

Supplementary Material

Data-Efficient Alignment in Medical Imaging via Reconfigurable Generative Networks

This document presents the supplementary materials omitted from the main paper due to space limitations.

1 Implementation Details

Hardware and environment. Our work was conducted on an open-source Ubuntu workstation (version 20.04.5), using Python 3.8 and PyTorch as the primary programming framework for deep learning. Experiments were performed on a GPU, NVIDIA RTX 3090, with 24GB of memory.

We use the CycleGAN, RegGAN, NICEGAN, UNIT, and shortest path regularization implementations from <https://github.com/Kid-Liet/Reg-GAN> and <https://github.com/Mid-Push/santa>. For all experiments, we set the initial network width NW_i to 0.1, the final network width NW_f to 0.7, and the growing frequency Δt to 1000. We use NVIDIA RTX3090 GPU (24GB) for all experiments.

Evaluation metrics. We use widely adopted three evaluation metrics as follows: Normalized Mean Absolute Error (NMAE)(\downarrow) which measures the normalized average deviation between true and predicted values. Lower NMAE indicates better data accuracy. Peak Signal-to-Noise Ratio (PSNR)(\uparrow) which compares signal to noise in reconstructed images, vital for high-detail datasets like MRI scans. Structural Similarity Index Measure (SSIM) (\uparrow) it evaluates texture, luminance, and structure similarities in images, essential for preserving details in medical imaging.

Training details. CycleGAN, RegGAN, NICEGAN, UNIT, and shortest path regularization were used in our study. We followed the settings of the experiment used in RegGAN. Specifically, all three models were trained using Adam optimizer with a learning rate $1e-4$ and $(\beta_1, \beta_2) = (0.5, 0.999)$. As the standard of the CycleGAN-like image-to-image translation model, batch size in all experiments was set to 1. Coefficients of \mathcal{L}_{adv} (for CycleGAN), \mathcal{L}_{cyc} (for CycleGAN), \mathcal{L}_{corr} (for registration network in RegGAN), \mathcal{L}_{smooth} (for registration network in RegGAN) were set to 1, 10, 20, 10, respectively. Note that our proposed R-CCL loss shares the same coefficient with \mathcal{L}_{cyc} . For full and 10% training dataset experiments, we train each model for 80/400 epochs, which is over 460/230K iterations while warmup training epoch, gradual growing training epoch, and fine-tuning epoch are set to 5/50, 35/150, 40/200, respectively.

Noise data simulation. For introducing Gaussian and Speckle noise, we use a Gaussian distribution with a mean value of 0 and a standard deviation of 0.1. For Salt and Pepper (S&P) noise, the probability for both salt and pepper disturbances is set at 0.5%.

2 Reconfig-MIT Performance under Various Noise Conditions

We further extend our analysis of Reconfig-MIT by adding different noises to the data. **Noise**, refers to unwanted variations or artifacts in the images. These variations can result from the imaging process itself, patient motion, or differences in imaging equipment. Noise can obscure or distort the relevant features in the images, making it difficult for a model to learn accurate representations and perform precise translations. The simulated noises are as follows:

- 1) Gaussian noise: Add random Gaussian noise to the pixel intensities, simulating sensor noise or image degradation.
- 2) Salt and pepper (S&P) noise: Randomly replace a certain percentage of pixels with either the minimum or maximum intensity value, simulating dead pixels or impulsive noise.
- 3) Speckle noise: Add multiplicative noise to the images, simulating noise common in ultrasound imaging.
- 4) Poisson noise: Modify the pixel intensities according to a Poisson distribution, simulating photon counting noise in imaging devices.

Table 1: Comparative results of CycleGAN (C), RegGAN (C+R), and Reconfig-MIT (Gr+R-CCL) under four different noise conditions, tested on a limited dataset comprising only 10% of the total data.

Models	Metrics	Noise Gaussian	Noise Poisson	Noise S&P	Noise Speckle
C	NMAE↓	0.086	0.087	0.092	0.087
	PSNR↑	24.369	24.280	23.328	24.102
	SSIM↑	0.790	0.820	0.811	0.707
C+R	NMAE↓	0.085	0.084	0.089	0.085
	PSNR↑	24.452	24.427	23.361	24.607
	SSIM↑	0.790	0.827	0.822	0.792
C+R+Gr+R-CCL	NMAE↓	0.085	0.083	0.086	0.083
	PSNR↑	24.532	24.719	24.032	24.653
	SSIM↑	0.798	0.830	0.830	0.793

These results shown in Table 1 underline the robustness of Reconfig-MIT in handling different types of noise, including Gaussian, Poisson, S&P, and Speckle, even in scenarios with limited data. The consistently superior performance of Reconfig-MIT across all metrics affirms its effectiveness and adaptability in medical image-to-image translation tasks. This emphasizes the potential of Reconfig-MIT as a valuable tool in clinical applications, especially where image noise is a significant challenge.

The introduction of S&P noise impacts the entirety of the image. Reconfig-MIT yields the highest NMAE and lowest PSNR performance compared to the other three types of noise. This is primarily due to the extremity of S&P noise values, where *salt* is represented as 1 and *pepper* as -1. These values deviate significantly from the original image patterns, posing a challenge for the model to differentiate them.

3 Performance on Unpaired Setting

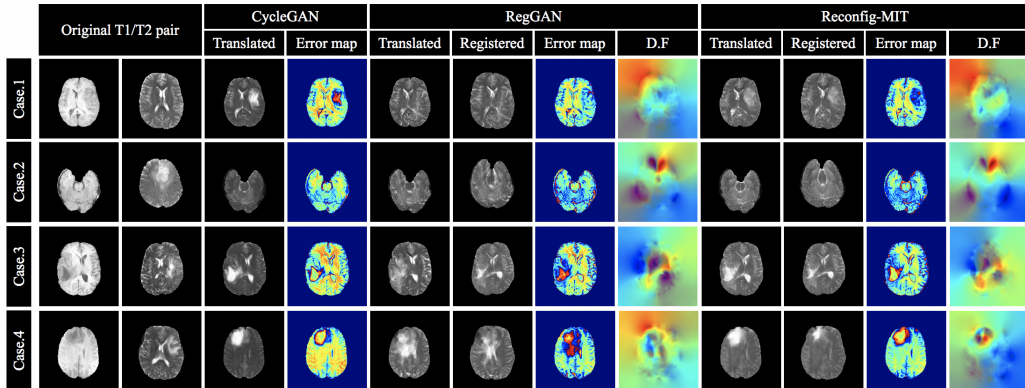


Figure 1: The translation results, error maps, and deformation field (D.F.) of CycleGAN (C), RegGAN (C+R), and Reconfig-MIT (Gr+R-CCL) on unpaired, 10% training dataset with misalignment level M.1.

The section outlines the performance of CycleGAN, RegGAN, and Reconfig-MIT enhanced with R-CCL on a uniquely challenging dataset, an unpaired set consisting of only 10% of the original training data, complicated further with a misalignment level denoted as M.1 (see Figure 1). These circumstances present significant obstacles for these models, namely handling complex image transformations and mitigating discrepancies caused by the M.1 level misalignment, all while working with limited data availability. We carefully examined the translation results, scrutinized error maps, and analyzed the deformation fields (D.F.) generated by each model.

4 Performance with Existing GAN for Image-to-Image Translation

Table 2: Integration of Reconfig-MIT (**Gr** only) with existing GANs architectures, UNIT (**U**), Registration Network (**R**), NICEGAN (**N**) for 10% training data

Misalignment (M) level	Metrics	N	N+R	N+R+Gr	U	U+R	U+R+Gr
M.0	NMAE↓	0.091	0.089	0.088	0.085	0.083	0.081
	PSNR↑	23.884	23.935	24.177	24.265	24.664	24.935
	SSIM↑	0.824	0.825	0.829	0.828	0.836	0.840
M.1	NMAE↓	0.092	0.090	0.088	0.084	0.080	0.079
	PSNR↑	23.638	24.076	24.168	24.001	24.782	25.139
	SSIM↑	0.827	0.828	0.830	0.837	0.844	0.848
M.5	NMAE↓	0.090	0.088	0.087	0.089	0.083	0.081
	PSNR↑	23.241	23.510	23.494	24.147	24.290	24.656
	SSIM↑	0.827	0.830	0.835	0.828	0.839	0.844

The results in Table 2 show the consistent performance of the proposed method (**N+R+Gr** and **U+R+Gr**), including registration and Reconfig-MIT, across all misalignment levels even when the data is reduced to 10%. This illustrates the method’s robustness to limited data scenarios, which is crucial in medical image translation applications where data scarcity is a common challenge.

In all cases, the incorporation of registration and Reconfig-MIT improves the metrics (decrease in NMAE and increase in PSNR and SSIM), demonstrating their effectiveness in preserving image quality and structural similarity despite data limitations and varying levels of misalignment.

An interesting observation is that the proposed method shows better or comparable performance (**U+R+Gr**) when the misalignment level is increased to **M.5**. This suggests the ability of the Reconfig-MIT model to produce reliable translations in more challenging conditions.

Overall, these results reinforce the viability of the proposed Reconfig-MIT model and its adaptability to various misalignment levels and data-limited scenarios in the context of image-to-image translation.

For **N+R+Gr**, we set the initial density of the model to 0.3 and the final density to 0.9.