OXFORD JOURNAL
OF STUDENT
SCHOLARSHIP

A Unified Deep Learning Framework for Magnetic Resonance Imaging Contrast Translation

Rachel Yang
rachel.yangg6@gmail.com

ABSTRACT

Despite advances in medical imaging, important limitations remain, including long acquisition times, patient discomfort, hardware constraints, and high costs, all of which can delay or limit accurate diagnosis. Magnetic Resonance Imaging (MRI) contrast translation, such as generating T1-weighted images from T2-weighted images or vice versa, is being investigated as a way to generate a missing contrast from an acquired contrast. In this work, we developed a unified deep learning-based image-to-image translation framework for bidirectional MRI contrast translation using a U-Net architecture. The model was trained using a combination of L1, gradient, perceptual, and frequency losses to preserve both image contrast and anatomical structures. Paired T1- and T2-weighted brain MRI data were used, with 30 patients for loss-weight selection, 110 patients for final training, and 10 patients held out from the start as an independent test set. Controlled experiments systematically assessed the effects of network design and training strategies on image quality, including structural accuracy and reconstruction fidelity. The proposed method achieved high image quality, with Structural Similarity Index Measure (SSIM) values of up to 0.9594 for T1-to-T2 translation and 0.9931 for T2-to-T1 translation, and Peak Signal-to-Noise Ratio (PSNR) values of 27.17 dB and 28.42 dB, respectively. These results demonstrate the technical feasibility of bidirectional MRI contrast synthesis on healthy-subject brain images. Possible reductions in scan time, cost, or clinical workload were not evaluated here and should be regarded as future applications rather than demonstrated benefits.

INTRODUCTION

Magnetic resonance imaging (MRI) is an important diagnostic tool in healthcare because it can generate multiple tissue contrasts from a single anatomical region, each sensitive to different physical properties of tissue[1]. MRI diagnosis often relies on multiple image contrasts, as each provides complementary information about tissue characteristics and pathology. Among the most commonly used contrasts are T1- and T2-weighted images. T1-weighted images are acquired using short echo times (TE) and repetition times (TR), in which tissues with high fat content appear bright, where fluid-filled regions appear dark[2]. In contrast, T2-weighted images are acquired using long TE and TR, producing the opposite signal pattern, where fat appears dark and fluid appears bright[3].

Since each contrast provides valuable and complementary information, multiple contrasts are often acquired for brain imaging in clinical practice, including T1-weighted imaging, T2-weighted imaging, fluid-attenuated inversion recovery (FLAIR)[4], proton density (PD)[5,6], diffusion-weighted imaging (DWI)[7,8], and susceptibility weighted imaging (SWI)[9]. However, acquiring multiple contrasts is a slow and costly process. A series of MRI sequences must be performed sequentially to obtain contrast-specific

images[10]. Because each contrast requires a unique pulse sequence, acquiring multiple contrasts can be time-consuming, increasing scan time and patient burden. Significant efforts have been devoted to accelerating MRI acquisition and reducing scan time, primarily through k-space undersampling techniques, including parallel imaging[11], compressed sensing[12], and deep learning-based methods[13].

Synthetic MRI represents another approach for generating new MRI contrasts from acquired MRI data. Examples include generating post-contrast T1-weighted images from pre-contrast T1-weighted, T2-weighted, and diffusion-weighted images[14], as well as synthesizing image contrasts in musculoskeletal MRI using mathematical image transformation, physics-based, or data-driven approaches[15]. MRI image contrast translation is a deep learning-based technique that generates a target MR image from one or more source images, reducing the need for multiple image acquisitions. Recent advances in deep learning have demonstrated strong performance in generating high-quality synthetic image contrasts with promising clinical potential[16-18]. On the paired multi-contrast (PMC) dataset, Bagheri and Uludag reported a U-Net baseline for supervised T1/T2 translation[19].

In this study, we propose a deep learning approach based on a convolutional neural network to predict one MRI contrast from another. We develop a unified deep learning framework with an optimized loss function for cross-contrast MRI image translation. The present experiments are limited to image-similarity metrics in healthy subjects. Any future use to shorten scan time, improve patient comfort, or reduce imaging cost would require additional clinical validation and is not claimed as a result of this study.

METHODS

Dataset

We used a publicly available paired multi-contrast MRI dataset (PMC) containing T1-, T2-, and PD-weighted scans acquired at 1.5T and 3T for supervised image-to-image translation[19]. In this study, only the T1- and T2-weighted images acquired at 3T were used. The dataset included 181 patients, each providing paired T1- and T2-weighted 3D volumes. In PMC, images from the same patient are already rigidly registered to a common orientation and skull-stripped by the dataset authors[19]; we did not apply additional registration between paired T1 and T2 slices.

All splitting was performed at the patient level, not the slice level, so that slices from the same patient never appeared in more than one of the training, validation, or test sets. Before any architecture comparison or loss-weight selection, 10 patients were reserved as a locked independent test set. These 10 patients were never used for training, validation, early stopping, architecture choice, or loss-weight selection. Of the remaining 171 patients, 30 were used for the development phase. These 30 patients were divided with an 80/20 patient-level split into 24 training patients and 6 validation patients. After the loss configuration was fixed using only this development set, the final models were trained on 110 patients from the non-test pool (the original 30 development patients plus 80 additional patients), again using an 80/20 patient-level train/validation split solely to adjust the learning rate and apply early stopping. All quantitative results in Table 1 were obtained from the same 10 held-out test patients.

To construct a two-dimensional dataset from the three-dimensional volumes, five central slices were extracted from each patient's scan at consistent anatomical positions: the mid-volume slice and the two adjacent slices on each side, using the same relative slice indices for every patient and both contrasts. Central slices were used to maximize anatomically relevant brain tissue and to avoid superior and inferior slices that contain predominantly background. Generalizability beyond these central axial levels needs further evaluation. The five-slice design was limited by computational resources, and further work will validate whole-brain coverage. Pixel values were converted to tensors scaled to the [0, 1] range, followed by normalization to the [-1, 1] intensity range. All images were resized to 256×256 pixels. Data augmentation was applied only to training patients and included horizontal flipping, rotation, and scaling.

Model Architecture

Cross-contrast translation was performed using a U-Net based convolutional neural network[20]. The generator is a 2D encoder-decoder with eight downsampling stages and eight upsampling stages. The encoder uses 4×4 convolutions with stride 2. Feature channels increase as 64, 128, 256, and 512, and then remain at 512 through the later encoder stages and the 1×1 bottleneck. Encoder layers use instance normalization, except the first and last downsampling stages, and a LeakyReLU activation (negative slope 0.2). The decoder upsamples with bilinear interpolation, followed by a 3×3 convolution, instance normalization, and ReLU; dropout (rate 0.5) is applied in the first three decoder stages. Skip connections concatenate encoder feature maps with the corresponding decoder maps at the same spatial resolution. The decoder output is concatenated with the original input and passed through a 3×3 convolution with a Tanh activation to produce a single-channel image in the [-1, 1] range. A single network with this architecture was trained jointly for both T1-to-T2 and T2-to-T1 translation, alternating the two directions during each training epoch; the same weights were used for both directions. As shown in Figure 1, this shared model provides a unified framework for bidirectional translation. During training, the predicted image is compared with the corresponding ground-truth image of the target contrast to compute the loss function.

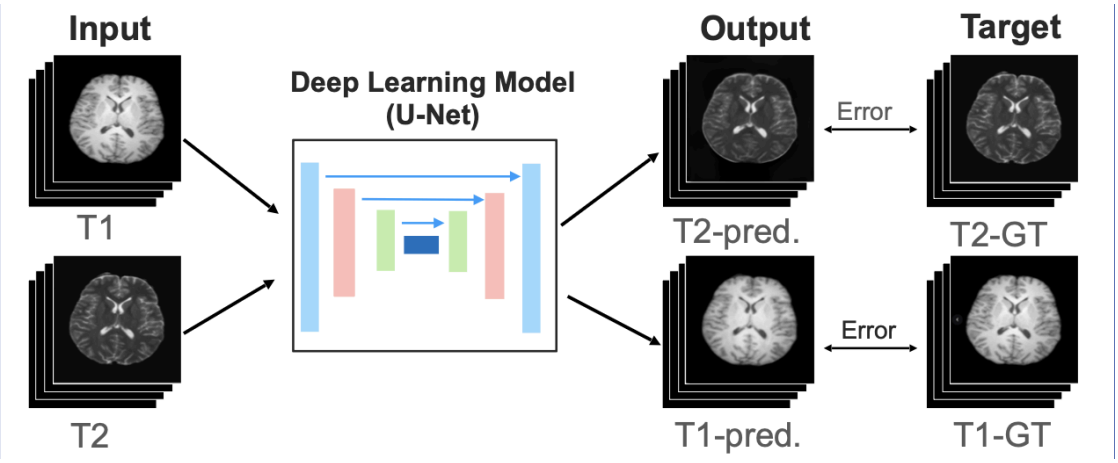

Figure 1: Paired T1- and T2-weighted images from the same patient are used as inputs and targets for a U-Net model. The network performs bidirectional cross-contrast translation (T1-to-T2 and T2-to-T1), and

the predictions are compared with the corresponding ground-truth (GT) images to compute the loss for training and evaluation.

Loss Functions

A baseline model was first trained using a standard L1 loss between the predicted and ground-truth images. L1 loss measures pixel-wise differences by minimizing the mean absolute error at each pixel[21]. Although models trained with L1 loss generally achieve accurate overall reconstruction, the resulting images may appear blurry because L1 loss prioritizes average pixel-level agreement rather than the preservation of fine details. Three additional loss functions were systematically evaluated in combination with L1 loss: perceptual loss, gradient loss, and frequency loss.

Gradient and frequency losses were calculated as L1 distances between quantities derived directly from the predicted and ground-truth images. Perceptual loss compares higher-level visual features, such as shapes and overall structural information[22]. It was computed as the L1 distance between feature maps from a frozen, ImageNet-pretrained VGG-19 network (first 36 convolutional layers). Single-channel MRI images were repeated to three channels and rescaled from [-1, 1] to [0, 1] before ImageNet normalization. Gradient loss emphasizes edges and sharp intensity transitions[23]; it was defined as the L1 distance between neighboring-pixel spatial derivatives of the prediction and the ground truth in the horizontal and vertical directions. Frequency loss helps maintain fine textures and detailed image patterns[24]; it was defined as the L1 distance between the magnitude spectra of the two-dimensional Fourier transforms of the prediction and the ground truth.

For each additional loss component, weights of 0.2, 0.4, 0.6, and 0.8 were evaluated on the 30-patient development set only. The selected weights were 0.4 for perceptual loss, 0.6 for gradient loss, and 0.3 for frequency loss. The composite loss used for final training was $L = L_L1 + 0.4\ L_perc + 0.6\ L_grad + 0.3\ L_freq$. The 10-patient test set was not used to choose these weights.

Model Training

Training was performed on 2D slices from the patient-level 80/20 train/validation splits described above, with a batch size of 16. The model was optimized using the Adam optimizer with an initial learning rate of 1×10^{-4}, which was adaptively adjusted based on the validation loss. Training was conducted for up to 100 epochs with early stopping (patience = 20) using validation loss from the development or final-training validation patients only. The independent 10-patient test set was not used to choose the architecture, loss weights, learning-rate schedule, or early-stopping checkpoint. Model performance on the locked test set was evaluated using quantitative metrics[25], [26], including structural similarity index (SSIM), peak signal-to-noise ratio (PSNR), and mean absolute and mean squared error (MAE and MSE). SSIM and PSNR assess structural similarity and image fidelity, with higher values indicating better performance, while MAE and MSE measure residual errors, where lower values indicate improved reconstruction accuracy. The model was implemented using the Pytorch framework in Python and trained on Google Colab with GPU acceleration. The computational environment included 12 GB of system RAM, a 15 GB GPU, and 120 GB of disk storage. The training time ranged from 5 to 12 hours. After

parameter optimization using the 30-patient development set, the final model was trained using data from 110 patients and evaluated once on the same 10-patient independent test set.

RESULTS

Loss Function Weights

The optimal loss function weights were identified through systematic evaluation of different combinations of perceptual, gradient, and frequency loss components on the 30-patient development set. The best-performing combination was achieved with weights of 0.4 for perceptual loss, 0.6 for gradient loss, and 0.3 for frequency loss. This optimized loss configuration was incorporated into a final composite loss function for training on the 110-patient cohort. The 10-patient test set was not consulted during this selection.

T1-to-T2 Translation

Predicted T2-weighted images generated from T1-weighted inputs are shown in Figure 2, alongside the corresponding input images and ground-truth T2-weighted targets. As demonstrated in the magnified views, the model trained with the composite loss function produces sharper and more detailed reconstructions compared with other models, particularly the L1-only baseline.

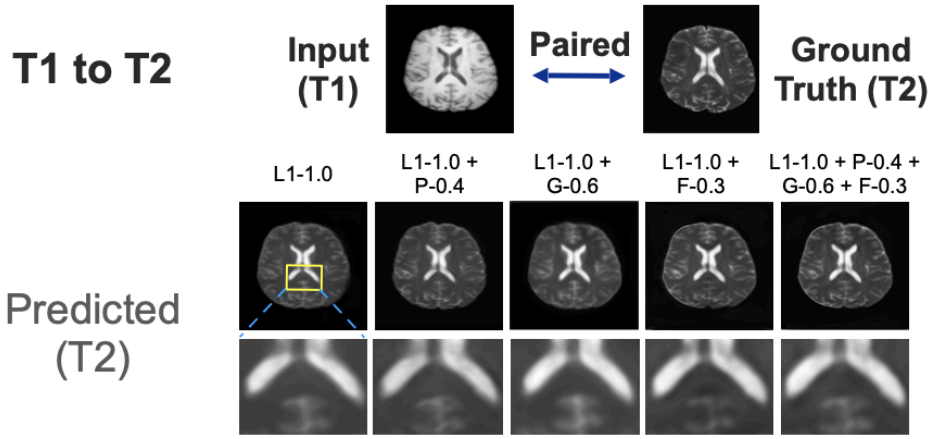

Figure 2: Paired T1-weighted input and ground-truth T2-weighted images are shown alongside model-generated T2-weighted outputs obtained using different loss function combinations (L1, L1+perceptual loss, L1+gradient loss, L1+frequency loss, and composite loss). Zoomed-in regions highlight differences in structural preservation and fine anatomical details among the models.

T2-to-T1 Translation

Predicted T1-weighted images generated from T2-weighted inputs are shown in Figure 3, alongside the corresponding input T2-weighted images and ground-truth T1-weighted targets for reference. Consistent with the T1-to-T2 translation results, the model trained with the composite loss function produces sharper reconstructions and better preserves fine structural details compared with models using other loss functions. However, the overall performance of all models is high, and the differences between methods are subtle and may not be readily distinguishable through visual inspection alone.

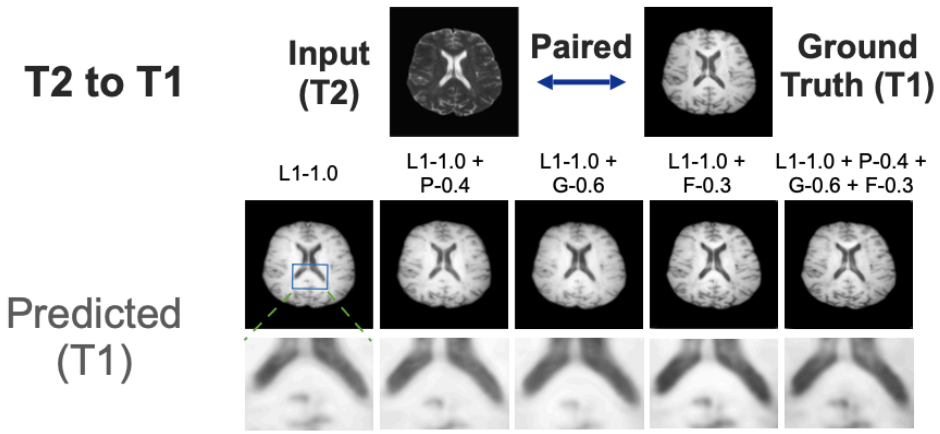

Figure 3: Paired T2-weighted input and ground-truth T1-weighted images are shown alongside model-generated T1-weighted outputs obtained using different loss function combinations (L1, L1+perceptual loss, L1+gradient loss, L1+frequency loss, and combined loss). Zoomed-in regions highlight differences in structural preservation and fine anatomical details among the models.

Quantitative Evaluation

The proposed model generates T1-weighted and T2-weighted images with strong structural similarity and high peak signal-to-noise ratio (PSNR) relative to the ground truth, while maintaining low prediction errors. These results indicate that the model captures both global contrast and fine anatomical details. Table 1 reports mean ± standard deviation of SSIM, PSNR, MAE, and MSE across the 10 independent test patients for models trained with 30 or 110 patients. We interpret differences by their absolute magnitude relative to this patient-to-patient variability, and by whether the ranking of loss configurations is consistent in the 30-patient and 110-patient experiments.

On the 110-patient models, the composite loss improved T1-to-T2 translation relative to the L1-only U-Net by 0.0213 in mean SSIM and 1.62 dB in mean PSNR, and reduced MAE from 0.0028 to 0.0020. These changes are larger than the corresponding T2-to-T1 differences among the top models and support a meaningful gain over the L1-only baseline in the more difficult T1-to-T2 direction. For T2-to-T1 translation, the L1+frequency model achieved slightly better mean SSIM (0.9935 versus 0.9931), PSNR

(28.61 dB versus 28.42 dB), MAE, and MSE than the composite model. Those differences are minor (0.0004 in SSIM and 0.19 dB in PSNR). T2-to-T1 metrics are already very high (mean SSIM of 0.993), so there is little remaining room for improvement in that direction. The composite model was selected to balance the bidirectional task: it produced the larger gain on the more difficult T1-to-T2 direction while remaining close to the best T2-to-T1 configuration. It was not selected because it was the best model on every individual metric.

Table 1: Evaluation results of models in T1-to-T2 and T2-to-T1 translation tasks.

Loss Function	**T1 → T2**				**T2 → T1**			
	SSIM	**PSNR (dB)**	**MAE**	**MSE**	**SSIM**	**PSNR (dB)**	**MAE**	**MSE**
L1 - 30 pairs	0.9200 ± 0.0216	24.85 ± 1.67	0.0036 ± 0.0016	0.0335 ± 0.0130	0.9828 ± 0.0078	24.63 ± 1.53	0.0037 ± 0.0016	0.0298 ± 0.0047
L1 - 110 pairs	0.9381 ± 0.0082	25.55 ± 0.89	0.0028 ± 0.0006	0.0258 ± 0.0057	0.9906 ± 0.0029	26.99 ± 1.04	0.0021 ± 0.0005	0.0214 ± 0.0025
L1 + Perceptual (0.4) - 30 pairs	0.9319 ± 0.0220	25.39 ± 1.90	0.0032 ± 0.0017	0.0331 ± 0.0142	0.9847 ± 0.0066	25.06 ± 1.47	0.0033 ± 0.0014	0.0276 ± 0.0038
L1 + Perceptual (0.4) - 110 pairs	0.9352 ± 0.0170	25.35 ± 1.27	0.0030 ± 0.0009	0.0324 ± 0.0093	0.9874 ± 0.0034	25.75 ± 1.05	0.0027 ± 0.0007	0.0250 ± 0.0033
L1 + Gradient (0.6) - 30 pairs	0.9357 ± 0.0181	25.50 ± 1.67	0.0031 ± 0.0013	0.0319 ± 0.0125	0.9867 ± 0.0069	25.68 ± 1.78	0.0030 ± 0.0015	0.0268 ± 0.0041
L1 + Gradient (0.6) - 110 pairs	0.9317 ± 0.0164	24.91 ± 1.14	0.0033 ± 0.0010	0.0335 ± 0.0072	0.9903 ± 0.0036	26.84 ± 1.12	0.0021 ± 0.0007	0.0227 ± 0.0029
L1 + Frequency (0.3) - 30 pairs	0.9480 ± 0.0143	26.31 ± 1.36	0.0025 ± 0.0009	0.0295 ± 0.0115	0.9917 ± 0.0018	27.54 ± 1.05	0.0018 ± 0.0004	0.0237 ± 0.0028
L1 + Frequency (0.3) - 110 pairs	0.9447 ± 0.0140	25.75 ± 0.82	0.0027 ± 0.0005	0.0307 ± 0.0043	**0.9935 ± 0.0019**	**28.61 ± 1.07**	**0.0014 ± 0.0003**	**0.0185 ± 0.0022**
L1 + P0.4 + G0.6 + F0.3 - 30 pairs	0.9485 ± 0.0146	26.34 ± 1.49	0.0025 ± 0.0010	0.0297 ± 0.0123	0.9913 ± 0.0023	27.35 ± 1.10	0.0019 ± 0.0005	0.0237 ± 0.0037
L1 + P0.4 + G0.6 + F0.3 - 110 pairs	**0.9594 ± 0.0139**	**27.17 ± 1.20**	**0.0020 ± 0.0006**	**0.0256 ± 0.0064**	0.9931 ± 0.0019	28.42 ± 1.16	0.0015 ± 0.0004	0.0198 ± 0.0026

Values are mean ± standard deviation over 10 independent test patients. Each loss combination was trained on 30 or 110 patients. The composite model (last row) was selected to balance both translation directions: it improved the more difficult T1-to-T2 task while keeping T2-to-T1 metrics close to the already high L1+frequency values.

The focus of this study is not a systematic comparison with other published translation models. We aimed to demonstrate that optimizing a composite loss function can improve bidirectional U-Net performance. For a general performance reference on the same public dataset, Bagheri and Uludag reported a U-Net baseline for 1.5T T1/T2 translation on PMC, with mean PSNR of 27.93 dB for T1-to-T2 and 26.97 dB for T2-to-T1[19]. Our 3T composite model reached a mean PSNR of 27.17 dB for T1-to-T2 and 28.42 dB for T2-to-T1, a similar numerical range. Field strength, slice selection, and data splits differ, so this comparison is contextual rather than a matched re-implementation. Within the present study, the L1-only U-Net remains the controlled baseline; the composite model improved on that baseline most clearly for T1-to-T2 translation.

Effect of Training Dataset Size

The proposed model achieved strong performance despite being trained initially on a relatively small dataset of 30 patients. After determining the optimal weights for each loss function using this development set, a larger training dataset was used to assess the effect of training-set size. Two training scenarios were compared: models trained using images from 30 patients and models trained using images from 110 patients. Both were evaluated on the same locked set of 10 test patients that had not been used for loss-weight selection. As shown in Table 1, increasing the training dataset size improved overall model performance across the evaluated metrics.

DISCUSSION

The most important finding of this study is that the model trained using the composite loss function, which combines L1, perceptual, gradient, and frequency losses, achieved the most consistent and optimal performance in both translation directions. Each individual loss function evaluates different aspects of the synthesized images; therefore, the composite loss function integrates complementary information from all components. By combining these loss functions, the model achieves a balance between preserving structural accuracy, fine details, texture information, and overall image fidelity in the generated image. These experiments are a loss-function study rather than a systematic comparison of our architecture with other published models. A contextual comparison of general performance with the PMC U-Net baseline of Bagheri and Uludag is reported in the Results.

Although the L1+frequency loss slightly outperformed the composite loss in T2-to-T1 translation on all four metrics, the difference was minor (0.0004 in SSIM and 0.19 dB in PSNR for the 110-patient models). T2-to-T1 scores were already very high, so further gains in that direction are limited. We selected the composite model to balance the bidirectional task: a larger improvement on T1-to-T2 translation, the more difficult direction, without giving up the already high T2-to-T1 performance.

T2-to-T1 translation consistently outperformed T1-to-T2 translation across all evaluated loss functions. This difference may be attributed to the characteristics of image intensity distributions and the behavior of evaluation metrics. T1-weighted images typically contain large regions with high signal intensity, making differences between the ground-truth and synthesized images less apparent in metric calculations. In contrast, T2-weighted images often contain larger low-intensity regions, where small discrepancies between the generated and ground-truth images may have a greater impact on quantitative evaluation metrics[27].

To evaluate the different models, we initially trained them on image pairs from 30 patients and monitored validation performance within that development cohort. A smaller dataset was used during this phase to optimize the loss-function weights, as shorter training times enabled more efficient systematic evaluation of different parameter combinations. The 10-patient test set was not used to choose the loss weights or the final configuration. Although a larger dataset can improve model performance, it also increases computational requirements, including training time and memory consumption, requiring a balance between model accuracy and computational efficiency. After the loss configuration was fixed, a larger dataset of 110 patients was used for final training, and the locked test set was used only for the results reported in Table 1.

Limitations and Future Work

Although the model was evaluated using a larger training dataset than the initial optimization dataset (110 patients compared to 30 patients), further training using the entire 181-patient dataset remains necessary to fully evaluate the potential of the proposed approach. During data preprocessing, five central 2D slices were extracted from each patient volume at matched anatomical positions rather than using the entire 3D volume. Two-dimensional training with five central slices was used because of limited computational resources. Generalizability to other axial levels needs further evaluation, and further work will validate whole-brain coverage and 3D approaches, which may also better preserve anatomical continuity across slices. Additionally, evaluation using larger, multi-institutional datasets would be valuable to determine whether the model can generalize across different scanners, imaging protocols, and magnetic field strengths.

The model that we trained only used T1- and T2-weighted images. One direction for future work is extending the model to other contrasts. FLAIR is clinically valuable for detecting white matter lesions, and DWI offers unique insight for stroke diagnosis by tracking the movement of water molecules[28]. A variety of contrasts provides a fuller picture of the patient, aiding in accurate diagnosis and testing whether our method generalizes beyond the specific properties of T1 and T2. Another consideration is the model's capability to accurately predict T1- and T2-weighted images containing abnormalities. The dataset used in this study includes images from healthy subjects, but future work should use a dataset containing images with tissue abnormalities to expand the scope of the model's translational ability. If this model were to be deployed in a clinical setting, this would be essential, as MRI images function mainly as a diagnostic tool.

The metrics we used (SSIM, PSNR, MAE, MSE) measure pixel and structural similarity or errors to the ground truth, but they do not answer the question of whether a radiologist would consider the synthesized image accurate for diagnosis. If a synthesized image smooths or misses a subtle detail, it could be dangerous for the patient, not just a concern related to image quality[29]. A next step would be to have radiologists review synthesized images compared with real images and rate their similarity, or attempt to distinguish real images from synthetic images. Radiologist validation would confirm diagnostic confidence and safety by adding human judgement that the evaluation metrics cannot capture.

CONCLUSION

We developed a U-Net-based model for bidirectional MRI contrast translation between T1- and T2-weighted images and systematically evaluated the effect of incorporating perceptual, gradient, and frequency loss functions alongside an L1 pixel-wise loss. Our results show that a composite loss combining all three supplementary loss functions produces sharper and more structurally accurate reconstructions compared with the L1-only model, particularly for T1-to-T2 translation. The composite model was selected to balance the bidirectional task: it improved T1-to-T2 translation, the more difficult direction, while remaining close to the best T2-to-T1 scores, which were already very high. These results highlight the importance of careful loss function design for high-quality medical image synthesis. They do not demonstrate a reduction in MRI scan time, cost, or patient burden; such clinical applications remain a goal for future work that includes pathologic cases, whole-brain evaluation, and radiologist review. We also found that using a larger dataset has the potential to improve model performance; however, the training time is significantly longer.

ABOUT THE AUTHOR

Rachel Yang is a 12th-grade student at Scarsdale High School in Scarsdale, New York. Her research focuses on deep learning methods for magnetic resonance imaging contrast translation.

REFERENCES

1. V. Grover, J. Tognarelli, M. Crossey, J. Cox, S. Taylor-Robinson, M. McPhail. Magnetic resonance imaging: principles and techniques: lessons for clinicians. Journal of Clinical and Experimental Hepatology. Vol. 5, pg. 246–255, 2015. https://doi.org/10.1016/j.jceh.2015.08.001.
2. W. Nitz, P. Reimer. Contrast mechanisms in MR imaging. European Radiology. Vol. 9, pg. 1032–1046, 1999. https://doi.org/10.1007/s003300050789.
3. S. Currie, N. Hoggard, I. Craven, M. Hadjivassiliou, I. Wilkinson. Understanding MRI: basic MRI physics for physicians. Postgraduate Medical Journal. Vol. 89, pg. 209–223, 2013. https://doi.org/10.1136/postgradmedj-2012-131342.
4. M. Saranathan, P. Worters, D. Rettmann, B. Winegar, J. Becker. Physics for clinicians: fluid-attenuated inversion recovery (FLAIR) and double inversion recovery (DIR) imaging. Journal of Magnetic Resonance Imaging. Vol. 46, pg. 1590–1600, 2017. https://doi.org/10.1002/jmri.25737.
5. H. Jara, O. Sakai, E. Farrher, A.-M. Oros-Peusquens, N. Shah, D. Alsop, K. Keenan. Primary multiparametric quantitative brain MRI: state-of-the-art relaxometric and proton density mapping techniques. Radiology. Vol. 305, pg. 5–18, 2022. https://doi.org/10.1148/radiol.211519.

6. L. Zrinzo, A. Papadaki, H. Akram, M. Hariz, M. T. Krueger. Proton density MR imaging: a simple and powerful sequence to visualise anatomical structures in functional neurosurgery. Stereotactic and Functional Neurosurgery. 2026. https://doi.org/10.1159/000552105.
7. J. Raya, A. Duarte, N. Wang, V. Mazzoli, D. Jaramillo, A. Blamire, O. Dietrich. Applications of diffusion-weighted MRI to the musculoskeletal system. Journal of Magnetic Resonance Imaging. Vol. 59, pg. 376–396, 2024. https://doi.org/10.1002/jmri.28870.
8. T. Tamada, Y. Ueda, Y. Ueno, Y. Kojima, A. Kido, A. Yamamoto. Diffusion-weighted imaging in prostate cancer. Magnetic Resonance Materials in Physics, Biology and Medicine. Vol. 35, pg. 533–547, 2022. https://doi.org/10.1007/s10334-021-00957-6.
9. T. Martin-Noguerol, E. Santos-Armentia, A. Ramos, A. Luna. An update on susceptibility-weighted imaging in brain gliomas. European Radiology. Vol. 34, pg. 6763–6775, 2024. https://doi.org/10.1007/s00330-024-10703-w.
10. F. Goncalves, S. Serai, G. Zuccoli. Synthetic brain MRI: review of current concepts and future directions. Topics in Magnetic Resonance Imaging. Vol. 27, pg. 387–393, 2018. https://doi.org/10.1097/rmr.0000000000000189.
11. J. Hamilton, D. Franson, N. Seiberlich. Recent advances in parallel imaging for MRI. Progress in Nuclear Magnetic Resonance Spectroscopy. Vol. 101, pg. 71–95, 2017. https://doi.org/10.1016/j.pnmrs.2017.04.002.
12. O. Jaspan, R. Fleysher, M. Lipton. Compressed sensing MRI: a review of the clinical literature. British Journal of Radiology. Vol. 88, pg. 20150487, 2015. https://doi.org/10.1259/bjr.20150487.
13. R. Heckel, M. Jacob, A. Chaudhari, O. Perlman, E. Shimron. Deep learning for accelerated and robust MRI reconstruction. Magnetic Resonance Materials in Physics, Biology and Medicine. Vol. 37, pg. 335–368, 2024. https://doi.org/10.1007/s10334-024-01173-8.
14. C. Chen, C. Raymond, W. Speier, X. Jin, T. Cloughesy, D. Enzmann, B. Ellingson, C. Arnold. Synthesizing MR image contrast enhancement using 3D high-resolution ConvNets. IEEE Transactions on Biomedical Engineering. Vol. 70, pg. 401–412, 2023. https://doi.org/10.1109/tbme.2022.3192309.
15. T. Hilbert, P. Omoumi, M. Raudner, T. Kober. Synthetic contrasts in musculoskeletal MRI: a review. Investigative Radiology. Vol. 58, pg. 111–119, 2023. https://doi.org/10.1097/rli.0000000000000917.
16. E. Moya-Sáez, Ó. Peña-Nogales, R. de Luis-García, C. Alberola-López. A deep learning approach for synthetic MRI based on two routine sequences and training with synthetic data. Computer Methods and Programs in Biomedicine. Vol. 210, pg. 106371, 2021. https://doi.org/10.1016/j.cmpb.2021.106371.
17. S. Qiu, S. Ma, L. Wang, Y. Chen, Z. Fan, F. Moser, M. Maya, P. Sati, N. Sicotte, A. Christodoulou, Y. Xie, D. Li. Direct synthesis of multi-contrast brain MR images from MR multitasking spatial factors using deep learning. Magnetic Resonance in Medicine. Vol. 90, pg. 1672–1681, 2023. https://doi.org/10.1002/mrm.29715.
18. A. Osman, N. Tamam. Deep learning-based convolutional neural network for intramodality brain MRI synthesis. Journal of Applied Clinical Medical Physics. Vol. 23, pg. e13530, 2022. https://doi.org/10.1002/acm2.13530.
19. F. Bagheri, K. Uludag. PMC: paired multi-contrast MRI dataset at 1.5T and 3T for supervised Image2Image translation. Proceedings of the Machine Learning for Cognitive and Mental Health

Workshop (ML4CMH), AAAI 2024. Vol. 3649, pg. 29–34, 2024. https://ceur-ws.org/Vol-3649/Paper21.pdf.

20. O. Ronneberger, P. Fischer, T. Brox. U-Net: convolutional networks for biomedical image segmentation. Medical Image Computing and Computer-Assisted Intervention – MICCAI 2015. Vol. 9351, pg. 234–241, 2015. https://doi.org/10.1007/978-3-319-24574-4_28.
21. J. Terven, D.-M. Cordova-Esparza, J.-A. Romero-González, A. Ramírez-Pedraza, E. A. Chávez-Urbiola. A comprehensive survey of loss functions and metrics in deep learning. Artificial Intelligence Review. Vol. 58, pg. 195, 2025. https://doi.org/10.1007/s10462-025-11198-7.
22. H. Takeshima. Deep learning and its application to function approximation for MR in medicine: an overview. Magnetic Resonance in Medical Sciences. Vol. 21, pg. 553–568, 2022. https://doi.org/10.2463/mrms.rev.2021-0040.
23. L. L. Iglesias, P. S. Bellón, A. P. del Barrio, P. M. Fernández-Miranda, D. R. González, J. A. Vega, A. A. G. Mandly, J. A. P. Blanco. A primer on deep learning and convolutional neural networks for clinicians. Insights into Imaging. Vol. 12, pg. 117, 2021. https://doi.org/10.1186/s13244-021-01052-z.
24. O. Yadav, K. Ghosal, S. Lutz, A. Smolic. Frequency-domain loss function for deep exposure correction of dark images. Signal, Image and Video Processing. Vol. 15, pg. 1829–1836, 2021. https://doi.org/10.1007/s11760-021-01915-4.
25. J. McNaughton, J. Fernandez, S. Holdsworth, B. Chong, V. Shim, A. Wang. Machine learning for medical image translation: a systematic review. Bioengineering. Vol. 10, pg. 1078, 2023. https://doi.org/10.3390/bioengineering10091078.
26. H. M. S. S. Herath, H. M. K. K. M. B. Herath, N. Madusanka, B.-I. Lee. A systematic review of medical image quality assessment. Journal of Imaging. Vol. 11, pg. 100, 2025. https://doi.org/10.3390/jimaging11040100.
27. Z. Yu, X. Han, S. Zhang, J. Feng, T. Peng, X.-Y. Zhang. MouseGAN++: unsupervised disentanglement and contrastive representation for multiple MRI modalities synthesis and structural segmentation of mouse brain. IEEE Transactions on Medical Imaging. Vol. 42, pg. 1197–1209, 2023. https://doi.org/10.1109/TMI.2022.3225528.
28. N. Nagaraja. Diffusion weighted imaging in acute ischemic stroke: a review of its interpretation pitfalls and advanced diffusion imaging application. Journal of the Neurological Sciences. Vol. 425, pg. 117435, 2021. https://doi.org/10.1016/j.jns.2021.117435.
29. R. Acs, H. Zhuang. A review on cross-contrast MRI image synthesis through deep learning. Discover Imaging. Vol. 2, pg. 9, 2025. https://doi.org/10.1007/s44352-025-00012-3.