

Style Transfer of MRI Images using CycleGAN and Dual-GAN

Nitish Reddy Gaddam

Department of Applied Data Science
nitishreddy.gaddam@sjsu.edu

Nikhil Mylarusetty

Department of Applied Data Science
nikhil.mylarusetty@sjsu.edu

Dharanidhar Reddy Banala

Department of Applied Data Science
dharnidharreddy.banala@sjsu.edu

Manidedeepya Chennapragada

Department of Applied Data Science
mandedeepya.chennapragada@sjsu.edu

Naga Pavan Kumar Kodeboina

Department of Applied Data Science
nagapavankumar.kodeboina@sjsu.edu

Kalyan Vikkurthi

Department of Applied Data Science
kalyan.vikkurthi@sjsu.edu

Abstract—The integration of deep learning in medical imaging has ushered in a transformative era for diagnostic precision. This project addresses the challenge of subjective interpretations in Magnetic Resonance Imaging (MRI) scans, particularly T1 and T2 weighted images, by leveraging Generative Adversarial Networks (GANs). Employing a modified U-Net architecture, the project aims to generate artificial MRI images with varied contrast levels, enhancing diagnostic accuracy and revolutionizing the interpretation of medical imaging. Motivated by the imperative to improve diagnostic consistency, the project's innovation lies in the translation between T1 and T2 weighted images, offering a more comprehensive view for analysis. The methodology involves the use of CycleGAN, DualGAN, and MUNIT for style transfer, providing a tool to augment radiologists' abilities and ultimately improve patient care outcomes.

I. INTRODUCTION

The goal of this research is to bridge the interpretation gap in medical imaging so that patients can be diagnosed with the highest degree of accuracy and receive the best possible care. This study is a significant step towards bettering healthcare outcomes since it integrates the most recent developments in artificial intelligence with medical knowledge. However, radiologists may interpret these images differently, resulting in disparate diagnosis for identical illnesses. This disparity poses a serious challenge to medical diagnosis. This project focuses on improving the accuracy of MRI analyses by specifically addressing the interpretation of T1 and T2 weighted MRI images. T1-weighted images are known for their clarity in showing fine anatomical details, particularly in the brain. T2-weighted images, in contrast, are better at highlighting variations in tissue, such as areas with fluid or swelling. These two types of images provide different yet complementary views of the same body structures. The objective of this work is to use Generative Adversarial Networks (GANs), a type of deep learning technology, to increase the diagnostic precision of MRI images. To obtain a broader view, this involves a sophisticated technique of merging the distinctive features of T1 and T2 images. CycleGAN, DualGAN, and MUNIT are examples of the sophisticated AI models used in the study to accomplish this goal. Creating a tool to assist radiologists in making more reliable and consistent diagnoses from MRI

scans is the goal here, not merely polishing the technology. In order to effectively diagnose and treat patients, the study aims to reduce picture interpretation variability by merging the features of T1 and T2 images. The main objective of this research is to diagnose patients as accurately as possible using medical imaging in order to give them the best care possible. This study integrates historical medical knowledge with the latest developments in artificial intelligence with the aim of enhancing healthcare results.

II. RELATED WORK

A notable study in this evolving landscape is that of [1]Le et al. (2021), which specifically addressed the challenges of low-field MRI systems. These systems, while cost-effective and space-efficient, suffer from inherent signal-to-noise ratio limitations. Le et al. introduced an innovative approach by generating training data that simulated scanner-specific images using a signal encoding matrix. This methodology incorporated modeled imaging gradients and fields from publicly available high-field MRI databases. By employing a stacked U-Net architecture, their approach focused on reducing noise and eliminating artifacts caused by various factors such as the inhomogeneous B0 field and nonlinear gradients. The outcome of their study was significant; the trained models markedly improved the visualization of fine structures in MRI images, as evidenced by increased metrics like Peak Signal-to-Noise Ratio (PSNR), Structural Similarity Index (SSIM), and Normalized Root Mean Squared Error (NRMSE). This work not only contributes to the ongoing efforts in enhancing low-field MR images but also aligns closely with contemporary projects, such as those exploring the use of advanced deep learning models like DualGAN and CycleGAN in MRI imaging. These recent endeavors, including the work on DeLaBE and the methodologies developed by Le et al., collectively represent a cutting-edge frontier in MRI image enhancement, promising significant improvements in diagnostic capabilities and patient outcomes. The study by [2]Campá-Usuga et al. (2021) represents a significant advancement in medical imaging, specifically in enhancing the quality of breast Dynamic Contrast-Enhanced Magnetic Resonance Imaging (DCE-MRI)

crucial for breast cancer diagnosis. This research, utilizing Convolutional Autoencoders (CAE), addresses the challenge posed by the varying resolution capabilities of MRI equipment, which can affect diagnostic accuracy. By synthetically degrading images from the Qin breast dataset and subsequently employing CAE for reconstruction, the study demonstrated notable improvements in image quality, as measured by metrics like MSE, PSNR, and SSIM. This work is particularly relevant to projects focusing on the use of advanced deep learning techniques, such as DualGAN and CycleGAN, for MRI image enhancement. The methodology of Campáz-Usuga et al. aligns with the broader trend of leveraging deep learning to overcome hardware limitations in medical imaging, enhancing diagnostic capabilities and patient care. Their approach underscores the growing potential and impact of computational sciences in medical image synthesis, offering insights beneficial for similar deep learning applications in MRI imaging. The research by [3] Bhan, Kapoor, and Gulati (2021) is a pivotal study in the realm of early Parkinson's disease (PD) diagnosis using deep learning techniques, particularly relevant to projects involving medical image analysis and enhancement. Focusing on MRI as a non-invasive diagnostic tool, their work involved preprocessing and image enhancement, with a specific emphasis on the substantia nigra (SN) region for detecting neuronal loss characteristic of PD. Utilizing various convolutional neural network (CNN) models, including custom CNN, VGG-16, ResNet-34, VGG-19, ResNet-50, and improved AlexNet, the study achieved significant success in differentiating between PD and control subjects, with the ResNet-50 model showing an impressive validation accuracy of 98.91. This approach highlights the potential of deep learning in early and accurate PD diagnosis. This study, with its extensive dataset and focus on CNN models, aligns well with MRI image enhancement projects like ours using advanced deep learning models. The insights from their research offer valuable guidance for similar applications in medical diagnostics, emphasizing the role of deep learning in enhancing image quality and diagnostic precision in neurodegenerative diseases. [4]Kim's (2020) study innovatively tackles the issue of image degradation in ultrasound displays during zooming, a challenge pertinent to the broader field of medical imaging, including areas like MRI image enhancement in your project. Traditional zoom algorithms often result in poor image quality, but Kim's application of Very Deep Convolution Networks (VDSR), trained on a mixed dataset of natural and ultrasound images, marks a significant advancement. The study demonstrated that using VDSR significantly improves the resolution and contrast of ultrasound images, overcoming the limitations of conventional zoom techniques. This approach, quantitatively validated through enhanced peak signal-to-noise ratios, showcases the potential of deep learning to refine image quality, offering insights relevant to enhancing MRI imaging techniques in our project. Kim's work highlights the versatility and effectiveness of deep learning methods in various medical imaging contexts, aligning with the objectives of applying advanced AI techniques for diagnostic image improvement. In

our project, we draw significant insights from [5]Rizaldi et al. (2022)'s study on classifying COVID-19 lung X-rays using deep learning. Their use of image enhancement techniques like gamma correction, CLAHE, and DCP, in combination with the ResNet50 CNN, mirrors our approach in enhancing MRI images with models like DualGAN and CycleGAN. The effectiveness of their method, particularly the success of gamma correction in significantly improving diagnostic accuracy, is directly relevant to our work. This research not only bolsters our methodology but also exemplifies the potential of integrating advanced image processing techniques with deep learning to enhance medical imaging and support clinical decision-making in critical health scenarios.

III. METHODOLOGY

A. Dataset Description

1) *Data Collection*: The dataset has been downloaded from [6], and the MRI Dataset is uploaded to google drive for shared access for the team members and is structured to facilitate easy access to images for subsequent processing.

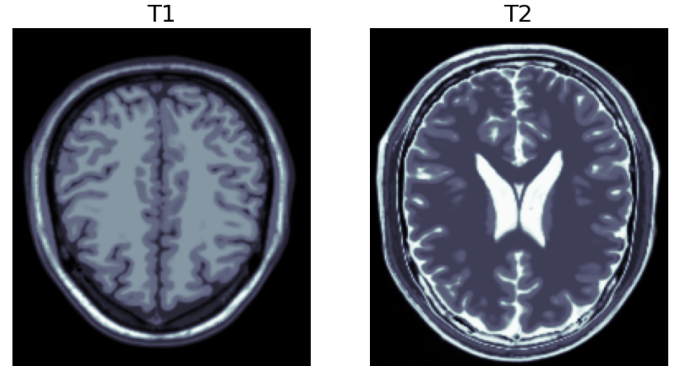


Fig. 1: T1 and T2

2) *Data Pre-processing*: The pre-processing of the dataset involves several steps. Firstly, images from the TrainT1 and TrainT2 folders are loaded using file paths obtained from the provided Google Drive path. Subsequently, data augmentation is applied to enhance the diversity of the dataset. This augmentation includes horizontal and vertical flips, 90-degree rotations, brightness adjustments, contrast adjustments, and combined brightness and contrast adjustments. The augmented images are then saved in separate directories corresponding to TrainT1 and TrainT2.

3) *Data Transformation*: The transformation of the dataset includes loading images, normalization, reshaping, and additional augmentation. Initially, images are loaded, and their pixel values are normalized to fall within the range $[-1, 1]$. The images are reshaped to have a single channel, representing grayscale. Further augmentation is performed using TensorFlow's RandomFlip to randomly flip images horizontally. TensorFlow functions are employed for pre-processing, incorporating random horizontal flipping, resizing, random cropping, and normalization.

4) *Data Preparation*: After transformation, the dataset is ready for preparation. The dataset is split into training, validation, and test sets, maintaining a random yet representative distribution of classes in each set. TensorFlow Dataset objects are created for each set, incorporating specified buffer sizes and batch sizes. The prepared dataset is now suitable for training deep learning models, with augmentation ensuring increased robustness and generalization capabilities.

B. Proposed Model Architecture

1) *CycleGAN*: CycleGAN architecture diagram, a sophisticated type of Generative Adversarial Network used for image-to-image translation tasks without the need for paired examples. It consists of two sets of models: discriminators and generators for two distinct domains A and B. The generators translate images from one domain to the other, while discriminators aim to distinguish between real and generated images where the outputs of the discriminator decision score between 0 and 1, where 0 means "fake" and 1 means "real". The key feature of CycleGAN is the cycle consistency, where an image from domain A is translated to domain B and then back to domain A, ensuring the translation process is meaningful and accurate by maintaining the essence of the original image.

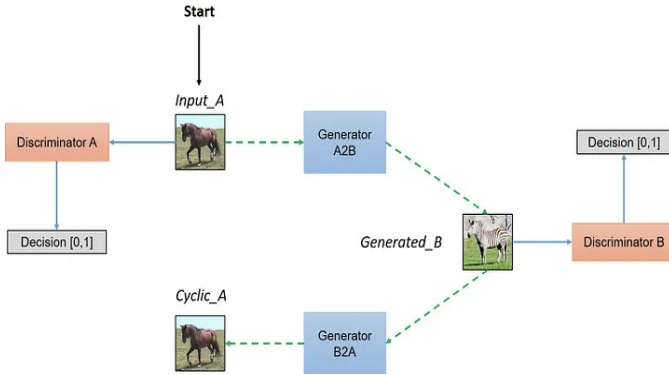


Fig. 2: CycleGAN Architecture

Generator (U-Net Architecture) The generator model adopts the U-Net architecture, consisting of two main sections: down-sampling (encoder) and up-sampling (decoder). In the down-sampling section, multiple downsample blocks are employed, each comprising a convolutional layer with instance normalization and LeakyReLU activation. The number of filters in these layers progressively increases, starting from 128 and doubling with each step. Down-sampling is achieved through convolutional layers with a stride of 2. The up-sampling section mirrors the down-sampling architecture, utilizing upsample blocks with transposed convolutional layers, instance normalization, and ReLU activation. The number of filters decreases progressively. Skip connections are established between corresponding layers in the down-sampling and up-sampling sections, facilitating the preservation of spatial information. The final output is obtained by concatenating up-sampled features with skip connection features from the

down-sampling path and feeding them into the last transposed convolutional layer, which produces the transformed output.

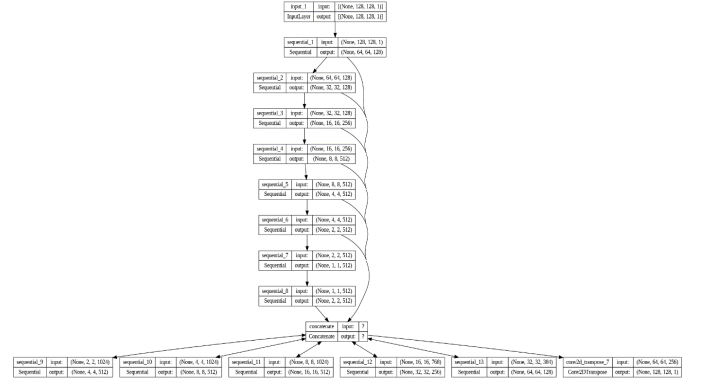


Fig. 3: Generator

Discriminator The discriminator model employs a PatchGAN architecture for distinguishing between real and generated images. It includes multiple convolutional layers with leaky ReLU activations, with the number of filters increasing progressively, similar to the generator. Instance normalization is applied at certain convolutional layers for improved stability during training. The final layer performs patch-wise classification, determining the authenticity of each patch in the input image.

2) *DualGAN*: The DualGAN architecture consists of two generator models, and two discriminator models. Similar to CycleGAN in this project both generator models follow a U-Net architecture with down-sampling and up-sampling components. The down-sampling path includes convolutional layers by gradually increasing the number of filters. The up-sampling path mirrors this structure, progressively reducing the number of filters. The discriminator models adopt a PatchGAN architecture, utilizing multiple convolutional layers with activations and normalization for stability. The final layer performs patch-wise classification to distinguish real and generated images. The choice of activation functions and normalization techniques contributes to the stability and effectiveness of the DualGAN architecture.

3) *Design Choices*: The weights are initialized using a random normal distribution with a mean of 0.0 and a standard deviation of 0.02. Instance normalization is applied in both the generator and discriminator to enhance model stability. LeakyReLU activation is used in the generator, while the discriminator employs leaky ReLU for introducing non-linearity. Skip connections play a crucial role in facilitating high-resolution information flow between encoder and decoder components. This architecture aims to capture and transform domain-specific features effectively, leveraging skip connections for enhanced information flow. The two models with this design, demonstrates proficiency in image-to-image translation tasks, ensuring the successful generation of realistic and domain-specific outputs

4) *Key Difference Between DualGAN and CycleGAN Architectures*: While both DualGAN and CycleGAN architectures

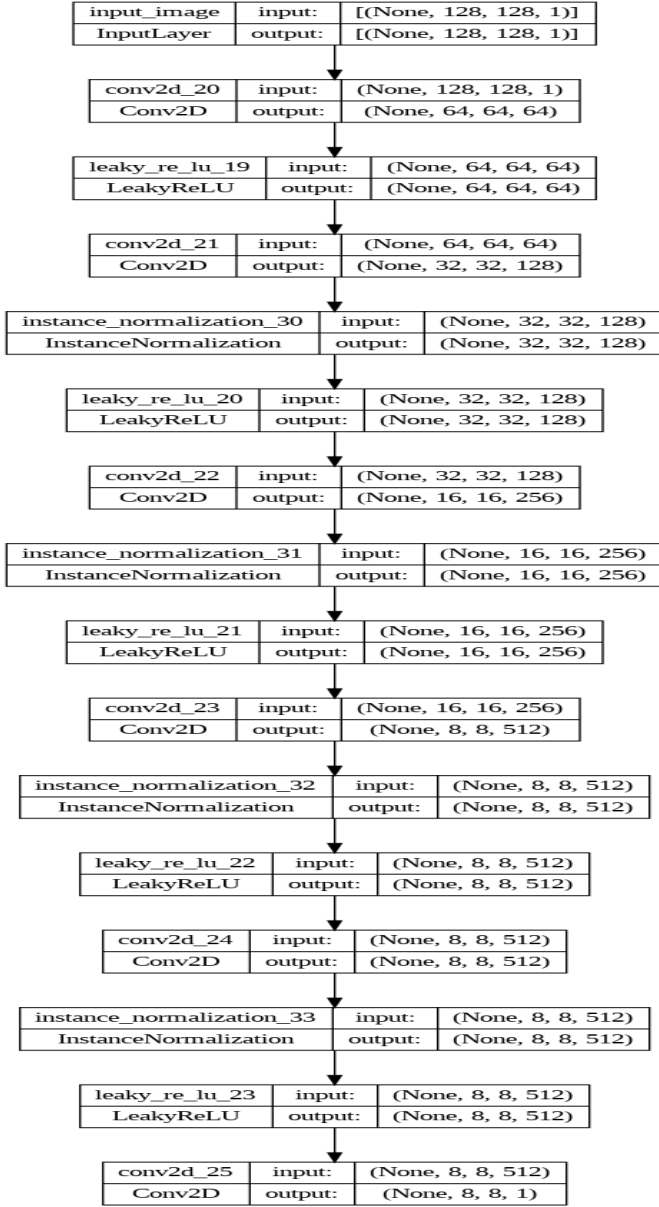


Fig. 4: Discriminator

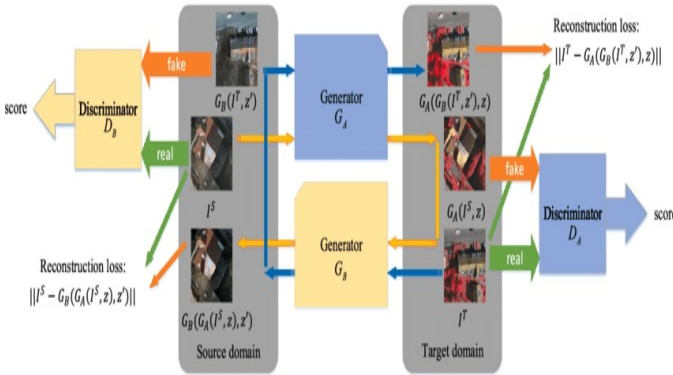


Fig. 5: DualGAN Architecture

share a common foundation based on U-Net generators and PatchGAN discriminators, a key difference lies in the formulation of the loss functions and the training objectives.

CycleGAN Specifics: On the other hand, CycleGAN relies on traditional adversarial losses for both discriminators. The generator adversarial loss is calculated using a standard GAN formulation, and the cycle consistency loss remains consistent with the L1 norm. Identity mapping losses are also introduced to further guide the generators to preserve input characteristics. CycleGAN adopts a more conventional GAN loss formulation compared to DualGAN.

DualGAN Specifics In DualGAN, the loss functions are tailored to employ the least squares (LS) metric. The discriminator LS loss is used for both, emphasizing the minimization of squared differences between real and generated outputs. The generator LS loss is designed to minimize the squared difference between the generated output and the target (real) output. Additionally, a total cycle consistency loss, calculated using the L1 norm, contributes to the preservation of image content during translation. These LS-based loss functions aim to enhance stability during training, potentially offering improvements in convergence and quality of generated images.

TABLE I: Comparison of DualGAN and CycleGAN

Aspect	DualGAN	CycleGAN
Loss Functions	Least Squares (LS) metric	Adversarial losses with GAN formulation
Discriminator Loss	LS loss for both discriminators	Adversarial loss using GAN formulation
Generator Loss	LS loss minimizing squared differences	Adversarial loss minimizing discrimination
Cycle Consistency Loss	L1 norm for total cycle consistency	L1 norm for cycle consistency
Stability and Convergence	LS-based losses aim for stability	Traditional GAN formulation for convergence
Image Content Preservation	LS and L1 contribute to content fidelity	L1 norm and identity mapping for preservation

C. Experimental Setup

The hyperparameters chosen for both DualGAN and CycleGAN models adhere to common practices in GAN literature. A learning rate of $2e-4$ is employed to facilitate stable training and prevent overshooting, and a batch size of 1 is used to enhance diversity in generated samples and mitigate mode collapse. Leaky ReLU activation functions are applied to introduce non-linearity and address the vanishing gradient problem in the generator models. The number of layers and neurons in the generators is determined empirically during model design, considering the complexity of the task and dataset. No explicit regularization parameters are used, but techniques like dropout or batch normalization are incorporated based on experimental findings. The Adam optimizer is selected for its adaptive learning rate capabilities and efficiency in GAN training.

IV. RESULTS

The comparative evaluation of DualGAN and CycleGAN for MRI image style transfer between T1 and T2 demonstrates

TABLE II: Hyperparameters for DualGAN and CycleGAN Models

Hyperparameter	DualGAN Value	CycleGAN Value
Learning Rate	2×10^{-4}	2×10^{-4}
Batch Size	1	1
Number of Epochs	Not Specified	Not Specified
Activation Function	Leaky ReLU (Generator)	Leaky ReLU (Generator)
Number of Layers (Generator)	15	15
Number of Neurons (Generator)	Varied for Each Layer	Varied for Each Layer
Number of Layers (Discriminator)	5	5
Number of Neurons (Discriminator)	Varied for Each Layer	Varied for Each Layer
Regularization Parameters	None	None
Optimization Algorithm	Adam	Adam
Optimization Variations	None	None

TABLE III: Software Libraries/Frameworks Used in the Analysis

Library/Framework	Purpose
skimage.transform	Image processing and transformation
PIL (Python Imaging Library)	Image processing and handling
imageio	Handling images and creating animations
matplotlib	Data visualization and plotting
os	File operations and handling directories
numpy	Array operations and numerical computing
tensorflow	Deep learning framework for model development
keras.utils	Utility functions for Keras (part of TensorFlow)
tensorflow_addons (tfa)	Additional functionalities for TensorFlow
tf.data.AUTOTUNE	Automatic tuning of data loading performance
warnings	Handling warnings during runtime

distinct performance characteristics. Notably, CycleGAN outperforms DualGAN in achieving effective style transfer. Figures 1 and 2 show the results of how the style transfer takes place from T1 to T2 and vice-versa for CycleGAN and DualGAN respectively.

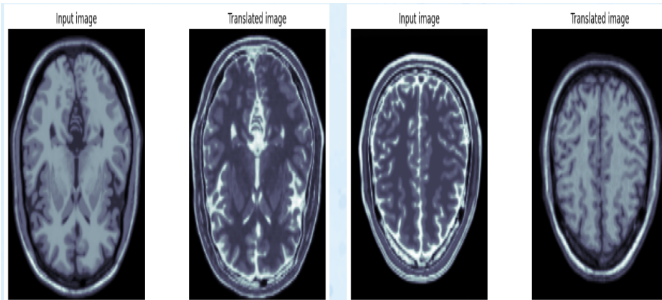


Fig. 6: CycleGAN Result

A. Evaluation

The evaluation metrics, namely Structural Similarity Index (SSIM) and Peak Signal-to-Noise Ratio (PSNR), were em-

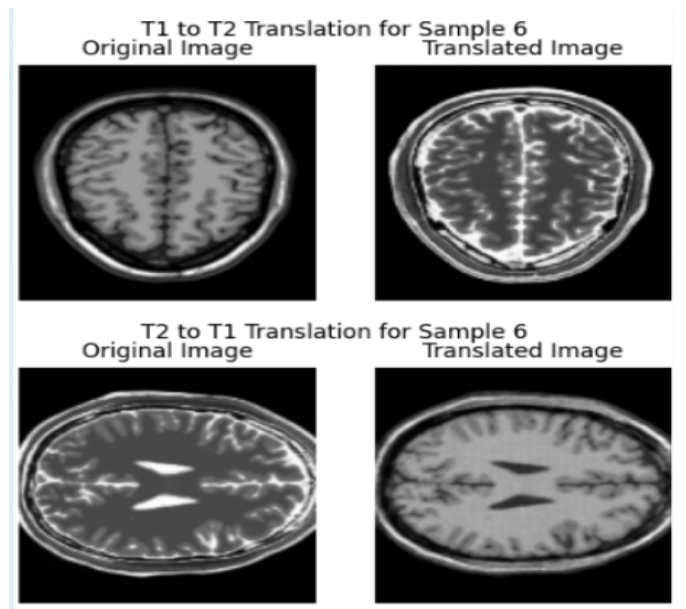


Fig. 7: DualGAN Result

ployed to quantitatively assess the fidelity of the generated images. Both models exhibited poor performance according to these quantitative measures as shown in figures 3 and 4, indicating challenges in preserving structural similarity and minimizing reconstruction errors during the style transfer process.

```
Calculating metrics for T1 to T2 translation...
Average PSNR over 100 samples: 8.09
Average SSIM over 100 samples: 0.00
```

Fig. 8: Evaluation for CycleGAN

```
Calculating metrics for T2 to T1 translation...
Average PSNR over 100 samples: 10.91
Average SSIM over 100 samples: 0.18
```

Fig. 9: Evaluation for DualGAN

It is crucial to emphasize, however, that these metrics serve as auxiliary assessments and may not fully capture the nuanced intricacies of medical image interpretation. The primary and authentic evaluation remains within the purview of expert radiologists, who possess the clinical acumen necessary for comprehensive assessment. Recognizing the limitations of quantitative metrics, particularly in the context of medical image analysis, the study underscores the imperative for subjective evaluation by radiological professionals. Visual inspection by expert practitioners is essential to discern clinically relevant details and nuances that may not be adequately captured

by numerical metrics alone. While SSIM and PSNR metrics provide quantitative benchmarks, the study underscores the importance of a multidimensional evaluation approach that incorporates both computational assessments and expert clinical judgment. Despite the quantitative challenges encountered, the overarching objective of this research is to contribute to the development of robust methodologies for MRI image style transfer, with the ultimate aim of facilitating improved diagnostic processes in collaboration with the expertise of radiologists.

V. CONCLUSION / FUTURE SCOPE

This project lays the foundation for numerous future avenues and advancements in the realm of MRI image style transfer. Firstly, there is substantial scope for refining and optimizing the existing DualGAN and CycleGAN architectures to enhance their performance metrics, considering the observed limitations in SSIM and PSNR outcomes. The incorporation of more sophisticated deep learning techniques, such as attention mechanisms and advanced loss functions, may contribute to the amelioration of image fidelity during style transfer. Additionally, exploring alternative generative models beyond GANs, such as variational autoencoders or conditional GAN variants, could offer new perspectives for achieving more accurate and clinically relevant results. Furthermore, the integration of domain-specific knowledge into the model training process may yield significant improvements. Collaborations with domain experts, including radiologists and medical imaging specialists, could guide the development of models that better align with the subtle intricacies and diagnostic considerations inherent in MRI data. Incorporating feedback from clinical practitioners during the iterative model development process is essential for tailoring the technology to real-world medical scenarios. The project's future trajectory also involves addressing the interpretability and explainability challenges associated with deep learning models in medical image analysis. Developing methodologies to elucidate the decision-making processes of the models and providing clinicians with insights into the features influencing the style transfer outcomes will foster trust and adoption in clinical settings.

REFERENCES

- 1 D. B. T. Le, M. Sadinski, A. Nacev, R. Narayanan and D. Kumar, "Deep Learning-based Method for Denoising and Image Enhancement in Low-Field MRI," 2021 IEEE International Conference on Imaging Systems and Techniques (IST), Kaohsiung, Taiwan, 2021, pp. 1-6, doi: 10.1109/IST50367.2021.9651441.
- 2 P. Campáz-Usuga, R. D. Fonnegra and C. Mera, "Quality Enhancement of Breast DCE-MRI Images Via Convolutional Autoencoders," 2021 IEEE 2nd International Congress of Biomedical Engineering and Bioengineering (CI-IBBI), Bogota D.C., Colombia, 2021, pp. 1-4, doi: 10.1109/CI-IBBI54220.2021.9626097.
- 3 A. Bhan, S. Kapoor and M. Gulati, "Diagnosing Parkinson's disease in Early Stages using Image Enhancement, ROI Extraction and Deep Learning Algorithms," 2021 2nd International Conference on Intelligent Engineering and Management (ICIEM), London, United Kingdom, 2021, pp. 521-525, doi: 10.1109/ICIEM51511.2021.9445381.
- 4 J. Seok Kim, "Improved image resolution during zooming in ultrasound image using deep learning technique," 2020 IEEE International Ultrasonics Symposium (IUS), Las Vegas, NV, USA, 2020, pp. 1-3, doi: 10.1109/IUS46767.2020.9251607.
- 5 S. T. Rizaldi, Mustakim, I. Permana and M. Afdal, "Image Enhancement on Deep Learning Algorithm for COVID-19 Lung X-Ray Classification," 2022 International Symposium on Information Technology and Digital Innovation (ISITDI), Padang, Indonesia, 2022, pp. 11-15, doi: 10.1109/ISITDI55734.2022.9944501.
- 6 <https://www.kaggle.com/datasets/ilknuricke/neurohackinginrimages?0.0.zip>