Cardiac PET Attenuation Correction and Inter-modal Medical Image Translation

John Boby Mesadieu

A thesis submitted to the University of Ottawa
in partial fulfillment of the requirements for the
Master of Science in Mathematics and Statistics

Department of Mathematics and Statistics
Faculty of Science
University of Ottawa

Thesis Supervisors:
Dr. Tanya Schmah
Dr. Robert DeKemp

Abstract

Positron Emission Tomography (PET) is a nuclear imaging technique that visualizes metabolic activity in the body by detecting gamma rays emitted from a radiotracer. For accurate quantification of radiotracer uptake, attenuation correction (AC) is essential to compensate for the absorption of photons by different tissues. Traditionally, Computed Tomography (CT) provides ideal attenuation maps for PET correction due to its direct relationship with tissue density. In hybrid PET/Magnetic Resonance Imaging (MRI) systems, however, MRI-based attenuation correction is challenging as MRI cannot directly visualize bone structures critical for proper photon attenuation estimation. This thesis addresses this limitation by developing a novel approach for generating synthetic CT images from MRI data for improved PET AC in cardiac imaging, with applications extending beyond AC to radiation therapy planning and other inter-modal medical image translation tasks.

This research makes three key contributions: (1) a CT template and multi-modal registration pipeline to establish spatial correspondence between MR and CT images, (2) a Kernel Ridge Regression (KRR) implementation serving as a baseline, and (3) a Dual Contrast CycleGAN (DC-CycleGAN) model incorporating a novel μ-map loss for bidirectional MR-CT translation. The dataset consisted of paired MR and CT images from 10 subjects, with 8 used for training and 2 for testing.

Results demonstrated that the DC-cycleGAN model using in-phase Dixon MR images (GAN-In) achieved superior performance in MR-to-CT translation, with a 39.5% reduction in Mean Absolute Error (MAE) and a 37.1% improvement in Structural Similarity Index compared to the KRR baseline. When applied to PET reconstruction, the synthetic CT-based AC (GAN-AC) achieved substantial improvements: SUVmean bias was reduced by 93% (from 18.85% to 1.27%), SUVmax bias by 89%, Root Mean Square Error by 45%, and MAE by 47% compared to the vendor-provided MR-based AC (MR-AC). The proposed implementation as a web-based application will enable researchers at The Royal's Institute of Mental Health Research to utilize this technology for improved diagnostic precision in PET/MR imaging.

Acknowledgements

I am deeply grateful to my supervisor, Dr. Tanya Schmah, for her exceptional mentorship and steadfast support throughout my Master's journey. Her deep knowledge, thoughtful guidance, and encouraging presence made this experience incredibly meaningful and transformative. Dr. Schmah's insightful feedback, intellectual curiosity, and genuine kindness consistently pushed me to grow, both as a student and as a researcher. I am especially thankful for the critical financial support she provided and for always looking out for opportunities to help me succeed. Her support made it possible for me to stay focused and committed to my research, even when resources were limited.

I would also like to sincerely thank my co-supervisor, Dr. Robert DeKemp, whose expertise in medical imaging and thoughtful suggestions greatly enriched this work.

I'm particularly thankful for the contributions of my collaborators at The Royal's Institute of Mental Health Research. A special thank you goes to Katie Dinelle and Owen Clarkin, whose technical insights and collaborative spirit were invaluable throughout this project.

This thesis is dedicated to the memory of my father, who passed away on May 28, 2019. His absence leaves the biggest void in me, but his values, wisdom, and unwavering belief in my potential continue to inspire and guide me every day. I owe profound thanks to my mother, Rosaire Poteau, whose love, sacrifices, and unwavering encouragement have been my foundation. To my brothers, Vladimir and Eventz Mesadieu, your support and motivation, especially during difficult moments, meant the world to me. I'm also grateful to my friend Mackenson Jean, whose friendship and moral support have been a steady source of strength throughout this journey. Finally, I'd like to thank the University of Ottawa and the Department of Mathematics and Statistics for providing a supportive academic environment and the resources necessary to complete this research.

Contents

List of Tables

List of Figures

List of Abbreviations

Abbreviation	Definition
AC	Attenuation Correction
ANTs	Advanced Normalization Tools
CNN	Convolutional Neural Network
CNR	Contrast-to-Noise Ratio
CT	Computed Tomography
CTAC	CT-based Attenuation Correction
CV	Coefficient of Variation
DC-cycleGAN	Dual Contrast Cycle-Consistent Generative Adversarial Network
Dixon	Dual-echo MR sequence for fat-water separation
FDG	Fluorodeoxyglucose
FOV	Field of View
GAN	Generative Adversarial Network
GPU	Graphics Processing Unit
HU	Hounsfield Units
IMHR	Institute of Mental Health Research
KRR	Kernel Ridge Regression
LAC	Linear Attenuation Coefficient
LOO-CV	Leave-One-Out Cross-Validation
MAE	Mean Absolute Error
MI	Mutual Information
MLAA	Maximum Likelihood reconstruction of Activity and Attenuation
MRAC	MR-based Attenuation Correction
MR/MRI	Magnetic Resonance Imaging
MSE	Mean Squared Error
NCC	Normalized Cross-Correlation
PET	Positron Emission Tomography
PSNR	Peak Signal-to-Noise Ratio

Abbreviation	Definition
RF	Radio Frequency
RMSE	Root Mean Square Error
ROI	Region of Interest
SNR	Signal-to-Noise Ratio
SSIM	Structural Similarity Index
SUV	Standardized Uptake Value
T1	Longitudinal relaxation time
T2	Transverse relaxation time
TOF	Time of Flight
UOHI	University of Ottawa Heart Institute
UTE	Ultra-short Echo Time
VRAM	Video Random Access Memory
ZTE	Zero Echo Time
μ-map	Attenuation map (mu-map)

Chapter 1

Introduction

1.1 Overview of Medical Imaging Modalities

Medical imaging plays a pivotal role in modern diagnostics, providing non-invasive insights into anatomical and functional aspects of the human body. This thesis focuses on three key modalities: Positron Emission Tomography (PET), Magnetic Resonance Imaging (MRI), and Computed Tomography (CT), each offering unique strengths that are often combined to enhance diagnostic accuracy.

Positron Emission Tomography (PET) is a nuclear medicine imaging technique that visualizes metabolic processes by detecting gamma rays emitted from a radiotracer injected into the body. PET is widely used in oncology, neurology, and cardiology to assess tissue function, such as glucose metabolism in tumors or neurotransmitter activity in the brain [1]. However, PET images require attenuation correction to account for photon absorption by tissues, a process that relies on anatomical information from complementary modalities like CT or MRI.

Magnetic Resonance Imaging (MRI) excels in providing high-resolution anatomical images with excellent soft tissue contrast, making it ideal for imaging the brain, muscles, and organs. Unlike CT, MRI does not involve ionizing radiation, using magnetic fields and radio waves to generate images based on the relaxation properties of hydrogen nuclei in tissues [2]. However, MRI struggles to directly visualize cortical bone due to its low proton density and rapid signal decay, posing challenges for applications requiring precise tissue density information.

Computed Tomography (CT) uses X-rays to produce detailed images of internal structures, particularly effective for visualizing bone and high-density tissues. CT images are quantified in Hounsfield Units (HU), which directly correlate with tissue density and can be converted into linear attenuation coefficients for PET attenuation

correction [3]. Despite its strengths, CT involves ionizing radiation, making it less suitable for repeated imaging or pediatric applications compared to MRI.

The integration of these modalities, such as in PET/CT and PET/MR systems, leverages their complementary strengths to provide comprehensive diagnostic information, with PET/MR being the focus of this thesis due to its unique challenges in attenuation correction.

1.2 Background on PET/MR Imaging

PET combined with MRI has emerged as a powerful hybrid imaging modality, offering both functional and anatomical insights in a single examination. PET/MR scanners, such as the Siemens system used at The Royal's Institute of Mental Health Research (IMHR), provide superior soft tissue contrast through MRI while simultaneously capturing metabolic information via PET, making them particularly valuable for applications in neurology, oncology, and psychiatry [1]. Unlike PET/CT systems, which use CT for anatomical reference and attenuation correction, PET/MR systems rely on MR images for these purposes, reducing radiation exposure and enabling simultaneous acquisition of PET and MR data.

1.3 PET Attenuation and Its Significance

Attenuation correction (AC) is a critical step in PET imaging to ensure accurate quantification of radiotracer uptake. During a PET scan, emitted 511 keV photons are absorbed or scattered by tissues, reducing the number of detected events and skewing the reconstructed image. AC compensates for this loss by using a μ-map, which quantifies the linear attenuation coefficients of tissues at 511 keV, to adjust the PET data during reconstruction [4]. Accurate μ-maps are essential for reliable Standardized Uptake Value (SUV) measurements, which are widely used in clinical settings to quantify metabolic activity, such as in tumor staging or monitoring treatment response. Errors in AC can lead to significant biases in SUV, with studies reporting discrepancies of up to 20% in brain imaging when bone attenuation is underestimated [5]. This underscores the importance of precise μ-map generation, particularly in hybrid imaging systems like PET/MR, where direct measurement of attenuation properties is challenging.

1.4 PET/MR vs PET/CT Attenuation Correction

In PET/CT systems, AC is relatively straightforward, as CT images provide Hounsfield Units (HU) that can be directly converted into μ-maps for 511 keV photons using established methods [3]. The high density contrast in CT allows for accurate differentiation of air, soft tissue, and bone, ensuring reliable μ-maps for PET reconstruction. However, the ionizing radiation associated with CT makes it less ideal for repeated imaging, particularly in sensitive populations such as pediatric patients.

In contrast, PET/MR systems face significant challenges in AC due to the inherent limitations of MRI. Traditional MR-based AC methods, such as those implemented in the Siemens PET/MR scanner at IMHR, rely on segmentation of MR images into discrete tissue classes (e.g., air, soft tissue, bone) to estimate μ-maps. These methods often use Dixon or ultra-short echo time (UTE) sequences to differentiate tissues based on MR signal characteristics [6]. However, a major limitation is the inability of MR to directly visualize cortical bone, which has a very short T2 relaxation time and thus appears as a signal void, similar to air. This leads to misclassification of bone as air or soft tissue, resulting in underestimation of attenuation coefficients in regions like the skull and spine. Such errors can significantly affect PET quantification, particularly in brain imaging, where bone attenuation is substantial, leading to SUV errors of up to 20% [5]. To address these challenges, researchers have explored methods to generate pseudo-CT images from MR data, which can then be converted into μ-maps using established HU-to-μ conversion techniques, leveraging advances in machine learning and deep learning [7].

1.5 Proposed Approach

This thesis proposes the use of a Dual Contrast cycleGAN (DC-cycleGAN), a deep learning-based architecture, for MR-to-CT image translation to improve AC in PET/MR imaging. DC-cycleGAN builds on the CycleGAN framework [8] and incorporates dual contrast mechanisms [9], enabling bidirectional translation (MR-to-CT and CT-to-MR). This approach ensures cycle consistency and improved synthesis quality through a novel mumap loss that directly optimizes the accuracy of derived μ-maps. By generating high-quality pseudo-CT images, DC-cycleGAN aims to produce more accurate μ-maps, particularly in bone-rich regions, thereby enhancing PET reconstruction accuracy.

To establish a baseline for comparison, we also implemented Kernel Ridge Regression (KRR), a machine learning technique that uses a Gaussian kernel to learn a non-linear mapping from MR intensities to CT intensities at the voxel or small patch

level [10].

The dataset consists of MR images from 10 subjects acquired using the PET/MR scanner at IMHR and corresponding CT images from the same subjects acquired using the PET/CT scanner at the University of Ottawa Heart Institute (UOHI). Of these, 8 subjects were used for training and 2 for testing, after aligning them to a common space. The PET reconstructed image using the pseudo-μ-map derived from the pseudo-CT generated by DC-cycleGAN, as well as the PET reconstructed image using the traditional Siemens MR-based AC method, are both compared to the reference PET reconstruction based on the true CT-derived μ-map, which serves as the gold standard.

1.6 Thesis Objectives

The primary objective of this thesis is to develop and evaluate methods for generating pseudo-CT images from MR data to enable AC correction in PET/MR imaging. In this work, we want to:

- Develop a CT template and a multi-modal registration framework to spatially align MR and CT images, ensuring consistent anatomical correspondence for effective model training and evaluation.
- Investigate the efficacy of KRR as a baseline method for unidirectional MR-to-CT translation.
- Develop a DC-cycleGAN model, with a novel mumap loss, to perform bidirectional MR-CT translation, optimizing the generation of pseudo-CT images and pseudo-μ-maps for PET attenuation correction.
- Compare the proposed methods against the traditional MR-based AC approach, assessing their impact on PET reconstruction accuracy using metrics such as MAE, SSIM, and SUV errors.

1.7 Thesis Structure

This thesis is organized as follows:

- **Chapter 2: Literature Review** provides an overview of existing methods for AC in PET/MR, including segmentation-based, atlas-based, and machine learning-based approaches, highlighting their strengths and limitations.

Chapter 3: Methodology details the preprocessing steps for MR and CT images, the CT template creation and registration framework, the proposed KRR and DC-cycleGAN methods and the AC workflow using pseudo-CT-derived μ-maps.

Chapter 4: Results and Discussion presents the preprocessing, CT template creation, and registration results, the performance of the synthetic images generated by the KRR and DC-cycleGAN models, and compares the MR-based and the pseudo-CT (DC-cycleGAN)-based PET reconstructions to the gold standard true CT-based PET reconstruction.

- **Chapter 5: Conclusion and Future Work** summarizes the findings, discusses the implications for PET/MR imaging at IMHR, and suggests directions for future research.

Chapter 2

Literature review

Accurate AC is crucial for quantitative positron emission tomography (PET), especially in cardiac PET/MR imaging, where precise assessment of myocardial perfusion and viability is essential [11, 12]. Combined PET/MR offers advantages like superior soft tissue contrast and reduced radiation exposure [13], but these benefits hinge on accurate AC. Inaccurate AC can lead to significant errors in standardized uptake values (SUVs), mimicking pathological conditions and potentially resulting in misdiagnoses [14, 15].

While MR-based AC methods, including segmentation-based and atlas-based techniques, exist, they have limitations [16]. Deep learning-based methods have emerged as a promising alternative, capable of learning complex mappings between MR and CT images or directly predicting attenuation maps [17]. This offers the potential for improved accuracy and robustness in cardiac PET/MR imaging.

This literature review examines AC methods in cardiac PET/MR, focusing on deep learning approaches like CNNs, GANs, and CycleGANs. The review will: (1) examine the evolution of PET AC techniques; (2) analyze current approaches to MR-based AC; (3) explore recent developments in deep learning for medical image translation and their application to AC; and (4) identify gaps in the current literature that justify the research objectives outlined in Chapter 1. By critically analyzing the existing literature, this review aims to establish the theoretical foundation for the development of improved AC methods in cardiac PET/MR imaging.

2.1 Evolution of Attenuation Correction in PET Imaging

AC is essential for accurate PET quantification. Methods have evolved from transmission-based approaches to integrated CT-based techniques.

2.1.1 Early Transmission-Based Methods

Early methods used external positron-emitting sources to directly measure tissue attenuation at 511 keV [18]. Transmission scans with ^{68}Ge/^{68}Ga sources, acquired before or after emission scans, estimated attenuation along lines of response [19].

These approaches were limited by long acquisition times (15-30 minutes) [18], statistical noise from low source activity [19], emission contamination in post-injection scans [18], and potential misregistration from patient movement [20].

For cardiac imaging, these limitations proved particularly problematic. Extended scan times increased motion artifacts, while noise compromised myocardial perfusion quantification. Misregistration at the heart-lung boundary created artifacts resembling perfusion defects [20].

Despite refinements like segmented attenuation maps and post-reconstruction filtering [19], these fundamental constraints ultimately led to the development of PET/CT systems.

2.1.2 Attenuation Correction in PET/CT

PET/CT scanners derive attenuation maps from CT images by converting CT coefficients to 511 keV equivalents [21]. CTAC advantages include faster scan times and better signal-to-noise ratios [22], higher spatial resolution [23], and elimination of emission contamination [21]. Despite these benefits, CTAC challenges involve increased radiation exposure [22], energy transformation errors [23], and temporal misalignment from cardiac/respiratory motion [20, 24]. Cardiac PET/CT is particularly susceptible to respiratory artifacts, with misregistration in up to 40% of studies causing perfusion defects [20]. Mitigation strategies include respiratory-gated CT and registration correction software [24]. While CTAC improves on previous methods, radiation concerns and misregistration issues persist. These limitations have motivated radiation-free approaches in PET/MR, building upon lessons from earlier AC techniques.

2.2 Challenges in MR-Based Attenuation Correction for Cardiac PET

2.2.1 Fundamental and Technical Limitations

MR-based AC (MRAC) in cardiac PET/MR is hindered by the disconnect between MR signal, which reflects proton density and tissue relaxation, and PET photon attenuation, which depends on electron density [25]. This lack of direct correlation complicates mapping MR images to linear attenuation coefficients (LACs). Standard MR sequences struggle to differentiate tissues with similar MR properties but distinct attenuation characteristics, such as cortical bone and air, leading to quantification errors in the myocardium [26]. Thoracic cage bones, lung-tissue interfaces, and air cavities are particularly problematic, with studies indicating 10-15% activity underestimation near bony structures [25]. MR hardware, including RF coils and patient tables, introduces attenuation not captured in images, potentially causing 5-20% underestimation of myocardial activity [27]. Limited MR field of view (FOV) leads to truncation artifacts, and B0 field inhomogeneities further impair tissue classification [28].

2.2.2 Physiological Motion and Cardiac Device Artifacts

Physiological motion, including respiratory and cardiac movements, causes spatial and temporal mismatches between MR-derived attenuation maps and PET emission data [29]. Diaphragmatic excursion and cardiac contraction blur attenuation maps, with misregistrations as small as 1-2 cm causing significant perfusion defects and quantification errors [15]. Differences between breath-hold MR and free-breathing PET exacerbate these issues. Implanted cardiac devices create susceptibility artifacts, leading to signal voids and geometric distortions that misclassify tissues as air, potentially underestimating activity by up to 40% near metallic implants [28]. Safety concerns also limit PET/MR use in patients with older implants, necessitating innovative solutions to improve AC accuracy [15].

2.3 Current Strategies for PET/MR Attenuation Correction

Several approaches for MR-based AC (MRAC) have been developed, each with advantages and limitations. This section examines the primary methodologies currently

employed for PET/MR AC.

2.3.1 Segmentation-Based Methods

Segmentation-based methods classify MR images into tissue types with predefined attenuation coefficients assigned to each class [11]. Dixon MR sequences commonly distinguish between air, lung, fat, and soft tissue [11].

These methods offer computational efficiency, robustness to certain MR artifacts, and adaptability to include additional tissue classes. However, they ignore continuous variation in attenuation properties [25], cannot reliably distinguish bone from soft tissue with standard Dixon sequences [28], face challenges at tissue interfaces [28], and are susceptible to propagating segmentation errors.

In cardiac applications, segmentation-based methods typically show errors of 5-10% compared to CT-based AC [28], with more pronounced errors at the lateral and inferior myocardial walls [15].

2.3.2 Atlas-Based Methods

Atlas-based approaches use paired MR and CT images to generate patient-specific attenuation maps by registering the patient's MR image to atlas MR images and transforming corresponding atlas CT images to create a pseudo-CT [30].

Implementations include single-atlas, multi-atlas with weighted averaging, and patch-based methods matching local image features. These approaches provide continuous linear attenuation coefficients, improved bone representation even when not directly visible in MR, and maintain structural integrity in the predicted maps.

Limitations include heavy dependence on registration accuracy [31], challenges with patients whose anatomy deviates from the atlas population [32], computational intensity, and the need for representative atlas databases.

For cardiac imaging, atlas methods face respiratory and cardiac motion challenges but can reduce quantification errors compared to segmentation approaches, particularly in regions with significant bone content [33].

2.3.3 Specialized MR Sequences (UTE/ZTE)

Ultra-short Echo Time (UTE) and Zero Echo Time (ZTE) sequences capture signals from tissues with short T2* relaxation times, such as cortical bone [34]. The approach involves acquiring UTE/ZTE sequences, identifying bone based on signal patterns, and generating attenuation maps.

These techniques directly visualize cortical bone [35], improve air-bone distinction, and derive bone information directly from patient images. Drawbacks include longer acquisition times [35], limited field of view, vulnerability to motion artifacts [34], and challenges in accurately quantifying bone density.

In cardiac PET/MR, UTE/ZTE applications remain limited as bone visualization benefits must be weighed against increased acquisition time and motion artifact susceptibility [31].

2.3.4 Emission-Based Methods (MLAA)

Emission-based methods derive attenuation information directly from PET data. Maximum Likelihood reconstruction of Activity and Attenuation (MLAA) simultaneously estimates radiotracer distribution and attenuation maps through alternating optimization steps [36].

These methods operate independently from anatomical imaging, reflect actual attenuation properties during PET acquisition, and offer potential for more consistent corrections. However, they are highly sensitive to emission data noise [36], suffer from activity-attenuation cross-talk [37], face convergence issues, and perform substantially better with high-quality time-of-flight information.

In cardiac applications, MLAA confronts challenges from heterogeneous tracer uptake and motion. Hybrid approaches combining MLAA with MR-derived constraints have been proposed as potential solutions [37].

2.4 Overview of Deep Learning for Medical Image Translation

Medical image translation converts images from one modality to another, such as generating synthetic CT images from MR data. Deep learning approaches address the mismatch between MR signal and electron density information [38].

Convolutional neural networks (CNNs) learn hierarchical representations from image data. Key developments include encoder-decoder architectures with U-Net incorporating skip connections to preserve spatial information [38], Generative Adversarial Networks (GANs) enabling realistic image generation [39], and Cycle-Consistent Adversarial Networks (CycleGANs) enforcing cycle consistency for training from unpaired data [40].

For PET/MR AC, these frameworks have been applied through segmentation-based, atlas-based, direct attenuation map prediction, and MR-to-CT synthesis

methods.

2.4.1 Segmentation-Based Approaches using Deep Learning

Segmentation-based attenuation correction approaches segment MR images into discrete tissue classes (typically air, soft tissue, and bone) and assign predetermined attenuation coefficients to each tissue type [41]. Deep learning improves segmentation-based AC by enhancing tissue classification accuracy from MR images through complex feature representations [42]. The workflow involves training neural networks like U-Net to segment MR images into tissue classes, processing patient-specific MR images to generate tissue probability maps, and assigning predefined attenuation coefficients to each class.

Jang et al. [42] reported improved bone delineation compared to conventional methods. For cardiac applications, deep learning approaches better delineate boundaries between heart, lungs, and surrounding tissues compared to traditional segmentation-based methods.

Limitations include continued use of discrete tissue classes with uniform attenuation coefficients, substantial labeled training data requirements, and persistent challenges with respiratory and cardiac motion sensitivity.

2.4.2 Atlas-Based Approaches using Deep Learning for Registration

Atlas-based attenuation correction methods generate pseudo-CT images by using databases containing paired CT and MR images from a population to predict attenuation maps for new patients, typically through registration of the patient's MR image to pre-existing atlases [41, 43]. Atlas-based registration methods have traditionally relied on pre-defined CT templates and intensity-based algorithms to align a patient's anatomy with reference atlases. While effective, these conventional approaches often depend heavily on voxel intensity similarities, which can limit their accuracy in complex cases. Recent advancements in deep learning have significantly improved these methods by enabling more accurate and flexible image registration. Deep learning-based approaches include learning-based registration, where neural networks are trained to directly predict deformation fields between images, and feature-based registration, which involves extracting distinctive anatomical landmarks or texture patterns to guide spatial alignment. These feature-based techniques enhance the robustness of registration by focusing on meaningful image characteristics, often leading to better performance than traditional methods.

Burgos et al. [32] demonstrated the effectiveness of deep learning-enhanced atlas approaches for brain PET/MR. For cardiac applications, these methods offer advantages in handling anatomical variability and enhancing motion robustness.

Cardiac-specific challenges include requirements for diverse training data, continued dependence on atlas database representativeness, and difficulties accurately aligning tissue boundaries.

2.4.3 Direct Deep Learning-Based Attenuation Map Prediction

Direct attenuation map prediction trains deep networks to generate continuous-valued attenuation maps directly from MR images [44]. Implementation typically involves encoder-decoder architectures like U-Net, training with paired MR and CT-derived attenuation maps, and combinations of pixel-wise, perceptual, and adversarial loss functions.

Hwang et al. [44] demonstrated this approach for brain PET/MR, achieving substantial improvements over conventional methods such as Dixon-based segmentation techniques and atlas-based approaches that classify tissues into discrete categories. For cardiac applications, direct prediction offers advantages through continuous attenuation coefficients and end-to-end optimization.

Significant challenges include limited generalization due to the dependence on the representativeness of the training data, sensitivity to respiratory and cardiac motion, and the requirement for paired MR-CT training data.

2.4.4 Deep Learning-Based MR to Pseudo-CT Synthesis

MR to pseudo-CT synthesis generates synthetic CT images from MR input, which can then be converted to attenuation maps [40]. Various architectures have been employed including standard CNNs like U-Net, GANs generating more realistic images [39], CycleGANs enabling training from unpaired data [40], and U-Net with attention mechanisms focusing on relevant image regions [45].

Loss functions crucial for synthesized CT quality include several components: L1/L2 losses, which measure differences between predicted and ground truth values (L1 computes the mean absolute error $|y - \hat{y}|$, while L2 computes the mean squared error $(y - \hat{y})^2$); adversarial losses, which train the generator to produce images that fool the discriminator into classifying them as real; cycle consistency losses, which ensure that translating an image from domain A to B and back to A reconstructs the original image; and perceptual losses, which compare high-level features extracted

from pre-trained neural networks to encourage structural similarity. We discuss more about the loss terms in section 3.3.2.

Cardiac imaging applications face challenges from motion artifacts, complex thoracic anatomy with diverse tissue types, significant inter-subject variability, and the need to accurately represent sharp transitions at air-tissue and bone-tissue boundaries.

Despite these challenges, MR to pseudo-CT synthesis has shown promising results for cardiac applications [45]. The field continues evolving with emerging approaches focusing on multi-sequence integration, motion-aware synthesis, uncertainty quantification, and semi-supervised learning.

2.5 Generative Adversarial Networks in Medical Imaging

2.5.1 Principles of GANs

Generative Adversarial Networks (GANs) have revolutionized the field of medical image synthesis since their introduction by Goodfellow et al. in 2014 [46]. The fundamental principle of GANs involves a two-player adversarial game between a generator network and a discriminator network. The generator creates synthetic images, while the discriminator attempts to distinguish between real and synthetic images.

GANs operate through a minimax game between two neural networks: the generator G learns a mapping from random noise z to data space G(z), while the discriminator D learns to distinguish between real data x and generated samples G(z). Mathematically, this is formulated as:

$$\min_G \max_D V(D,G) = \mathbb{E}_{x\sim p_{data}(x)}[\log D(x)] + \mathbb{E}_{z\sim p_z(z)}[\log(1 - D(G(z)))] \tag{2.1}$$

where the generator minimizes this objective while the discriminator maximizes it. Through iterative training, the generator improves its ability to produce increasingly realistic images that can fool the discriminator.

In the context of medical imaging, GANs offer several advantages for MR-CT translation:

- The ability to learn complex, non-linear mappings between imaging modalities
- Preservation of anatomical structure and fine details

- Generation of realistic images with proper texture and statistical properties
- Less reliance on perfect alignment between training image pairs

However, GANs also present challenges, particularly in terms of training stability. Mode collapse (where the generator produces a limited variety of outputs), vanishing gradients (where gradients become extremely small during backpropagation, preventing effective learning in deeper network layers), and oscillatory training behavior (where the generator and discriminator fail to reach equilibrium, instead constantly fluctuating without stable improvement) are common issues. Various architectural and algorithmic innovations have been proposed to address these challenges, including Wasserstein GANs, spectral normalization, and progressive growing techniques [47].

2.5.2 CycleGANs for Unpaired Image Translation

A significant advancement in GAN architecture for medical imaging is the Cycle-Consistent GAN (CycleGAN), introduced by Zhu et al. This architecture enables translation between image domains without requiring paired training examples, which is particularly valuable in medical imaging where acquiring perfectly paired MR and CT data from the same patient is challenging due to anatomical deformations and respiratory motion between scans.

CycleGAN employs two generators: $G_{\text{MR}\to\text{CT}}$ for translating MR to CT and $G_{\text{CT}\to\text{MR}}$ for the reverse direction. Additionally, two discriminators, D_{CT} and D_{MR}, assess the realism of generated images in their respective domains. The key innovation is the cycle-consistency loss, which ensures that translating an image from one domain to another and back again should yield the original image:

$$\begin{aligned}\mathcal{L}_{cyc}(G_{\text{CT}\to\text{MR}}, G_{\text{MR}\to\text{CT}}) = \mathbb{E}_{\text{MR}}[||G_{\text{CT}\to\text{MR}}(G_{\text{MR}\to\text{CT}}(\text{MR})) - \text{MR}||_1] \\ + \mathbb{E}_{\text{CT}}[||G_{\text{MR}\to\text{CT}}(G_{\text{CT}\to\text{MR}}(\text{CT})) - \text{CT}||_1]\end{aligned} \tag{2.2}$$

In practice, these expected values are computed empirically by averaging over mini-batches of training samples. The L1 norm $\|\cdot\|_1$ operates on the pixel-wise differences between images, treating each pixel location independently. While individual pixels are univariate (single intensity values), the overall loss considers the full multivariate image structure by summing across all pixel locations in the image.

Dong et al. [48] applied 3D CycleGAN for MR-based AC in brain PET imaging, demonstrating its effectiveness in generating pseudo-CT images from standard T1-weighted MR sequences (where image contrast is primarily determined by the

T1 relaxation time of tissues, emphasizing differences in longitudinal magnetization recovery). The 3D dimension refers to the spatial structure of the input data: 2D approaches process individual image slices (height × width), while 3D approaches process entire volumetric data (height × width × depth), where depth corresponds to the number of consecutive slices in the imaging volume. For medical imaging, 3D approaches can better preserve spatial relationships between adjacent slices. Their approach achieved mean absolute errors of approximately 100 Hounsfield Units (HU), which quantify X-ray attenuation in CT images on a standardized scale where water = 0 HU, air = -1000 HU, and dense bone ≈ +1000 HU, compared to true CT, with PET quantification errors below 3% in most brain regions. These quantification errors refer to inaccuracies in measuring radiotracer uptake (typically expressed as Standardized Uptake Values or SUVs), which result from improper AC. Accurate AC is essential for reliable quantification of metabolic activity. A key advantage was the ability to train on unpaired datasets, eliminating the need for difficult-to-acquire, perfectly co-registered MR-CT pairs. For unpaired training, CycleGAN relies solely on the cycle consistency loss and adversarial losses, learning domain translation without direct pixel-wise supervision. The cycle consistency loss ensures that G(F(x)) ≈ x and F(G(y)) ≈ y, providing indirect supervision. When paired data is available, additional supervised losses can be incorporated, such as L1 loss between G(x) and the corresponding ground truth y, combining the benefits of both supervised and unsupervised learning.

Recent enhancements to CycleGAN for medical applications include attention mechanisms to focus on anatomically relevant features, shape-consistency constraints to preserve critical structures, and the incorporation of physical models to ensure the generated images conform to known physical properties of the imaging modalities [47]. Attention mechanisms are implemented using learned weight matrices that compute relevance scores for different spatial locations, enabling the network to focus on anatomically important regions. This is typically achieved by computing attention weights $\alpha_{ij} = \text{softmax}(f(h_i, h_j))$, where h_i and h_j are hidden representations at different spatial locations.

2.5.3 Advanced GAN Architectures for Medical Image Synthesis

Dual Contrast Mechanisms in cycleGANs

A notable advancement in GAN architectures for medical image synthesis is the Dual Contrast cycleGAN (DC-cycleGAN), proposed by Wang et al. [9]. This ap-

proach extends the standard CycleGAN framework by incorporating a dual contrast mechanism that enhances the discriminator's ability to distinguish between real and synthetic images.

In traditional CycleGANs, the discriminator only evaluates whether an image is real or generated using convolutional neural networks with binary classification at the output layer, typically employing sigmoid activation and binary cross-entropy loss to produce a final probability score. The DC-cycleGAN introduces an additional task for the discriminator: distinguishing between real images from the target domain and real images from the source domain. In domain translation, the source domain refers to the input modality (e.g., MR images) while the target domain refers to the desired output modality (e.g., CT images). This dual contrast mechanism encourages the generator to not only produce realistic images but also to capture the specific characteristics that differentiate between imaging modalities.

Formally, the dual contrast loss is defined as:

$$\mathcal{L}_{\mathrm{DC}}(D_{\mathrm{CT}}) = \mathbb{E}_{\mathrm{CT},\mathrm{CT}'}[\log D_{\mathrm{CT}}(\mathrm{CT}) + \log(1 - D_{\mathrm{CT}}(\mathrm{CT}'))] \tag{2.3}$$

where CT' represents negative samples randomly selected from the source domain. A similar loss is applied for the MR discriminator.

Talebianet al. [49] demonstrated the efficacy of 3D DC-cycleGAN for multi-contrast MRI synthesis, showing superior performance compared to standard CycleGAN in preserving fine structural details and tissue contrast. The authors reported improvements in SSIM of up to 8% and reductions in MAE of approximately 15% when using the dual contrast mechanism.

For PET/MR AC, the dual contrast approach offers significant advantages in generating pseudo-CT images that accurately represent bone structures, which are critical for proper AC. The enhanced discriminative capability helps the generator produce more realistic bone textures and densities, leading to more accurate μ-maps and improved PET quantification.

2.6 Image Registration in Multi-Modal Cardiac Imaging

Accurate spatial alignment of imaging modalities is essential for AC in cardiac PET/MR, as misalignments cause quantification errors. Image registration is a computational process that finds the optimal spatial transformation to align corresponding anatomical structures between two or more images. The process involves three key components: a similarity metric (such as mutual information or normalized

cross-correlation) that quantifies how well images are aligned, a transformation model (rigid, affine, or deformable) that defines the allowable spatial mappings, and an optimization algorithm that searches for transformation parameters that maximize image similarity. This enables integration of anatomical and functional information in a unified spatial framework.

Registration methods include rigid registration, which applies only translations (shifts in x, y, z directions) and rotations around three axes, preserving distances and angles between all points in the image; affine registration, which additionally allows scaling, shearing, and reflection transformations; and deformable registration, which uses spatially varying transformations modeled through spline functions or vector fields to account for local tissue deformations and cardiac motion [50]. Template-based registration creates a standardized anatomical reference (template) by averaging multiple aligned images, then registers new patient images to this template using iterative optimization algorithms that minimize intensity differences or maximize statistical similarity measures, ensuring consistent alignment across modalities such as CT and MR in PET/CT and PET/MR workflows [51].

Cardiac PET/MR registration faces challenges, including respiratory motion causing heart displacement [20], cardiac contractile motion altering internal structures [50], differing temporal resolutions between MR single-phase captures and PET multi-cycle accumulation [52], and contrast differences between MR, PET, and CT.

Misregistration results in perfusion defects and reduced uptake [20]. Strategies like motion-averaged imaging, respiratory gating, motion modeling, and deep learning-based registration mitigate these issues. Template-based methods, often using motion-averaged data, enhance alignment consistency for AC [51, 20]. Evolving techniques, including integrated motion compensation in PET reconstruction, underscore the importance of spatial alignment for reliable diagnostics.

2.7 Research Gaps and Motivation

Accurate pseudo-CT generation from MR data for cardiac PET/MR AC remains challenging, resulting in PET quantification errors. Segmentation and atlas-based methods fail to capture continuous attenuation variations and are prone to registration errors and inter-subject variability [53, 54]. UTE/ZTE sequences, despite recent advances, are hindered by motion artifacts and long acquisition times [55, 56]. Emission-based approaches suffer from noise and dependence on high-quality TOF data [57].

Deep learning methods for pseudo-CT generation show promise but lack cardiac-specific development and rely on scarce paired MR-CT training data [58, 59]. Voxel-

based models excel at local details but miss global anatomical context, while GAN-based methods often fail to preserve precise cardiac features [60].

This thesis addresses these gaps by developing novel methods for pseudo-CT generation, which are then transformed into attenuation maps, as outlined in Chapter 1:

1. **CT Template and Multi-Modal Registration Framework**: Develop a CT template and registration pipeline to spatially align MR and CT images, ensuring consistent anatomical correspondence for model training and evaluation.

2. **Kernel Ridge Regression (KRR) for MR-to-CT Translation**: Investigate KRR as a baseline for unidirectional MR-to-CT translation, capturing localized anatomical patterns.

3. **Dual Contrast cycleGAN (DC-cycleGAN) with Mumap Loss**: Develop a DC-cycleGAN model, with a novel mumap loss, for bidirectional MR-CT translation, optimizing pseudo-CT and pseudo-μ-map generation for PET AC.

4. **Comparison with Traditional Methods**: Evaluate the proposed methods against traditional MR-based AC, using metrics like MAE, SSIM, and SUV errors to assess PET reconstruction accuracy.

These innovations enhance PET quantification accuracy and diagnostic confidence [61], enabling broader adoption of PET/MR for cardiac monitoring [62]. The methodologies may also improve radiation therapy planning and other inter-modal image translation tasks, unlocking the full potential of PET/MR imaging.

2.8 Conclusion

This review examined the evolution of AC methods in PET imaging, focusing on cardiac PET/MR. The transition to PET/MR introduced challenges due to the limitations of MR for characterizing photon attenuation.

Key challenges include the mismatch between MR signal and photon attenuation, limitations of standard MR sequences, the impact of MR hardware and limited field of view, physiological motion, and artifacts from cardiac devices [25]. Deep learning techniques show promise for improving segmentation accuracy, enhancing registration, predicting attenuation maps, and synthesizing pseudo-CT images [38, 17].

Nevertheless, knowledge gaps remain, particularly in cardiac applications, motion compensation, and combining local and global information [40]. This thesis aims to address these gaps through a comprehensive approach integrating multiple methodologies.

By improving the quantitative accuracy of cardiac PET/MR, this research could enhance diagnostic confidence and facilitate wider adoption of this modality. The proposed methods hold promise for broader applications in medical image translation. Robust AC methods remain critical for realizing the full diagnostic potential of hybrid imaging technologies.

Chapter 3

Methodology

This chapter describes the data acquisition process and the preprocessing steps implemented to prepare the imaging data for training. The dataset was obtained through a validation study titled "Optimization of Attenuation Correction for FDG PET-MR Imaging" conducted at The University of Ottawa Heart Institute (UOHI) and the University of Ottawa Institute of Mental Health Research (IMHR) [63]. The following sections detail the dataset characteristics, the preprocessing pipelines developed for both MR and CT modalities, the image registration framework, and the two image translation approaches employed in this thesis.

3.1 Dataset Description and Preprocessing

This study used FDG-PET-CT and FDG-PET-MR scans from 10 participants, with each participant undergoing two separate imaging visits at respective institutions. While this sample size is modest compared to large-scale computer vision datasets, it is comparable to other medical imaging studies in PET/MR attenuation correction research, where acquiring paired, high-quality medical data is challenging due to patient scheduling, safety considerations, institutional logistics, and the cost of dual-modality imaging. The use of 8 subjects for training and 2 for testing follows established practices in medical imaging research for proof-of-concept validation.

Participants were selected based on distinct known imaging patterns that provide essential validation scenarios for attenuation correction methods:

- Controls (participants with no history of myocardial infarction or cardiac disease, confirmed by normal perfusion patterns on previous cardiac scans) - these establish baseline performance for the algorithms

- Subjects with previous myocardial infarction demonstrating clear mismatch in FDG uptake between PET and CT, indicating viable but underperfused myocardium - these present challenging cases with metabolically heterogeneous tissue that test the algorithm's ability to accurately correct for attenuation in regions with varying uptake patterns

Data Structure and Dimensions

The dataset comprises paired MR and CT images from the same patients acquired on separate scanners: PET/MR (Siemens mMR) at IMHR and PET/CT (GE Discovery 690) at UOHI. Each subject contributes approximately 126 axial slices with 192×128 pixel resolution, resulting in 3D volumes of 192×128×126 voxels with voxel spacing of 2.604×2.604×3.12 mm. MR intensities are arbitrary units ranging from 0-4095 (12-bit), while CT values are in Hounsfield Units ranging from -1000 (air) to +3000 (dense bone). After preprocessing and normalization, MR and CT intensities are scaled to [0,1], maintaining the relative intensity relationships essential for cross-modal translation.

The analysis operates on quantitative imaging data with different dimensionality approaches: For the KRR method, univariate intensity values are extracted from each voxel along with multivariate neighborhood features (mean, standard deviation, range from 3×3×3 patches), creating 4-dimensional feature vectors per voxel. For the GAN-based approach, the data is inherently multivariate, processing full 2D image slices (192×128 pixels, resized to 256×256 for network input) where each pixel represents a spatial location with associated intensity values.

The PET/CT scans were performed on a GE Discovery 690 PET-CT system at the University of Ottawa Heart Institute (UOHI), while PET/MR scans used a Siemens mMR system at The Royal's Institute of Mental Health Research (IMHR). For this thesis, we specifically utilized in-phase and out-of-phase MR images (Dixon sequences where in-phase images are acquired when fat and water protons have the same phase, while out-of-phase images are acquired when they have opposite phases, enabling fat-water separation for tissue classification).

The dataset adheres to standardized quality control procedures, including consistent participant preparation (8-hour fasting except water) and glucose loading protocols prior to imaging. This standardization ensures data consistency, which is crucial for evaluating AC methods and facilitating reproducibility of the research.

3.1.1 MR Preprocessing Pipeline

The MR preprocessing pipeline contains three primary steps: bias field correction, registration to a template, and intensity normalization with rescaling. Each step addresses specific challenges in MR imaging to ensure robust data for subsequent analysis.

1.

3.1.2 Bias Field Correction

: The N4 bias field correction algorithm is applied to mitigate intensity inhomogeneity inherent in MR images. This inhomogeneity is primarily caused by Radio Frequency coil sensitivity variations and magnetic field non-uniformities in the scanner. This correction is critical, as bias field artifacts can adversely affect the accuracy of downstream image processing and analysis tasks.

2. **Spatial Registration**: Images are spatially registered to a common template space to ensure anatomical alignment across subjects, facilitating consistent comparisons and analyses.

3. **Intensity Normalization and Rescaling**: Due to the arbitrary intensity scales in MR images, intensity normalization and rescaling are performed to standardize image intensities. This process follows a two-step approach:

1. **Relative Intensity Normalization**: Using liver tissue as a reference, the image intensities are normalized as follows:

$$I_{\text{normalized}} = I_{\text{original}} \times \frac{\mu_{\text{liver, all}}}{\mu_{\text{liver, individual}}}$$

where:

- I_{original}: The original intensity value of a voxel in the MR image.
- $I_{\text{normalized}}$: The intensity value after relative normalization.
- $\mu_{\text{liver, all}}$: The mean intensity of liver tissue across all subjects.
- $\mu_{\text{liver, individual}}$: The mean intensity of liver tissue for an individual subject.

The mean liver intensity is calculated by manually segmenting liver regions in each MR image, then computing the average intensity value across all voxels within the segmented liver region.

2. **Global Min-Max Normalization**: The normalized intensities are rescaled to the range [0, 1] using:

$$I_{\text{final}} = \frac{I_{\text{normalized}} - \min(I_{\text{global}})}{\max(I_{\text{global}}) - \min(I_{\text{global}})}$$

where:

- I_{final}: The final intensity value after rescaling.
- $I_{\text{normalized}}$: The intensity value from the previous normalization step.
- $\min(I_{\text{global}})$: The global minimum intensity across the dataset.
- $\max(I_{\text{global}})$: The global maximum intensity across the dataset.

The liver is selected as the reference region for intensity normalization due to its homogeneous tissue composition and stable MR signal characteristics in healthy subjects. Its large size and consistent appearance make it an ideal reference structure. Additionally, liver tissue is less prone to pathological changes compared to other organs, ensuring reliable intensity standardization across subjects.

3.1.3 CT Preprocessing

CT image preprocessing involves registration, intensity normalization, and additional steps to ensure compatibility with MR images and suitability for deep learning pipelines. Although CT values (Hounsfield Units, HU) are standardized across scanners and institutions, spatial registration and further processing are necessary for anatomical correspondence and consistent intensity ranges.

The preprocessing pipeline includes the following steps:

1. **Manual Bed Scanner Removal**: The scanner bed is manually segmented and removed from the CT images using 3D Slicer to eliminate artifacts that could interfere with subsequent analysis.

2. **Spatial Registration**: Images are registered to a common template space to ensure anatomical alignment with corresponding MR images, facilitating cross-modal comparisons.

3. **Value Clipping**: CT intensities are clipped to the range [-1000, +1000] HU to focus on the typical density range of human tissues:

$$I_{\text{clipped}} = \begin{cases} -1000 & \text{if } I_{\text{original}} < -1000 \\ I_{\text{original}} & \text{if } -1000 \leq I_{\text{original}} \leq 1000 \\ 1000 & \text{if } I_{\text{original}} > 1000 \end{cases}$$

where I_{original} is the original CT intensity value in Hounsfield Units, and I_{clipped} is the intensity after clipping.

4. **Linear Scaling**: The clipped intensities are scaled to the range [-1, 1] for deep learning compatibility:

$$I_{\text{scaled}} = \frac{I_{\text{clipped}}}{1000}$$

where I_{scaled} is the final scaled intensity value.

This pipeline preserves relative tissue density information while standardizing the intensity range. The clipping range of [-1000, +1000] HU encompasses the typical density range of human tissues. The HU ranges for common tissue types are summarized in Table 3.1.

Table 3.1: Hounsfield Unit (HU) ranges for common tissue types in CT imaging with corresponding scaled values.

Tissue Type	**HU Range**	**Scaled Range [-1, 1]**
Air	< -950	< -0.95
Lung	-950 to -650	-0.95 to -0.65
Fat	-100 to -50	-0.1 to -0.05
Soft Tissue	-50 to 100	-0.05 to 0.1
Bone	> 300	> 0.3

3.2 Template Creation and Registration

This section outlines the creation of an unbiased CT template and the registration process to align CT and MR images into a common template space.

3.2.1 CT Template Generation

The CT template is generated using the Advanced Normalization Tools (ANTs) framework through a multi-step process, as illustrated in Steps 1 and 2 of Figure 3.1. First, each CT image is registered to its corresponding MR image to align them into the same patient space. This registration employs a three-stage approach: rigid, affine, and deformable transformations, as described in section 2.6. Next, the registered CT images from all subjects are averaged to create an unbiased CT template, providing a common anatomical reference for subsequent registration.

3.2.2 Registration Framework

The registration framework is a critical component of the methodology, designed to align all original CT and MR images to a common CT template space, as illustrated in Figure 3.1. This process ensures accurate spatial correspondence across modalities, which is essential for effective model training and subsequent PET AC. The framework was specifically developed with the AC of PET images from the PET/MR scanner at The Royal's Institute of Mental Health Research (IMHR) in mind.

The registration process is structured in three distinct steps, as shown in Figure 3.1. In Step 1, each CT image is registered to its corresponding MR image. This involves a three-stage registration process comprising rigid, affine, and deformable transformations to account for differences in patient positioning and scanner-specific distortions, particularly since the CT and MR images were acquired during separate visits (PET/CT at UOHI and PET/MR at IMHR). The result is a set of CT images aligned with their corresponding MR images, ensuring anatomical consistency between the two modalities.

In Step 2, the CT images that have been registered to their corresponding MR images are used to create a CT template. This is achieved by averaging the registered CT images, producing a standardized CT template that represents a common anatomical reference space. The template serves as a fixed target for subsequent registrations, mitigating variability across subjects and imaging sessions, which is crucial for consistent model performance in downstream tasks like synthetic CT generation.

In Step 3, all original MR and CT images are registered to the CT template created in Step 2. The MR images are directly transformed to the template space using a similar three-stage registration process, ensuring they align with the template's anatomical structure. For the CT images, a chain of transformations is applied: first, the CT is registered to its corresponding MR (as in Step 1), and then the resulting transform is composed with the MR-to-template transform to map the CT to the template space (CT$\rightarrow$MR$\rightarrow$Template). This multi-step approach ensures that both MR and CT images are spatially aligned in the same template space, facilitating accurate cross-modality synthesis and PET AC later.

From a data analysis perspective, each imaging modality represents a multivariate dataset where each voxel location (x,y,z) has associated intensity values. The registration process aligns these multivariate spatial datasets to ensure corresponding anatomical structures occupy the same voxel coordinates across modalities.

This registration workflow and template creation process were specifically chosen to support the AC needs of the PET/MR scanner at The Royal's Institute of Mental Health Research (IMHR). The framework ensures that new MR images acquired at IMHR can be consistently aligned to a common space (template) for standardized

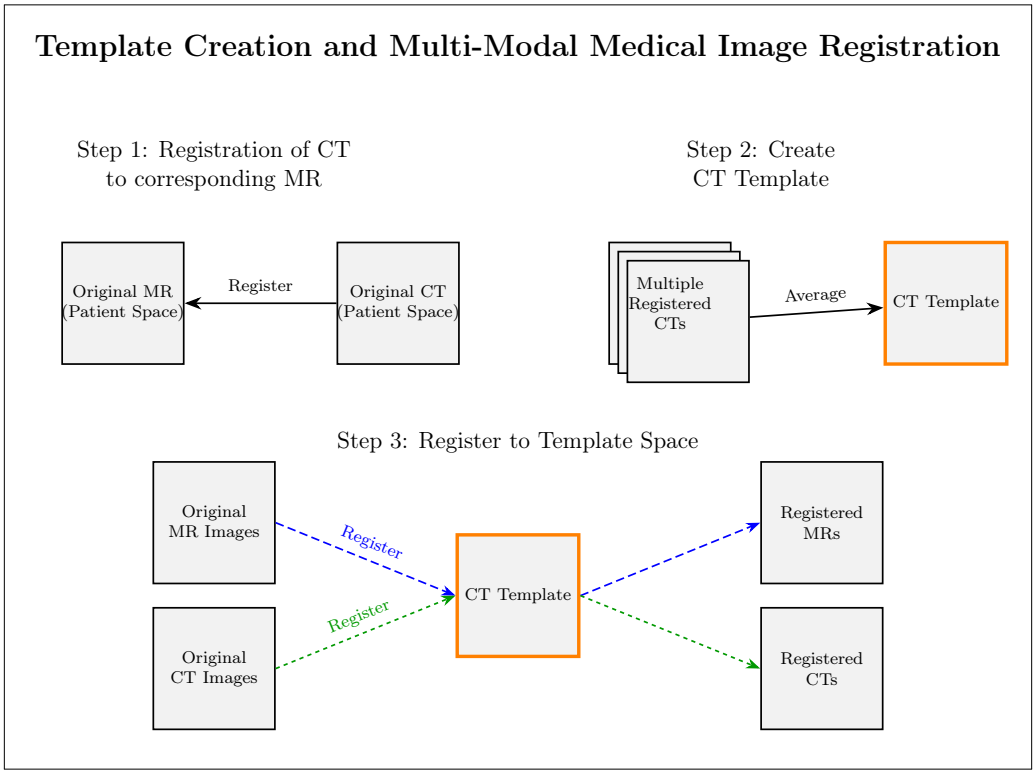

Figure 3.1: Workflow for template creation and CT/MR image registration in the template space.

processing, while allowing the final attenuation maps to be generated in the original MR space. Specifically, when IMHR researchers acquire a new MR image (without corresponding CT) for which a synthetic CT is needed, the process will involve registering this MR image to the pre-established CT template using the three-stage registration process. The registered MR image, now in template space, is then used to generate a synthetic CT in the same template space. Subsequently, the synthetic CT is mapped back to the original MR space by applying the inverse of the MR-to-template transformation. This step is critical because the MR and PET data are inherently aligned in the PET/MR scanner, as they are acquired simultaneously during the same session. In the MR space, the synthetic CT is used to generate (μ-maps) that aligns with the PET data, which is then applied for PET AC.

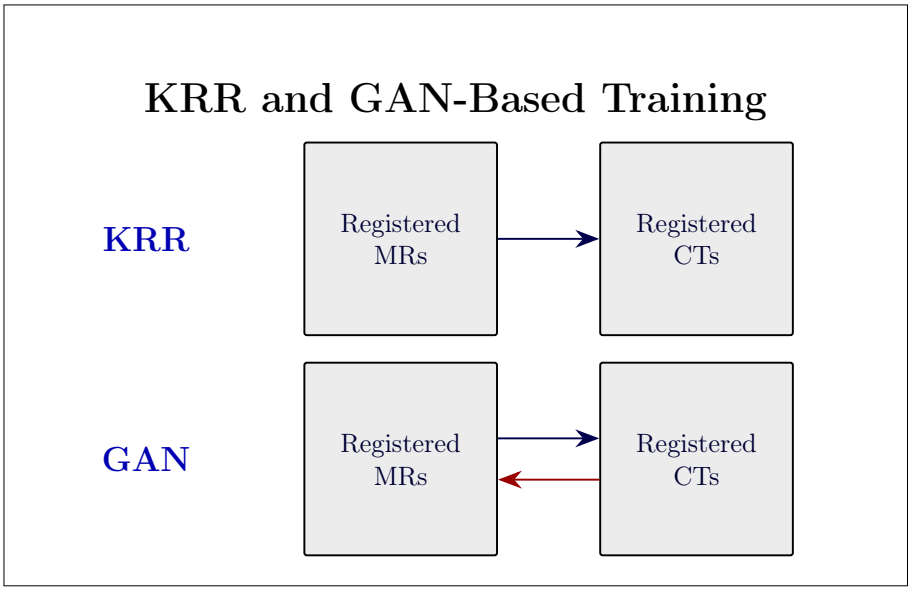

Figure 3.2: Kernel Ridge Regression (KRR) and Generative Adversarial Network (GAN) training. The first row illustrates the unidirectional training of KRR from Registered MRs to Registered CTs. The second row depicts the bidirectional mapping between Registered MRs and Registered CTs using a GAN.

3.3 Image Translation Framework

Having established a common spatial framework, the next step is to develop methods for translating between MR and CT image contrasts. Image translation models are used to 1) generate synthetic CT images, which are then transformed into 2) μ-maps for AC of PET images. This section describes two approaches for translating between MR and CT images, as illustrated in Figure 3.2: a single-voxel/small patch translation using Kernel Ridge Regression (KRR) for unidirectional MR-to-CT mapping, and a deep learning-based approach using a dual contrast cycleGAN architecture for bidirectional MR-to-CT and CT-to-MR mapping. Both models were trained using data from 8 subjects and evaluated on the remaining 2 subjects to assess generalization performance.

3.3.1 Single-voxel/Small Patch Translation

As a baseline approach to investigate simpler methods for MR-to-CT image translation, we employ KRR, a machine learning technique that combines ridge regression

with the kernel trick to model non-linear relationships [10]. This method operates at the voxel level (where a voxel (volumetric pixel) represents a discrete 3D unit of space, analogous to pixels in 2D images but with additional depth dimension), treating each voxel's intensity and its local neighborhood as an independent translation problem. KRR provides a baseline method for generating synthetic CT images, offering a point of comparison with the more complex GAN-based approach.

Algorithm Design and Implementation

KRR, as described by Vovk (2013), learns a non-linear mapping between MR and CT intensities by solving the optimization problem in its dual form:

$$\min_{\alpha} \left(\|y - K\alpha\|^2 + \lambda \alpha^T K \alpha \right)$$

where:

- y: Target CT intensity values, a vector of length n, where n is the number of training voxels ($n = 10\ 000$).
- K: Kernel matrix of size $n \times n$, with $K_{ij} = k(x_i, x_j)$, computed using MR intensity vectors x_i and x_j.
- $k(x_i, x_j)$: Kernel function measuring similarity between MR intensity vectors.
- α: Dual coefficients, a vector of length n, where each α_i determines the influence of the i-th training voxel on predictions.
- λ: Regularization parameter balancing model fit and complexity.

The loss term $\|y - K\alpha\|^2$ ensures the predicted CT values closely match the actual values, while the regularization term $\lambda \alpha^T K \alpha$ prevents overfitting by penalizing large coefficients, with K scaling the penalty based on voxel similarities.

We use a Gaussian kernel, also known as the Radial Basis Function (RBF) kernel, defined as:

$$k(x_i, x_j) = \exp\left(-\frac{\|x_i - x_j\|^2}{2\sigma^2} \right)$$

where $\|x_i - x_j\|$ is the Euclidean distance between MR intensity vectors, and σ is a hyperparameter controlling the kernel's width, selected using 3-fold cross-validation with RandomizedSearchCV optimizing for R^2 score. The Gaussian kernel assigns

higher similarity to MR intensity vectors that are closer in value, making it effective for capturing localized patterns in image data, such as anatomical similarities.

The dual coefficients α are computed via a closed-form solution:

$$\alpha = (K + \lambda I)^{-1} y$$

where I is the $n \times n$ identity matrix. Each α_i reflects the importance of the corresponding training voxel in predicting CT intensities, with larger values indicating greater influence. During prediction, the CT intensity for a new MR intensity vector x' is computed as:

$$\hat{y} = \sum_{i=1}^{n} \alpha_i k(x_i, x')$$

This weighted sum relies on the similarities $k(x_i, x')$, ensuring that training voxels with MR intensities similar to x' contribute more to the prediction.

For feature extraction, we use not only the raw MR intensity but also three additional patch-based features derived from 3×3 neighborhood patches: the mean intensity, standard deviation, and intensity range (maximum - minimum value). These patch-based features help capture local tissue context and structural patterns that a single intensity value cannot represent.

The analysis uses imaging data of the torso (neck to abdomen), with each individual contributing multivariate information: spatial coordinates (x,y,z), intensity values, and neighborhood statistics. The KRR model processes this multivariate data at the voxel level across the entire imaging volume, not limited to specific organs."

To enhance robustness, we implement an ensemble of five independent KRR models, each trained on random subsets of voxels. For computational efficiency, we limit each model to a maximum of n = 10,000 randomly selected voxels accross the training subjects, leading to a $10,000 \times 10,000$ kernel matrix K and a 10,000-dimensional α. This ensemble approach averages predictions across models, reducing variance and improving stability.

3.3.2 DC-cycleGAN for Bidirectional MR-CT Image Translation

To enable bidirectional MR-to-CT and CT-to-MR translation, we employ a Dual Contrast cycleGAN (DC-cycleGAN), a specialized variant of the CycleGAN introduced by Zhu et al. (2017) [8], as shown in Figure 3.3, and extended with dual contrast mechanisms by Wang et al. (2023) [9], as illustrated in Figure 3.4. Unlike

the unidirectional KRR method, DC-cycleGAN simultaneously learns mappings between MR and CT domains, with a key objective of generating pseudo-CT images and pseudo-μ-maps that closely resemble true CT images and μ-maps for accurate PET AC, achieved by introducing a novel mumap loss component alongside cycle consistency and dual contrast mechanisms to enhance synthesis quality.

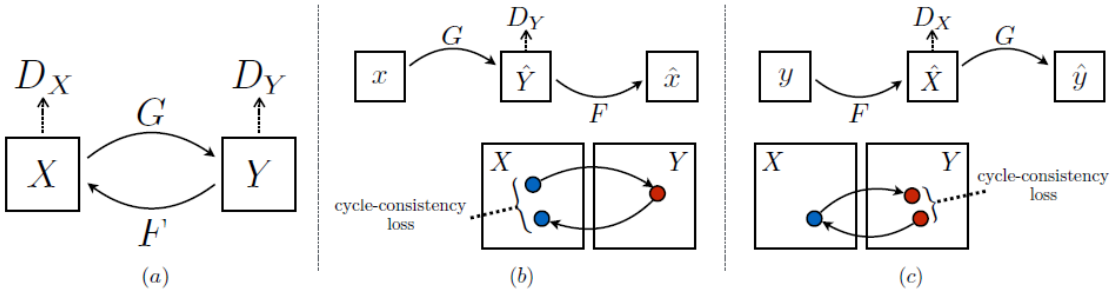

Figure 3.3: CycleGAN architecture, reproduced from Zhu et al. (2017) [8]. The figure shows: (a) the model with two mapping functions $G : X \rightarrow Y$ and $F : Y \rightarrow X$, and associated adversarial discriminators D_Y and D_X. D_Y encourages G to translate X into outputs indistinguishable from domain Y, and vice versa for D_X and F. To further regularize the mappings, two cycle consistency losses are introduced that capture the intuition that if we translate from one domain to the other and back again, we should arrive at where we started: (b) forward cycle-consistency loss: $x \rightarrow G(x) \rightarrow F(G(x)) \approx x$, and (c) backward cycle-consistency loss: $y \rightarrow F(y) \rightarrow G(F(y)) \approx y$.

Network Architecture

The DC-cycleGAN architecture consists of two generators and two discriminators, designed to handle the structural and intensity variations between MR and CT modalities.

Generators

- **Structure**: Following Wang et al. (2023), the generators are ResNet-based with 9 residual blocks, including stride-2 convolutions for downsampling and fractionally strided convolutions for upsampling, optimized for medical image synthesis.

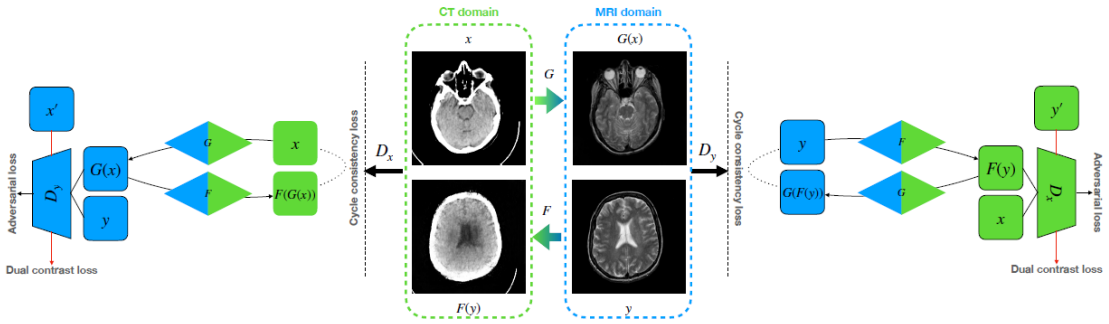

Figure 3.4: DC-cycleGAN architecture for bidirectional CT-to-MR synthesis, reproduced from Wang et al. (2023) [9]. The figure illustrates the proposed DC-cycleGAN model for medical image synthesis from unpaired data. It receives real CT/MR images through a generator to synthesize MR/CT images and discriminators distinguishing real images from generated and real images from the source domain. Note that x' and y' are negative samples that are randomly selected from source images.

- **Function**: Generator G translates MR to CT (producing pseudo-CT images), while F translates CT to MR (producing pseudo-MR images).

Discriminators

- **Structure**: PatchGAN discriminators classify 70×70 pixel patches as real or fake, enhanced with dual contrast loss as per the DC-cycleGAN framework.
- **Function**: Discriminator D_{CT} distinguishes real CT images from pseudo-CT images, and D_{MR} distinguishes real MR images from pseudo-MR images.

Loss Functions and Training

The DC-cycleGAN is trained using a combination of standard losses (adversarial, cycle consistency, dual contrast) and a novel mumap loss to achieve AC objectives for PET imaging.

Standard DC-cycleGAN Losses

- **Adversarial Loss**: Ensures generated images are indistinguishable from real ones:

$$\mathcal{L}_{\text{GAN}}(G, D_{\text{CT}}, \text{MR}, \text{CT}) = \mathbb{E}_{\text{ct}}[\log D_{\text{CT}}(\text{ct})] + \mathbb{E}_{\text{mr}}[\log(1 - D_{\text{CT}}(G(\text{mr})))] \quad (3.1)$$

The expected values $\mathbb{E}_{\text{ct}}[\cdot]$ and $\mathbb{E}_{\text{mr}}[\cdot]$ are computed empirically by averaging over mini-batches of training samples from the CT and MR domains respectively. The discriminator outputs a single probability score per image, and the loss is averaged across all samples in the batch.

A similar loss applies for F and D_{MR}.

- **Cycle Consistency Loss**: Ensures reconstruction of original images:

$$\mathcal{L}_{\text{cyc}}(G, F) = \mathbb{E}_{\text{mr}}[\|F(G(\text{mr})) - \text{mr}\|_1] + \mathbb{E}_{\text{ct}}[\|G(F(\text{ct})) - \text{ct}\|_1] \tag{3.2}$$

  Weighted by $\lambda_{\text{cyc}} = 10$.

- **Dual Contrast Loss**: Uses negative samples[1] (x', y') to enhance discrimination, improving synthesis quality in medical imaging [9]:

$$\mathcal{L}_{\text{DC}}(D_{\text{CT}}) = \mathbb{E}_{\text{ct,ct}'}[\log D_{\text{CT}}(\text{ct}) + \log(1 - D_{\text{CT}}(\text{ct}'))] \tag{3.3}$$

  A similar loss applies for D_{MR}.

Novel Mumap Loss for Attenuation Correction

Attenuation maps (μ-maps) are critical for PET imaging, as they quantify the linear attenuation coefficients (μ) of tissues for 511 keV photons, enabling accurate correction of photon attenuation during image reconstruction. To ensure that pseudo-CT images generated by the DC-cycleGAN model produce accurate μ-maps for PET/MR AC, we introduce a novel mumap loss ($\mathcal{L}_{\text{mumap}}$).

This μ-map loss represents a key modification to the standard CycleGAN framework, specifically leveraging the paired nature of our dataset. While standard CycleGAN does not include direct comparison between synthesized and real images, our paired data structure allows us to implement this direct supervision signal that is crucial for optimizing AC accuracy. Rather than viewing paired data as a limitation, this approach demonstrates how to effectively utilize available paired samples for clinically-relevant training objectives. As discussed in the conclusion, the framework's design allows for future extension to mixed datasets containing both paired and unpaired examples, where the cycle consistency component would handle unpaired samples while the μ-map loss provides targeted supervision for paired data. This flexibility makes the approach particularly valuable for medical imaging applications where paired data availability varies across institutions.

[1] A negative sample refers to an image from the source domain (MR) that is used to help the discriminator learn to distinguish not only between real and generated CT images, but also between CT images and unrelated MR images.

μ-Map Generation

Both true and pseudo-μ-maps are derived from CT images (or pseudo-CT images), where Hounsfield Units (HU) represent the linear attenuation coefficients of tissues relative to water and are directly encoded as pixel values in the CT image data. These HU values are then converted to linear attenuation coefficients (μ) for 511 keV photons using a piecewise linear function, as proposed by Carney et al. (2006) [3]:

$$\mu(\text{HU}) = \begin{cases} 9.6 \times 10^{-5} \times (\text{HU} + 1000) & \text{if HU} \leq 50 \\ 5.10 \times 10^{-5} \times (\text{HU} + 1000) + 4.71 \times 10^{-2} & \text{if HU} > 50 \end{cases}$$

This function ensures that:

- Air (HU $= -1000$) has negligible attenuation ($\mu \approx 0\,\text{cm}^{-1}$).
- Water (HU $= 0$) has a standard attenuation coefficient ($\mu \approx 0.096\,\text{cm}^{-1}$).
- Higher-density tissues, such as bone, have appropriately higher attenuation coefficients (e.g., $\mu \approx 0.149\,\text{cm}^{-1}$ for HU $= 1000$).

These parameters were selected based on the calibration provided by Carney et al. (2006), ensuring accurate representation of tissue attenuation properties for our dataset.

Mumap Loss Definition

The mumap loss measures the L1 distance between pseudo-μ-maps (derived from pseudo-CT images) and true μ-maps (derived from actual CT images):

$$\mathcal{L}_{\text{mumap}}(G) = \mathbb{E}_{\text{mr,ct}}[\|\mu(G(\text{mr})) - \mu(\text{ct})\|_1] \tag{3.4}$$

where:

- $G(\text{mr})$: The pseudo-CT image generated from the MR image by the DC-cycleGAN model.
- ct: The corresponding true CT image.
- $\mu(\cdot)$: The μ-map conversion function defined above.

This loss encourages the generator G to produce pseudo-CT images that, when converted to μ-maps, closely resemble the true μ-maps, thereby improving the accuracy of AC in PET reconstruction. By directly comparing the attenuation properties derived from synthesized and true CT images from the same subject, this loss function provides a strong supervision signal that would not be possible with traditional unpaired CycleGAN implementations.

Integration

The mumap loss is added to the DC-cycleGAN objective with a weight of $\lambda_{\text{mumap}} = 15$, determined through empirical validation on the training set to balance the novel mumap loss with standard CycleGAN losses (adversarial, cycle consistency, and dual contrast losses). This weighting prevents any single loss component from dominating training while ensuring the model learns both general domain translation and task-specific attenuation correction.

The total loss function is:

$$\begin{aligned}\mathcal{L} = &\mathcal{L}_{\text{GAN}}(G, D_{\text{CT}}, \text{MR}, \text{CT}) + \mathcal{L}_{\text{GAN}}(F, D_{\text{MR}}, \text{CT}, \text{MR}) \\ &+ \lambda_{\text{cyc}}\mathcal{L}_{\text{cyc}}(G, F) + \mathcal{L}_{\text{DC}}(D_{\text{CT}}) + \mathcal{L}_{\text{DC}}(D_{\text{MR}}) + \lambda_{\text{mumap}}\mathcal{L}_{\text{mumap}}(G)\end{aligned} \tag{3.5}$$

The model was trained on paired MR and CT images from 8 subjects, using 2D slices extracted from the registered volumes. The pairing allows for both cycle consistency training (characteristic of CycleGAN) and direct supervision through our novel μ-map loss.

The remaining 2 subjects were used for evaluation, assessing the quality of synthetic CT images and their derived μ-maps through metrics like MAE and SSIM.

Application of CycleGAN to Paired Data

While CycleGAN was originally designed for unpaired image-to-image translation, this research applies a modified DC-cycleGAN architecture to paired medical imaging data (though not perfectly aligned). Previous research has shown that supervised approaches can outperform unsupervised ones when paired data is available. Wang (2023) notes: "In Li et al. (2020), a proposed method compared the performance of the classic CycleGAN, an unsupervised medical image synthesis model, with U-net, a supervised learning model. The results indicated that U-net outperformed CycleGAN. However, it is important to note that U-net requires paired CT-MR images, while CycleGAN does not have this requirement" [9].

Despite this potential performance gap between unsupervised CycleGAN and supervised methods like U-net, our approach offers several advantages: (1) it maintains the robust domain translation capabilities of CycleGAN (the ability to learn meaningful mappings between different imaging modalities even when training data contains anatomical variations and noise); (2) it allows for direct supervision through the novel μ-map loss, addressing the typical limitations of unsupervised methods; and (3) it provides a framework that could be extended to mixed datasets (both paired and unpaired) in future work.

Unlike traditional applications of CycleGAN where no direct evaluation against ground truth is possible, our paired dataset enables direct comparison between synthesized CT images and registered "true" CT images. This allows for quantitative evaluation of translation performance and is leveraged in our training through the μ-map loss term (Section 3.3.2), which directly compares the attenuation properties of synthesized and real CT images.

3.3.3 Computational Environment

All experiments were conducted on poincare, a workstation with dual NVIDIA GeForce RTX 3090 GPUs (24GB VRAM each) running CUDA 12.4. The KRR and DC-cycleGAN models were implemented in PyTorch, with training taking approximately 3 up to 12 hours depending on model complexity and convergence behavior. Image registration was performed using the Advanced Normalization Tools (ANTs) framework, leveraging CPU parallelization for the computationally intensive template creation and multi-modal registration tasks. This hardware configuration allowed for efficient processing of the large 3D medical volumes while handling the memory demands of both registration and deep learning components.

3.4 Attenuation Correction for PET Imaging

This section outlines the attenuation correction (AC) process for PET images acquired from the PET/MR scanner (Siemens) at The Royal's Institute of Mental Health Research (IMHR), comparing the traditional MR-based method (from vendor) with our proposed approach using the DC-cycleGAN model to generate pseudo-CT images, as illustrated in Figure 3.5.

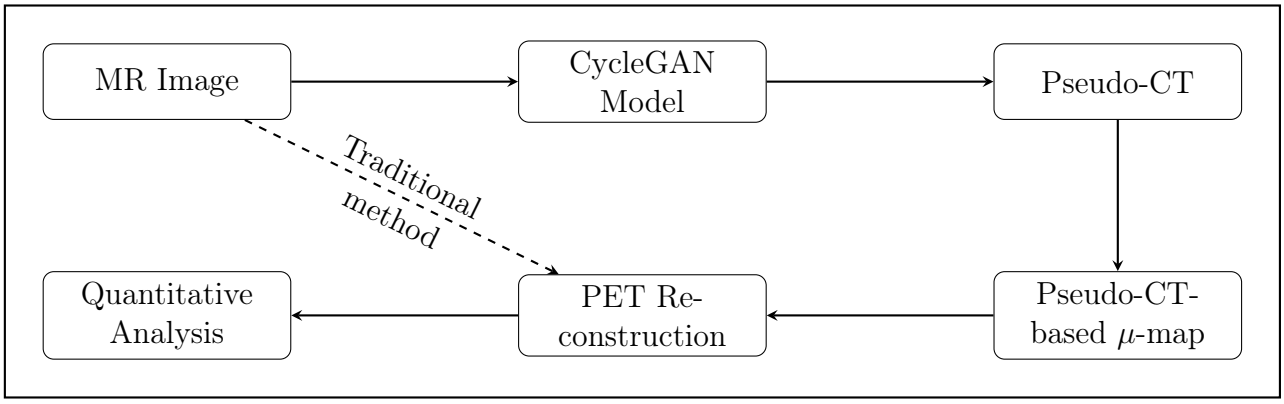

Figure 3.5: Workflow diagram showing the methodology from input MR image through CycleGAN model, pseudo-CT generation, μ-map creation, and PET reconstruction. The dashed line represents the traditional MR-based μ-map approach for comparison.

3.4.1 Traditional MR-Based Attenuation Correction

At The Royal's IMHR, the Siemens PET/MR scanner employs a traditional MR-based attenuation correction method. This approach derives μ-maps directly from MR images using segmentation-based techniques, typically classifying tissues into categories (e.g., air, soft tissue, bone) based on MR intensity values. These μ-maps are then used in PET reconstruction to correct for photon attenuation. However, this method often struggles with accurately distinguishing bone structures due to the lack of direct bone signal in MR images, potentially leading to errors in attenuation correction.

3.4.2 Proposed Pseudo-CT-Based Attenuation Correction

Our proposed method leverages the DC-cycleGAN model to generate pseudo-CT images from MR images. The workflow, depicted in Figure 3.5, proceeds as follows:

1. **Pseudo-CT Generation**: The DC-cycleGAN model translates an input MR image into a pseudo-CT image, capturing anatomical details such as bone structures more accurately than MR-based methods.

2. **μ-Map Creation**: The pseudo-CT image is converted into a pseudo-μ-map using HU conversion, as described in Section 3.3.2.

3. **PET Reconstruction**: The pseudo-μ-map is used in the PET reconstruction process to correct for photon attenuation, producing corrected PET images.

4. **Quantitative Analysis**: The reconstructed PET images undergo quantitative analysis to evaluate the accuracy of AC, using metrics such as SUV errors compared to ground truth CT-based reconstructions.

This approach improves upon the traditional MRAC method by utilizing pseudo-CT images, which better represent bone structures and tissue densities, leading to more accurate μ-maps and enhanced PET reconstruction quality.

Chapter 4

Results

This chapter presents the experimental results of the image translation pipeline for MR-CT translation. It covers the preprocessing steps, including registration and template creation, followed by the evaluation of synthetic CT generation and PET reconstruction performance.

4.1 Dataset Partitioning

As detailed in the methodology chapter, the dataset includes paired CT and MR scans from 10 subjects across 2 institutions. Subject demographics are shown in Table 4.1. Data partitioning refers to dividing the 10 subjects into training (8 subjects) and testing (2 subjects) sets, ensuring no overlap between training and evaluation data to provide unbiased performance assessment.

4.2 Preprocessing Results

The goal of preprocessing is to get MR and CT images ready for the models by making sure they are high quality and properly aligned. This involves steps like registration and template creation to put both MR and CT images into the same template space, using custom ANTs and Python scripts and established medical imaging tools like Mango, Vinci and 3D Slicer. This section looks at how well these preprocessing steps worked.

Table 4.1: Subject Demographics and Dataset Partitioning

Subject ID	Age	Sex	Clinical Condition	Usage
MRACCARD011	58	F	Cardiac	Testing
MRACCARD016	58	M	Cardiac	Training
MRACCARD017	61	M	Cardiac	Testing
MRACCARD019	48	M	Cardiac	Training
MRACCARD020	51	M	Cardiac	Training
MRACNORM002	59	F	Normal	Training
MRACNORM004	57	F	Normal	Training
MRACNORM007	71	M	Normal	Training
MRACNORM009	25	F	Normal	Training
MRACNORM021	57	M	Normal	Training

4.2.1 Original Image Characteristics

Before preprocessing, MR and CT images exhibit modality-specific characteristics that affect subsequent processing steps. Figure 4.1 displays representative slices from both modalities.

Figure 4.1 demonstrates the diversity of contrast mechanisms. Top row: In-phase MR images showing coronal (left), axial (middle), and sagittal (right) views. Middle row: Out-of-phase MR images in corresponding views. Bottom row: CT images in the same anatomical planes. The comparison highlights the complementary nature of these imaging modalities. In-phase and out-of-phase MR sequences provide different soft tissue contrasts, with out-of-phase images particularly useful for detecting fat-water interfaces. CT images offer superior bone visualization and anatomical reference. Each modality presents unique preprocessing challenges including intensity variations in MR data, presence of non-anatomical structures in CT acquisitions such as the scanner bed, and varying spatial orientations across all modalities and field-of-view coverage. A robust preprocessing pipeline must address these differences to establish reliable anatomical correspondence between the datasets before performing cross-modality image translation.

4.2.2 Bias Field Correction Performance

N4 bias field correction effectively reduced intensity inhomogeneities in MR images, as shown in Figure 4.2. The correction enhances tissue homogeneity, improving contrast between different tissue types. This improved contrast is critical for accurate

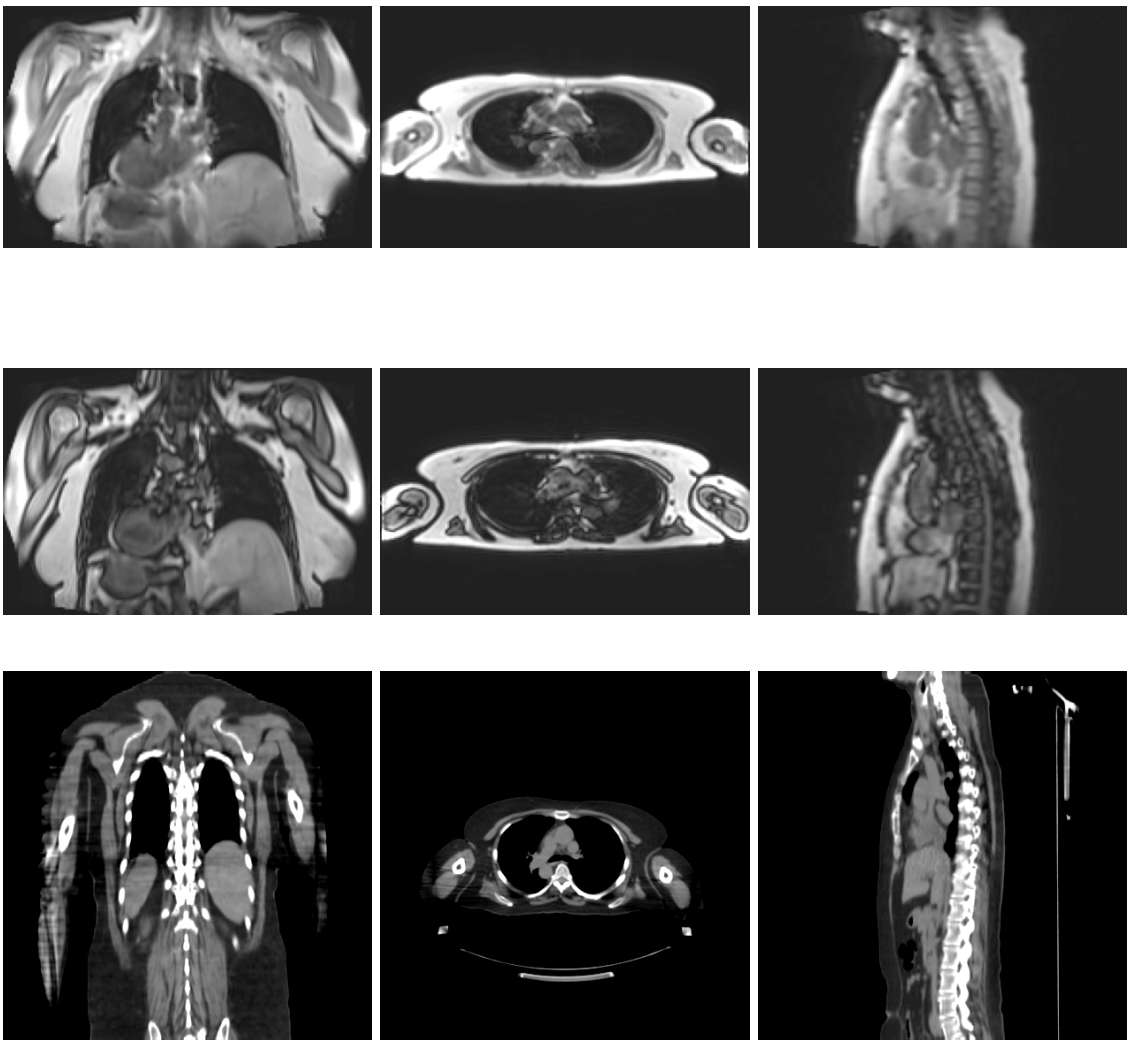

Figure 4.1: Original MR and CT images before preprocessing. Top row: In-phase MR images showing coronal (left), axial (middle), and sagittal (right) views. Middle row: Out-of-phase MR images in corresponding views. Bottom row: Corresponding CT views.

registration, as it enhances the visibility of anatomical landmarks that guide the registration process.

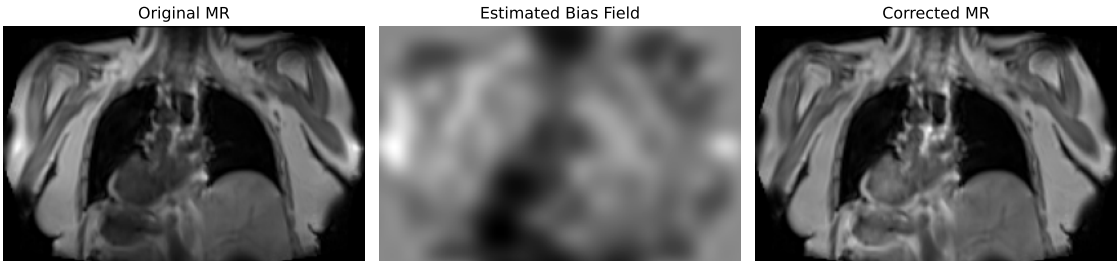

Figure 4.2: Bias field correction visualization (coronal slices). (left) Original MR image with intensity inhomogeneity. (middle) Estimated bias field. (right) Corrected MR image with improved homogeneity.

4.2.3 Intensity Normalization Effectiveness

Intensity normalization standardized MR intensities using liver tissue as a reference, as shown in Table 4.2. The coefficient of variation (CV) decreased from 0.327 to 0.309, a 5.6% improvement. While there is no established threshold for meaningful improvement in medical imaging preprocessing, this reduction demonstrates enhanced consistency that contributes to more reliable cross-subject comparisons. This is an essential aspect of structural MRI analysis where intensity standardization helps reduce variability [64, 65].

Table 4.2: Liver-Based Intensity Normalization: Patient-by-Patient Results

Patient ID	Liver Avg	Mean Intensity (Original)	Mean Intensity (Normalized)	% Change
016	200.0	112.7	102.3	-9.2%
011	160.0	85.3	96.8	+13.4%
017	185.0	47.6	46.7	-1.9%
019	225.0	62.2	50.2	-19.3%
020	220.0	99.8	82.3	-17.5%
002	180.0	130.3	131.4	+0.8%
004	165.0	72.0	79.2	+10.0%
007	170.0	59.1	63.0	+6.8%
009	125.0	57.8	84.0	+45.2%
021	185.0	63.0	61.8	-1.9%
Global Statistics				
Mean	181.5	CV: 0.327	CV: 0.309	Improvement: 5.6%

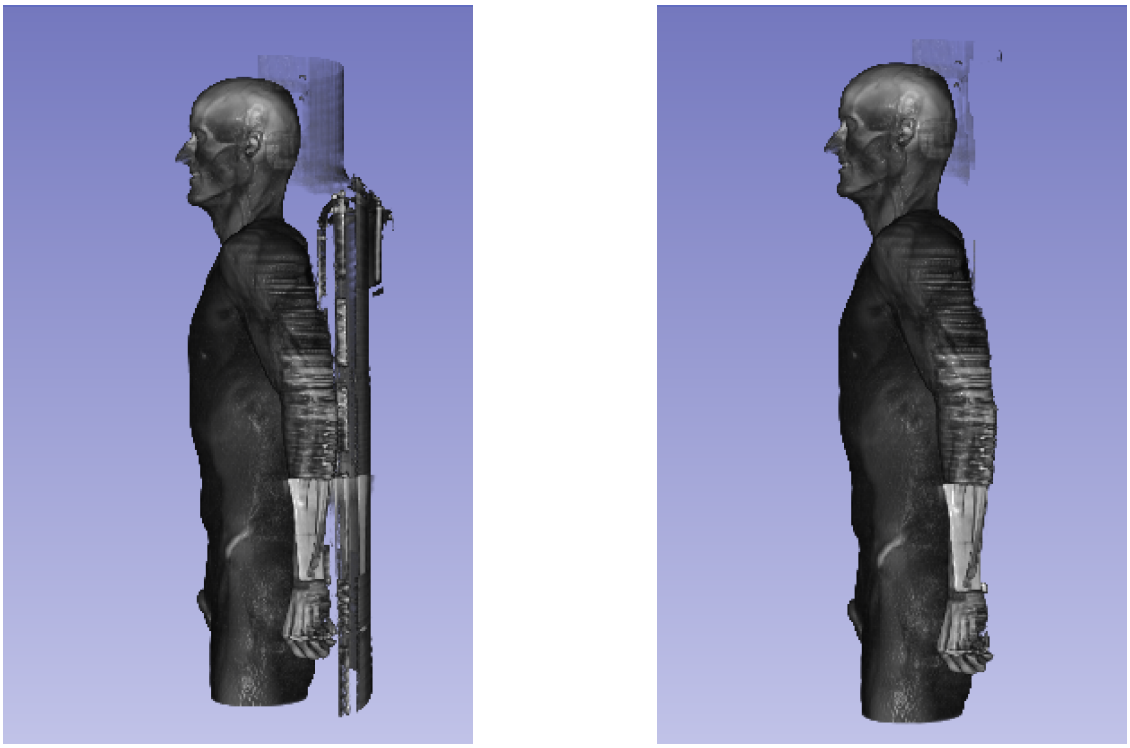

Figure 4.3: CT bed removal comparison showing 3D lateral views. (left) Original CT with scanner bed. (right) Processed CT after manual bed removal in 3D Slicer.

4.2.4 CT Bed Removal

A critical preprocessing step for CT images is the removal of the scanner bed, which can introduce artifacts in registration and affect the quality of the template. The bed structure appears as a high-density region that does not correspond to patient anatomy and could bias intensity-based registration algorithms. Figure 4.3 illustrates the before and after results of the bed removal process.

The bed removal was performed manually using built-in features of the 3D Slicer software, which allowed for precise segmentation and removal of non-anatomical structures while preserving patient tissue.

4.2.5 Template Creation Results

The CT template, created from CT images registered to MR (step 1 of Figure 3.1) using custom ANTs scripts, was evaluated using standard metrics for template quality

in medical imaging. These metrics assess anatomical accuracy, tissue differentiation, and alignment with individual subjects, aligning with practices in atlas-based studies for PET/MR AC [66, 67, 68].

Figure 4.4: Final CT template after 6 iterations, showing (left) axial view, (middle) coronal view, and (right) sagittal view.

The template converged after 6 iterations, producing a high-quality anatomical reference with clear details in axial, coronal, and sagittal views (Figure 4.4). Quantitative metrics (Table 4.3) include entropy, contrast-to-noise ratio (CNR), left-right symmetry index, average template-to-subject normalized cross-correlation (NCC), and signal-to-noise ratio (SNR) by tissue type.

Table 4.3: Quantitative Assessment of CT Template Quality

Metric	**Value**
Entropy	4.26
Contrast-to-Noise Ratio	2.38
Left-Right Symmetry Index	0.93
Average Template-to-Subject NCC	0.64 ± 0.05
Signal-to-Noise Ratio by Tissue Type	
Air	30.32
Lung	9.80
Fat	5.06
Soft Tissue	0.48
Bone	3.21

The entropy of 4.26 indicates good template quality, representing appropriate anatomical complexity and detail preservation in the averaged template. Higher

entropy values suggest better preservation of anatomical features and tissue boundaries during the template creation process. The CNR of 2.38 suggests adequate tissue differentiation, meeting the threshold of above 2 reported for medical imaging segmentation tasks [66]. The left-right symmetry index of 0.93, close to 1.0, reflects excellent anatomical symmetry, consistent with high-quality brain atlases used in PET/MR [68]. The NCC of 0.64 $\pm$ 0.05 demonstrates moderate to good alignment with individual subjects, as values above 0.5 are considered acceptable in atlas-based registration [67]. Tissue-specific SNR values show high clarity for air (30.32) and lung (9.80), with bone (3.21) within the reported range of 2–4 and soft tissue (0.48) slightly below the typical 0.5–1.0, indicating limited differentiation for soft tissue [68].

The template will serve as a good reference for mapping CT and MR images, showing clear anatomical detail and strong alignment, though improvements may still be needed for better soft tissue differentiation.

4.2.6 Registration Results

The registration pipeline (CT-to-MR followed by CT and MR to the CT template) was evaluated qualitatively and quantitatively. Figure 4.5 shows overlay visualizations for CT-to-MR, CT-to-template, and MR-to-template registrations, comparing alignments before and after registration.

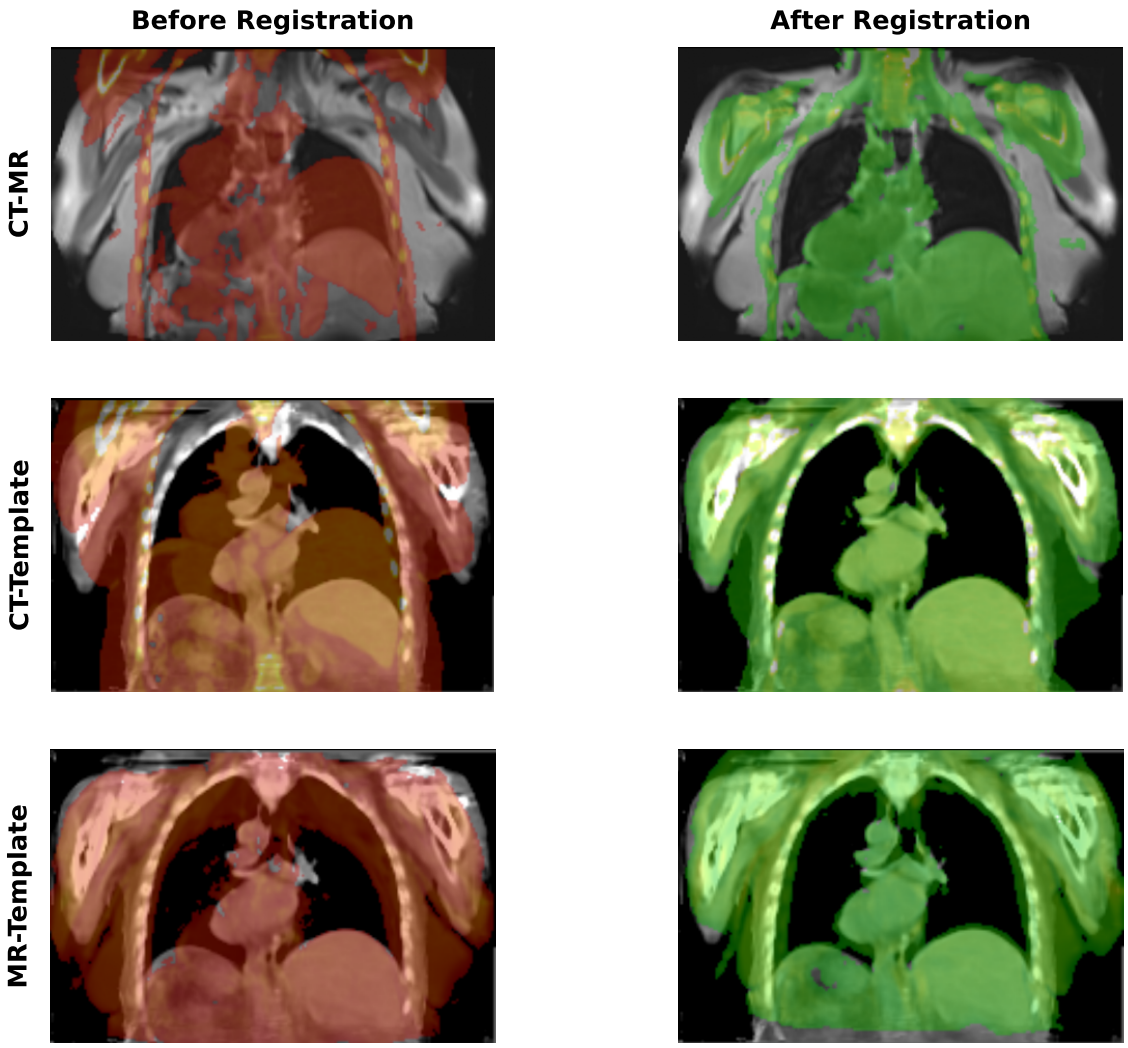

Figure 4.5: Registration overlay comparisons. Rows show (top to bottom): CT-to-MR, CT-to-template, and MR-to-template registrations. Columns show (left) before registration and (right) after registration. Post-registration overlays demonstrate improved alignment of anatomical structures.

Qualitatively, the overlays reveal significant alignment improvements. Pre-registration images show misalignments in structures like the heart and shoulders, while post-registration overlays exhibit more precise correspondence, particularly in thoracic regions.

Quantitative metrics (Table 4.4) validate these findings. CT-to-template registration achieved an NCC of 0.81 ± 0.05 and mutual information (MI) of 1.15 ± 0.00, indicating strong alignment. The high NCC value suggests excellent structural correspondence between the registered images and the template. While no universal threshold exists for MI values, successful registration typically achieves MI $>$ 1.0, with values above 1.15 indicating good alignment quality.

MR-to-template registration yielded an NCC of 0.66 ± 0.03 and a MI of 1.08 ± 0.01, reflecting the inherent challenges of cross-modality alignment. This moderate value is expected due to the different image characteristics between MR and CT.

Table 4.4: Registration Accuracy Metrics

Registration Type	**NCC**	**MI**
CT-to-Template (Direct)	0.81 ± 0.05	1.15 ± 0.00
MR-to-Template (Composed)	0.66 ± 0.03	1.08 ± 0.01

The combination of visual assessment and quantitative metrics confirms that the registration framework successfully establishes spatial correspondence between individual subject images and the template space, enabling subsequent multi-modality analysis.

The final outcome of the preprocessing and registration workflow is a set of spatially normalized CT and MR images aligned to a common template space. Figures 4.6 and 4.7 demonstrate the quality of the final registered images.

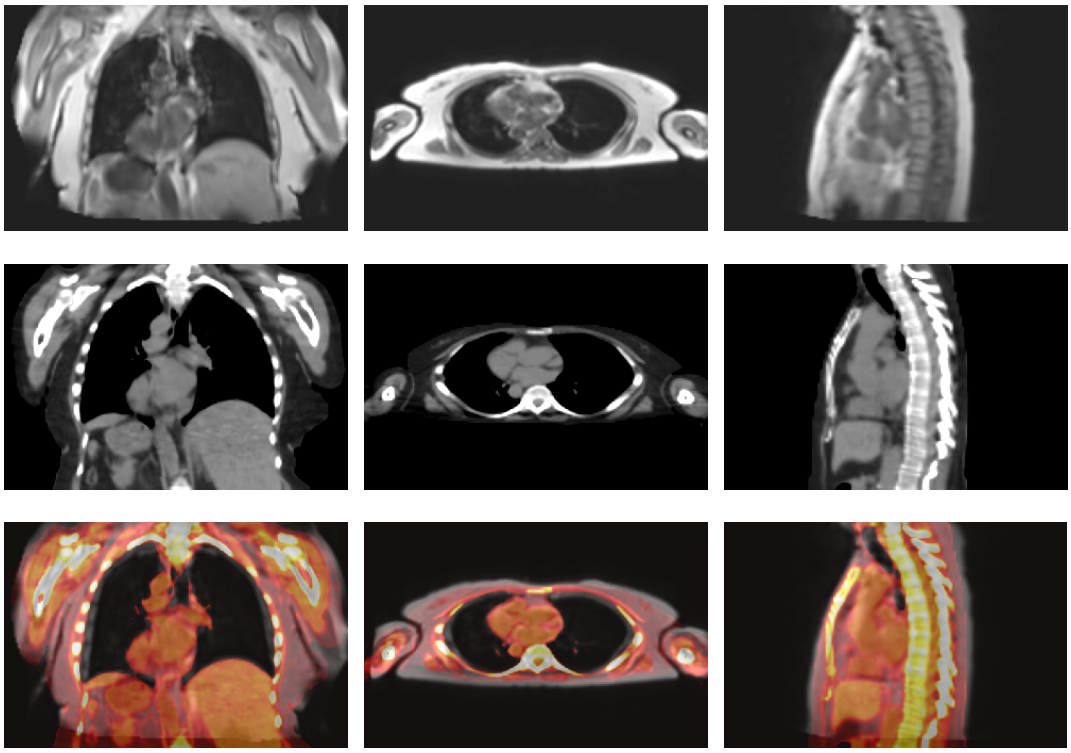

Figure 4.6: Images in template space after complete preprocessing and registration. Top row: MR (In-Phase) images showing coronal (left), axial (middle), and sagittal (right) views. Middle row: Corresponding CT views in the same template space. Bottom row: Overlay visualization of CT (hot colormap) on MR (grayscale), demonstrating the spatial correspondence achieved between modalities.

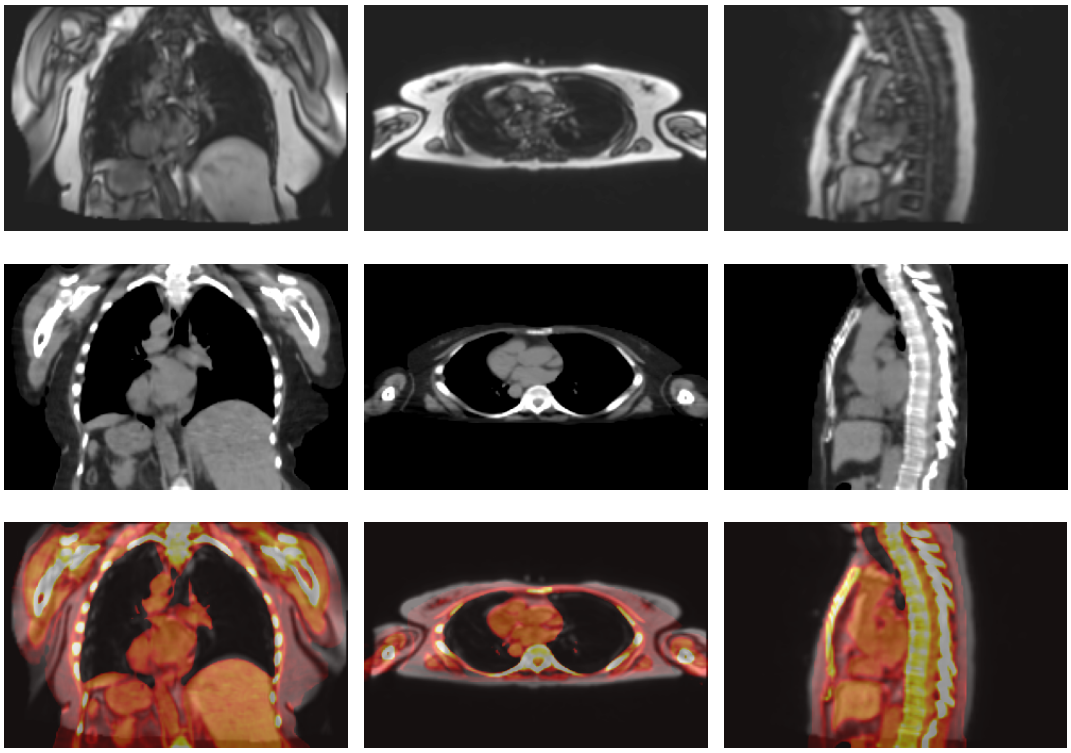

Figure 4.7: Images in template space after complete preprocessing and registration. Top row: MR (Out-of-phase) images showing coronal (left), axial (middle), and sagittal (right) views. Middle row: Corresponding CT views in the same template space. Bottom row: Overlay visualization of CT (hot colormap) on MR (grayscale), demonstrating the spatial correspondence achieved between modalities.

The overlay visualization (bottom rows) confirms successful alignment between modalities. This spatial correspondence is important for the models to learn accurate mappings between anatomical structures without being confounded by misalignments.

Image Dimension and Slice-based Training Strategy

Prior to image registration, considerable variation was observed in the original dimensions of the acquired images across the ten study subjects. The CT images maintained consistent in-plane dimensions of 512×512 voxels with spatial resolution of 1.37×1.37 mm, but exhibited substantial variability in the axial dimension, ranging from 47 to 335 slices with a uniform slice thickness of 3.27 mm. In contrast, the

MR images demonstrated greater dimensional consistency, with all subjects having dimensions of 192 × 128 × 126 voxels. The MR voxel sizes were predominantly 2.6 × 3.9 × 2.6 mm, with one exception (patient002) that had a voxel size of 2.6 × 3.12 × 2.6 mm. Following the registration process, both MR and CT images for all subjects were standardized to dimensions of 192 × 128 × 126 voxels with an isotropic voxel size of 2.604 × 2.604 × 3.9 mm, ensuring spatial correspondence across all imaging datasets. This standardization step was crucial for our model training approach, which operated on a slice-by-slice basis rather than utilizing full 3D volumes. Both our GAN-based models and the KRR model were trained using individual 2D axial slices extracted from the registered volumes. Each of the 126 axial slices per volume was processed independently during both training and inference phases, effectively transforming our 3D registration problem into a series of 2D translation tasks.

4.3 Translation results and PET reconstruction evaluation

With preprocessing complete and all images aligned to a common template space, the stage is set for cross-modality synthesis. This section presents both quantitative and qualitative evaluation of synthetic CT images generated from MRI across different models. We also evaluate how well our approach works for PET imaging by taking the synthetic CT from our best-performing model and using it for PET AC. We then compare both these resulting PET reconstructed images and those obtained by the MR-based AC method against true CT-based AC PET images, which serve as the reference standard.

While the registration process establishes general anatomical correspondence between MR and CT volumes, the models evaluated have different sensitivities to alignment precision. The KRR model (KRR-In) relies more heavily on precise voxel-to-voxel correspondence (meaning that each spatial location (x,y,z) in the MR image directly maps to the same spatial location in the CT image, enabling direct pixel-wise comparison and training at matching anatomical positions), as it learns direct intensity mappings between spatially aligned input-output pairs. In contrast, the GAN approach has the advantage of not requiring perfect voxel-to-voxel alignment. Instead, GANs can learn the underlying relationship between tissue characteristics in MR and their corresponding CT appearances, even with some degree of alignment imprecision.

Figure 4.8 illustrates the different masking strategies employed for the models.

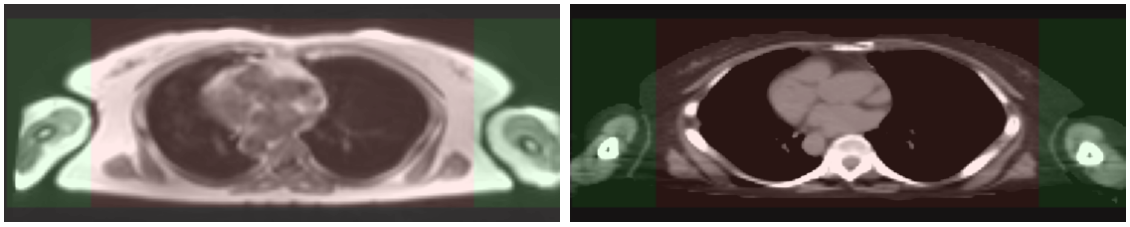

Figure 4.8: Axial slices of registered MR (left) and CT (right) images with mask overlays. The full-body rectangular mask (green) encompasses the entire patient anatomy and was used for GAN training. The thoracic cage mask (blue) focuses on regions with high registration accuracy and was used for the KRR model training, leveraging the precise alignment in this region.

The thoracic cage mask (blue) focuses on regions with higher registration accuracy, which benefits the KRR model that requires precise voxel-to-voxel correspondence. The rectangular body mask (green) encompasses the entire patient anatomy and was used for GAN training, allowing the model to learn from a broader range of tissues despite varying degrees of registration precision across different body regions.

Four methods were evaluated: a baseline KRR model trained on in-phase MR images (KRR-In), and three CycleGAN-based models using in-phase MR (GAN-In), out-of-phase MR (GAN-Out), and a multi-channel approach combining both in-phase and out-of-phase MR sequences (GAN-In-Out). Performance was assessed using MAE, MSE, Peak Signal-to-Noise Ratio (PSNR), and SSIM, calculated on normalized intensity values ([-1, 1]). The synthetic CT images are evaluated against the real CT images that have been registered to the same template space, with results summarized as mean and standard deviation across a cohort of 2 test patients, including a detailed analysis of performance across different tissue types.

4.3.1 Comparison of Synthetic CT Generation Methods

The evaluation quantifies both the voxel-level accuracy and structural integrity of synthetic CT images compared to ground-truth CT scans. Four complementary metrics were employed: MAE and MSE to assess direct voxel intensity differences in HU; PSNR to evaluate overall image quality with emphasis on noise reduction; and SSIM to measure preservation of anatomical structures critical for accurate AC. These metrics were computed across 2 test patients, with tissue-specific analysis performed to identify performance variations across different anatomical regions. This evaluation

approach aligns with established methodologies in the field of PET/MR AC [69].

Quantitative Evaluation

Table 4.5 presents a comprehensive evaluation of the four synthetic CT generation methods across 2 test patients. Part A displays the mean and standard deviation of quantitative metrics, while Part B provides statistical significance through pairwise Wilcoxon signed-rank tests.

Table 4.5: Quantitative evaluation and statistical comparison of synthetic CT generation models using Wilcoxon signed-rank tests

A. Slice-level Metrics (Mean ± Std across patients)				
Model	**MAE ↓**	**MSE ↓**	**PSNR ↑**	**SSIM ↑**
KRR-In	0.185 ± 0.011	0.073 ± 0.006	17.388 ± 0.374	0.367 ± 0.025
GAN-In-Out	0.116 ± 0.009 †	0.044 ± 0.005 †	19.596 ± 0.526 †	0.503 ± 0.028 †
GAN-Out	0.150 ± 0.037 †	0.071 ± 0.027	17.834 ± 1.743	0.408 ± 0.087 †
GAN-In	0.112 ± 0.007 †‡§	0.045 ± 0.003 †‡	19.537 ± 0.337 †‡	0.497 ± 0.031 †‡§

B. Statistical Significance (p-values) from Pairwise Wilcoxon Signed-Rank Tests				
Comparison	**MAE**	**MSE**	**PSNR**	**SSIM**
KRR-In vs GAN-In-Out	< 0.001 ***	< 0.001 ***	< 0.001 ***	< 0.001 ***
KRR-In vs GAN-Out	< 0.001 ***	0.942	0.152	< 0.001 ***
KRR-In vs GAN-In	< 0.001 ***	< 0.001 ***	< 0.001 ***	< 0.001 ***
GAN-In-Out vs GAN-Out	< 0.001 ***	< 0.001 ***	< 0.001 ***	0.295
GAN-In-Out vs GAN-In	0.004 **	0.336	0.063	< 0.001 ***
GAN-Out vs GAN-In	< 0.001 ***	< 0.001 ***	< 0.001 ***	< 0.001 ***

↓ Lower is better, ↑ Higher is better
Statistical significance levels: $^{*}p < 0.05$, $^{**}p < 0.01$, $^{***}p < 0.001$, n.s. not significant
† Significantly better than KRR-In ($p < 0.05$)
‡ Significantly better than GAN-Out ($p < 0.05$)
§ Significantly better than GAN-In-Out ($p < 0.05$)

Statistical analysis using Wilcoxon signed-rank tests on slice-level metrics (where each metric was computed individually for each 2D slice, resulting in approximately 126 slices $\times$ 2 test subjects $=$ 252 data points per metric) revealed significant differences in model performance (Table 4.5). All GAN-based approaches significantly outperformed the kernel ridge regression baseline in terms of MAE and SSIM ($p < 0.001$). The GAN-In model emerged as the strongest performer, showing significant improvements over both GAN-Out ($p < 0.001$ across all metrics) and GAN-In-Out in MAE ($p = 0.004$) and SSIM ($p < 0.001$). While GAN-In and GAN-In-Out

performed similarly in terms of MSE ($p = 0.336$) and PSNR ($p = 0.063$), the significant advantages in the other metrics suggest that GAN-In provides more consistent results.

Notably, the GAN-Out model exhibited inconsistent performance across test patients. This suggests that out-phase MR images alone may not provide sufficiently stable features for reliable CT synthesis across different subjects. In contrast, the models incorporating in-phase MR images (GAN-In and GAN-In-Out) demonstrated more consistent performance. Overall, the GAN-In model, utilizing only in-phase MR images, emerged as the most robust approach.

Tissue-Specific Analysis

To understand performance variations across anatomical structures, images were segmented into five tissue classes based on normalized intensity ranges: Air (< -0.95), Lung (-0.95 to -0.65), Fat (-0.1 to -0.05), Soft Tissue (-0.05 to 0.1), and Bone (> 0.3). Table 4.6 presents the MAE, MSE, PSNR, and SSIM for each method across these tissue classes, with the best results in **bold**.

The tissue-specific analysis reveals distinct performance patterns across anatomical structures. GAN-In-Out excelled in Air and Lung tissues, achieving significant reductions in MAE (80.0% for Air, 37.7% for Lung compared to KRR-In) and improvements in SSIM (30.5% for Air, 50.8% for Lung). For Air, GAN-In-Out's MAE of 0.045 and SSIM of 0.792 indicate excellent accuracy and structural fidelity, likely due to the uniform intensity of air regions. In Lung tissue, GAN-In-Out's low MAE (0.139) and high PSNR (18.729) suggest robust handling of complex lung structures.

GAN-In dominated in Fat and Soft Tissue, with MAE reductions of 16.0% for Fat and 30.9% for Soft Tissue compared to KRR-In, and SSIM improvements of 13.8% and 163.5 %, respectively. The high PSNR (23.737) and SSIM (0.419) for Soft Tissue highlight GAN-In's ability to capture fine structural details. For Bone, GAN-In achieved the best MAE (0.345), MSE (0.176), and PSNR (13.633), with a 44.3% MAE reduction over KRR-In, while GAN-In-Out had the best SSIM (0.137). Bone tissue exhibited the highest errors across all models (e.g., KRR-In MAE of 0.619), reflecting the challenge of generating bone structures from MR images due to their high intensity variability and lack of direct MR signal.

Table 4.6: Summary of MAE, MSE, PSNR, and SSIM Metrics for Different Models by Tissue Type (Mean ± Std). Best results are in **bold**.

Tissue	Model	MAE	MSE	PSNR	SSIM
Air	KRR-In	0.225 ± 0.005	0.064 ± 0.004	17.953 ± 0.247	0.607 ± 0.010
Air	GAN-In-Out	**0.045 ± 0.006**	**0.019 ± 0.004**	**23.233 ± 0.957**	**0.792 ± 0.042**
Air	GAN-Out	0.068 ± 0.018	0.038 ± 0.018	20.804 ± 2.207	0.722 ± 0.025
Air	GAN-In	0.109 ± 0.078	0.075 ± 0.063	19.928 ± 5.303	0.690 ± 0.142
Lung	KRR-In	0.223 ± 0.035	0.112 ± 0.026	15.641 ± 1.044	0.299 ± 0.040
Lung	GAN-In-Out	**0.139 ± 0.000**	**0.054 ± 0.001**	**18.729 ± 0.046**	**0.451 ± 0.007**
Lung	GAN-Out	0.196 ± 0.008	0.098 ± 0.006	16.105 ± 0.277	0.308 ± 0.019
Lung	GAN-In	0.265 ± 0.129	0.167 ± 0.114	15.171 ± 3.627	0.296 ± 0.112
Fat	KRR-In	0.119 ± 0.003	0.038 ± 0.003	20.247 ± 0.342	0.240 ± 0.044
Fat	GAN-In-Out	0.131 ± 0.011	0.042 ± 0.008	19.878 ± 0.848	0.092 ± 0.052
Fat	GAN-Out	0.162 ± 0.080	0.073 ± 0.051	18.818 ± 3.717	0.186 ± 0.166
Fat	GAN-In	**0.100 ± 0.014**	**0.027 ± 0.008**	**21.849 ± 1.297**	**0.207 ± 0.033**
Soft Tissue	KRR-In	0.097 ± 0.011	0.036 ± 0.006	20.560 ± 0.768	0.159 ± 0.018
Soft Tissue	GAN-In-Out	0.093 ± 0.006	0.032 ± 0.005	21.020 ± 0.725	0.251 ± 0.055
Soft Tissue	GAN-Out	0.122 ± 0.048	0.049 ± 0.030	20.157 ± 3.121	0.162 ± 0.107
Soft Tissue	GAN-In	**0.067 ± 0.019**	**0.020 ± 0.011**	**23.737 ± 2.551**	**0.419 ± 0.076**
Bone	KRR-In	0.619 ± 0.052	0.489 ± 0.100	9.216 ± 0.900	-0.024 ± 0.002
Bone	GAN-In-Out	0.357 ± 0.041	0.206 ± 0.058	13.055 ± 1.258	**0.137 ± 0.010**
Bone	GAN-Out	0.487 ± 0.104	0.336 ± 0.139	11.164 ± 1.909	0.061 ± 0.028
Bone	GAN-In	**0.345 ± 0.043**	**0.176 ± 0.032**	**13.633 ± 0.796**	0.123 ± 0.041

4.3.2 Visual Comparison of Synthetic CT Slices

To complement the quantitative metrics, a visual comparison of synthetic CT slices is presented in Figure 4.9. The figure shows axial, coronal, and sagittal views of a true CT slice alongside the corresponding synthetic CT slices generated by each model (KRR-In, GAN-In-Out, GAN-Out, GAN-In). Intensities are normalized to the range [-1, 1], with the colorbar indicating values from -0.20 to 0.20.

The visual comparison highlights the strengths and weaknesses of each model across different anatomical regions. GAN-In produces synthetic CT slices that closely resemble the true CT, especially in the sagittal view, where bone structures (e.g., spine) and soft tissues are accurately captured, supporting its low MAE for Bone (0.345) and relatively high SSIM for Soft Tissue (0.419), as reported in Table 4.6. GAN-In-Out also performs well, particularly in the axial and coronal views, where lung and air regions are well-defined with clear boundaries. This aligns with its superior quantitative performance in Air (MAE 0.045, SSIM 0.792) and Lung (MAE

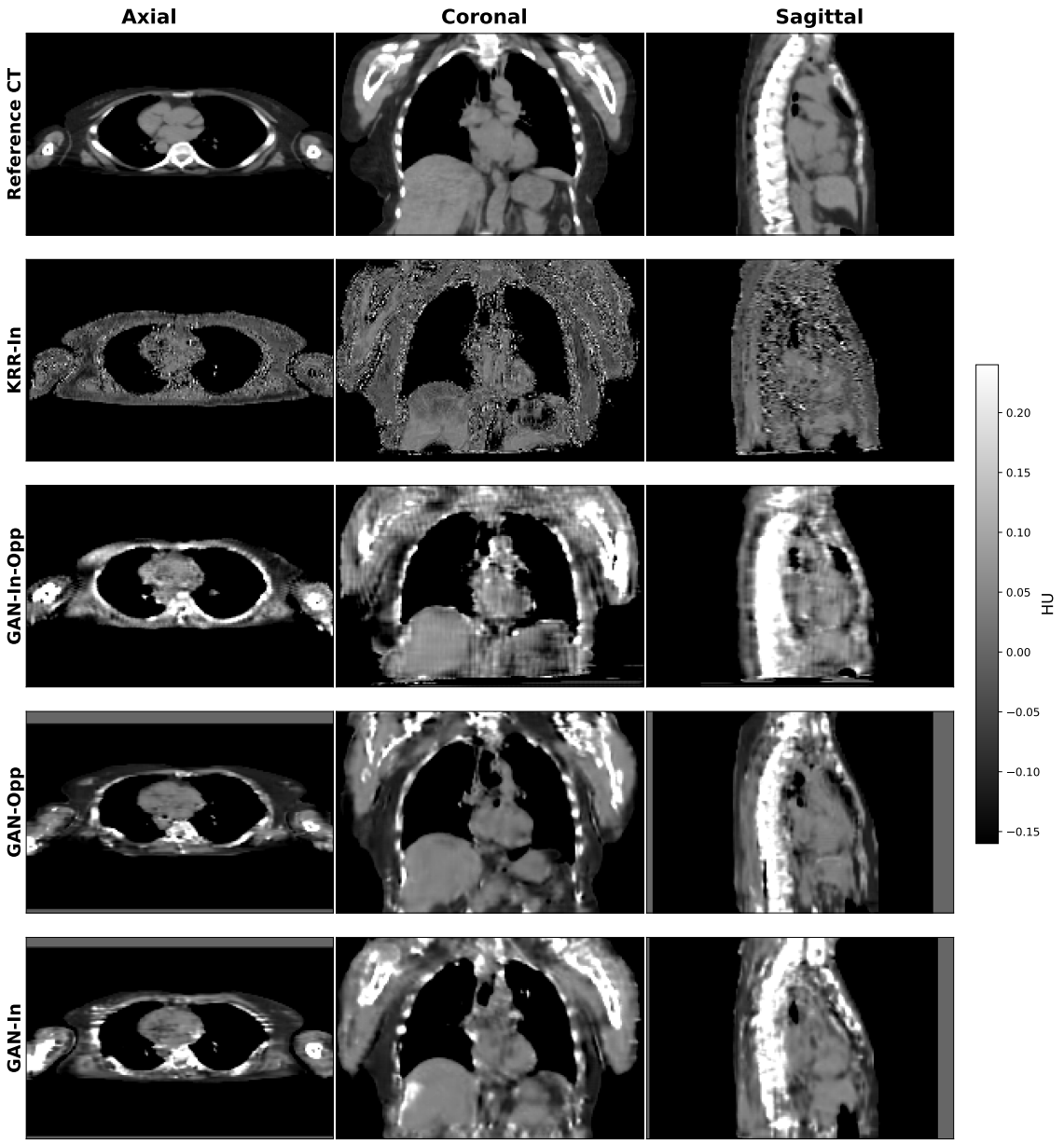

Figure 4.9: Axial, coronal, and sagittal views of a true CT slice (top row) compared to synthetic CT slices generated by KRR-In, GAN-In-Out, GAN-Out, and GAN-In models. Intensities are normalized to [-1, 1], with a colorbar ranging from -0.20 to 0.20.

0.139, SSIM 0.451) tissues.

In contrast, KRR-In exhibits significant blurring across all views, particularly in air and bone regions, consistent with its higher errors (MAE 0.225 for Air, 0.619 for Bone). GAN-Out shows moderate performance but with visible noise, especially in bone regions (sagittal view), reflecting its higher variability (e.g., MAE standard deviation of 0.104 for Bone). These visual observations reinforce the quantitative findings, demonstrating GAN-In-Out and GAN-In as the most promising models for whole-body synthetic CT generation, with complementary strengths in different tissue types.

4.3.3 Evaluation of GAN-In and GAN-Out Models for Synthetic MR Generation

Table 4.7: Comparison of GAN-In and GAN-Out Model Performance (Mean ± Standard Deviation)

Model (Phase)	**MAE**	**MSE**	**PSNR**	**SSIM**
GAN-In (In)	**0.045 ± 0.005**	**0.006 ± 0.001**	**28.56 ± 0.79**	**0.693 ± 0.022**
GAN-Out (Out)	0.060 ± 0.010	0.008 ± 0.001	26.93 ± 0.61	0.556 ± 0.095

Note: Bold values indicate better performance. Lower values are better for MAE and MSE, higher values are better for PSNR and SSIM. Each model is evaluated on its respective target MR phase.

One of the objectives outlined in the first chapter was to generate synthetic MR images from CT scans, even though these images will not be used for AC. While CT-to-MR translation may seem less clinically relevant than MR-to-CT for our specific application, it serves several important purposes. First, it validates the bidirectional capabilities of the CycleGAN architecture, ensuring the model learns meaningful domain mappings in both directions rather than just memorizing one-way transformations. Second, medical image translation extends well beyond AC applications (synthetic MR generation could be valuable for research applications, data augmentation, or retrospective studies where MR acquisition was not feasible or available.) To achieve this, we evaluated the GAN-In and GAN-Out models for synthetic MR image generation, both synthesizing from CT scans using their respective trained CT-to-MR generators. Specifically, the GAN-In model's MR-to-CT generator produces synthetic In-phase MR images, while the GAN-Out model's generator produces synthetic Out-of-phase MR images.

Quantitative assessment using MAE, MSE, PSNR, and SSIM metrics across the test patients revealed superior performance of the GAN-In model. As shown in Table 4.7, GAN-In demonstrated lower error rates, achieving an MAE of 0.045 ($\pm$0.005) compared to 0.060 ($\pm$0.010) for GAN-Out, and an MSE of 0.006 ($\pm$0.001) compared to 0.008 ($\pm$0.001) for GAN-Out.

Image quality metrics reinforced these findings, with GAN-In achieving higher PSNR (28.56 dB $\pm$0.79 versus 26.93 dB $\pm$0.61) and significantly better SSIM scores (0.693 $\pm$0.022 versus 0.556 $\pm$0.095). These results suggest that synthesizing In-phase MR images from CT may be inherently more feasible than generating Out-of-phase MR, possibly due to the closer relationship between tissue properties in CT and In-phase MR compared to the fat-water separation characteristics in Out-of-phase MR. The larger standard deviation in GAN-Out's SSIM scores also indicates less consistent performance across patients, suggesting out-of-phase MR synthesis may be more sensitive to inter-patient variability.

Figure 4.10 shows the synthetic MR images generated by the GAN-In and GAN-Out models alongside the reference CT and true MR scans. The results demonstrate that both models successfully capture the anatomical structures present in the true MR images while maintaining consistent tissue contrast, with the GAN-In model performing better.

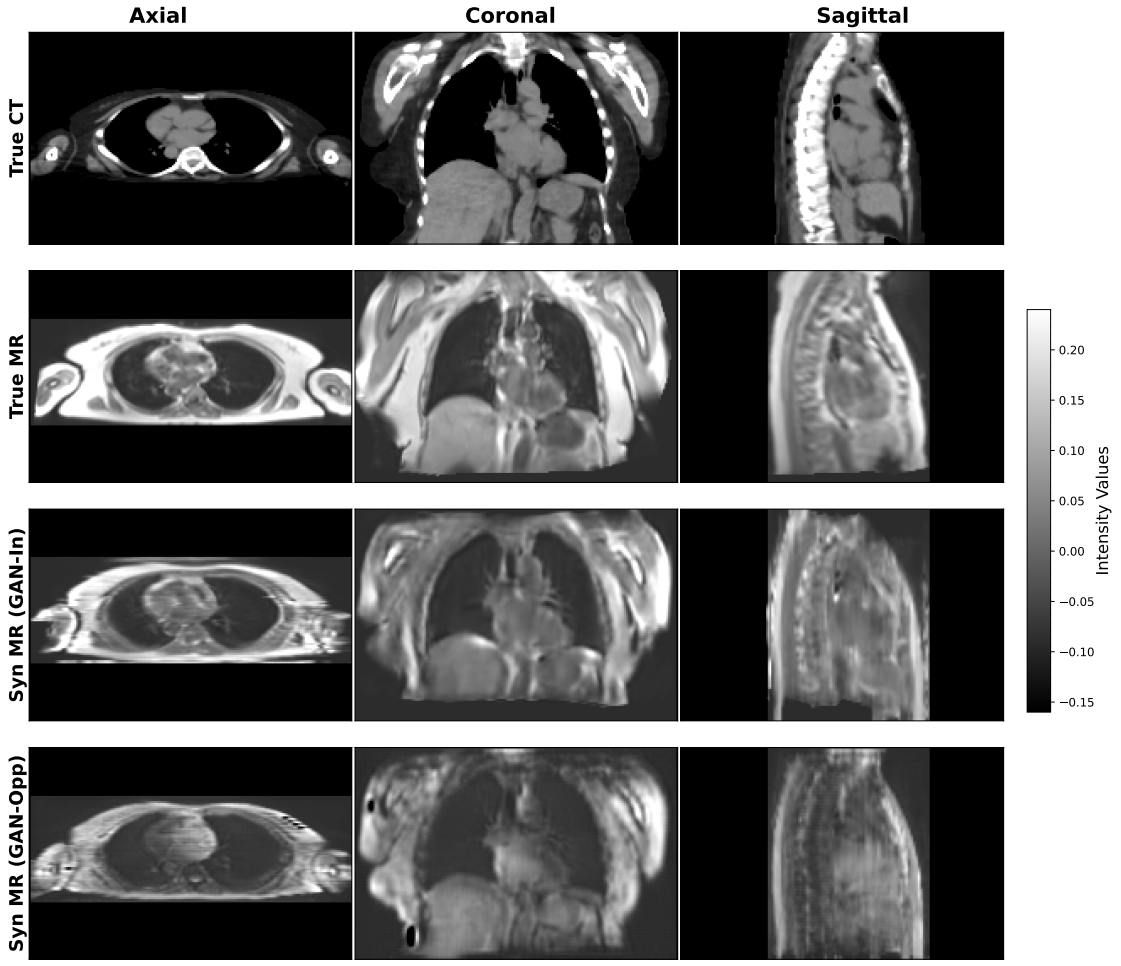

Figure 4.10: Comparison of true and synthetic MR images. The figure shows axial, coronal, and sagittal views of (top to bottom): reference CT scan, true MR scan, synthetic MR (In) generated by GAN-In model, and synthetic MR (Out) generated by GAN-Out model. The synthetic MR images demonstrate the capability of the GAN models to transform CT data into MR-like representations while preserving anatomical information.

4.4 Comparison of Attenuation Maps for PET Reconstruction

Figure 4.11 illustrates the three different attenuation maps used for PET reconstruction: the ground truth CT-based μ-map, our GAN-generated pseudo CT μ-map (based on GAN-In), and the standard MR scanner-based μ-map. The GAN-In model, which demonstrated superior performance in synthetic CT generation, particularly for soft tissue and bone structures, was employed to produce the pseudo CT μ-map. As evident in Figure 4.11, this approach successfully captures the detailed attenuation properties across different tissue types, closely resembling the ground truth CT-based μ-map. In contrast, the MR scanner-based μ-map, while capturing general anatomical structures, lacks the detailed bone representations visible in both CT-based maps. The visual similarity between the true and pseudo CT μ-maps in all three anatomical planes supports our quantitative findings about the GAN-In model's effectiveness, particularly for bone structures which are critical for accurate PET AC. The impact of these different μ-maps on final PET quantification will be examined in subsequent analyses.

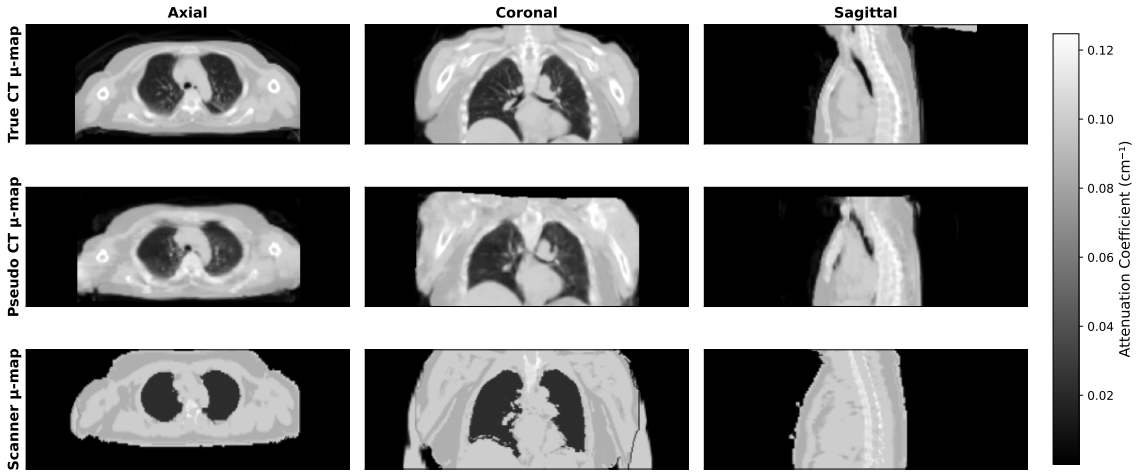

Figure 4.11: Comparison of attenuation maps (μ-maps; subject 011) used for PET/MR AC. The three rows show the true CT-based μ-map (top), pseudo CT-based μ-map generated using GAN-In (middle), and MR-AC μ-map (bottom). Each column represents different anatomical views: axial (left), coronal (center), and sagittal (right).

4.5 Quantitative PET Analysis

To properly evaluate our GAN-based synthetic CT approach in a clinical context, we employed a Leave-One-Out Cross-Validation (LOO-CV) strategy across 8 patients. This methodology involved training on 7 patients while reserving one for testing, repeating this process 8 times so that each patient served as a test case exactly once. We chose this approach to provide statistical robustness and minimize potential bias from patient selection.

Each test patient underwent PET reconstruction using three distinct AC approaches: CT-based AC (CT-AC) serving as our reference standard, synthetic CT from our best-performing GAN-In model (GAN-AC), and the standard vendor MR-based AC (MR-AC). We spatially aligned all μ-maps prior to PET reconstruction to eliminate potential misalignment artifacts that could confound our results.

To quantify the improved performance of GAN-AC seen qualitatively in the previous section, we evaluate difference metrics between GAN-AC and our reference CT-AC and compare them with the same difference metrics between MR-AC and the reference CT-AC. The metrics, detailed in Appendix A.1.3 are: SUVmean Bias (%) to assess systematic over- or underestimation of average uptake values, SUVmax Bias (%) to evaluate peak uptake accuracy, RMSE (SUV) for root mean square error, and MAE (SUV) for mean absolute error.

To evaluate the statistical significance of the observed improvements, we performed paired statistical tests comparing GAN-AC and MR-AC methods across the 8 patients. However, it is important to note that within our LOO-CV framework, while the 8 test subjects are independent, the training sets are different (though overlapping) for each test subject, meaning our tests only approximate tests of performance of a single model. This is a widely used procedure in machine learning evaluation [70], though one result of Nadeau and Bengio (2003) is that uncorrected statistical tests in a cross-validation framework lead to artificially low p-values and thus inflated Type I errors, so our significance results should be treated with some caution.

Paired tests were appropriate because we compared two different AC methods applied to the same patient's PET data, creating naturally paired observations that control for patient-specific factors such as anatomy, pathology, and tracer distribution. Prior to testing, normality of the paired differences was assessed using the Shapiro-Wilk test, which is appropriate for small sample sizes (n=8). For metrics with normally distributed differences (SUVmean Bias, SUVmax Bias, RMSE), paired t-tests were employed as they provide optimal statistical power when normality assumptions are met. For metrics that violated normality assumptions (MAE), the

non-parametric Wilcoxon signed-rank test was used, which makes no distributional assumptions and is robust for small samples. Statistical significance was set at $p < 0.05$.

Table 4.8 summarizes the quantitative comparison between methods across all 8 test patients relative to the CT-AC reference. We examined SUVmean Bias (%), SUVmax Bias (%), RMSE (SUV), and MAE (SUV). Values represent mean ± standard deviation across the patient cohort.

Table 4.8: Quantitative metrics for PET reconstruction using different AC methods with LOO-CV across 8 patients. Values are presented as mean ± standard deviation relative to the reference CT-AC.

Metric	**GAN-AC**	**MR-AC**	**Improvement**	**p-value**
SUVmean Bias (%)	1.27 ± 20.8	18.85 ± 21.5	93% reduction	0.020*
SUVmax Bias (%)	-1.17 ± 12.5	13.11 ± 15.5	89% reduction	0.050*
RMSE (SUV)	409.4 ± 305.2	741.1 ± 395.5	45% reduction	<0.001**
MAE (SUV)	172.5 ± 147.1	326.8 ± 227.5	47% reduction	0.008**

Paired t-test was used for normally distributed differences, Wilcoxon signed-rank test for non-normal distributions.
* $p < 0.05$, ** $p < 0.01$

The statistical analysis confirmed that GAN-AC achieved significantly better performance than MR-AC across all evaluated metrics. SUVmean Bias showed significant improvement ($p = 0.020$), as did SUVmax Bias ($p = 0.050$), demonstrating that our synthetic CT approach produces statistically significant reductions in both average and peak uptake bias. Error metrics RMSE and MAE showed highly significant improvements ($p < 0.001$ and $p = 0.008$, respectively), confirming that the observed reductions in quantification errors are statistically robust and not due to chance variation across the patient cohort.

Our LOO-CV results reveal substantial improvements with GAN-AC compared to vendor MR-AC across all evaluated metrics. The most striking finding concerns SUVmean Bias, where GAN-AC achieved 1.27% ± 20.8% versus MR-AC's 18.85% ± 21.5%, representing a statistically significant 93% reduction ($p = 0.020$). This near-zero bias suggests our synthetic CT approach produces PET images with minimal systematic error relative to the CT-AC gold standard.

Similarly, SUVmax Bias showed a statistically significant 89% reduction with GAN-AC (-1.17% ± 12.5%) compared to MR-AC (13.11% ± 15.5%, $p = 0.050$). While the slight negative value indicates minor underestimation of peak uptake, this

magnitude remains clinically acceptable and markedly outperforms the substantial overestimation we observed with MR-AC. Accurate peak uptake quantification is clinically important as SUVmax reflects the most metabolically active regions [71].

Error metrics further reinforced GAN-AC's superiority with highly significant improvements. RMSE decreased 45% (409.4 ± 305.2 vs 741.1 ± 395.5 SUV, $p < 0.001$), while MAE improved 47% (172.5 ± 147.1 vs 326.8 ± 227.5 SUV, $p = 0.008$). These statistically significant reductions in both systematic and random errors indicate our GAN approach delivers more reliable PET quantification across diverse patient anatomies and pathologies. The improvements align well with previous findings that synthetic CT methods reduce SUV errors compared to MR-based approaches [72], demonstrating performance consistent with reported synthetic CT advantages [73]. Notably, GAN-AC's consistent standard deviations suggest robust performance across patients, while MR-AC's larger variability indicates less predictable results - likely due to well-documented limitations in air and bone regions where Dixon-based methods struggle [74].

Figures 4.12 and 4.13 provide visual support for these quantitative findings across axial, coronal, and sagittal views from a representative patient. Figure 4.12 shows PET slices from each AC method. GAN-AC closely matches the CT-AC reference across all anatomical planes, exhibiting similar SUV uptake patterns in cardiac and surrounding tissues. A subtle intensity reduction is visible, consistent with the slight underestimation observed in our LOO-CV results (SUVmean Bias 1.27%). Conversely, MR-AC demonstrates moderate overestimation with brighter uptake in cardiac and spinal regions across views, aligning with its positive SUVmean Bias of 18.85%.

Figure 4.13 displays difference maps relative to CT-AC reference. The (GAN-AC) - (CT-AC) maps show predominantly small differences throughout the image, reflecting the excellent quantitative accuracy achieved through our LOO-CV methodology. Meanwhile, (MR-AC) - (CT-AC) maps exhibit more pronounced positive differences (red) in cardiac and spinal regions. The higher error variability in MR-AC becomes evident through the mixture of positive and negative differences, particularly visible in axial and sagittal views. These visual observations support our quantitative LOO-CV findings and highlight GAN-In synthetic CT's superior performance for PET AC, especially in high-uptake regions like the heart where MR-AC continues showing limitations in air and bone regions [75].

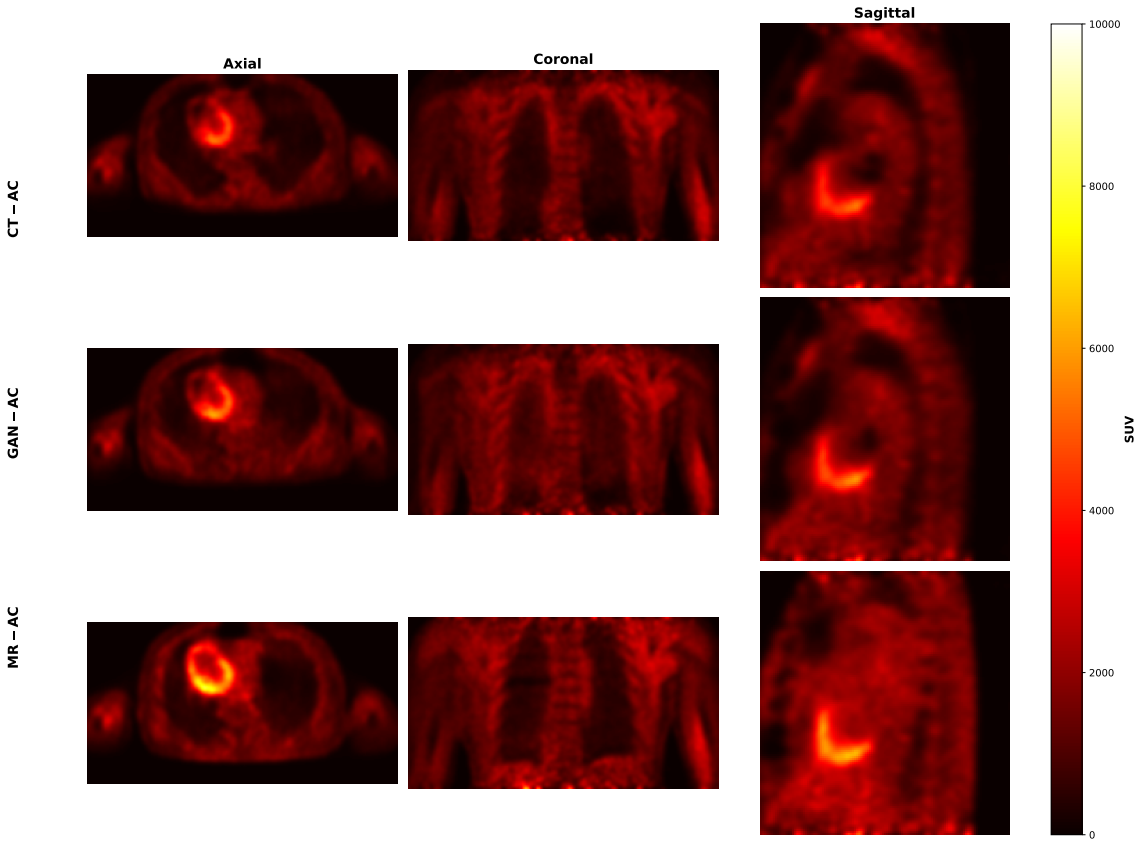

Figure 4.12: PET slices for subject 019 in axial, coronal, and sagittal views, reconstructed using different AC methods: CT-AC (Reference), GAN-AC, and MR-AC.

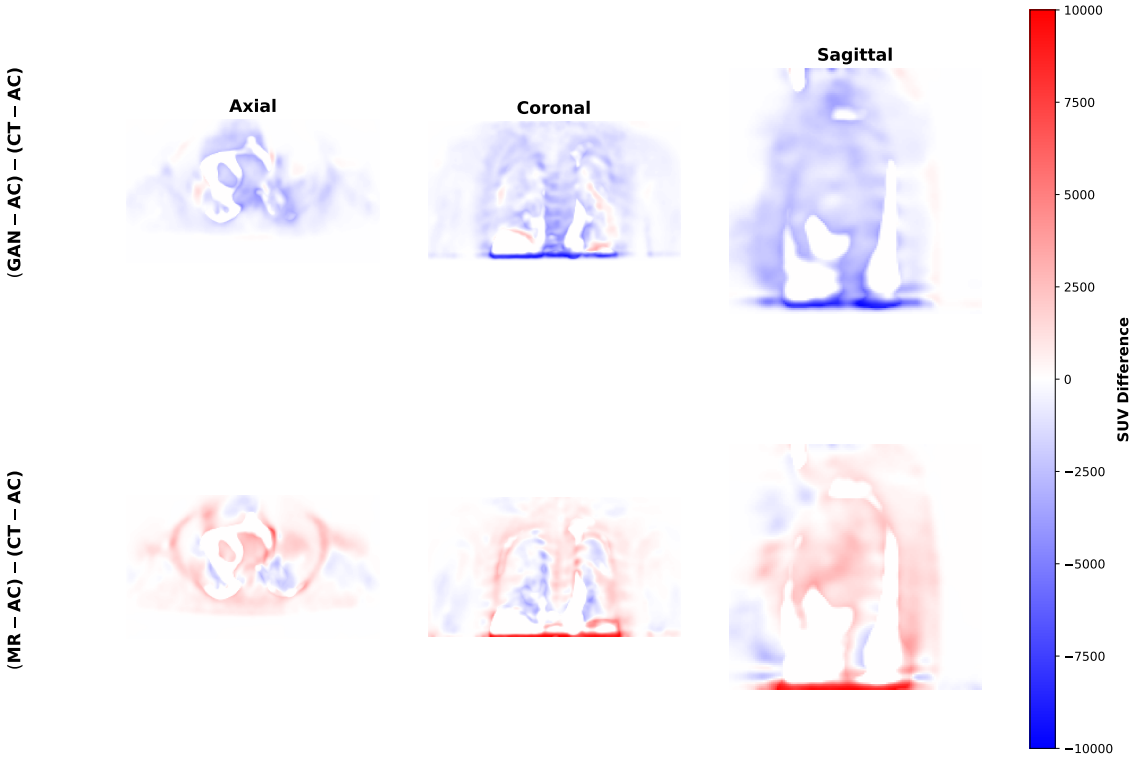

Figure 4.13: Difference maps of reconstructed PET images for subject 019 in axial, coronal, and sagittal views, comparing GAN-AC and MR-AC against the CT-AC reference. Blue indicates underestimation, red indicates overestimation, and white indicates minimal difference.

Chapter 5

Conclusion and Future Work

5.1 Summary of Research Objectives

This thesis aimed to address the challenges in PET/MR AC by developing and evaluating methods for generating pseudo-CT images from MR data, with a focus on improving PET quantification accuracy in cardiac imaging. The primary motivation stemmed from the need to replace the vendor-provided AC method at The Royal's IMHR, which uses a segmentation-based approach that struggles with poor bone estimation, leading to suboptimal PET/MR imaging results. This issue is compounded by broader limitations in existing methods, such as segmentation and atlas-based approaches that fail to capture continuous attenuation variations, and deep learning methods that lack cardiac-specific development due to scarce paired MR-CT data. By introducing a novel framework involving a CT template, multi-modal registration, and a DC-cycleGAN model with a novel μ-map loss, this work sought to enhance the accuracy of PET reconstruction and unlock the potential of PET/MR imaging for clinical applications.

5.2 Key Findings and Contributions

The proposed methodology successfully achieved the outlined objectives. A CT template and multi-modal registration framework were developed to align MR and CT images, ensuring anatomical correspondence for model training and evaluation. The baseline KRR model was investigated for unidirectional MR-to-CT translation, providing a reference for comparison. The DC-cycleGAN model, incorporating a novel mumap loss, was implemented for bidirectional MR-CT translation, generat-

ing pseudo-CT images which are converted into pseudo-μ-maps for PET AC. The evaluation of this method against the traditional MR-based AC by the vendor demonstrated significant improvements. Notably, the GAN-In model, utilizing in-phase MR images, emerged as the best performer, achieving a 39.5% reduction in MAE (0.112 vs. 0.185) and a 37.1% improvement in SSIM (0.503 vs. 0.367) compared to KRR-In for synthetic CT generation. Tissue-specific analysis revealed GAN-In's strengths in fat, soft tissue, and bone, while GAN-In-Out excelled in air and lung tissues.

For PET reconstruction, we conducted a comprehensive LOO-CV study across 8 patients to evaluate the clinical potential of our GAN-based approach (GAN-In). The results were highly encouraging: our synthetic CT-based AC (GAN-AC) using the GAN-In model substantially outperformed the vendor's standard MR-AC, relative to CT-AC considered as the gold standard method for AC. Statistical analysis using paired t-tests and Wilcoxon signed-rank tests confirmed the significance of all observed improvements. We observed a statistically significant 93% reduction in SUVmean bias (1.27% vs. 18.85%, $p = 0.020$), a significant 89% reduction in SUVmax bias (-1.17% vs. 13.11%, $p = 0.050$), and highly significant improvements in both RMSE (45% reduction: 409.4 vs. 741.1 SUV, $p < 0.001$) and MAE (47% reduction: 172.5 vs. 326.8 SUV, $p = 0.008$). These findings demonstrate that our approach not only achieves superior accuracy but also provides the consistency needed for reliable clinical application.

The GAN-In model, which uses only in-phase MR images, delivered consistently strong results across all metrics. Meanwhile, the more complex multi-channel GAN-In-Out model (which combines both in-phase and out-of-phase MR inputs) also performed well but didn't offer any statistically significant improvements over GAN-In. This suggests that in-phase MR images alone provide the key anatomical details necessary for accurate CT synthesis, and that adding out-of-phase information offers little additional benefit in our current implementation.

These findings contribute to the field of medical imaging by providing a robust framework for PET/MR AC. The significant improvements in synthetic CT generation and PET reconstruction accuracy, particularly in high-uptake regions like the heart, address the research gap in cardiac-specific pseudo-CT generation. The proposed methods enhance PET quantification accuracy, which is critical for diagnostic confidence in cardiac monitoring, and have potential applications in radiation therapy planning and other inter-modal image translation tasks, aligning with the broader motivation of advancing PET/MR imaging.

5.3 Limitations

Despite the above achievements, the study faced several limitations. Misalignment and field-of-view coverage issues between MR and CT images, due to their acquisition from different scanners, impacted registration accuracy. The registration process, both from CT to MR and to the template, could benefit from further parameter tuning. Similarly, the CT template creation process, which relied on registered CT-to-MR images, was constrained by alignment issues and would improve with better parameter optimization. The dataset size was limited to 10 subjects, potentially affecting the generalizability of the models. Additionally, the DC-cycleGAN model's performance could be enhanced with more extensive parameter tuning, which was mainly constrained by time. Furthermore, the study employed only 2D methods for image synthesis; training the models using 3D approaches may yield improved spatial consistency and overall performance.

5.4 Future Directions

Future work should focus on addressing the above limitations to further improve the proposed framework. Expanding the dataset to include a larger and more diverse cohort of subjects would enhance model generalizability and robustness. Improved alignment between MR and CT images, through further advanced registration techniques and parameter optimization, could enhance both template creation and overall model performance. Similarly, further tuning of the DC-cycleGAN model's hyperparameters could lead to better synthetic CT generation, particularly in challenging regions like bone.

While our multi-channel GAN-In-Out model didn't significantly outperform the single-channel GAN-In model, alternative methods for combining in-phase and out-of-phase MR images deserve closer examination. Future work could explore different fusion techniques at various network depths, attention mechanisms that dynamically weight the importance of each channel, or specialized architectures designed to better leverage the complementary information between in-phase and out-of-phase sequences. Such approaches might unlock additional performance gains that weren't realized in the current implementation.

5.4.1 Extensions to Mixed Datasets, Other Modalities and Regions

While this work utilized paired data with the DC-cycleGAN architecture, a significant advantage of this approach is its potential extension to mixed datasets containing both paired and unpaired examples. The cycle consistency component would allow learning from unpaired samples, while the μ-map loss could provide direct supervision for paired examples. This flexibility makes the approach particularly promising for medical imaging applications where paired data may be limited or inconsistently available across institutions or cohorts.

Furthermore, extending the framework to other imaging modalities or anatomical regions beyond cardiac PET/MR could broaden its impact, potentially benefiting areas like neuroimaging or oncology.

5.4.2 Clinical Implementation

To facilitate practical application, the implementation of the GAN-In model as a web-based application will enable IMHR researchers to upload MRI images and generate synthetic CT images internally via a browser interface, hosted on a secure institutional server. This ensures ease of use, compatibility with their clinical workflows, and data privacy, enabling IMHR to leverage the model for PET/MR AC and related research.

In conclusion, this thesis provides a promising step forward in PET/MR AC, leveraging deep learning to improve imaging accuracy. The proposed framework, with its CT template, registration pipeline, and DC-cycleGAN model, lays a foundation for future advancements in medical imaging. By addressing the identified limitations and exploring the suggested research directions, this work has the potential to further enhance diagnostic accuracy and enable broader adoption of PET/MR imaging in clinical settings.

References

[1] Bernd J. Pichler et al. "PET/MRI: The next generation of multimodality imaging?" In: *Seminars in Nuclear Medicine* 40.3 (2010), pp. 199–208 (cit. on pp. 1, 2).

[2] Mark A. Brown and Richard C. Semelka. *MRI: Basic Principles and Applications*. 5th. Wiley-Blackwell, 2014 (cit. on p. 1).

[3] Jonathan P. J. Carney et al. "Method for transforming CT images for attenuation correction in PET/CT imaging". In: *Medical Physics* 33.4 (2006), pp. 976–983 (cit. on pp. 2, 3, 33).

[4] Dale L. Bailey. *Positron Emission Tomography: Basic Sciences*. Springer, 1996. Chap. Attenuation Correction (cit. on p. 2).

[5] Flemming L. Andersen et al. "Combined PET/MR imaging in neurology: MR-based attenuation correction implies a strong spatial bias when ignoring bone". In: *NeuroImage* 84 (2014), pp. 206–216 (cit. on pp. 2, 3).

[6] Vincent Keereman et al. "MRI-based attenuation correction for PET/MRI using ultrashort echo time sequences". In: *Journal of Nuclear Medicine* 51.5 (2010), pp. 812–818 (cit. on p. 3).

[7] Fang Liu et al. "Deep learning MR imaging–based attenuation correction for PET/MR imaging". In: *Radiology* 286.2 (2018), pp. 676–684 (cit. on p. 3).

[8] Jun-Yan Zhu et al. "Unpaired Image-to-Image Translation using Cycle-Consistent Adversarial Networks". In: *Proceedings of the IEEE International Conference on Computer Vision* (2017), pp. 2223–2232. URL: https://arxiv.org/abs/1703.10593 (cit. on pp. 3, 29, 30).

[9] Jiayuan Wang, Q M Jonathan Wu, and Farhad Pourpanah. "DC-cycleGAN: Bidirectional CT-to-MR synthesis from unpaired data". In: *Computerized Medical Imaging and Graphics* (2023), p. 102249. URL: https://%doi.org/10.1016/j.compmedimag.2023.102249 (cit. on pp. 3, 15, 29, 31, 32, 34).

[10] Vladimir Vovk. "Kernel Ridge Regression". In: *Empirical Inference: Festschrift in Honor of Vladimir N. Vapnik*. Ed. by Bernhard Schölkopf, Zhiyuan Luo, and Vladimir Vovk. Berlin, Heidelberg: Springer, 2013, pp. 105–116 (cit. on pp. 4, 28).

[11] A. Martinez-Möller et al. "Tissue Classification as a Potential Approach for Attenuation Correction in Whole-Body PET/MRI: Evaluation with PET/CT Data". In: *Journal of Nuclear Medicine* 50.4 (2009), pp. 520–526 (cit. on pp. 6, 9).

[12] C. N. Ladefoged et al. "A multi-centre evaluation of eleven clinically feasible brain PET/MRI attenuation correction techniques using a large cohort of patients". In: *NeuroImage* 147 (2017), pp. 346–359 (cit. on p. 6).

[13] C. Rischpler et al. "Hybrid PET/MR imaging of the heart: potential, initial experiences, and future prospects". In: *Journal of Nuclear Medicine* 54.3 (2013), pp. 402–415 (cit. on p. 6).

[14] S. H. Keller et al. "Image artifacts from MR-based attenuation correction in clinical, whole-body PET/MRI". In: *MAGMA* 26.1 (2013), pp. 173–181 (cit. on p. 6).

[15] F. Nensa et al. "Hybrid cardiac imaging using PET/MRI: a joint position statement by the European Society of Cardiovascular Radiology (ESCR) and the European Association of Nuclear Medicine (EANM)". In: *European Journal of Hybrid Imaging* 2 (2018), p. 14 (cit. on pp. 6, 8, 9).

[16] D. Hwang et al. "Improving the Accuracy of Simultaneously Reconstructed Activity and Attenuation Maps Using Deep Learning". In: *Journal of Nuclear Medicine* 59.10 (2018), pp. 1624–1629 (cit. on p. 6).

[17] T. J. Bradshaw et al. "Feasibility of Deep Learning-Based PET/MR Attenuation Correction in the Pelvis Using Only Diagnostic MR Images". In: *Tomography* 4.3 (2018), pp. 138–147 (cit. on pp. 6, 18).

[18] D. L. Bailey. "Transmission scanning in emission tomography". In: *European Journal of Nuclear Medicine* 25.7 (1998), pp. 774–787 (cit. on p. 7).

[19] H. Zaidi and B. Hasegawa. "Determination of the attenuation map in emission tomography". In: *Journal of Nuclear Medicine* 44.2 (2003), pp. 291–315 (cit. on p. 7).

[20] K. L. Gould et al. “Frequent diagnostic errors in cardiac PET/CT due to misregistration of CT attenuation and emission PET images: a definitive analysis of causes, consequences, and corrections”. In: *Journal of Nuclear Medicine* 48.7 (2007), pp. 1112–1121 (cit. on pp. 7, 17).

[21] P. E. Kinahan et al. “Attenuation correction for a combined 3D PET/CT scanner”. In: *Medical Physics* 25.10 (1998), pp. 2046–2053 (cit. on p. 7).

[22] T. Schepis et al. “Attenuation correction in cardiac PET/CT with three different CT protocols: a comparison with conventional PET”. In: *European Journal of Nuclear Medicine and Molecular Imaging* 34.12 (2007), pp. 1991–2000 (cit. on p. 7).

[23] C. Burger et al. “PET attenuation coefficients from CT images: experimental evaluation of the transformation of CT into PET 511-keV attenuation coefficients”. In: *European Journal of Nuclear Medicine and Molecular Imaging* 29.7 (2002), pp. 922–927 (cit. on p. 7).

[24] A. Martinez-Möller et al. “Artifacts from misaligned CT in cardiac perfusion PET/CT studies: frequency, effects, and potential solutions”. In: *Journal of Nuclear Medicine* 48.2 (2007), pp. 188–193 (cit. on p. 7).

[25] Vincent Keereman et al. “MRI-based attenuation correction for PET/MRI using ultrashort echo time sequences”. In: *Journal of Nuclear Medicine* 51.5 (2013), pp. 812–818 (cit. on pp. 8, 9, 18).

[26] G. Wagenknecht et al. “MRI for attenuation correction in PET: methods and challenges”. In: *MAGMA* 26.1 (2013), pp. 99–113 (cit. on p. 8).

[27] G. Delso et al. “The effect of limited MR field of view in MR/PET attenuation correction”. In: *Medical Physics* 38.7 (2011), pp. 4173–4182 (cit. on p. 8).

[28] J. M. C. Lau et al. “Evaluation of attenuation correction in cardiac PET using PET/MR”. In: *Journal of Nuclear Cardiology* 24.3 (2017), pp. 839–846 (cit. on pp. 8, 9).

[29] S. Fürst et al. “Motion Correction Strategies for Integrated PET/MR”. In: *Journal of Nuclear Medicine* 56.2 (2015), pp. 261–269 (cit. on p. 8).

[30] Matthias Hofmann et al. “MRI-based attenuation correction for whole-body PET/MRI: quantitative evaluation of segmentation-and atlas-based methods”. In: *Journal of Nuclear Medicine* 52.9 (2011), pp. 1392–1399 (cit. on p. 9).

[31] D. Izquierdo-Garcia et al. “Impact of MR-based attenuation correction on neurological PET studies”. In: *NeuroImage* 111 (2015), pp. 622–629 (cit. on pp. 9, 10).

[32] Ninon Burgos et al. "Attenuation correction synthesis for hybrid PET-MR scanners: application to brain studies". In: *IEEE Transactions on Medical Imaging* 33.12 (2014), pp. 2332–2341 (cit. on pp. 9, 12).

[33] F. M. Bengel et al. "Cardiac positron emission tomography". In: *Journal of the American College of Cardiology* 54.1 (2013), pp. 1–15 (cit. on p. 9).

[34] C. Catana et al. "Toward implementing an MRI-based PET attenuation-correction method for neurologic studies on the MR-PET brain prototype". In: *Journal of Nuclear Medicine* 51.9 (2010), pp. 1431–1438 (cit. on pp. 9, 10).

[35] F. Wiesinger et al. "Zero TE MR bone imaging in the head". In: *Magnetic Resonance in Medicine* 79.3 (2018), pp. 1704–1714 (cit. on p. 10).

[36] A. Rezaei et al. "Simultaneous reconstruction of activity and attenuation in time-of-flight PET". In: *IEEE Transactions on Medical Imaging* 31.12 (2012), pp. 2224–2233 (cit. on p. 10).

[37] A. Salomon et al. "Simultaneous reconstruction of activity and attenuation for PET/MR". In: *IEEE Transactions on Medical Imaging* 30.3 (2011), pp. 804–813 (cit. on p. 10).

[38] F. Liu et al. "Deep Learning MR Imaging–based Attenuation Correction for PET/MR Imaging". In: *Radiology* 286.2 (2018), pp. 676–684 (cit. on pp. 10, 18).

[39] K. Armanious et al. "MedGAN: Medical image translation using GANs". In: *Computerized Medical Imaging and Graphics* 79 (2019), p. 101684 (cit. on pp. 10, 12).

[40] X. Dong et al. "Deep learning-based attenuation correction in the absence of structural information for whole-body positron emission tomography imaging". In: *Physics in Medicine & Biology* 65.5 (2020), p. 055011 (cit. on pp. 10, 12, 19).

[41] Claes Nøhr Ladefoged et al. "A review of PET attenuation correction methods for PET-MR". In: *EJNMMI Physics* 10.1 (2023), pp. 1–21 (cit. on p. 11).

[42] H. Jang et al. "Technical Note: Deep learning based MRAC using rapid ultrashort echo time imaging". In: *Medical Physics* 45.8 (2018), pp. 3697–3704 (cit. on p. 11).

[43] Andrew P Leynes et al. "Attenuation Correction of PET/MR Imaging". In: *Seminars in nuclear medicine* 47.5 (2017), pp. 556–565 (cit. on p. 11).

[44] D. Hwang et al. “Improving the Accuracy of Simultaneously Reconstructed Activity and Attenuation Maps Using Deep Learning”. In: *Journal of Nuclear Medicine* 59.10 (2018), pp. 1624–1629 (cit. on p. 12).

[45] K. Baes, D. Robben, and P. Suetens. “Attention-based 3D U-Net with learned upsampling for brain MR-to-CT synthesis”. In: *Medical Imaging with Deep Learning (MIDL)* (2020), pp. 49–60 (cit. on pp. 12, 13).

[46] Ian J. Goodfellow et al. “Generative Adversarial Nets”. In: *Advances in Neural Information Processing Systems* 27 (2014), pp. 2672–2680 (cit. on p. 13).

[47] S Kevin Zhou et al. “A review of deep learning in medical imaging: Imaging traits, technology trends, case studies with progress highlights, and future promises”. In: *Proceedings of the IEEE* 110.1 (2022), pp. 138–157 (cit. on pp. 14, 15).

[48] Xin Dong et al. “MR-based attenuation correction for brain PET using 3D cycle-consistent adversarial network”. In: *Annals of Translational Medicine* 8.7 (2020), p. 462 (cit. on p. 14).

[49] Mina Talebian et al. “Assessing the efficacy of 3D Dual-CycleGAN model for multi-contrast MRI synthesis”. In: *Egyptian Journal of Radiology and Nuclear Medicine* 55.1 (2024), p. 36 (cit. on p. 16).

[50] C. Kolbitsch et al. “Cardiac and Respiratory Motion Correction for Simultaneous Cardiac PET/MR”. In: *Journal of Nuclear Medicine* 58.5 (2017), pp. 846–852 (cit. on p. 17).

[51] M. Hofmann et al. “MRI-based attenuation correction for PET/MRI: A novel approach combining pattern recognition and atlas registration”. In: *Journal of Nuclear Medicine* 49.11 (2008), pp. 1875–1883 (cit. on p. 17).

[52] J. Ouyang, Q. Li, and G. El Fakhri. “Magnetic resonance-based motion correction for positron emission tomography imaging”. In: *Seminars in Nuclear Medicine* 43.1 (2013), pp. 60–67 (cit. on p. 17).

[53] J. M. Lau et al. “Cardiac applications of PET/MR imaging”. In: *Magnetic Resonance Imaging Clinics of North America* 24.1 (2016), pp. 159–168 (cit. on p. 17).

[54] N. Burgos et al. “Attenuation correction in PET/MR: Evaluation of atlas-based methods”. In: *IEEE Transactions on Medical Imaging* 33.11 (2014), pp. 2166–2177 (cit. on p. 17).

[55] F. Wiesinger et al. "Zero TE-based pseudo-CT image conversion in the head and its application in PET/MR attenuation correction and MR-guided radiation therapy planning". In: *Magnetic Resonance in Medicine* 76.5 (2016), pp. 1440–1451 (cit. on p. 17).

[56] A. P. Leynes et al. "Zero-echo-time and Dixon deep pseudo-CT (ZeDD CT)". In: *Journal of Nuclear Medicine* 59.5 (2018), pp. 792–798 (cit. on p. 17).

[57] A. Mehranian, H. Zaidi, and A. J. Reader. "Joint estimation of activity and attenuation in whole-body TOF PET/MRI". In: *Journal of Nuclear Medicine* 57.5 (2016), pp. 780–786 (cit. on p. 17).

[58] F. Liu et al. "Deep learning-based MR-to-CT synthesis using unpaired data". In: *Simulation and Synthesis in Medical Imaging*. Vol. 11037. 2018, pp. 14–23 (cit. on p. 17).

[59] K. Armanious et al. "MedGAN: Medical image translation using GANs". In: *Computerized Medical Imaging and Graphics* 79 (2020), p. 101684 (cit. on p. 17).

[60] X. Dong et al. "Synthetic CT generation from non-attenuation corrected PET images for whole-body PET imaging". In: *Physics in Medicine and Biology* 64.21 (2019), p. 215016 (cit. on p. 18).

[61] F. Nensa et al. "Hybrid cardiac imaging using PET/MRI: a joint position statement by the European Society of Cardiovascular Radiology and the European Association of Nuclear Medicine". In: *European Journal of Hybrid Imaging* 2.1 (2018), p. 14 (cit. on p. 18).

[62] F. Nensa, A. Demircioglu, and C. Rischpler. "Clinical applications of PET/MRI in cardiology". In: *Journal of the American College of Cardiology: Cardiovascular Imaging* 12.7 (2019), pp. 1266–1279 (cit. on p. 18).

[63] The University of Ottawa Heart Institute and University of Ottawa Institute of Mental Health Research. "Optimization of Attenuation Correction for FDG PET-MR Imaging". Unpublished study protocol, Version 2, The University of Ottawa Heart Institute and University of Ottawa Institute of Mental Health Research. Oct. 2017 (cit. on p. 20).

[64] László G Nyúl, Jayaram K Udupa, and Xuan Zhang. "New variants of a method of MRI scale standardization". In: *IEEE Transactions on Medical Imaging* 19.2 (2000), pp. 143–150 (cit. on p. 41).

[65] Russell T Shinohara et al. "Statistical normalization techniques for magnetic resonance imaging". In: *NeuroImage: Clinical* 6 (2014), pp. 9–19 (cit. on p. 41).

[66] Abdel Aziz Taha and Allan Hanbury. "Metrics for evaluating 3D medical image segmentation: analysis, selection, and tool". In: *BMC Medical Imaging* 15.1 (2015), p. 29 (cit. on pp. 43, 44).

[67] Torsten Rohlfing et al. "Evaluation of atlas selection strategies for atlas-based image segmentation with application to confocal microscopy images of bee brains". In: *NeuroImage* 21.4 (2004), pp. 1428–1442 (cit. on pp. 43, 44).

[68] Ninon Burgos et al. "Robust CT synthesis for radiotherapy planning: application to the head and neck region". In: *IEEE Transactions on Medical Imaging* 33.9 (2014), pp. 1825–1834 (cit. on pp. 43, 44).

[69] Isaac Shiri et al. "Feasibility of Deep Learning-Guided Attenuation and Scatter Correction of Whole-Body 68Ga-PSMA PET Studies in the Image Domain". In: *European Journal of Nuclear Medicine and Molecular Imaging* 47.11 (2020), pp. 2533–2548. URL: https://%doi.org/10.1007/s00259-020-04852-5 (cit. on p. 51).

[70] Claude Nadeau and Yoshua Bengio. "Inference for the generalization error". In: *Machine Learning* 52.3 (2003), pp. 239–281 (cit. on p. 59).

[71] Martin A Lodge, Yalda Jahangiri, Maria A Chaudhry, et al. "Noise considerations for PET quantification using maximum and peak standardized uptake value". In: *Journal of Nuclear Medicine* 53.7 (2012), pp. 1041–1047 (cit. on p. 61).

[72] Isaac Shiri et al. "Feasibility of Deep Learning-Guided Attenuation and Scatter Correction of Whole-Body 68Ga-PSMA PET Studies in the Image Domain". In: *Clinical Nuclear Medicine* 46.8 (2021), pp. 609–615 (cit. on p. 61).

[73] Abolfazl Mehranian and Habib Zaidi. "Attenuation Correction of PET/MR Imaging". In: *Magnetic Resonance Imaging Clinics of North America* 25.2 (2017), pp. 281–297 (cit. on p. 61).

[74] Isaac Shiri et al. "Deep learning-guided estimation of attenuation correction factors from time-of-flight PET emission data". In: *Medical Image Analysis* 64 (2020), p. 101718 (cit. on p. 61).

[75] Andrew P Leynes et al. "Hybrid ZTE/Dixon MR-based attenuation correction for quantitative evaluation of PET images in the pelvis". In: *Journal of Nuclear Medicine* 58.2 (2017), pp. 201–207 (cit. on p. 61).

APPENDICES

Appendix A

Appendix

A.1 Evaluation Metrics

This section of the appendix presents the mathematical formulas for all evaluation metrics used in this thesis.

A.1.1 Image Quality Metrics for Synthetic CT Evaluation

Mean Absolute Error (MAE)

$$\text{MAE} = \frac{1}{N} \sum_{i=1}^{N} |y_i - \hat{y}_i| \tag{A.1}$$

where y_i is the true CT intensity value at voxel i, $\hat{y}_i$ is the synthetic CT intensity value at voxel i, and N is the total number of voxels.

Mean Squared Error (MSE)

$$\text{MSE} = \frac{1}{N} \sum_{i=1}^{N} (y_i - \hat{y}_i)^2 \tag{A.2}$$

Peak Signal-to-Noise Ratio (PSNR)

$$\text{PSNR} = 10 \log_{10} \left(\frac{\text{MAX}^2}{\text{MSE}} \right) \tag{A.3}$$

where MAX is the maximum possible intensity value in the image (1.0 for normalized images).

Structural Similarity Index (SSIM)

$$\text{SSIM}(x, y) = \frac{(2\mu_x\mu_y + C_1)(2\sigma_{xy} + C_2)}{(\mu_x^2 + \mu_y^2 + C_1)(\sigma_x^2 + \sigma_y^2 + C_2)} \tag{A.4}$$

where:

- μ_x is the mean intensity of image x
- μ_y is the mean intensity of image y
- σ_x^2 is the variance of image x
- σ_y^2 is the variance of image y
- σ_{xy} is the covariance of images x and y
- $C_1 = (k_1 L)^2$ and $C_2 = (k_2 L)^2$ are constants to stabilize the division (where L is the dynamic range of pixel values, typically $2^{\text{bits per pixel}} - 1$)

A.1.2 Registration Quality Metrics

Normalized Cross-Correlation (NCC)

$$\text{NCC} = \frac{\sum_{i=1}^{N}(A_i - \bar{A})(B_i - \bar{B})}{\sqrt{\sum_{i=1}^{N}(A_i - \bar{A})^2 \sum_{i=1}^{N}(B_i - \bar{B})^2}} \tag{A.5}$$

where A_i and B_i are intensity values of corresponding voxels in images A and B, and $\bar{A}$ and $\bar{B}$ are the mean intensity values of images A and B.

Mutual Information (MI)

$$\text{MI}(A, B) = H(A) + H(B) - H(A, B) \tag{A.6}$$

where:

- $H(A)$ is the entropy of image A: $H(A) = -\sum_a p_A(a) \log p_A(a)$
- $H(B)$ is the entropy of image B: $H(B) = -\sum_b p_B(b) \log p_B(b)$
- $H(A, B)$ is the joint entropy: $H(A, B) = -\sum_{a,b} p_{AB}(a, b) \log p_{AB}(a, b)$
- $p_A(a)$ and $p_B(b)$ are the marginal probability distributions
- $p_{AB}(a, b)$ is the joint probability distribution

Left-Right Symmetry Index

$$\text{Symmetry Index} = \frac{2 \times \text{NCC}(I, I_{\text{flipped}})}{1 + \text{NCC}(I, I_{\text{flipped}})} \tag{A.7}$$

where I is the original image and I_{flipped} is the left-right flipped version of the image.

A.1.3 PET Quantification Metrics

Relative SUV Error

$$\text{Relative SUV Error (\%)} = \frac{|\text{SUV}_{\text{evaluated}} - \text{SUV}_{\text{reference}}|}{\text{SUV}_{\text{reference}}} \times 100 \tag{A.8}$$

where $\text{SUV}_{\text{evaluated}}$ is the SUV value in the PET image reconstructed using the method being evaluated (e.g., GAN-AC or MR-AC) and $\text{SUV}_{\text{reference}}$ is the SUV value in the reference PET image (reconstructed using CT-AC).

SUVmean Bias

$$\text{SUVmean Bias (\%)} = \frac{\text{SUVmean}_{\text{evaluated}} - \text{SUVmean}_{\text{reference}}}{\text{SUVmean}_{\text{reference}}} \times 100 \tag{A.9}$$

where $\text{SUVmean}_{\text{evaluated}}$ is the mean SUV in the PET image reconstructed using the method being evaluated (e.g., GAN-AC or MR-AC) and $\text{SUVmean}_{\text{reference}}$ is the mean SUV in the reference PET image (reconstructed using CT-AC).

Mean Absolute Error (MAE) for SUV

$$\text{MAE} = \frac{1}{N} \sum_{i=1}^{N} |\text{SUV}_i - \text{SUV}_{\text{reference},i}| \tag{A.10}$$

where SUV_i is the SUV value at voxel i in the PET image reconstructed using the method being evaluated (e.g., GAN-AC or MR-AC) and $\text{SUV}_{\text{reference},i}$ is the SUV value at the same voxel in the reference PET image (reconstructed using CT-AC), and N is the total number of voxels.

SUVmax Bias

$$\text{SUVmax Bias (\%)} = \frac{\text{SUVmax}_{\text{evaluated}} - \text{SUVmax}_{\text{reference}}}{\text{SUVmax}_{\text{reference}}} \times 100 \tag{A.11}$$

where $\text{SUVmax}_{\text{evaluated}}$ is the maximum SUV in a region of interest in the PET image reconstructed using the method being evaluated (e.g., GAN-AC or MR-AC) and $\text{SUVmax}_{\text{reference}}$ is the maximum SUV in the same region in the reference PET image (reconstructed using CT-AC).

Root Mean Square Error (RMSE) for SUV

$$\text{RMSE} = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(\text{SUV}_i - \text{SUV}_{\text{reference},i})^2} \tag{A.12}$$

where SUV_i is the SUV value at voxel i in the PET image reconstructed using the method being evaluated (e.g., GAN-AC or MR-AC) and $\text{SUV}_{\text{reference},i}$ is the SUV value at the same voxel in the reference PET image (reconstructed using CT-AC), and N is the total number of voxels.

A.1.4 Additional Statistical Metrics

Coefficient of Variation (CV)

$$\text{CV} = \frac{\sigma}{\mu} \tag{A.13}$$

where σ is the standard deviation and μ is the mean.

Signal-to-Noise Ratio (SNR)

$$\text{SNR} = \frac{\mu_{\text{signal}}}{\sigma_{\text{background}}} \tag{A.14}$$

where μ_{signal} is the mean intensity in the signal region and $\sigma_{\text{background}}$ is the standard deviation of intensity in the background region.

Contrast-to-Noise Ratio (CNR)

$$\text{CNR} = \frac{|\mu_A - \mu_B|}{\sqrt{\sigma_A^2 + \sigma_B^2}} \tag{A.15}$$

where μ_A and μ_B are the mean intensities of regions A and B, and σ_A and σ_B are the standard deviations of regions A and B.

Wilcoxon Signed-Rank Test

$$W = \sum_{i=1}^{N} [sgn(x_{2,i} - x_{1,i}) \cdot R_i] \tag{A.16}$$

where:

- $x_{1,i}$ and $x_{2,i}$ are the measurements for the i-th pair
- $sgn()$ is the sign function: $sgn(x) = \begin{cases} 1 & \text{if } x > 0 \\ 0 & \text{if } x = 0 \\ -1 & \text{if } x < 0 \end{cases}$
- R_i is the rank of the absolute difference $|x_{2,i} - x_{1,i}|$
- N is the number of pairs

The test statistic W is compared to critical values from the Wilcoxon signed-rank table, or for larger samples, a z-score can be calculated as:

$$z = \frac{W - \frac{N(N+1)}{4}}{\sqrt{\frac{N(N+1)(2N+1)}{24}}} \tag{A.17}$$

where the z-score follows an approximately normal distribution for $N > 20$.

The null hypothesis is typically that the median difference between pairs is zero, and is rejected when $|z| > z_{\alpha/2}$ for a significance level α.

A.2 Detailed Model Implementation

A.2.1 DC-CycleGAN Architecture and Training Algorithm

The Dual-Contrast CycleGAN (DC-CycleGAN) with μ-map loss is implemented as an enhanced version of the traditional CycleGAN for medical image translation between CT and MR modalities, with a focus on preserving attenuation properties critical for PET reconstruction.

Pseudocode for DC-CycleGAN Training with μ-map Loss

Algorithm 1.1 DC-CycleGAN Training with μ-map Loss

Input: CT dataset $\mathcal{X}$, MR dataset $\mathcal{Y}$, batch size b, number of training epochs N_{epoch}
Input: Generators $G_{CT\to MR}$, $G_{MR\to CT}$, Discriminators D_{CT}, D_{MR}
Input: Loss weights: λ_{cycle}, $\lambda_{\mu map}$, λ_{DC}
Input: Learning rates: lr_G, lr_D, decay epoch N_{decay}

Initialize generators $G_{CT\to MR}$, $G_{MR\to CT}$ and discriminators D_{CT}, D_{MR}
Initialize optimizers for generators and discriminators
Initialize CT-to-μmap converter $\mathcal{M}_{CT\to\mu}$
Initialize mixed precision scaler if using mixed precision training
for $epoch = 1$ to N_{epoch} **do**
  Set networks to training mode
  for each batch (x, y) of size b from datasets $\mathcal{X}$ and $\mathcal{Y}$ **do**
    // Training generators
    Zero generator gradients
    $y_{fake} \leftarrow G_{CT\to MR}(x)$ {Generate fake MR from CT}
    $x_{fake} \leftarrow G_{MR\to CT}(y)$ {Generate fake CT from MR}
    $x_{rec} \leftarrow G_{MR\to CT}(y_{fake})$ {Reconstruct CT}
    $y_{rec} \leftarrow G_{CT\to MR}(x_{fake})$ {Reconstruct MR}
    // Calculate generator losses
    $\mathcal{L}_{GAN}^{CT\to MR} \leftarrow MSE(D_{MR}(y_{fake}), 1)$ {GAN loss for fake MR}
    $\mathcal{L}_{GAN}^{MR\to CT} \leftarrow MSE(D_{CT}(x_{fake}), 1)$ {GAN loss for fake CT}
    $\mathcal{L}_{GAN} \leftarrow \frac{1}{2}(\mathcal{L}_{GAN}^{CT\to MR} + \mathcal{L}_{GAN}^{MR\to CT})$
    $\mathcal{L}_{cycle}^{CT} \leftarrow L1(x_{rec}, x)$ {Cycle consistency loss for CT}
    $\mathcal{L}_{cycle}^{MR} \leftarrow L1(y_{rec}, y)$ {Cycle consistency loss for MR}
    $\mathcal{L}_{cycle} \leftarrow \frac{1}{2}(\mathcal{L}_{cycle}^{CT} + \mathcal{L}_{cycle}^{MR})$
    if using μ-map loss **then**
      $\mu_{real} \leftarrow \mathcal{M}_{CT\to\mu}(x)$ {Convert real CT to μ-map}
      $\mu_{fake} \leftarrow \mathcal{M}_{CT\to\mu}(x_{fake})$ {Convert fake CT to μ-map}
      $\mathcal{L}_{\mu map} \leftarrow L1(\mu_{fake}, \mu_{real})$ {μ-map loss}
    else
      $\mathcal{L}_{\mu map} \leftarrow 0$
    end if
    $\mathcal{L}_G \leftarrow \mathcal{L}_{GAN} + \lambda_{cycle} \cdot \mathcal{L}_{cycle} + \lambda_{\mu map} \cdot \mathcal{L}_{\mu map}$
    Update generators using $\mathcal{L}_G$
    // Training discriminators
    Zero discriminator gradients
    // Train CT discriminator
    $\mathcal{L}_{real}^{CT} \leftarrow MSE(D_{CT}(x), 1)$ {Real CT loss}
    $\mathcal{L}_{fake}^{CT} \leftarrow MSE(D_{CT}(x_{fake}.detach()), 0)$ {Fake CT loss}
    if using dual contrast **then**
      $\mathcal{L}_{DC}^{CT} \leftarrow MSE(D_{CT}(y), 0)$ {CT discriminator should classify MR as fake}
      $\mathcal{L}_D^{CT} \leftarrow \frac{\mathcal{L}_{real}^{CT} + \mathcal{L}_{fake}^{CT} + \lambda_{DC} \cdot \mathcal{L}_{DC}^{CT}}{2 + \lambda_{DC}}$
    else
      $\mathcal{L}_D^{CT} \leftarrow \frac{1}{2}(\mathcal{L}_{real}^{CT} + \mathcal{L}_{fake}^{CT})$
    end if
    // Train MR discriminator
    $\mathcal{L}_{real}^{MR} \leftarrow MSE(D_{MR}(y), 1)$ {Real MR loss}
    $\mathcal{L}_{fake}^{MR} \leftarrow MSE(D_{MR}(y_{fake}.detach()), 0)$ {Fake MR loss}

A.2.2 Hyperparameter Configuration

The implementation uses the following hyperparameters for training the DC-CycleGAN with μ-map loss:

Table A.1: Hyperparameters for DC-CycleGAN with μ-map Loss

Parameter	**Value**	**Description**
Epochs	400	Number of training epochs
Batch Size	8	Batch size per GPU
Generator Learning Rate	0.0002	Initial learning rate for generators
Discriminator Learning Rate	0.0002	Initial learning rate for discriminators
Optimizer	Adam	Optimization algorithm
Adam β_1	0.5	First moment decay for Adam optimizer
Adam β_2	0.999	Second moment decay for Adam optimizer
Decay Epoch	25	Epoch from which to start learning rate decay
LR Decay	Linear	Learning rate decay schedule
λ_{cycle}	35.0	Weight for cycle consistency loss
$\lambda_{\mu map}$	17.5	Weight for μ-map loss
λ_{DC}	4.0	Weight for dual contrast loss
Residual Blocks	9	Number of residual blocks in generators
Image Size	256Ã—256	Input/output image dimensions
Channels	1	Number of input/output channels (grayscale)
Sample Interval	10	Interval between saving sample images (epochs)
FID Interval	5	Interval between FID evaluations (epochs)
Mixed Precision	Enabled	Use of mixed-precision training

A.2.3 Network Architecture

Generator Architecture

The Generator network follows a structure of encoding, transformation, and decoding:

Generator Architecture

- **Initial convolution block:**
  ReflectionPad2d(3), Conv2d($1 \to 64$, 7×7), InstanceNorm2d, ReLU
- **Downsampling ($\times 2$):**
  Conv2d($64 \to 128$, 3×3, stride=2), InstanceNorm2d, ReLU
  Conv2d($128 \to 256$, 3×3, stride=2), InstanceNorm2d, ReLU
- **Residual blocks ($\times 9$):**
  ResidualBlock(256) containing:
  ReflectionPad2d(1), Conv2d($256 \to 256$, 3×3), InstanceNorm2d, ReLU
  ReflectionPad2d(1), Conv2d($256 \to 256$, 3×3), InstanceNorm2d
  Skip connection (Addition)
  ...8 more identical blocks...
- **Upsampling ($\times 2$):**
  ConvTranspose2d($256 \to 128$, 3×3, stride=2), InstanceNorm2d, ReLU
  ConvTranspose2d($128 \to 64$, 3×3, stride=2), InstanceNorm2d, ReLU
- **Output layer:**
  ReflectionPad2d(3), Conv2d($64 \to 1$, 7×7), Tanh

Discriminator Architecture

- **Initial layer without normalization:**
  Conv2d($1 \to 64$, 4×4, stride=2, padding=1), LeakyReLU(0.2)
- **Downsampling blocks:**
  Conv2d($64 \to 128$, 4×4, stride=2, padding=1), InstanceNorm2d, LeakyReLU(0.2)
  Conv2d($128 \to 256$, 4×4, stride=2, padding=1), InstanceNorm2d, LeakyReLU(0.2)
  Conv2d($256 \to 512$, 4×4, stride=2, padding=1), InstanceNorm2d, LeakyReLU(0.2)
- **Output layer for patch discrimination:**
  ZeroPad2d((1, 0, 1, 0)), Conv2d($512 \to 1$, 4×4, padding=1)

Complete DC-CycleGAN Architecture

The complete DC-CycleGAN combines these components:

```
DC-CycleGAN(
  G_CT->MR (CT->MR Generator): Generator as described above
  G_MR->CT (MR->CT Generator): Generator as described above
  D_CT (CT Discriminator): Discriminator as described above
  D_MR (MR Discriminator): Discriminator as described above

  mumap Converter: Differentiable piecewise linear function
)
```

This architecture is trained using the combined losses described in the previous sections, with the μ-map loss providing an additional constraint to preserve attenuation properties in the generated CT images.

A.3 Registration Parameters and Template Creation Process

This section documents the key registration parameters and template creation process used in this thesis. The registration framework was implemented using the Advanced Normalization Tools (ANTs).

A.3.1 Registration Parameter Specifications

A.3.2 CT Template Creation Process

The CT template was created using the ANTs multivariate template construction tool with the following process:

1. **Initialization:** Collection of 10 registered CT-to-MR images and creation of configuration files with random seed (12345) for reproducibility.
2. **Initial Template Formation:** Creation of an initial template by rigidly aligning all input images to a reference image and averaging them.
3. **Iterative Refinement:** Three outer iterations of the following process:
    - Registration of each input image to the current template using SyN transformation
    - Calculation of a shape update field by averaging all deformation fields

Table A.2: Key Registration Parameters

Parameter	**CT-to-MR**	**CT Template Creation**	**Registration to Template**
Transform model	SyN[0.1,3,0] (Symmetric Normalization)		
Similarity metric	Mutual Information (MI), then Cross-Correlation (CC)	Cross-Correlation (CC) with 4-voxel radius	Cross-Correlation (CC)
Multi-resolution levels	4	4	4
Shrink factors	8×4×2×1	8×4×2×1	8×4×2×1
Smoothing sigmas	3×2×1×0 vox	3×2×1×0 vox	3×2×1×0 vox
Registration approach	Two-stage: Affine then Deformable	Iterative template construction	Composition of transforms
Iterations	[1000×500×250×100,1e-6,10] for affine [100×70×50×20,1e-6,10] for deformable	3 outer iterations	Same as CT-to-MR
Preprocessing	N4BiasFieldCorrection for MR Intensity rescaling [0,1]	N/A	N/A

- Application of the inverse average deformation to maintain an unbiased template
- Template update by averaging the transformed input images

4. **Convergence:** The process continued until the maximum number of iterations (3) was reached.
5. **Final Processing:** The completed template was saved as `final_ct_template.nii.gz` for subsequent registrations.

A.3.3 Complete Registration Pipeline

The complete registration pipeline consisted of the following steps:

1. **CT-to-MR Registration:** Each patient's CT was registered to their corresponding MR image using a two-stage approach (affine followed by SyN).
2. **CT Template Creation:** A population template was created from the registered CT images as described above.
3. **Registration to Template:** A hierarchical approach registered all modalities to the template:
   - CT-to-MR registered images to template space
   - Original CT images to template via composed transforms
   - MR in-phase images directly to template
   - MR opposed-phase (template space) alignment via rigid registration to in-phase images

This registration framework established the spatial correspondence necessary for model training and evaluation, ensuring that MR and CT images were properly aligned in a common template space.