

Original Research Article

Image Modality Translation between MRI and CT Using Diffusion and Score-Matching Algorithms

Kharananda Sharma^{1*}, Tshetiz Dahal², Bimal Nepal³, Jarina Tamang⁴¹Assistant Professor, Department of Radio-diagnosis and Imaging, Sikkim Manipal Institute of Medical Sciences, Tadong East Sikkim, India²General Physician, Clinical Researcher, Department of Medicine and Surgery, Lugansk State Medical University, Lypnia St. Rivne, Ukraine³PhD Research Scholar, Department of Radiology and Imaging Technology, College of Allied and Healthcare Sciences, Teerthankar Mahaveer University Moradabad, Uttar Pradesh, India⁴Senior Resident, Department of Radio diagnosis and Imaging, Sikkim Manipal Institute of Medical Sciences, Tadong East Sikkim, India**Article History**

Received: 07.09.2025

Accepted: 01.11.2025

Published: 07.11.2025

Journal homepage:<https://www.easpublisher.com>**Quick Response Code**

Abstract: MRI and CT are the two most widely used medical imaging techniques. Often, doctors need images from both modalities to diagnose conditions accurately and plan treatments like radiation therapy. However, using both MRI and CT can be costly, and it often results in misaligned images. A practical alternative is to use computational methods to convert images from one modality to another—particularly converting MRI images into CT images. In this study, we explore a deep learning approach using diffusion models and score-matching techniques to address this challenge. Specifically, we adapt denoising diffusion probabilistic models and score-matching strategies, apply four different sampling methods, and compare their performance with that of traditional models like generative adversarial networks (GANs) and convolutional neural networks (CNNs). Our results show that diffusion and score-matching models generate synthetic CT images with higher quality than CNN and GAN approaches. We also assess the uncertainty in these models using Monte Carlo simulations and further improve the final image quality by averaging the Monte Carlo outputs. Overall, our research suggests that diffusion and score-matching models not only rival CNNs and GANs in generating cross-modality medical images but also offer a more mathematically grounded and reliable framework.

Keywords: Computed Tomography, Magnetic Resonance Imaging, Image Synthesis, Uncertainty Estimation, Diffusion Model, and Score-Matching Model.

Copyright © 2025 The Author(s): This is an open-access article distributed under the terms of the Creative Commons Attribution **4.0 International License (CC BY-NC 4.0)** which permits unrestricted use, distribution, and reproduction in any medium for non-commercial use provided the original author and source are credited.

INTRODUCTION

The two most popular medical imaging techniques are computed tomography (CT) and magnetic resonance imaging (MRI). While CT is the preferred technique for imaging hard tissues like bones and the interfaces between air, bone, and soft tissues, MRI provides excellent contrast images of soft tissues like organs and arteries. Clinical practice frequently uses multi-modality imaging using MRI and CT because of their complementing characteristics. For instance, radiation necessitates both CT and MRI since MRI defines soft tissues and malignancies, while CT offers an electron density distribution that is essential for treatment planning [1]. While CT and MRI scans are

currently done separately, which is not only costly but also causes nonrigid misalignment between MRI and CT pictures, simultaneous CT-MRI is still a study topic [2]. This issue might be resolved by creating a simultaneous CT-MRI device, and we have carried out research to suggest the best possible designs for such a device [3, 4].

One practical solution to the above described issue is medical image synthesis. This method simulates a mapping between a source image and an unidentified target image. Traditional image synthesis techniques concentrate on using a variety of models, including random forest and dictionary learning, to extract expert-defined features [5]. These techniques, however, are only applicable to manually created feature representations.

***Corresponding Author:** Kharananda Sharma

Assistant Professor, Department of Radio-diagnosis and Imaging, Sikkim Manipal Institute of Medical Sciences, Tadong East Sikkim, India

Deep learning has recently demonstrated enormous promise and remarkable success in medical image processing tasks, including artefact reduction [10, 11], super-resolution [8, 9], and denoising [6, 7]. Deep neural networks generate better feature representations and learn features in a data-driven manner than conventional techniques. Recent years have seen the publication of numerous deep learning-based cross-modality medical picture synthesis research, the majority of which are based on generative adversarial networks (GANs) and convolutional neural networks (CNNs) [12-32]. A new generating technique that has garnered a lot of interest in the field of medical imaging is diffusion and score-matching models. High-fidelity, realistic natural images can be produced by these models [33-35]. Diffusion and score-matching models are analytically principled, simple to train, and provide state-of-the-art image quality when compared to other generative model types like GANs and variational auto-encoders, which are challenging to train and interpret and don't always produce satisfactory image quality. Remarkably, a growing body of research indicates that in several image production tasks, diffusion and score-matching models outperform GANs and variational auto-encoders [36].

With a focus on mapping from MRI to CT images, we suggest using diffusion and score-matching models for image conversion between CT and MRI in this article. Our analysis is based on two models: the model that solves the stochastic differential equation (SDE) [35], and the denoising diffusion probabilistic model (DDPM) [34]. These models are contrasted with consistent network architecture CNN and GAN models. By averaging these random samples, we may further quantify the diffusion models' uncertainties from Monte-Carlo sampling results and get better results.

Synthesis of Deep Medical Images

Researchers used deep learning for medical imaging problems after being inspired by its success in the computational vision area. CNN and GAN models were suggested for medical picture synthesis. Nie *et al.*, [12], suggested a 3D fully convolutional network to create pelvic CT images from the matching MRI pictures in order to convert images between MRI and CT modalities. A cascaded GAN model with several consecutive generators and discriminators was then suggested [16]. A CNN model based on an encoder-decoder backbone was created by Bahrami *et al.*, [27], who also noted that the suggested model demonstrated a rapid rate of convergence with a small number of training subjects. Leynes *et al.*, [14], and Han *et al.*, [13], both used UNet to create synthetic CT images. Tao *et al.*, [30], and Emami *et al.*, [17], developed GAN models using ResNet. A conditional GAN was constructed by Boni *et al.*, [28] using multi-center pelvic datasets to produce

synthetic CT images. Additionally, CycleGAN was used to create synthetic MRI and CT images by Chartsias *et al.*, [15], Hiasa *et al.*, [18], Zhang *et al.*, [19], and Cai *et al.*, [25]. When converting brain CT scans into MRI pictures, Li *et al.*, [26], evaluated the effectiveness of the UNet, cycleGAN, and pix2pix models and discovered that UNet performed the best out of the three.

A conditional GAN with a fully convolutional network for liver PET image synthesis was proposed by Ben-Cohen *et al.*, [20], for image synthesis between CT and PET modalities. In order to minimise a complex objective function, Armanious *et al.*, [21], constructed a model named MedGAN using cascaded encoder-decoders. A multi-channel GAN model that can encode semantic information was created by Bi *et al.*, [22]. Ben-Cohen *et al.*, [29], improved lesion identification by creating synthetic PET images using a GAN model with a fully connected network. Wei *et al.*, [24], suggested a sketcher-refiner approach using two cascaded GANs for image synthesis between MRI and PET modalities. The initial GAN produces rough artificial images. The results are refined by the second GAN. Choi *et al.*, [23], used UNet as the generator to create a GAN model for MRI image synthesis. A model named BPGAN was introduced by Zhang *et al.*, [32], to create synthetic brain PET images. The bidirectional mapping generative adversarial network (BMGAN), a 3D end-to-end synthesis model created by Hu *et al.*, [31], jointly optimised the latent vector and picture context for brain MRI-to-PET image synthesis.

Models of Diffusion and Score-Matching

With remarkable generative capabilities for a variety of tasks, including image generation, super-resolution, and picture in-painting, diffusion and score-matching models are emerging as the most promising deep generative models [36]. A forward stage is typically used to add noise gradually, while a reverse stage is used to gradually denoise and recover the original sample. Stochastic differential equations (SDE) [35], noise conditioned score networks (NCSN) [33], and denoising diffusion probabilistic models (DDPM) [34], are currently representative frameworks in this category of picture generating techniques.

Denoising diffusion probabilistic models: DDPM includes several minor phases in its diffusion stage. Gaussian noise marginally taints a data sample, such an image, at each stage. We have $x_0 \sim q(x_0)$ if x_0 is an original image and $q(x_0)$ is the original distribution of x_0 . Following each diffusion phase, a series of progressively corrupted images $x_1, x_2, \dots, x_T$ can be calculated using the Markovian process as follows:

$$q(x_t|x_{t-1}) = \mathcal{N}(x_t; \sqrt{1 - \beta_t} \cdot x_{t-1}, \beta_t \cdot \mathbf{I}), \quad (1)$$

$$q(x_{1:T}|x_0) = \prod_{t=1}^T q(x_t|x_{t-1}), \quad (2)$$

$$q(x_t|x_0) = \mathcal{N}(x_t; \sqrt{\bar{\alpha}_t} \cdot x_0, (1 - \bar{\alpha}_t) \cdot \mathbf{I}). \quad (3)$$

Where $\mathcal{N}(x; \mu, \sigma)$ denotes a Gaussian distribution with mean μ and covariance σ , T is the total number of noising steps, and $\beta_t \in (0,1)$ is a hyper-parameter regulating the variance of incremental Gaussian noise. By parametrising $\alpha_t = 1 - \beta_t$ and $\bar{\alpha}_t = \sum_{i=1}^t \alpha_i$, we obtain x_T turns into an isotropic Gaussian distribution when $T \rightarrow \infty$. To recover an original image,

DDPM completes a denoising process in its reverse stage [37], states that if β_t is small, then every step $q(x_{t-1}|x_t)$ is also a Gaussian distribution. Next, we may estimate the mean $\mu_\theta(x_t, t)$ and the covariance $\Sigma_\theta(x_t, t)$ by training a neural network p_θ to mimic each reverse diffusion step:

$$p_\theta(x_{0:T}) = p(x_T) \prod_{t=1}^T p_\theta(x_{t-1}|x_t), \quad (4)$$

$$p_\theta(x_{t-1}|x_t) = \mathcal{N}(x_{t-1}; \mu_\theta(x_t, t), \Sigma_\theta(x_t, t)), \quad (5)$$

$$q(x_{t-1}|x_t, x_0) = \mathcal{N}(x_{t-1}; \tilde{\mu}(x_t, x_0), \tilde{\beta}_t \cdot \mathbf{I}), \quad (6)$$

$$\tilde{\mu}(x_t, x_0) = \frac{\sqrt{\bar{\alpha}_{t-1}}\beta_t}{1 - \bar{\alpha}_t} x_0 + \frac{\sqrt{\alpha_t}(1 - \bar{\alpha}_{t-1})}{1 - \bar{\alpha}_t} x_t, \quad (7)$$

$$\tilde{\beta}_t = \frac{1 - \bar{\alpha}_{t-1}}{1 - \bar{\alpha}_t} \beta_t. \quad (8)$$

Given that $x_0 = \frac{1}{\sqrt{\bar{\alpha}_t}}(x_t - \sqrt{1 - \bar{\alpha}_t}\epsilon_t)$, where $\epsilon_t \sim \mathcal{N}(0, \mathbf{I})$. Equation (7) can be rewritten as

$$\begin{aligned} \mu_\theta(x_t, t) &= \tilde{\mu}(x_t, \frac{1}{\sqrt{\bar{\alpha}_t}}(x_t - \sqrt{1 - \bar{\alpha}_t}\epsilon_\theta(x_t))) \\ &= \frac{1}{\sqrt{\alpha_t}}(x_t - \frac{1 - \alpha_t}{\sqrt{1 - \bar{\alpha}_t}}\epsilon_\theta(x_t, t)), \end{aligned} \quad (9)$$

Where the density function of x_T is denoted by $p(x_T)$. The reverse step is tractable conditioned on x_t and x_0 , per [34]:

$$\mathcal{L}_{VLB} = \mathcal{L}_0 + \sum_{t=1}^{T-1} \mathcal{L}_t + \mathcal{L}_T, \quad (10)$$

$$\mathcal{L}_0 = -\log p_\theta(x_0|x_1), \quad (11)$$

$$\mathcal{L}_t = KL(q(x_{t-1}|x_t, x_0)||p_\theta(x_{t-1}|x_t)), \quad (12)$$

$$\mathcal{L}_T = KL(q(x_T|x_0)||p_\theta(x_T)), \quad (13)$$

Optimising the variational lower bound (VLB) is the goal of training the noise estimation network ϵ_θ (added noise ϵ_t in x_t):

Where the Kullback-Leibler divergence between two probability distributions is indicated by the

letter KL. Because $q(x_T|x_0)$ has no learnable parameters and x_T is a Gaussian noise, $\mathcal{L}_T$ is constant and may be disregarded.

Table 1: Both sampling & Training Method of our Conditional DDPM Proposal

Algorithm 1 Training	Algorithm 2 Sampling
1: repeat 2: $(x_0^i, y^i) \sim p(x, y)$ 3: $t \sim \mathcal{U}(\{1, \dots, T\})$ 4: $\epsilon \sim \mathcal{N}(0, I)$ 5: Take gradient descent step on $\nabla_{\theta} \ \epsilon_t - \epsilon_{\theta}(y^i, \sqrt{\bar{\alpha}_t}x_0^i + \sqrt{1 - \bar{\alpha}_t}\epsilon_t, t)\ ^2$ 6: until converged	Require: N: Number of steps 1: $x_T \sim \mathcal{N}(0, \mathbf{I})$ 2: for $i = N, \dots, 1$ do 3: $z \sim \mathcal{N}(0, \mathbf{I})$ if $i > 1$, else $z = 0$ 4: $x_{i-1} = \frac{1}{\sqrt{\alpha_i}}(x_i - \frac{1-\alpha_i}{\sqrt{1-\alpha_i}}\epsilon_{\theta}(y, x_i, t)) + \sigma_i z$ 5: end for 6: return x_0

L0 is calculated from $N(x_0; \mu\theta(x_1, 1), \Sigma\theta(x_1, 1))$ in [34], and the loss term in (12) can be simplified and reparameterized as:

$$x_t = x_{t-1} + \frac{\gamma}{2} \nabla_x \log p(x) + \sqrt{\gamma} \cdot \omega_t, \forall t \in \{1, \dots, T\}, \quad (16)$$

$$\mathcal{L}_t^{simple} = \mathbb{E}_{t \sim [1, T], x_0, \epsilon_t} [\|\epsilon_t - \epsilon_{\theta}(\sqrt{\bar{\alpha}_t}x_0 + \sqrt{1 - \bar{\alpha}_t}\epsilon_t, t)\|^2]. \quad (14)$$

The final simplified objective function is

$$\mathcal{L}_{simple} = \mathcal{L}_t^{simple} + C, \quad (15)$$

Where C is a constant that is unaffected by the parameter vector θ . 2) Noise conditioned score network: Langevin dynamics uses only the score function $\nabla_x \log p(x)$ to generate samples from a probability density function $p(x)$. With π serving as a prior distribution and a fixed step size $\gamma > 0$, the sampling procedure utilising the Langevin technique can be written as follows: where C is a constant that is unaffected by the parameter vector

θ . 2) Noise conditioned score network: Langevin dynamics uses only the score function $\nabla_x \log p(x)$ to generate samples from a probability density function $p(x)$. With π serving as a prior distribution and a fixed step size $\gamma > 0$, the sampling procedure utilising the Langevin technique can be written as follows:

$$\mathcal{L}_{sm} = \mathbb{E}_{x \sim p(x)} \|s_{\theta}(x) - \nabla_x \log p(x)\|_2^2. \quad (17)$$

Where ω_t is a member of $N(0, I)$. When T approaches ∞ and $\gamma \rightarrow 0$, the distribution of $x_T = p(x)$. To estimate the score, a neural network s_{θ} trained so that $s_{\theta}(x, t) = \nabla_x \log p(x)$. Ideally, score matching based on the following goal function can be used to train the network.

However, because it is difficult to determine the score $\nabla_x \log p(x)$, equation (17) is difficult to optimise. In

order to get around this problem, Song *et al.*, [33], suggested using Gaussian noises at various scales to alter the original data distribution: $\sigma_1 < \sigma_2 < \dots < \sigma_T$ so that $p_{\sigma_1}(x) \approx p(x_0)$ and $p_{\sigma_T}(x) \approx N(0, I)$. A score estimate NCSN is then trained: $s_{\theta}(x, \sigma_t) \approx \nabla_x \log p_{\sigma_t}(x), \forall t \in \{1, \dots, T\}$. Next, we have :

$$\nabla_{x_t} \log p_{\sigma_t}(x_t|x) = -\frac{x_t - x}{\sigma_t}. \quad (18)$$

Combining (17) and (18) on all $(\sigma_t)_{t=1}^T$, we have

$$\begin{aligned} \mathcal{L}_{dsm} = & \frac{1}{T} \sum_{t=1}^T \lambda(\sigma_t) \mathbb{E}_{p(x)} \mathbb{E}_{x_t \sim p_{\sigma_t}(x_t|x)} \|s_{\theta}(x_t, \sigma_t) \\ & + \frac{x_t - x}{\sigma_t}\|_2^2, \end{aligned} \quad (19)$$

Where the weighting factor is $\lambda(\sigma_t)$.

The Stochastic Differential Equation

(SDE) framework gradually converts the initial data distribution into a Gaussian distribution in the forward stage, much like DDPM and NCSN. The SDE approach deals with a continuous process, in contrast to

the other two approaches that divide the diffusion process into numerous discrete steps. Thus, the SDE approach can be considered as a generalisation of DDPM and NCSN methods. Let us use $p_t(x)$ for the probability density function of $x(t)$, and $p_{st}(x(t)|x(s))$ for the transition kernel from $x(s)$ to $x(t)$, where $0 \leq s < t$:

$$dx = f(x, t) \cdot dt + g(t) \cdot d\omega, \quad (20)$$

Where f and g are the drift and diffusion coefficients, respectively, and $t \sim U([0, T])$, where ω is the

Brownian motion. We also have the related reverse-time SDE:

$$dx = [f(x, t) + g(t)^2 \cdot \nabla_x \log p_t(x)]dt + g(t)d\hat{\omega}, \quad (21)$$

Where $t \sim U([T, 0])$, $\hat{\omega}$ denotes the Brownian motion when time is reversed, and $\nabla_x \log p_t(x)$ is the scoring function of the data distribution $p_t(x)$. In order

for $s_{\theta}(x, t) = \nabla_x \log p_t(x)$, a neural network s_{θ} is trained to estimate the score. The continuous form of (19) is the objective function, which may be written as:

$$\begin{aligned} \mathcal{L}_{dsm}^* = & \mathbb{E}_t [\lambda(t) \mathbb{E}_{x(0)} \mathbb{E}_{x(t)|x(0)} \|s_{\theta}(x_t, t) \\ & - \nabla_{x(t)} \log p_{0t}(x(t)|x(0))\|_2^2], \end{aligned} \quad (22)$$

where $x(0) \sim p_0(x)$, $t \sim U([0, T])$, $x(t) \sim p_{0t}(x(t)|x(0))$, and $\lambda(t)$ is a positive weighting function. In [22], the original score is substituted by $\nabla_x(t) \log p_{0t}(x(t)|x(0))$, as suggested in [38].

METHODOLOGY

Data

This study used co-registered T2w MRI and CT image pairings from 19 individuals from the Gold Atlas male pelvic dataset [39]. Three distinct departments were used to get the data. CT scans with pixel sizes ranging from 0.98mm×0.98mm to 1mm×1mm were acquired using a Siemens Somatom Definition AS+ scanner, a Toshiba Aquilion scanner, and a Siemens Emotion 6 scanner. A Siemens scanner with the TSE sequence, a GE Discovery 750w scanner with the FRFSE sequence, and a GE Signa PET/MR scanner with the FRFSE sequence were used to scan T2w MR

images. The pixel sizes ranged from 0.875mm × 0.875mm to 1.1mm × 1.1mm. Two patients with 135 image pairs were chosen at random for testing, and 17 patients with 1,416 image pairs were chosen for training. This study's image matrices are all 512. Every image was pre-processed for pixel intensity unification prior to training.

DDPM Conditioned

Here, we extend the conditional DDPM proposed by Saharia *et al.*, [40], to map across different imaging modes (CT and MRI) in order to implement image synthesis between CT and MRI. This allows us to construct diffusion and score-matching models conditioned on T2w images, rather than working within the same imaging mode (photographs). Our objective function of (14) is as follows, given the co-registered CT and T2w MRI pairs (x_i, y_i) K $i=1$, where K is the number of image pairs in the dataset:

$$\mathcal{L}_t^{simple} = \mathbb{E}_{t \sim [1, T], x_0, \epsilon_t} [\|\epsilon_t - \epsilon_\theta(\sqrt{\bar{\alpha}_t} x_0^i + \sqrt{1 - \bar{\alpha}_t} \epsilon_t, y, t)\|^2]. \quad (23)$$

Beginning with a Gaussian noise $x_T \sim N(0, I)$, the sampling procedure is a reverse Markovian process. The reverse procedure of (6) and (9) can be changed as :

$$q(x_{t-1} | x_t, x_0, y) = \mathcal{N}(x_{t-1}; \tilde{\mu}(x_t, x_0, y), \tilde{\beta}_t \cdot \mathbf{I}), \quad (24)$$

$$\tilde{\mu}_\theta(x_t, y, t) = \frac{1}{\sqrt{\alpha_t}}(x_t - \frac{1 - \alpha_t}{\sqrt{1 - \bar{\alpha}_t}} \epsilon_\theta(x_t, y, t)). \quad (25)$$

Table 1 I contains a list of the updated DDPM training and sample protocols. For the modified DDPM, there were 1,000 diffusion steps and 1,000 sampling steps. In order to denoise the reverse diffusion process, UNet [41], was used.

SDE Conditioned

Under the direction of a condition of interest, the reverse-time SDE should be solved in order to get the necessary sampling data. Classifier-free and classifier-guidance methods are two ways to enforce a requirement, either softly or aggressively. According to Song *et al.*, [35], the classifier-free technique conducts network training in a supervised fashion and includes a condition in the diffusion model training procedure. The classifier-guidance strategy is unsupervised, in contrast to the classifier-free approach. Using a proximal

optimisation step based on a physical measurement model for medical imaging, such as the Radon and Fourier transforms, Song *et al.*, [42], trained a network for unconditional score estimation before including conditional information into the sampling procedure. In order to control the reverse diffusion process of a pre-trained unconditional diffusion model, Dharwal *et al.*, [43], and Liu *et al.*, [44], added a classifier and utilized its gradients. T2w MRI images are used as the training condition in this investigation. Stated differently, we oversee both the forward and backward diffusion methods. Specifically, we adopted the variance exploding (VE) SDE setup described in the study [35], with $f=0$ and $g=\sigma$. Given the original expression of $g(t) = d[\sigma^2(t)]/dt$ and $\sigma(0) = 0$, we have $\sigma(t) = \sigma_0 e^{\int_0^t g(s) ds}$. Therefore, equation (20) can be rewritten as follows:

$$dx = \sigma^t \cdot d\omega, \quad (26)$$

$$p_{0t}(x(t)|x(0)) = \mathcal{N}(x(t); x(0), \sigma(t) \cdot \mathbf{I}), \quad (27)$$

$$\mathcal{L}_{dsm}^* = \mathbb{E}_t [\lambda(t) \mathbb{E}_{x(0)} \mathbb{E}_{(x(t)|x(0))} \|s_\theta(x(t), y, t) + \frac{x(t) - x(0)}{\sigma(t)}\|_2^2]. \quad (28)$$

The reverse-time SDE of (21) can be expressed as

$$dx = -\sigma^{2t} s_\theta(x(t), y, t) dt + \sigma^t d\bar{\omega}. \quad (29)$$

where $t \sim U([0, T])$. Since the perturbation kernel is Gaussian, its gradient is $p_{0t}(x(t)|x(0))$. The

formula is $\nabla_x p_{0t}(x(t)|x(0)) = -x(t) - x(0) \sigma(t)$. Consequently, the objective function turns into:

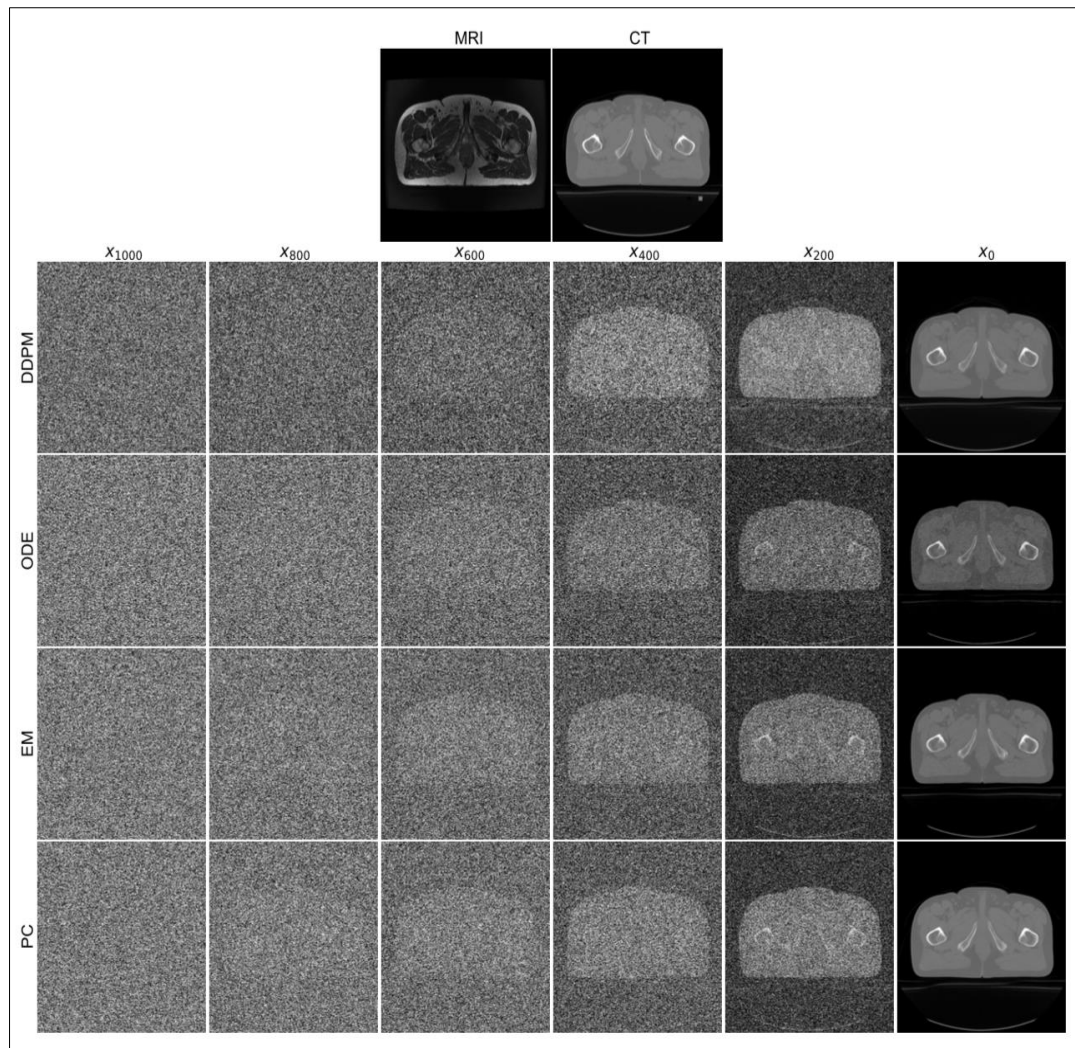


Figure 1: Results of reverse diffusion utilising various sampling techniques are shown in Fig. 1.

CT picture serves as the ground truth and a T2w MR image as the condition in the first row. Results from the DDPM, ODE, EM, and PC sampling methods are shown in the second through fifth rows, respectively.

The reverse-time SDE is solved numerically after a sample x_T is drawn from the prior distribution $p_T \sim \mathcal{N}(x(0), \sigma(T) \cdot \mathbf{I})$ in order to sample from the time dependent score-based model $s_\theta(x(t), y, t)$. We can approximate the prior distribution to $p_T \sim \mathcal{N}(0, \sigma(T) \cdot \mathbf{I})$ when $\sigma(T)$ is big since the mean value of the prior

distribution is near 0. Three sampling methods—the Euler-Maruyama (EM), Prediction-Corrector (PC), and probability flow ordinary differential equation (ODE) methods—were applied in this investigation.

The Euler-Maruyama Technique

A straightforward discretisation technique is used in the EM approach to solve the reverse-time SDE of (29), substituting a Gaussian noise $z \sim \mathcal{N}(0, \Delta t \cdot \mathbf{I})$ for dw and a tiny increment Δt for dt . Next, we have:

$$x_{t-\Delta t} = x_t + \sigma^{2t} s_\theta(x(t), y, t) \Delta t + \sigma^t \sqrt{\Delta t} z_t, \quad (30)$$

where $z_t \sim \mathcal{N}(0, \mathbf{I})$.

The Prediction-Correction Technique

The PC sampling technique switches between the prediction and correction phases in the prediction-correction approach. Any numerical solver for the reverse-time SDE with a set discretisation approach,

such the EM method of (30), can serve as the predictor. Any score-based Markov Chain Monte Carlo technique, including annealed Langevin dynamics, can be used as the corrector. A Langevin step size γ must be determined in order to implement annealed Langevin dynamics:

$$\gamma = \frac{(r\|z\|_2)^2}{\|s_\theta(x_i, y, \sigma_i)\|_2^2}, \quad (31)$$

where $z \sim N(0, I)$ and r are the signal-to-noise ratios. We can sample in accordance with the Langevin dynamics of (16) after determining the Langevin step size γ .

The ODE Approach for Probability Flow

For simplicity's sake, we refer to the probability flow ODE approach as ODE in this study. Any SDE in the form of (20) has a corresponding ODE.

$$dx = [f(x, t) - \frac{1}{2}g(t)^2 \nabla_x \log p_t(x)]dt, \quad (32)$$

Which is identical to the SDE's marginal probability density $p_t(t)$ trajectory. Therefore, calculating the aforementioned ODE in the reverse time direction is identical to sampling by solving the reverse-time SDE. The ODE sampling procedure begins with

deriving x_T from p_T , much like the EM and PC methods do. After that, we Obtain a sample from p_0 by integrating the ODE in the opposite time direction. The ODE equation in this instance is expressed as follows:

$$dx = -\frac{1}{2}\sigma^{2t}s_\theta(x(t), y, t)dt. \quad (33)$$

To solve (33), we employed the Explicit Runge-Kutta technique of order 5(4). Table II contains a list of the three approaches' training and sample protocols. To estimate scores, we employed UNet. We set the total

number of sampling steps to 1,000 for each of the comparative sampling techniques. Specifically, 500 prediction steps and 500 corrective steps were used in the PC sampling.

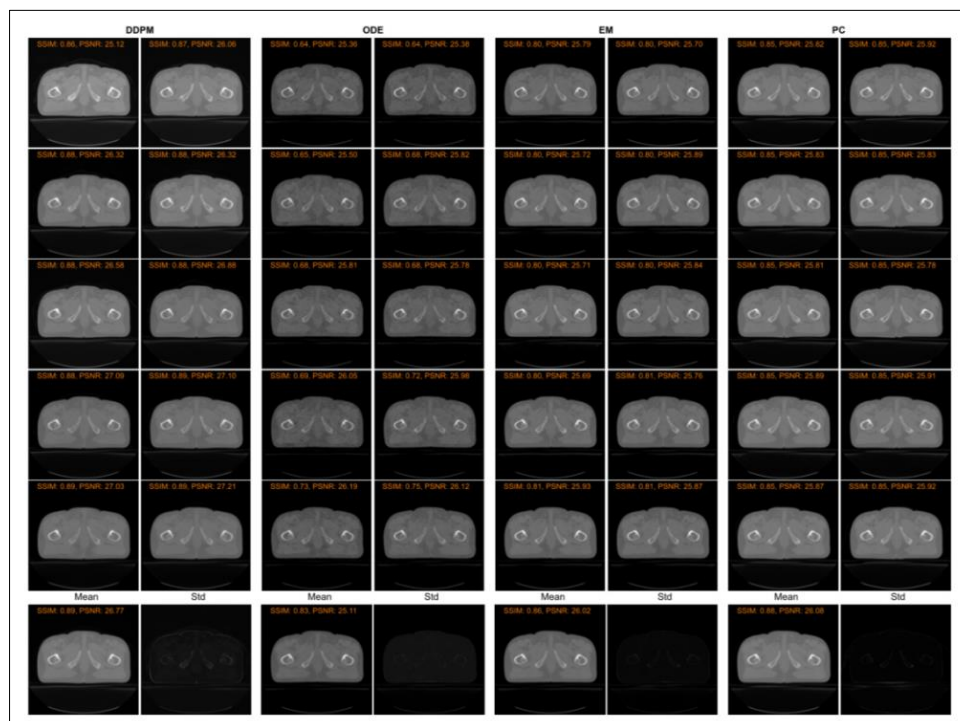


Figure 2: Monte Carlo sampling results are compared using several reverse techniques. The findings conditioned on the same MR picture using four distinct sampling procedures are displayed in the first five rows, correspondingly. The results, together with their pixel-by-pixel standard deviation maps, are averaged across all 10 MC sample results in the bottom row. Interestingly, every image is displayed in the [0, 1] range, with the exception of the standard deviation maps, which are displayed in the [0, 0.5] range.

Additional Techniques for Comparison

Additionally, we contrasted the diffusion and score-matching models with models based on CNN and GAN. UNet was taught to reduce MSE for CNN. With

UNet as the generator and MSE as the data fidelity metric, Wasserstein distance and gradient penalty (WGAN-GP) were added to GAN [45].

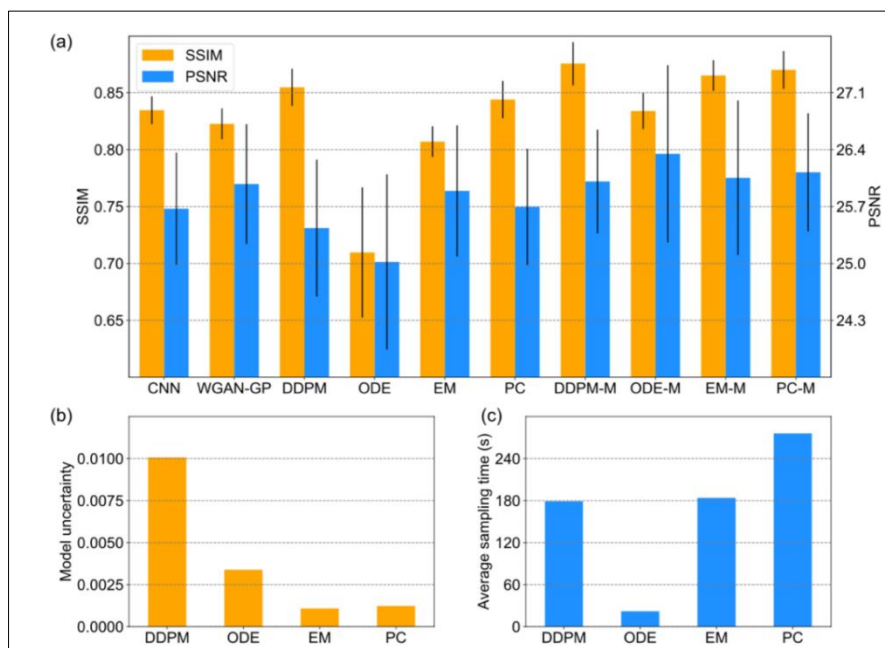


Figure 3: Statistical analysis. (a) The average SSIM deviations; (b) the corresponding average model uncertainty and PSNR scores from several CT image sampling techniques that were generated; and (c) the average utilising various techniques, where the error bars display typical sample times for a slice using various sampling techniques

Details of Implementation

We continuously employed UNet of the same architecture (UNet of CNN and WGAN-GP without time-embedding) and the aAdam optimiser with a learning rate of 10^{-4} , betas of 0.9 and 0.999, and eps of 10^{-8} for all the methods in our experiments, including conditional DDPM, conditional SDE, CNN, and WGAN-GP. After the loss did not drop by 1% in comparison to the average loss for the twenty preceding epochs, the training procedure was terminated after at least 100 epochs. Two was the set batch size. A 24GB Nvidia RTX Titan GPU running PyTorch 1.11 was used for all of the experiments. Following the release of this paper, we will post all of the study's codes to Github. The final outputs of the conditional DDPM and conditional SDE methods were susceptible to random fluctuations because noises were present during the sampling procedures. Therefore, we used the Monte Carlo (MC) approach to further examine the uncertainty of the diffusion and score-matching models. Ten copies of each target CT image were produced. We logged every sampling result in each instance and averaged the ten outcomes to get the MC result. Using the DDPM, ODE, EM, and ODE sampling techniques, we represent these averaged results as DDPM-M, ODE-M, EM-M, and PC-M, respectively. By examining the standard deviation map of the 10 sampling data in each configuration, the model uncertainty was exposed.

RESULTS

In order to assess picture quality and structural similarity index signal-to-noise ratio (PSNR) metrics, we used the measure (SSIM) and peak in this investigation.

Results of Diffusion and Score-Matching

Fig. 1 compares the intermediate outcomes from the reverse process using the DDPM, EM, PC, and ODE methods. It is discovered that all four approaches, conditioned on the corresponding T2w MR picture, ultimately produce a desired CT image ($\times 0$) from a Gaussian noise ($\times 1,000$). Noise is progressively eliminated as the reverse diffusion process continues, giving an image a more realistic appearance.

Estimation of Model Uncertainty

MC results using various sampling techniques are displayed in Fig. 2. It seems that DDPM generates outcomes with the highest SSIM and PSNR scores whereas the ODE approach produces results with the lowest scores. For those results obtained by averaging all ten MC samples conditioned on the same T2w MR image, ODE, EM, and PC techniques yield significantly better synthetic CT images with higher SSIM and PSNR scores than the equivalent individual result. The PC and EM approaches produce smaller standard deviations than the DDPM and ODE approaches in terms of the standard deviation map. Model uncertainty scores for the

sampling techniques are displayed quantitatively in Fig. 3(b). These scores were calculated by averaging over a standard deviation map, and they reveal that EM and PC

had lower model uncertainty scores than DDPM and ODE.

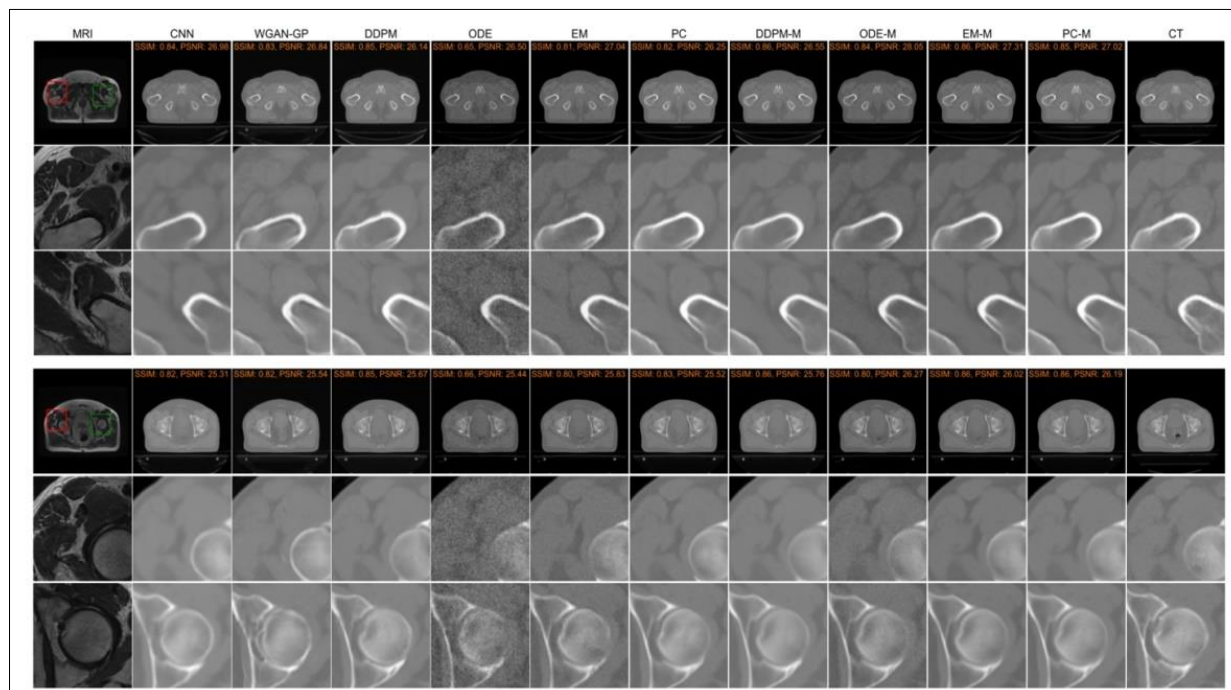


Figure 4: Two image synthesis results employing various techniques are compared. The entire image is displayed in the first row for each example, while the zoomed-in areas, delineated by red and green boxes, respectively, are displayed in the second and third rows.

Comparing GAN and CNN

We contrasted CNN and WGAN-GP based models on image synthesis between CT and MRI with diffusion and score-matching models. Figure 4 displays the qualitative findings. It is discovered that the CNN approach frequently produces results that are excessively smoothed. The WGAN-GP approach tends to produce artefact's even though it produces findings with details. There are significant artefact's on the left side of the spherical bone structure created by WGAN-GP in the bottom row of Figure 4. However, faithful details can be found in the diffusion and score-matching outcomes. Out of the diffusion and score-matching outcomes, the DDPM and PC techniques exhibit higher SSIM scores than the other two sampling methods, while the ODE approach produces less favourable results than the other three sampling methods. In terms of information fidelity, SSIM, and PSNR scores, it is discovered that all four types of averaged results—DDPM-M, ODE-M, EM-M, and PC-M—are superior to the corresponding individual sample results. In terms of SSIM and PSNR, the quantitative results in Fig. 3(a) likewise show that DDPM-M, ODE-M, EM-M, and PC-M scored higher than DDPM, ODE, EM, and PC, respectively. We also looked into each sampling method's inference speed. The time required to create a 512x512 synthetic CT picture using each sampling technique is contrasted in Fig. 3(c). Out of the four approaches, the ODE method is the

fastest. While the PC approach is the slowest, the sample times for DDPM and EM are similar.

DISCUSSION AND CONCLUSION

The training and sample procedures in this work use DDPM, which is time-discrete. SDE, on the other hand, is time-continuous during training and time-discrete during sampling. We looked into the DDPM and SDE techniques for creating fake CT images from provided T2w MRI data. Four distinct sample techniques—three SDE-based and one DDPM-based—were contrasted. Our findings show that all four sampling techniques are capable of producing realistic CT pictures and eliminating noise. Excellent results can be obtained for all four approaches after averaging the results of several Monte Carlo samplings. The ODE approach yields lower-quality sample results, whereas the other three approaches yield results that are similarly good. But when it comes to sampling speed, the PC technique is the slowest and the ODE method is noticeably faster than the others. In practice, sampling speed and quality must be balanced according to the needs of each application. We would suggest the EM approach as a solid option for SDE-based sampling due to its good sampling quality and comparatively quick sampling speed. We conclude that the time-discrete training may be a significant contributing factor to the DDPM model's uncertainty when examining the model uncertainty in Fig. 3(b) and the standard deviation maps

in Fig. 2, which show that DDPM has a higher model uncertainty than SDE-based sampling techniques.

We also assume that DDPM would behave more like a time-continuous model with less model uncertainty if the number of steps in the diffusion and reversal processes were increased. We have contrasted the diffusion and score-matching outcomes with those based on CNN and GAN. Because MSE is used as the objective function, CNN findings are more likely to be over-smoothed than other results [46]. Although GAN results are more detailed than CNN results, they are tainted by artefact's. These artefact's could result from the GAN model's low robustness and strong propensity for hallucinations. In contrast to CNN and GAN models, diffusion and score-matching models have the fundamental capacity to produce high-quality images and fit data distributions.

Nevertheless, diffusion and score-matching models have a well-known drawback: their reliance on a lengthy Markov chain for generation leads to a comparatively slow pace. We will continue investigating computational methods to greatly improve the sample procedures as a future course. A few methods have already been devised for that purpose [47–51], although the speed of diffusion sampling still lags well behind CNN and GAN inference. The unsupervised technique is more beneficial in medical imaging applications than the supervised approach since it does not require paired pictures. Nevertheless, there are drawbacks to the current unsupervised diffusion and score-matching models. The limitations in Dharwal and Liu's approaches [43, 44], are insufficient to access particular structures or contents in the images that are produced. However, Song's technique [42], has too strict of requirements and is difficult to use for CT-MRI synthesis because of the low degree of similarity between CT and MRI data. We will investigate ways to balance these conditional constraints for high-quality CT-MRI synthesis in subsequent research.

Finally, for image synthesis between CT and MRI, we have modified the newly developed diffusion and score-matching models. To sample CT pictures conditioned on an MRI image, the four strategies—DDPM, ODE, EM, and PC—have been implemented. Results produced with traditional CNN and GAN models have been favourably compared with the resulting CT images utilising various sampling techniques. Additionally, the diffusion and score-matching models' uncertainties have been measured. On the basis of this work, additional research is underway.

REFERENCES

1. M. Owringi, P. B. Greer, and C. K. Glide-Hurst, "MRI-only treatment planning: benefits and challenges," *Phys. Med. Biol.*, vol. 63, p. 05TR01, Feb. 2018. doi: 10.1088/1361-6560/aaaca4.
2. V. Kearney, B. P. Ziemer, A. Perry, T. Wang, J. W. Chan, L. Ma, O. Morin, S. S. Yom, and T. D. Solberg, "Attention-aware discrimination for MR-to-CT image translation using cycle-consistent generative adversarial networks," *Radiol. Artif. Intell.*, vol. 2, Mar. 2020. doi:10.1148/ryai.2020190027.
3. G. Wang, M. Kalra, V. Murugan, Y. Xi, L. Gjesteb, M. Getzin, *et al.*, Vision 20/20: Simultaneous CT-MRI—next chapter of multimodality imaging, *Med. Phys.*, vol. 42, no. 10, pp. 5879–5889, 2015. doi: 10.1118/1.4929559.
4. Y. Peng, M. Li, J. Grandinetti, G. Wang, and X. Jia, "Top-level design and simulated performance of the first portable CT-MR scanner," *arXiv preprint*, 2022. arXiv:2203.15989.
5. Yu, Y. Wang, L. Wang, D. Shen, and L. Zhou, "Medical image synthesis via deep learning," in *Deep learning in medical image analysis: challenges and applications*, pp. 23–44, Springer, Cham, 2020. doi:10.1007/978-3-030-33128-3.
6. Q. Yang, P. Yan, Y. Zhang, H. Yu, Y. Shi, X. Mou, *et al.*, "Low-dose CT image denoising using a generative adversarial network with Wasserstein distance and perceptual loss," *IEEE Trans. Med. Imaging*, vol. 37, pp. 1348–1357, Jun. 2018. doi: 10.1109/TMI.2018.2827462.
7. H. Chen, Y. Zhang, M. K. Kalra, F. Lin, Y. Chen, P. Liao, *et al.*, "Low-dose CT with a residual encoder-decoder convolutional neural network," *IEEE Trans. Med. Imaging*, vol. 36, pp. 2524–2535, Jun. 2017. doi: 10.1109/TMI.2017.2715284.
8. Q. Lyu, H. Shan, and G. Wang, "MRI super-resolution with ensemble learning and complementary priors," *IEEE Trans. Comput. Imaging*, vol. 6, pp. 615–624, Jan. 2020. doi: 10.1109/TCI.2020.2964201.
9. K. Hammernik, T. Klatzer, E. Kobler, M. P. Recht, D. K. Sodickson, T. Pock, *et al.*, "Learning a variational network for reconstruction of accelerated MRI data," *Magn. Reson. Med.*, vol. 79, pp. 3055–3071, Jun. 2018. doi: 10.1002/mrm.26977.
10. Y. Zhang and H. Yu, "Convolutional neural network based metal artifact reduction in X-ray computed tomography," *IEEE Trans. Med. Imaging*, vol. 37, pp. 1370–1381, Jun. 2018. doi: 10.1109/TMI.2018.2823083.
11. L. Gjesteb, Q. Yang, Y. Xi, H. Shan, B. Claus, Y. Jin, *et al.*, "Deep learning methods for CT image-domain metal artifact reduction," in *Proc. SPIE 2017 Developments in X-ray Tomography XI*, vol. 10391, pp. 147–152, Sep. 2017. doi: 10.1117/12.2274427.
12. Nie, X. Cao, Y. Gao, L. Wang, and D. Shen, "Estimating CT image from MRI data using 3D fully convolutional networks," in *Deep Learning and Data Labeling for Medical Applications. DLMIA LABELS 2016*, pp. 170–178, Springer, Cham, Sep. 2016. doi: 10.1007/978-3-319-46976-8_18.

13. X. Han, "MR-based synthetic CT generation using a deep convolutional neural network method," *Med. Phys.*, vol. 44, pp. 1408–1419, Feb. 2017. doi: 10.1002/mp.12155.
14. A. P. Leynes, J. Yang, F. Wiesinger, S. S. Kaushik, D. D. Shanbhag, Y. Seo, *et al.*, "Direct pseudoCT generation for pelvis PET/MRI attenuation correction using deep convolutional neural networks with multi-parametric MRI: zero echo-time and dixon deep pseudoCT (ZeDD-CT)," *J. Nucl. Med.*, Oct. 2017. doi: 10.2967/jnumed.117.198051.
15. Chartsias, T. Joyce, R. Dharmakumar, and S. A. Tsiftaris, "Adversarial image synthesis for unpaired multi-modal cardiac data," in *Int. workshop on simulation and synthesis in medical imaging. SASHIMI 2017*, pp. 3–13, Springer, Cham, Sep. 2017. doi: 10.1007/978-3-319-68127-61.
16. "Medical image synthesis with deep convolutional adversarial networks," *IEEE. Trans. Biomed. Eng.*, vol. 65, pp. 2720–2730, Dec. 2018, doi: 10.1109/TBME.2018.2814538, D. Nie, R. Trullo, J. Lian, L. Wang, C. Petitjean, S. Ruan, *et al.*
17. "Generating synthetic CTs from magnetic resonance images using generative adversarial networks," *Med. Phys.*, vol. 45, pp. 3627–3636, June 2018, doi: 10.1002/mp.13047, H. Emami, M. Dong, S. P. Nejad-Davarani, and C. K. Glide-Hurst.
18. "Cross-modality image synthesis from unpaired data using CycleGAN," by Y. Hiasa, Y. Otake, M. Takao, T. Matsuoka, K. Takashima, A. Carass, *et al.*, at the International Workshop on Simulation and Synthesis in Medical Imaging. Springer, Cham, September 2018, SASHIMI 2018, pp. 31–41, doi: 10.1007/978-3-030-00536-8 4.
19. Z. Zhang, L. Yang, and Y. Zheng, "Translating and segmenting multimodal medical volumes with cycle-and shape-consistency generative adversarial network," in *Computer Vision and Pattern Recognition. CVPR 2018*, pp. 9242–9251, 2018. doi: 10.1109/CVPR.2018.00963.
20. A. Ben-Cohen, E. Klang, S. P. Raskin, M. M. Amitai, and H. Greenspan, Virtual PET images from CT data using deep convolutional networks: initial results, in *Int. workshop on simulation and synthesis in medical imaging. SASHIMI 2017*, pp. 49–57, Springer, Cham, Sep. 2017. doi:10.1007/978-3-319-68127-6 6.
21. K. Armanious, C. Jiang, M. Fischer, T. K. ustner, T. Hepp, K. Nikolaou, *et al.*, "MedGAN: Medical image translation using GANs," *Comput. Med. Imaging Graph.*, vol. 79, p. 101684, Jan. 2020. doi: 10.1016/j.compmedimag.2019.101684.
22. L. Bi, J. Kim, A. Kumar, D. Feng, and M. Fulham, "Synthesis of positron emission tomography (PET) images via multi-channel generative adversarial networks (GANs)," in *Molecular imaging, and analysis of moving body organs, and stroke imaging and treatment. RAMBO CMMI SWITCH 2017*, pp. 43–51, Springer, Cham, Sep. 2017. doi: 10.1007/978-3-319-67564-0 5.
23. H. Choi and D. S. Lee, "Generation of structural MR images from amyloid PET: application to MR-less quantification," *J. Nucl. Med.*, vol. 59, pp. 1111–1117, Jul. 2018. doi: 10.2967/jnumed.117.199414.
24. W. Wei, E. Poirion, B. Bodini, S. Durrleman, N. Ayache, B. Stankoff, *et al.*, "Learning myelin content in multiple sclerosis from multimodal MRI through adversarial training," in *Medical Image Computing and Computer-Assisted Intervention. MICCAI 2018*, pp. 514–522, Springer, Cham, Sep. 2018. doi: 10.1007/978-3-030-00931-1 59.
25. J. Cai, Z. Zhang, L. Cui, Y. Zheng, and L. Yang, "Towards cross-modal organ translation and segmentation: A cycle-and shape-consistent generative adversarial network," *Med.ImageAnal.*, vol. 52, pp. 174–184, Feb 2019. doi: 10.1016/j.media.2018.12.002.
26. W. Li, Y. Li, W. Qin, X. Liang, J. Xu, J. Xiong, *et al.*, "Magnetic resonance image (MRI) synthesis from brain computed tomography (CT) images based on deep learning methods for magnetic resonance(MR)-guided radiotherapy," *Quant. Imaging Med. Surg.*, vol. 10, pp. 1223–1236, Jun. 2020. doi: 10.21037/qims-19-885.
27. A. Bahrami, A. Karimian, E. Fatemizadeh, H. Arabi, and H. Zaidi, "A new deep convolutional neural network design with efficient learning capability: Application to CT image synthesis from MRI," *Med.Phys.*, vol. 47, pp. 5158–5171, Jul. 2020. doi:10.1002/mp.14418.
28. K. N. B. Boni, J. Klein, L. Vanquin, A. Wagner, T. Lacornerie, D. Pasquier, *et al.*, "MR to CT synthesis with multicenter data in the pelvic area using a conditional generative adversarial network," *Phys. Med. Biol.*, vol. 65, p. 075002, Apr. 2020. doi: 10.1088/1361-6560/ab7633.
29. A. Ben-Cohen, E. Klang, S. P. Raskin, S. Soffer, S. Ben-Haim, E. Konen, *et al.*, "Cross-modality synthesis from CT to PET using FCN and GAN networks for improved automated lesion detection," *Eng.Appl.Artif.Intell.*, vol. 78, pp. 186–194, Feb. 2019. doi: 10.1016/j.engappai.2018.11.013.
30. L. Tao, J. Fisher, E. Anaya, X. Li, and C. S. Levin, "Pseudo CT image synthesis and bone segmentation from MR images using adversarial networks with residual blocks for MR-based attenuation correction of brain PET data," *IEEE Trans. Radiat. Plasma Med. Sci.*, vol. 5, pp. 193–201, Mar. 2020. doi: 10.1109/TRPMS.2020.2989073.
31. S. Hu, B. Lei, S. Wang, Y. Wang, Z. Feng, and Y. Shen, "Bidirectional mapping generative adversarial networks for brain MR to PET synthesis," *IEEE Trans. Med. Imaging*, vol. 41, pp. 145–157, Jan. 2022. doi: 10.1109/TMI.2021.3107013.
32. J. Zhang, X. He, L. Qing, F. Gao, and B. Wang, "BPGAN: Brain PET synthesis from MRI using

- generative adversarial network for multi-modal Alzheimer's disease diagnosis," *Comput. Methods Programs Biomed.*, vol. 217, p. 106676, Apr. 2022. doi: 10.1016/j.cmpb.2022.106676.
33. Y. Song and S. Ermon, "Generative modeling by estimating gradients of the data distribution," in *Advances in Neural Information Processing Systems*. NeurIPS 2019, vol. 32, pp. 11918–11930, Curran Associates, Inc., 2019.
34. J. Ho, A. Jain, and P. Abbeel, "Denoising diffusion probabilistic models," in *Advances in Neural Information Processing Systems*. NeurIPS 2020, vol. 33, pp. 6840–6851, Curran Associates, Inc., 2020.
35. Y. Song, J. Sohl-Dickstein, D. P. Kingma, A. Kumar, S. Ermon, and B. Poole, "Score-based generative modeling through stochastic differential equations," in *International Conference on Learning Representations*, ICLR 2021, OpenReview.net, 2021.
36. F.-A. Croitoru, V. Hondru, R. T. Ionescu, and M. Shah, "Diffusion models in vision: A survey," arXiv preprint, 2022. arXiv:2209.04747.
37. J. Sohl-Dickstein, E. Weiss, N. Maheswaranathan, and S. Ganguli, Deep unsupervised learning using nonequilibrium thermodynamics, in *International Conference on Machine Learning*. ICML 2015, vol. 37, pp. 2256–2265, PMLR, 2015.
38. P. Vincent, "A connection between score matching and denoising autoencoders," *Neural Comput.*, vol. 23, pp. 1661–1674, Jul. 2011. doi:10.1162/NECO.2011.00142.
39. T. Nyholm, S. Svensson, S. Andersson, J. Jonsson, M. Sohlén, C. Gustafsson, *et al.*, "MR and CT data with multiobserver delineations of organs in the pelvic area — part of the Gold Atlas project," *Med. Phys.*, vol. 45, pp. 1295–1300, Jan. 2018. doi: 10.1002/mp.12748.
40. Saharia, J. Ho, W. Chan, T. Salimans, D. J. Fleet, and M. Norouzi, Image super-resolution via iterative refinement," *IEEE Trans. Pattern Anal. Mach. Intell.*, 2022. doi: 10.1109/TPAMI.2022.3204461.
41. O. Ronneberger, P. Fischer, and T. Brox, "U-Net: Convolutional networks for biomedical image segmentation," in *Medical Image Computing and Computer-Assisted Intervention*. MICCAI 2015, pp. 234–241, Springer, Cham, 2015. doi: 10.1007/978-3-319-24574-4_28.
42. Y. Song, L. Shen, L. Xing, and S. Ermon, "Solving inverse problems in medical imaging with score-based generative models," in *International Conference on Learning Representations*, ICLR2022, OpenReview.net, 2022.
43. P. Dhariwal and A. Q. Nichol, "Diffusion models beat GANs on image synthesis," in *Advances in Neural Information Processing Systems*. NeurIPS 2021, vol. 34, pp. 8780–8794, Curran Associates, Inc., 2021.
44. X. Liu, D. H. Park, S. Azadi, G. Zhang, A. Chopikyan, Y. Hu, H. Shi, A. Rohrbach, and T. Darrell, "More control for free! Image synthesis with semantic diffusion guidance," arXiv preprint, 2021. arXiv:2112.05744.
45. Gulrajani, F. Ahmed, M. Arjovsky, V. Dumoulin, and A. Courville, Improved training of Wasserstein GANs, in *Advances in Neural Information Processing Systems*. NeurIPS 2017, vol. 30, pp. 5769–5779, Curran Associates, Inc., 2017.
46. Johnson, A. Alahi, and L. Fei-Fei, "Perceptual losses for real time style transfer and super-resolution," in *European Conference on Computer Vision*. ECCV 2016, pp. 694–711, Springer, Cham, 2016. doi: 10.1007/978-3-319-46475-6_43.
47. Song, C. Meng, and S. Ermon, "Denoising diffusion implicit models," in *International Conference on Learning Representations*. ICLR 2021, OpenReview.net, 2021.
48. Lu, Y. Zhou, F. Bao, J. Chen, C. Li, and J. Zhu, "DPM-Solver: A fast ODE solver for diffusion probabilistic model sampling in around 10 steps," arXiv preprint, 2022. arXiv:2206.00927.
49. H. Ma, L. Zhang, X. Zhu, and J. Feng, "Accelerating score-based generative models with preconditioned diffusion sampling," arXiv preprint, 2022. arXiv:2207.02196.
50. Z. Lyu, X. Xu, C. Yang, D. Lin, and B. Dai, "Accelerating diffusion models via early stop of the diffusion process," arXiv preprint, 2022. arXiv:2205.12524.
51. H. Chung, B. Sim, and J. C. Ye, "Come-closer-diffuse-faster: Accelerating conditional diffusion models for inverse problems through stochastic contraction," in *Computer Vision and Pattern Recognition*. CVPR 2022, pp. 12413–12422, 2022.

Cite This Article: Kharananda Sharma, Tshetiz Dahal, Bimal Nepal, Jarina Tamang (2025). Image Modality Translation between MRI and CT Using Diffusion and Score-Matching Algorithms. *EAS J Radiol Imaging Technol*, 7(6), 179-191.
