignores irrelevant background voxels. The overall architecture of the system can be divided into five major modules: data preparation module, data loading module, VNet enhancement module, training and optimization module, and evaluation module.

4.1 Data Preparation Module

At the beginning of the system, the preprocessed MRI data are organized into three separate directories: one for 3T T1w input images, one for 7T T1w target images, and one for the corresponding brain masks. Each subject has one 3T image, one 7T image, and one mask. The filename structure is kept consistent so that the system can correctly pair the input, target, and mask for each subject.

$$(X_{3T}^i, Y_{7T}^i, M^i) \quad (4.1)$$

where X_{3T}^i represents the preprocessed 3T T1w image of subject i , Y_{7T}^i represents the corresponding 7T T1w ground truth image, and M^i represents the brain mask.

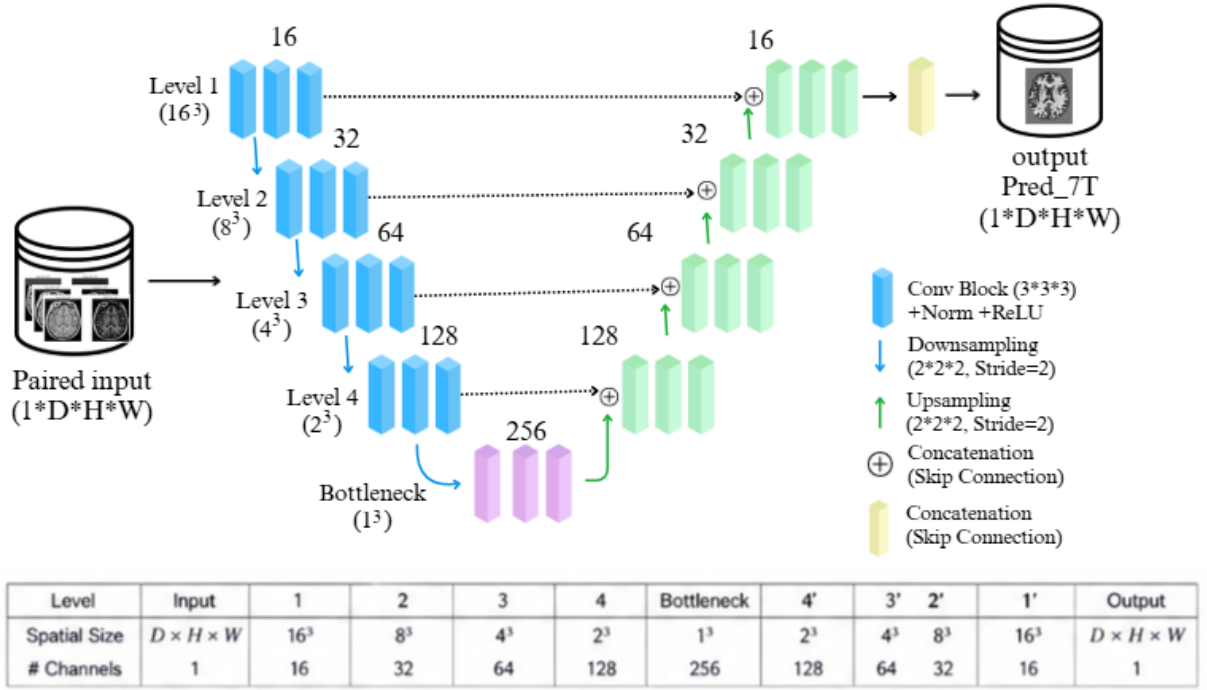


Figure 4.1: Architecture of VnetLS

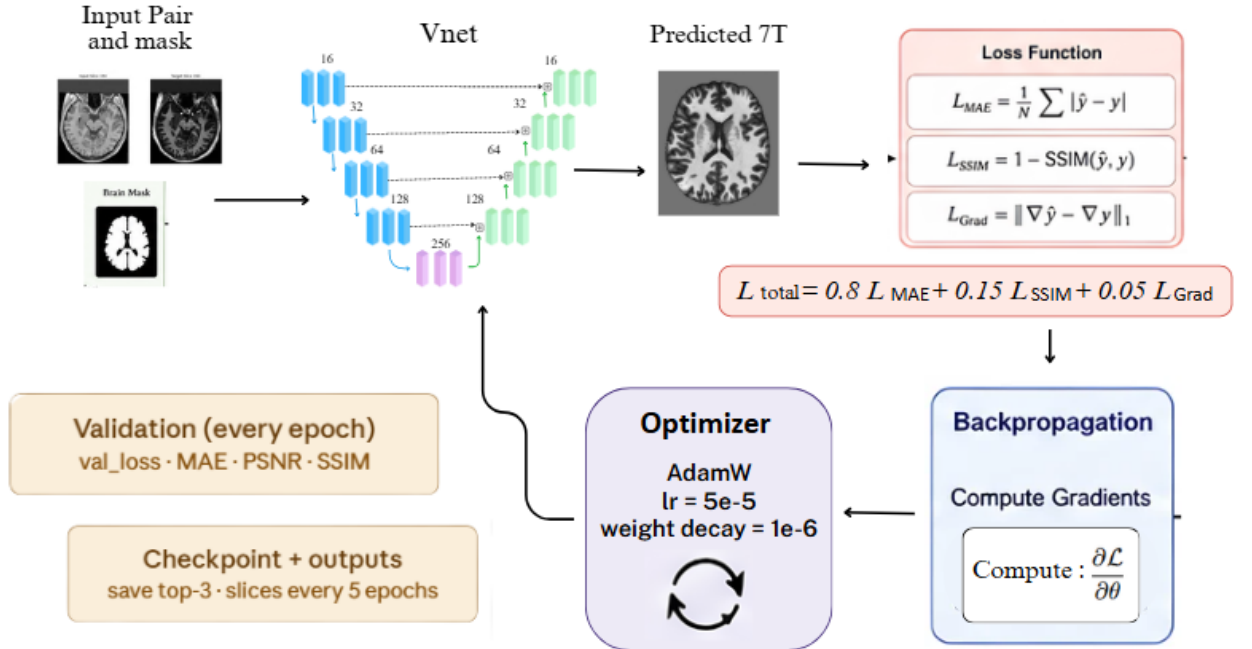


Figure 4.2: VnetLS pipeline

4.2 Data Loading and Padding Module

The system uses a custom MRI dataset loader to read the `.nii.gz` MRI volumes using NiBabel. Each 3D volume is loaded and converted into a PyTorch tensor. Since the model operates with 3D convolutional layers, the input is arranged in the form:

$$(C, D, H, W)$$

where C is the channel dimension, D is depth, H is height, and W is width.

A padding function is applied to ensure that each volume dimension is divisible by 16. This is necessary because VNet employs multiple downsampling and upsampling operations, and inconsistent dimensions can lead to shape mismatches during decoding.

The padding step ensures:

$$D, H, W \equiv 0 \pmod{16}$$

This guarantees that the MRI volumes remain compatible with the encoder–decoder structure of the model.

4.3 Dataset Splitting

The dataset is split subject-wise into training, validation, and testing sets. In the implemented system, 8 subjects are used for training, 1 subject is used for validation, and 1 subject is used for testing.

4.3.1 VNet-Based Enhancement Module

The core enhancement model is a 3D VNet architecture. The VNet receives the preprocessed 3T T1w MRI volume as input and generates a synthesized 7T-like MRI volume as output. VNet is suitable for this task because it processes MRI data as a full 3D volume instead of isolated 2D slices. This allows the network to learn spatial relationships across slices and preserve volumetric anatomical continuity.

4.3.2 Loss Function Design

The implemented system uses a combined loss function consisting of:

- Masked Mean Absolute Error (L1) Loss
- Structural Similarity Index Measure (SSIM) Loss
- Gradient-based Loss

The total loss is defined as:

$$\mathcal{L}_{total} = 0.8 \mathcal{L}_{MAE} + 0.15 \mathcal{L}_{SSIM} + 0.05 \mathcal{L}_{grad} \quad (4.2)$$

Masked MAE (L1) Loss

$$\mathcal{L}_{MAE} = \frac{\sum |\hat{Y}_{masked} - Y_{masked}|}{\sum M + \epsilon} \quad (4.3)$$

This loss measures the voxel-wise intensity difference only within the brain region.

SSIM Loss

Since normalized MRI intensities may contain negative values, the inputs are first scaled into a valid range:

$$X_{sim} = \text{clamp} \left(\frac{X + 4}{8}, 0, 1 \right) \quad (4.4)$$

The SSIM loss is then computed as:

$$\mathcal{L}_{SSIM} = 1 - SSIM(\hat{Y}, Y) \quad (4.5)$$

This loss encourages structural similarity between the synthesized image and the ground truth 7T image.

Gradient Loss

$$\mathcal{L}_{grad} = \frac{1}{3} (|\nabla_x \hat{Y} - \nabla_x Y| + |\nabla_y \hat{Y} - \nabla_y Y| + |\nabla_z \hat{Y} - \nabla_z Y|) \quad (4.6)$$

This term preserves edges, tissue boundaries, and fine anatomical structures.

4.4 Training and Optimization Module

The model is trained using the AdamW optimizer with learning rate and weight decay values loaded from the VNet configuration file. Training is performed using PyTorch Lightning.

- The input 3T volume is fed into the VNet model.
- The network generates a synthesized 7T-like output.
- A brain mask is applied to restrict computation to relevant regions.
- The total loss is computed using the combined loss function.
- Model parameters are updated via backpropagation.

4.5 Evaluation Module

After training, the model is evaluated using one unseen test subject. The predicted 7T-like output is compared with the real 7T ground truth. The main quantitative metrics are:

$$MSE = \frac{1}{N} \sum_{i=1}^N (\hat{Y}_i - Y_i)^2 \quad (4.7)$$

$$MAE = \frac{1}{N} \sum_{i=1}^N |\hat{Y}_i - Y_i| \quad (4.8)$$

$$PSNR = 10 \log_{10} \left(\frac{MAX^2}{MSE} \right) \quad (4.9)$$

$$SSIM = SSIM(\hat{Y}, Y) \quad (4.10)$$

Qualitative evaluation is performed by visually comparing the input 3T image, generated 7T-like image, and ground truth 7T image.

4.6 Website Implementation

The web application is implemented to make the 3T to 7T enhancement process easier to the user. In frontend React.js and Tailwind CSS is used for the interface. Ngrock is used from fast Api to call model api from colab. After fetching the uploaded .nii file, the API activate the peprocessing pipeline. After a successfull preprocessing of the image the predicted 7T images are generated and showed in the web application.

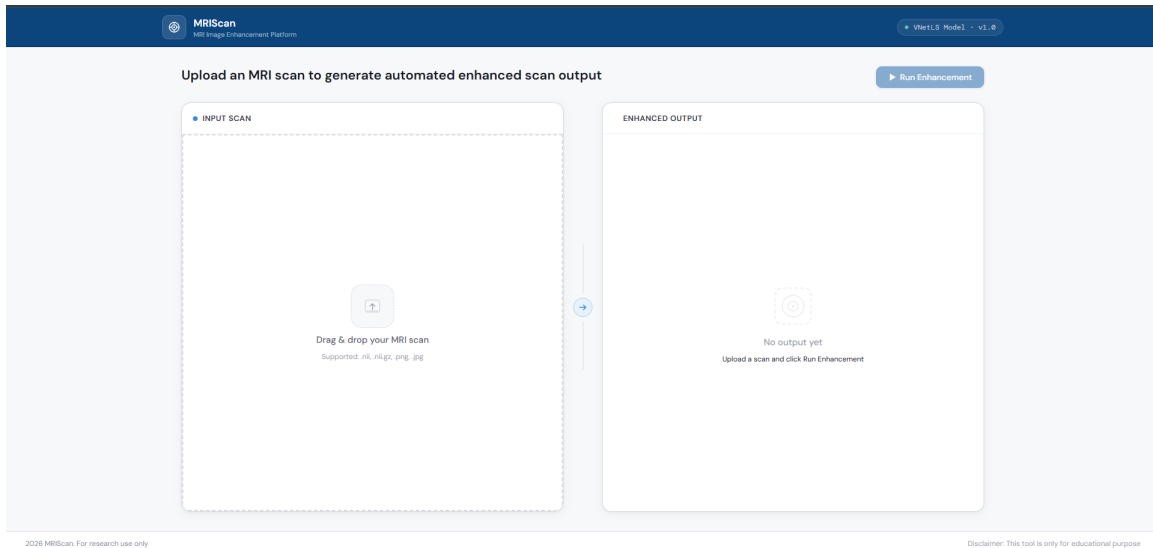


Figure 4.3: User Interface

The user uploads a NIfTI file via UploadPanel.jsx. App.jsx manages state and calls generate7TMRI() in mriApi.js, which fires two parallel fetch requests, POST predict and POST preview through a ngrock HTTPS proxy to the FastAPI server running on Google Colab. The VNetLS PyTorch model runs inference, nibabel handles NIfTI I/O, and matplotlib renders the middle axial slice as a PNG. Predict returns the full .nii.gz for download; Preview returns the PNG displayed immediately in OutputPanel.jsx with zoom and pan controls.

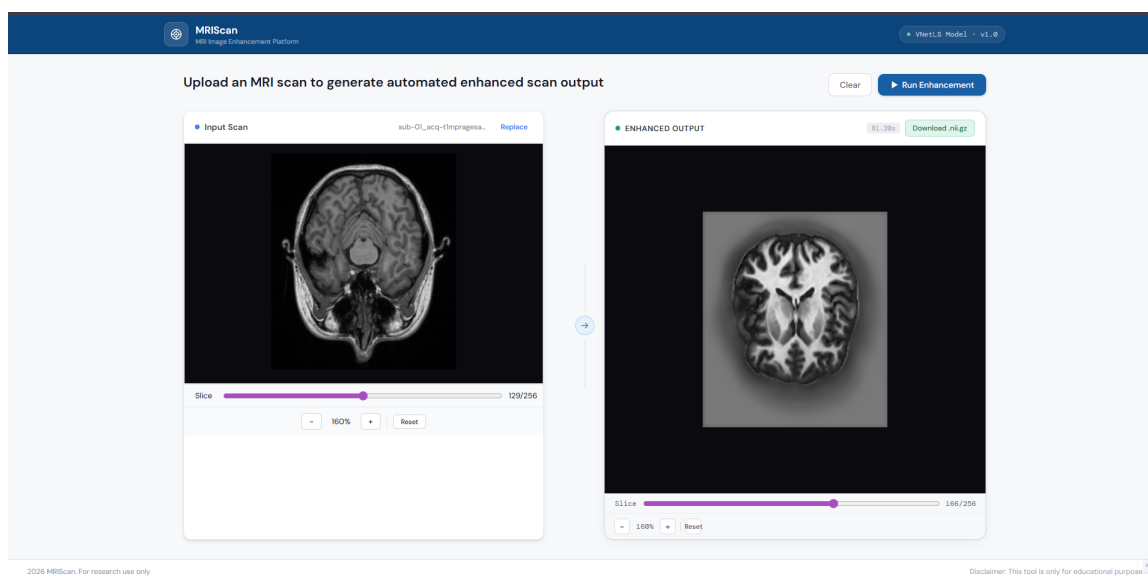


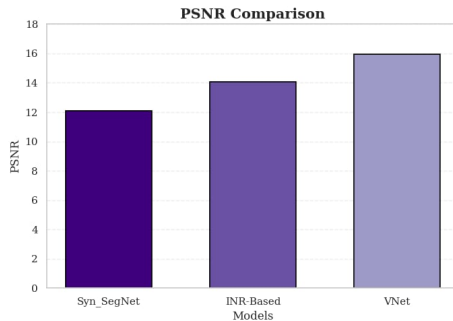
Figure 4.4: Showing The 3T Input & 7T Output

Chapter 5

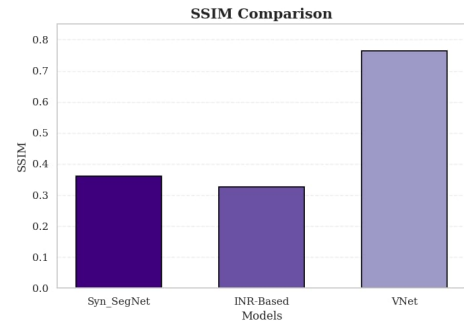
Result analysis

This chapter presents the results obtained from the proposed MRI enhancement system and provides a detailed analysis of the model performance. The evaluation is conducted using both quantitative metrics and qualitative visual assessment to determine how effectively the model enhances low-field (3T) MRI images into high-field (7T) images.

5.1 Model Selection and Performance Analysis



(a) PSNR Comparison Between Different Model



(b) SSIM Comparison Between Different Model

During the model selection phase, three different models: SynSegNet, INR-based model, and VNet were implemented and evaluated based on computational feasibility and suitability for 3D volumetric data. The performance of each model was assessed using standard image quality metrics, particularly PSNR and SSIM, and the results were visualized through graphs for analysis. From the obtained

results, SynSegNet achieved a PSNR of 12.07 and SSIM of 0.36, while the INR-based model produced a slightly improved PSNR of 14.08 but a lower SSIM of 0.327, indicating weaker structural preservation. In contrast, the VNet model significantly outperformed both alternatives, achieving the highest PSNR of 15.95 and higher SSIM of 0.7643. These results clearly indicate that while the INR-based model improves intensity reconstruction, it lacks consistency in preserving anatomical structure, and SynSegNet underperforms in both aspects. Additionally, considering the predefined selection criteria such as efficient handling of 3D volumetric data, GPU memory constraints, and overall enhancement capability, the VNet model provided the best balance between performance and computational feasibility. Therefore, based on both evaluation and practical considerations, the VNet-based 3D architecture was selected as the final model for the MRI enhancement task.

5.2 Quantitative Results

The table presents a comparative evaluation of multiple VNet configurations. The integration of both brain masking and combined loss functions yields the highest PSNR and SSIM with the lowest error values, indicating superior reconstruction quality. In contrast, models without masking or with limited loss design show noticeable performance degradation.

Table 5.1: Impact of skull stripping on model performance

Model	PSNR	SSIM	MAE	MSE
VNet with mask	8.5068	0.9566	0.2419	0.1410
VNet without mask	5.1011	0.8034	0.2404	0.3089
VNetLS with mask	20.7538	0.9426	0.2134	0.1034
VNetLS without mask	17.9412	0.7281	0.3361	0.3976

We first evaluated the VNet base model with and without skull masking. The results showed that the model performs better with masking. With mask, the PSNR increased to 8.5068 and SSIM reached 0.9565, while MAE and MSE decreased to 0.2419 and 0.1410. Without mask, the performance was much lower. It shows that skull stripping helps the model focus on important brain regions.

After that, we improved the base model by adding extra loss functions. This gave us the VNetLS model. We again tested it with both settings. For VNetLS

without mask, the PSNR was 17.9412 and SSIM was 0.7281. But with mask, the performance improved further, giving the best results: PSNR 20.7538, SSIM 0.9425, MAE 0.2133, and MSE 0.1034.

From both experiments, we clearly observed that masked input always performs better than unmasked input. So finally we selected VNetLS with mask for all further experiments.

Table 5.2: Comparison of model performance on different fine tuning experiments

Model	PSNR	SSIM	MAE	MSE
VNetLS with AdamW optimizer	21.0124	0.9428	0.2025	0.0974
VNetLS with decreasing learning rate	20.9616	0.9428	0.2032	0.0986
VNetLS with learning rate scheduling	20.7538	0.9429	0.2028	0.0980

After selecting VNetLS as our main model, we further fine-tuned it using different strategies to improve performance. We tested three approaches: using the AdamW optimizer, applying a decreasing learning rate, and using learning rate scheduling. From the results, all three configurations performed very closely, but slight differences can still be observed. The model with AdamW optimizer achieved the highest PSNR value (21.0124) and the lowest MSE (0.0974), indicating better reconstruction quality and reduced error. The learning rate scheduling approach showed a slightly better SSIM (0.9429) but its error values were slightly higher compared to AdamW.

Overall, even though the differences are small, the AdamW optimizer consistently provides the best balance across all evaluation metrics. Based on the quantitative comparison, we selected VNetLS with AdamW optimizer fine-tuning as our final model, as it delivers the most stable and reliable performance.

5.2.1 Result Analysis

To evaluate the performance of the proposed model, several standard image quality metrics were used, including Peak Signal-to-Noise Ratio (PSNR), Structural Similarity Index Measure (SSIM), Mean Absolute Error (MAE), and Mean Squared Error (MSE). These metrics measure the similarity between the generated 7T-like images and the ground truth 7T MRI images. The model achieved improved performance compared to the input 3T MRI images. PSNR values increased, indicating better

reconstruction quality and reduced noise. Similarly, SSIM values showed improvement, which reflects better preservation of structural and anatomical information. Lower MAE and MSE values indicate that the generated images are closer to the ground truth in terms of pixel-wise accuracy. Although the dataset size is limited, the results demonstrate that the proposed VNet-based model is capable of learning meaningful mappings between low-field and high-field MRI images.

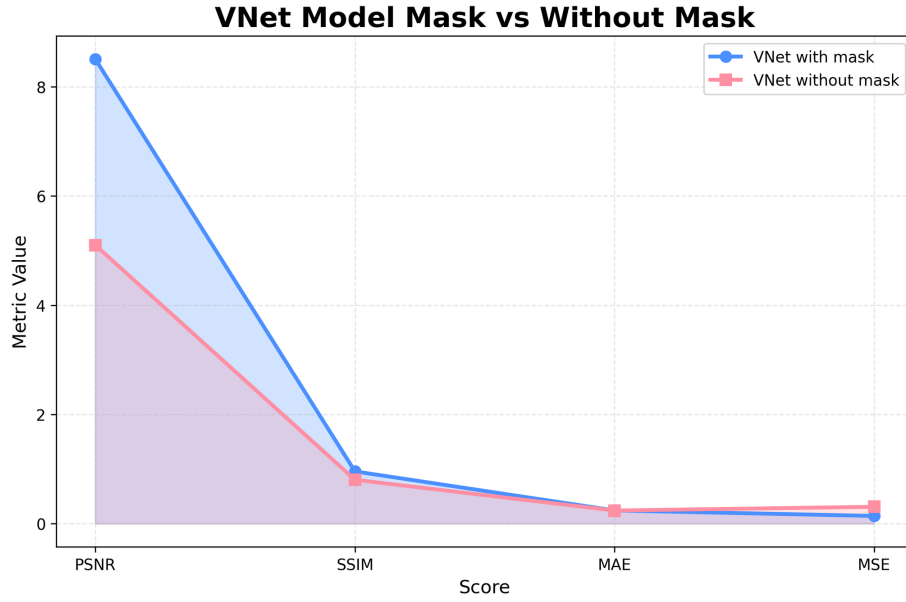


Figure 5.1: Vnet Mask vs Vnet Without Mask

The VNet model achieved highest PSNR (8.5) and SSIM (0.98), and the lowest MAE (0.24) and MSE (0.15). In comparison, the model without masking produced lower PSNR (5.1) and SSIM (0.82), with higher MAE (0.26) and MSE (0.30). These results show that brain masking and combined loss functions significantly improve reconstruction quality and model accuracy by reducing the negative influence of background regions.

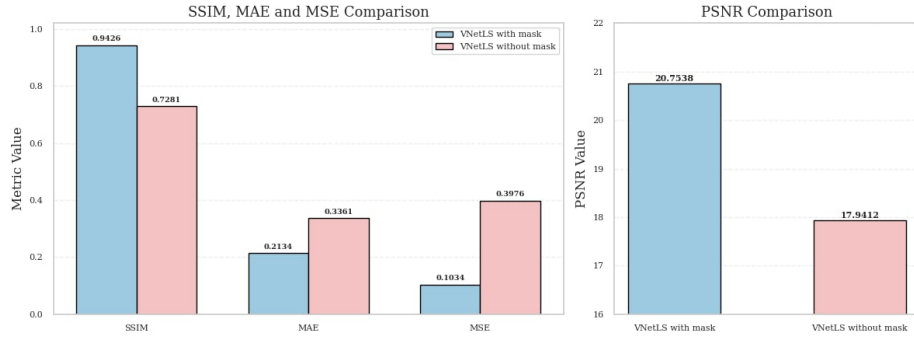


Figure 5.2: VnetLS Mask vs VnetLS Without Mask

The VNetLS model with brain masking achieved significantly better performance than the model without masking. With masking, the model obtained higher PSNR (20.7538) and SSIM (0.9426), along with lower MAE (0.2134) and MSE (0.1034). In contrast, the model without masking produced lower PSNR (17.9412) and SSIM (0.7281), while showing higher MAE (0.3361) and MSE (0.3976). These results demonstrate that brain masking effectively improves reconstruction quality by reducing the influence of irrelevant background regions.

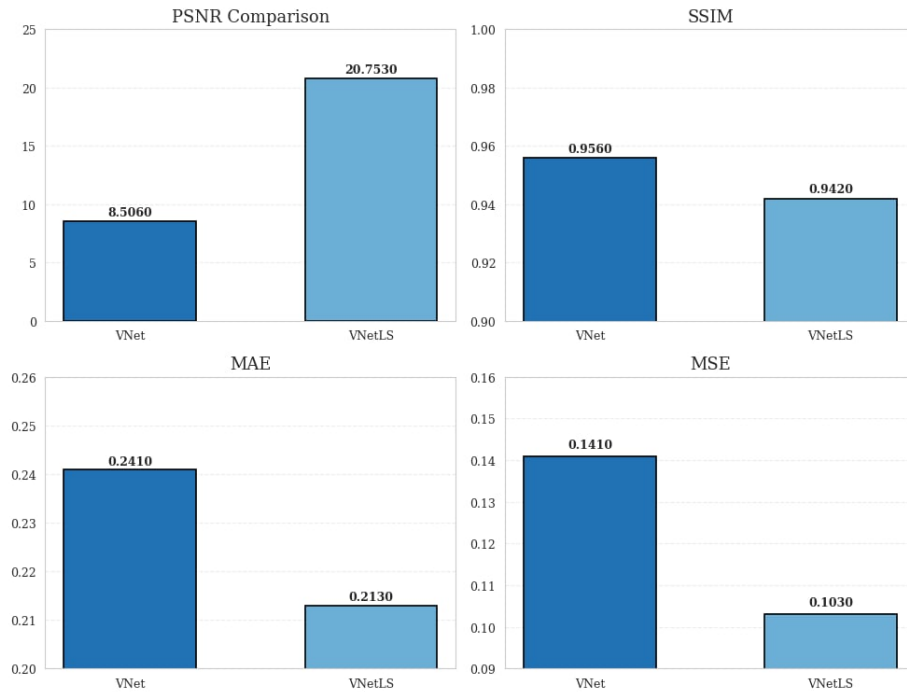


Figure 5.3: Vnet vs VnetLS

VNetLS achieves a much higher PSNR (20.75), which shows better reconstruction quality. It also reduces both MAE (0.213) and MSE (0.103), meaning the prediction errors are lower. The SSIM value is slightly lower in VNetLS, but the difference is very small and overall image structure preservation is still strong. So overall, VNetLS is a clear improvement over VNet. Therefore further experiments were conducted using the VNetLS model.

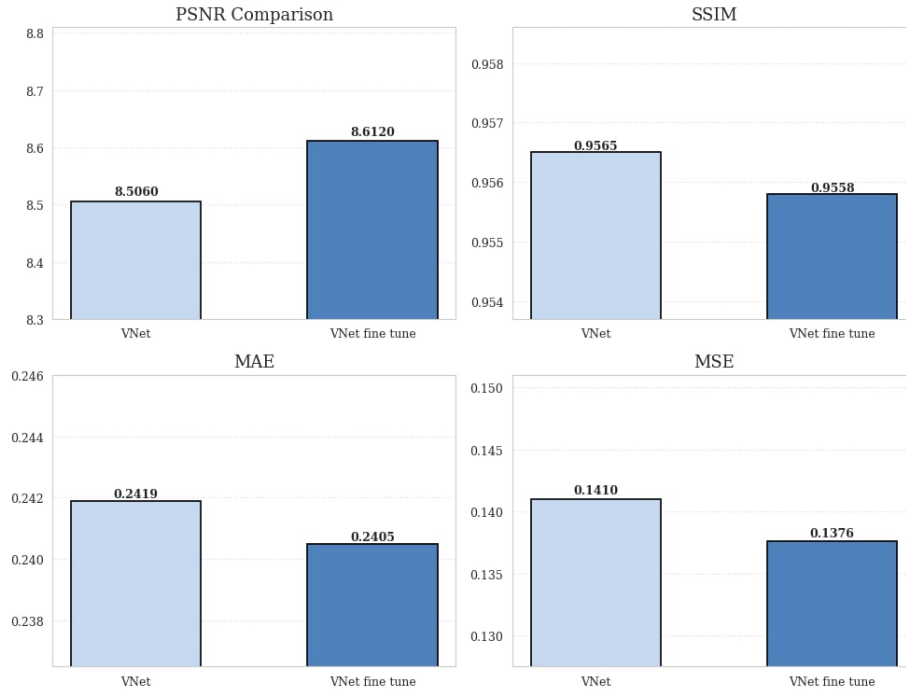


Figure 5.4: Vnet vs Vnet Fine Tune

The fine-tuned VNet version shows a slight improvement in PSNR(8.6120), which suggests better overall image reconstruction quality. It also achieves lower MAE and MSE values, indicating reduced pixel-level errors and more accurate reconstructions. While SSIM(0.9558) shows a very small drop, the change is negligible and the structural similarity remains consistently high. Overall, the fine-tuned VNet fine tune model demonstrates improved reconstruction performance, with better error reduction and slightly enhanced image quality compared to the baseline.

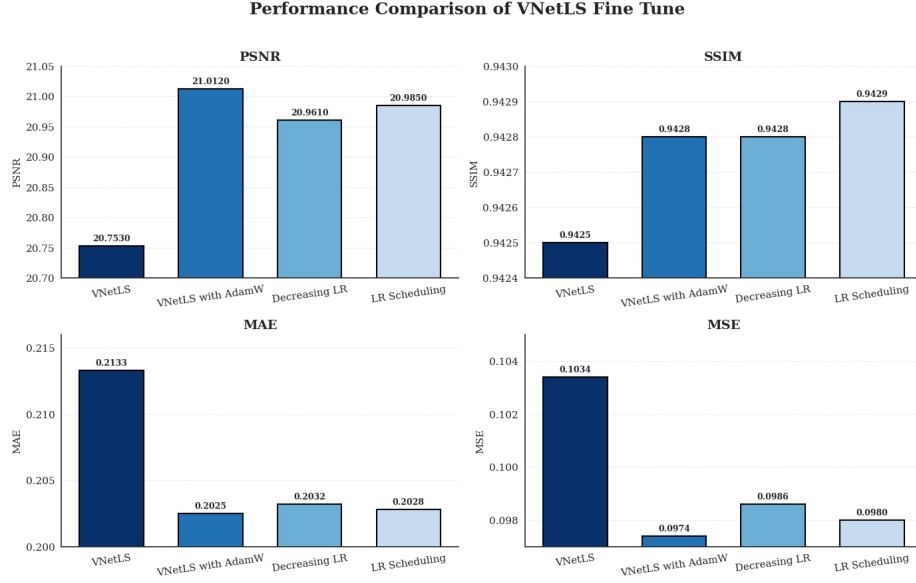


Figure 5.5: VNetLS vs VNetLS Fine Tune

The fine-tuned VNetLS model showed improved performance across all evaluation metrics. The VNetLS with AdamW optimizer achieved the highest PSNR (21.0120) and lowest MSE (0.0974), while LR scheduling produced the best SSIM (0.9429). Compared to the VNetLS model, which obtained PSNR (20.7530), SSIM (0.9425), MAE (0.2133), and MSE (0.1034), all fine-tuning strategies reduced reconstruction errors and improved image quality. Overall, optimizer tuning provided the most effective enhancement for the VNetLS model. So we selected VNetLS with AdamW as our final model

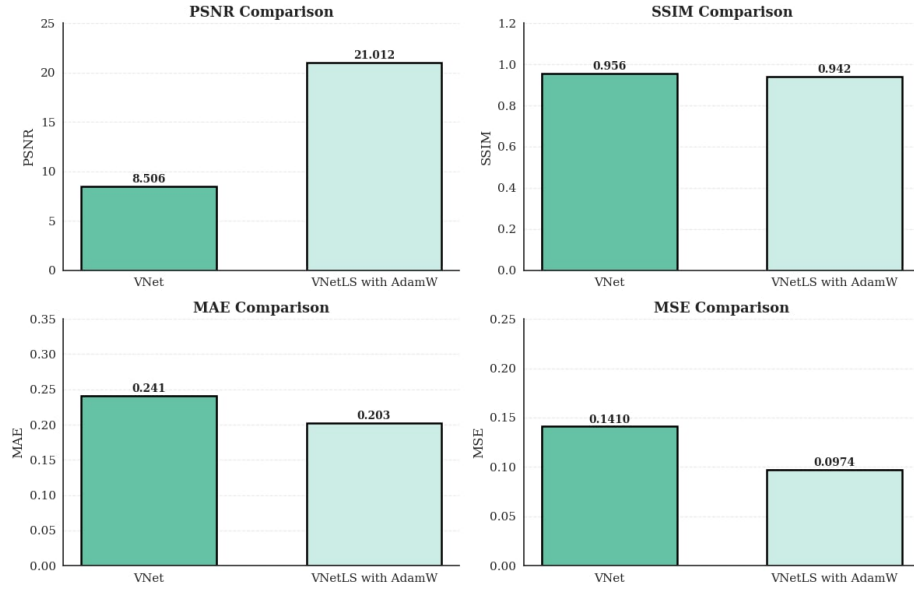


Figure 5.6: Vnet vs VNetLS with AdamW

The VNetLS with AdamW model shows better reconstruction performance than the baseline VNet model. It achieves a higher PSNR value (21.012), which means improved image quality and less distortion. It also has lower MAE (0.203) and MSE (0.0974), showing fewer errors in reconstruction. Overall, VNetLS with AdamW performs better than the original VNet model.

5.3 Qualitative Analysis

In addition to numerical evaluation, qualitative analysis was performed through visual inspection of the generated images. The synthesized 7T-like MRI images were compared with both the original 3T images and the corresponding ground truth 7T images. The visual results show that the enhanced images have:

- Reduced noise levels
- Improved contrast
- Sharper anatomical boundaries
- Better visibility of fine structures

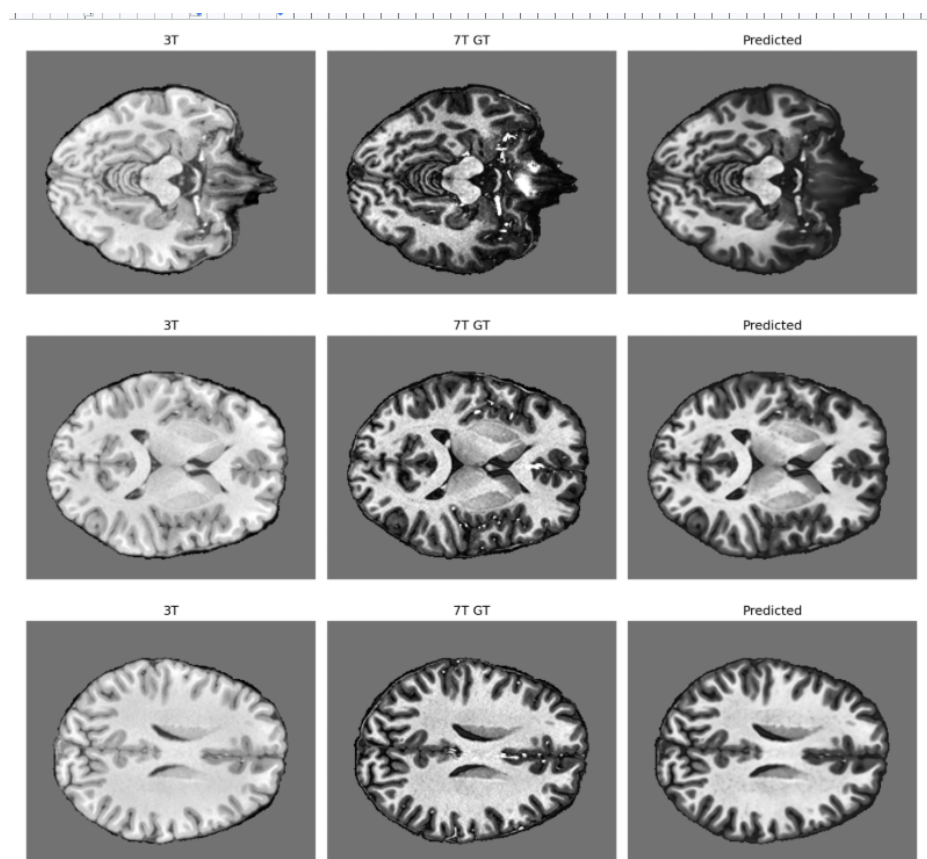


Figure 5.7: Predicted 7T Comparison with Target in validation

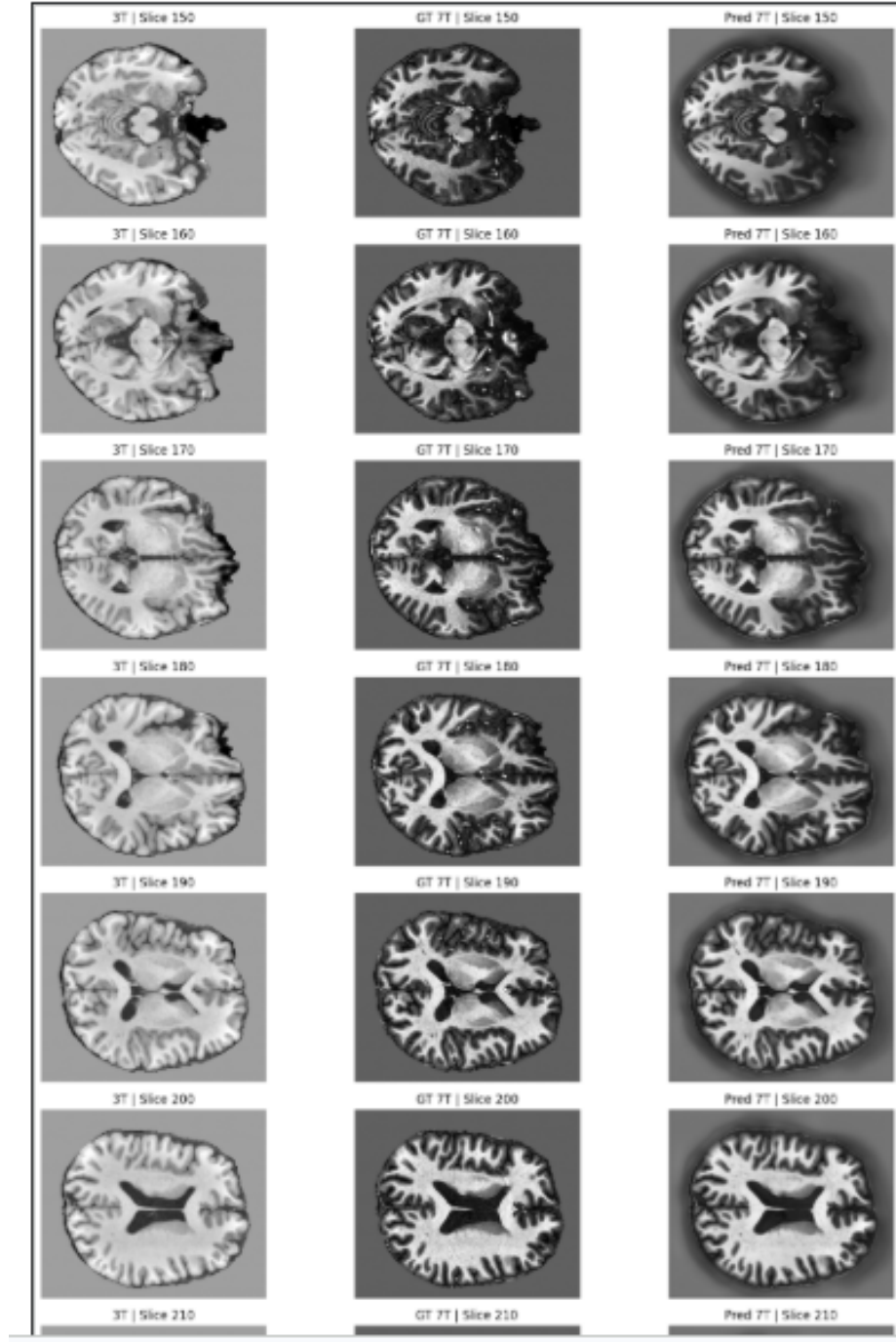


Figure 5.8: Predicted 7T Comparison with Target testing

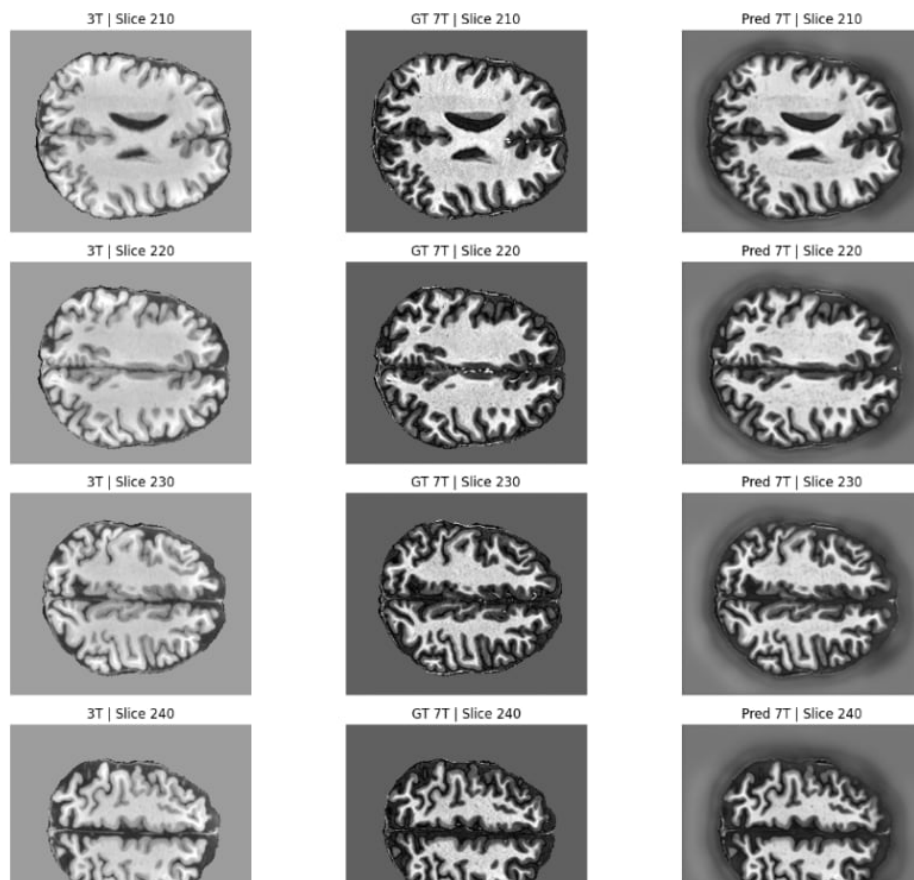


Figure 5.9: Predicted 7T Comparison with Target testing

Compared to the original 3T images, the generated outputs appear clearer and more detailed. Important brain structures such as cortical regions and tissue boundaries are more distinguishable. The model successfully preserves anatomical consistency while improving overall image quality.

The each 7T predicted images were also examined by expert where it is seen that among all the slices most of the slices are quite similar to the real 7T slices. Though there are more scope to make the images more clearer with better intensity such that each tissue will be clearly visible for better diagnosis.

5.3.1 Performance Comparison

The performance of the proposed model was also compared with baseline approaches. Traditional image enhancement techniques, such as filtering and contrast adjustment, showed limited improvement and often failed to preserve important anatomical details. Compared to these methods, the proposed deep learning approach provides better structural preservation, improved perceptual quality, realistic high-field image.

In conclusion, the use of a 3D VNet architecture allows the model to capture volumetric spatial information, resulting in consistent enhancement across slices.

Chapter 6

Plans and Deliverables

6.1 Budget

Table : Budget Breakdown for the Project

Category	Description	Cost (BDT)	Cost Source
Software Tools & Domain	DL frameworks, Python libraries, and IDEs	2,000	Free tools (e.g., TensorFlow, Jupyter Notebook IDEs), paid Colab
Cloud Computing	Storage and computational power	3,000	GoogleCloud, or similar platforms
Model Development	Colab subscription, training and optimization of DL models	30,000	GPU instances, cloud based solutions
Documentation and printing	Report printing, poster printing, binding	5,000	Local printing services
Hardware	Additional hardware (e.g., GPUs, laptops)	10,000	Local hardware or cloud-based GPU instances
Miscellaneous Expenses	Unexpected costs, storage devices	1,000	Pendrive
Total		= 51,000	

Figure 6.1: Total Budget

The table presents a detailed budget estimation for the project, categorizing expenses into software tools, cloud computing, model development, documentation, hardware, and miscellaneous costs. The highest cost is allocated to model development (30,000 BDT), primarily for GPU usage and training, followed by hard-

ware (10,000 BDT). Other costs include cloud services, printing, and software tools. The total estimated budget of the project is 51,000 BDT, indicating a moderate-cost implementation leveraging both free and paid resources.

6.1.1 Project Timeline



Figure 6.2: Project Timeline

The Gantt chart illustrates the timeline and progression of the project from June to May. The project begins with idea exploration and topic selection, followed by literature review and data collection. Mid-phase activities include preprocessing, model selection, and training. Later stages focus on testing, result analysis, and performance comparison. The final phase includes model refinement and report writing. The chart shows a structured and sequential workflow, ensuring systematic completion of the project within the defined timeline.

Chapter 7

Conclusions

7.1 Social, Legal, Ethical, and Environmental Issues

Deep learning techniques in medical imaging, particularly for MRI enhancement, involve a range of social, legal, ethical, and environmental considerations that must be carefully addressed. Although MRI is a non-invasive and widely accepted diagnostic modality, the use of medical imaging data and the operational footprint of MRI systems introduce responsibilities that extend beyond technical performance.

7.1.1 Social Impact

From a societal perspective, MRI enhancement technologies have the potential to improve access to high-quality diagnostic imaging, specifically in regions with limited medical resources. By maximizing the diagnostic value of existing MRI infrastructure, such approaches can support more equitable healthcare delivery and help reduce disparities between well-resourced and under-resourced medical facilities.

7.1.2 Ethical and Legal Considerations

MRI scans inherently contain sensitive patient information which makes data privacy and confidentiality a significant concern in medical image analysis. Inappropriate handling or storage of such data may lead to breaches of patient privacy and misuse of personal health information. In this project, anonymized datasets are utilized, and standard data protection practices are followed to minimize privacy

risks. Ensuring secure data handling and maintaining patient anonymity are essential, particularly when large datasets are used and evaluated in the models. MRI enhancement systems must preserve anatomical accuracy and diagnostic integrity. Enhancement algorithms should not introduce artificial structures, suppress subtle pathological features, or alter tissue characteristics in ways that could mislead clinical interpretation. Ethical development therefore requires careful validation using objective image quality metrics, combined with qualitative visual assessment that have been used in this research to ensure that improvements in image quality do not compromise clinical reliability. Legal consideration is significant in clinical use of MRI images, and the system must comply with healthcare regulations and medical device standards governing diagnostic software. Although our research is conducted within an academic and exploratory context, awareness of these regulatory frameworks informs a responsible research approach and helps avoid overstated claims regarding clinical applicability.

7.1.3 Environmental Issues

While MRI plays a vital role in modern healthcare, it is associated with a notable environmental footprint. MRI scanners are among the most energy-intensive imaging systems, requiring continuous electrical power to operate superconducting magnets, cooling mechanisms, and auxiliary electronics. Studies assessing the environmental impact of medical imaging services have shown that MRI contributes significantly to greenhouse gas emissions within healthcare facilities due to its high energy consumption. Another important environmental issue relates to the use of gadolinium-based contrast agents. After administration, a substantial portion of gadolinium is excreted by patients and enters wastewater systems. Conventional water treatment processes are not designed to effectively remove these compounds leading to detectable accumulation in natural water bodies. This raises concerns regarding long-term ecological effects and environmental sustainability. In addition, MRI systems depend on liquid helium, a finite and non-renewable resource, to maintain superconducting conditions. Helium loss during routine operation or maintenance represents both an economic and environmental challenge.. Although our research focuses on computational MRI enhancement rather than the development of imaging hardware, the environmental context remains directly relevant.

MRI-to-MRI enhancement techniques have the potential to indirectly reduce the environmental impact of MRI services by:

- Reducing repeat scans caused by inadequate image quality, thereby lowering scanner usage time and energy consumption
- Optimizing scanner utilization, enabling lower-field or more resource-efficient MRI systems to produce diagnostically useful images.

7.2 Brief Summary

This research presents a deep learning-based framework, VNetLS, for enhancing low-field 3T MRI images to high-quality 7T-like outputs. The study utilizes a structured preprocessing pipeline, including skull stripping, bias field correction, registration, and normalization, to ensure data consistency before model training. A 3D VNet encoder-decoder architecture is employed to effectively capture volumetric anatomical features. The key contribution lies in the integration of a multi-component loss function, combining MAE, SSIM, and gradient loss, which helps in the improvement process of image clarity, structural preservation, and edge information. Experimental results demonstrate that the proposed model outperforms baseline approaches in both PSNR and SSIM, indicating improved visual quality and anatomical fidelity. The literature analysis further supports that combining structural and perceptual loss functions within 3D architectures significantly enhances MRI reconstruction performance, though limited datasets and generalization remains prominent challenge in the study.

7.3 Future works

Future work will focus on improving the matrices vlaue and generalization of the model by increasing the number of paired 3T-7T MRI datasets and incorporating large-scale multi-center data, which can help the model learn more diverse anatomical and scanner variations. Additionally, integrating T2-weighted MRI data alongside T1-weighted images can further enhance the quality and accuracy of synthesized 7T outputs by using complementary structural information. Further

improvements may include extending the model to handle multi-modal inputs, optimizing computational efficiency, and validating performance on clinical datasets to ensure reliability in practical healthcare applications.

References

- Acs, R., & Zhuang, H. (2025). Semi-supervised synthesis of 7t mri from 3t using 3d fr-u-net with anatomical segmentation consistency assessment. *Plos one*, 20(11), e0333499.
- Asiri, A. A., Soomro, T. A., Shah, A. A., Pogrebna, G., Irfan, M., & Alqahtani, S. (2024). Optimized brain tumor detection: a dual-module approach for mri image enhancement and tumor classification. *IEEE access*, 12, 42868–42887.
- Chen, X., Qu, L., Xie, Y., Ahmad, S., & Yap, P.-T. (2023). A paired dataset of t1-and t2-weighted mri at 3 tesla and 7 tesla. *Scientific Data*, 10(1), 489.
- Chen, Z., Pawar, K., Ekanayake, M., Pain, C., Zhong, S., & Egan, G. F. (2023). Deep learning for image enhancement and correction in magnetic resonance imaging—state-of-the-art and challenges. *Journal of Digital Imaging*, 36(1), 204–230.
- Cui, Q., Tosun, D., Mukherjee, P., & Abbasi-Asl, R. (2024a). 7t mri synthesization from 3t acquisitions. , 35–44.
- Cui, Q., Tosun, D., Mukherjee, P., & Abbasi-Asl, R. (2024b). 7t mri synthesization from 3t acquisitions. In *International conference on medical image computing and computer-assisted intervention* (pp. 35–44).
- Diniz, E., Santini, T., Karim, H., Aizenstein, H. J., & Ibrahim, T. S. (2025). Cross-modality image translation of 3 tesla magnetic resonance imaging to 7 tesla using generative adversarial networks. *Human Brain Mapping*, 46(9), e70246.
- Islam, K. T., Ekanayake, M., & Chen, Z. (2025a). Ultra-low-field mri enhancement via inr-based style transfer. , 597–607.
- Islam, K. T., Ekanayake, M., & Chen, Z. (2025b). Ultra-low-field mri enhancement via inr-based style transfer. In *International conference on medical image computing and computer-assisted intervention* (pp. 597–607).

- Islam, K. T., Zhong, S., Zakavi, P., Chen, Z., Kavnoudias, H., Farquharson, S., ... others (2023). Improving portable low-field mri image quality through image-to-image translation using paired low-and high-field images. *Scientific Reports*, 13(1), 21183.
- Jeme V, J., & Jerome S, A. (2024a). Efficient mri image enhancement by improved denoising techniques for better skull stripping using attention module-based convolution neural network. *Computer Methods in Biomechanics and Biomedical Engineering: Imaging & Visualization*, 12(1), 2309874.
- Jeme V, J., & Jerome S, A. (2024b). Efficient mri image enhancement by improved denoising techniques for better skull stripping using attention module-based convolution neural network. *Computer Methods in Biomechanics and Biomedical Engineering: Imaging & Visualization*, 12(1), 2309874.
- Jiang, M., Jia, P., Huang, X., Yuan, Z., Ruan, D., Liu, F., & Xia, L. (2025a). Frequency-aware diffusion model for multi-modal mri image synthesis. *Journal of Imaging*, 11(5), 152.
- Jiang, M., Jia, P., Huang, X., Yuan, Z., Ruan, D., Liu, F., & Xia, L. (2025b). Frequency-aware diffusion model for multi-modal mri image synthesis. *Journal of Imaging*, 11(5), 152.
- Kalluvila, A. (2023). Super-resolution of brain mri via u-net architecture. *International Journal of Advanced Computer Science and Applications*, 14(5).
- Kawahara, D., & Nagata, Y. (2021). T1-weighted and t2-weighted mri image synthesis with convolutional generative adversarial networks. *reports of practical Oncology and radiotherapy*, 26(1), 35–42.
- Li, B. M., Castorina, L. V., Valdés Hernández, M. d. C., Clancy, U., Wiseman, S. J., Sakka, E., ... others (2022). Deep attention super-resolution of brain magnetic resonance images acquired under clinical protocols. *Frontiers in Computational Neuroscience*, 16, 887633.
- Li, Y., Yang, M., Bian, T., & Wu, H. (2024). Mri super-resolution analysis via mrir: Deep learning for low-field imaging. *Information*, 15(10), 655.
- Rasheed, Z., Ma, Y.-K., Ullah, I., Ghadi, Y. Y., Khan, M. Z., Khan, M. A., ... Shehata, A. M. (2023). Brain tumor classification from mri using image enhancement and convolutional neural network techniques. *Brain Sciences*, 13(9), 1320.
- Roy, M., Upadhyaya, B., Rai, J., & Sharma, K. (2024). Resolution enhancement of brain mri images using deep learning. *Engineering Proceedings*, 59(1), 158.

- Sarracanie, M., & Salameh, N. (2020). Low-field mri: how low can we go? a fresh view on an old debate. *Frontiers in Physics*, 8, 172.
- Shamaei, A., Stebner, A., Bosshart, S. L., Ospel, J., Ginde, G., Bento, M., & Souza, R. (2025a). Enhancing and accelerating brain mri through deep learning reconstruction using prior subject-specific imaging. *Magnetic Resonance Imaging*, 110558.
- Shamaei, A., Stebner, A., Bosshart, S. L., Ospel, J., Ginde, G., Bento, M., & Souza, R. (2025b). Enhancing and accelerating brain mri through deep learning reconstruction using prior subject-specific imaging. *Magnetic Resonance Imaging*, 110558.
- Solak, M., Tören, M., Asan, B., Kaba, E., Beyazal, M., & Çeliker, F. B. (2025). Generative adversarial network based contrast enhancement: synthetic contrast brain magnetic resonance imaging. *Academic radiology*, 32(4), 2220–2232.
- Sun, Y., Wang, L., Li, G., Lin, W., & Wang, L. (2025). A foundation model for enhancing magnetic resonance images and downstream segmentation, registration and diagnostic tasks. *Nature Biomedical Engineering*, 9(4), 521–538.
- Wang, J., Levman, J., Pinaya, W. H. L., Tudosiu, P.-D., Cardoso, M. J., & Marinescu, R. (2023). Inversesr: 3d brain mri super-resolution using a latent diffusion model. In *International conference on medical image computing and computer-assisted intervention* (pp. 438–447).
- Wang, Z., Yu, X., Wang, C., Chen, W., Wang, J., Chu, Y.-H., ... others (2025a). One for multiple: Physics-informed synthetic data boosts generalizable deep learning for fast mri reconstruction. *Medical Image Analysis*, 103616.
- Wang, Z., Yu, X., Wang, C., Chen, W., Wang, J., Chu, Y.-H., ... others (2025b). One for multiple: Physics-informed synthetic data boosts generalizable deep learning for fast mri reconstruction. *Medical Image Analysis*, 103, 103616.
- Xue, H., Hooper, S. M., Pierce, I., Davies, R. H., Stairs, J., Naegele, J., ... others (2025). Snraware: Improved deep learning mri denoising with signal-to-noise ratio unit training and g-factor map augmentation. *Radiology: Artificial Intelligence*, 7(6), e250227.
- Zhang, Z., Jing, P., Beitone, C., Huang, J., Gao, Z., Yang, G., & Lally, P. (2025a). Pretext task adversarial learning for unpaired low-field to ultra high-field mri synthesis. *arXiv preprint arXiv:2503.05339*.
- Zhang, Z., Jing, P., Beitone, C., Huang, J., Gao, Z., Yang, G., & Lally, P. (2025b). Pretext task adversarial learning for unpaired low-field to ultra high-field mri synthesis. *arXiv preprint arXiv:2503.05339*.