Unpaired Medical Image Translation and Detection:- *Leveraging CycleGAN for Medical Image Generation with Brain Tumor Detection*

(Team ID: 530)

Shaurya Thakur (E21CSEU0580) Shriyash Saxena (E21CSEU0605)

Introduction

Medical imaging is the foundation of modern diagnostic healthcare. It helps by providing critical insights into anatomical and pathological conditions of the subject. In the recent years, the field of medical imaging has been significantly influenced by the innovations in technology and primarily the Artificial Intelligence. Among various imaging techniques, Computed Tomography (CT) and Magnetic Resonance Imaging (MRI) are the one's which are fundamentally used for diagnostic analysis and evaluation of various medical conditions. But both of these modalities have a unique advantages and challenges. CT scans are expeditious, cost-effective and widely available. But they have limited ability to resolve soft tissue contrast and leads to higher radiation exposure to patients. On the other hand, MRI delivers rich and detailed image without radiation exposure. But MRI scans are expensive and of longer acquisition times. This motivates the research into synthesizing CT scan images into MRI scan images and vice-versa using Generative AI which helps in enhancing diagnostic capabilities, reducing radiation exposure and lowering the expense.

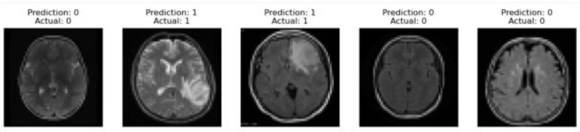


Fig 1: - Predicted and Actual values of brain tumor detection on various images.

Recent advancements in Artificial Intelligence, especially in Generative Adversarial Networks (GANs) have revolutionized the process of image synthesis and medical data augmentation. This research leverages "CycleGAN", a deep learning framework for unpaired image-to-image synthesis. CycleGan enables bidirectional translation between two images without requiring paired datasets which implies the suitability for medical Imaging applications, in the cases where pairs are often not available. This means the model will be able to generate MRI images from CT scan images and vice-versa. In this model, to maintain structural fidelity, we are training two generators (one for CT to MRI and other one for MRI to CT conversion) and two corresponding discriminators. CycleGan enforces cycle consistency such that the generated image must maintain similar properties to its original image. Moreover, the integration of diagnostic verification related to Brain Tumor, into the pipeline further augments clinical utility of the generated images.

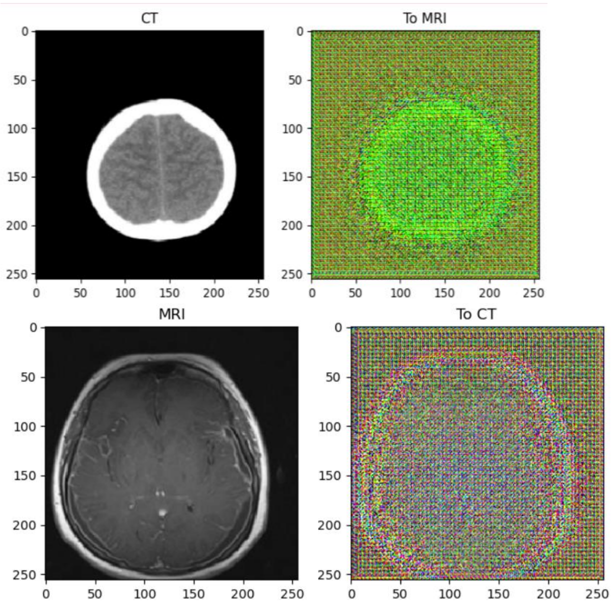


Fig 2: - Image Generation from CT to MRI and vice versa

The aim of this project is to leverage CycleGAN for generating MRI to CT scan images and vice-versa due to the purpose of medical data augmentation and lowering radiation exposure while incorporating a diagnostic module like Brain Tumor detection to show the use case of the research.

The generated images are evaluated both quantitively and qualitatively by using metrics such as the Structural Similarity Index Measure (SSIM) and Fréchet Inception Distance (FID). For instance, the SSIM is calculated using the below formula:-

$$\mathrm{SSIM}(x,y) = \frac{(2\mu_x\mu_y + c_1)(2\sigma_{xy} + c_2)}{(\mu_x^2 + \mu_y^2 + c_1)(\sigma_x^2 + \sigma_y^2 + c_2)}$$

Where ‘ux’ and ‘uy’ are the averages of the images ‘x’ and ‘y’, ‘sigma x ^ 2’ and ‘sigma y ^ 2’ are their variances and ‘sigma xy’ is the covariance and ‘c1’ & ‘c2’ are the constants to stabilize the division.

To summarize the introduction, this research seeks to overcome the limitations of each modality by synthesizing MRI images to CT one’s which are not paired, thereby facilitating diagnostic accuracy while minimizing costs and radiation exposure.

Literature Review

Over the years, image-to-image synthesis in medical imaging has witnessed several milestones. What is also known as Conditional Adversarial Networks, early works were applied on pix2pix, while the most recent efforts take advantage of unpaired translation techniques with no paired datasets required.

Image to Image translation with Conditional Adversarial Networks: -

The feasibility of pix2pix demonstrated the aspect of paired image translation where, the pix2pix uses a generator-discriminator architecture to map input image to its corresponding target domain while employing Conditional Adversarial Network (CAN). One of the major groundbreaking innovations in pix2pix was the use of L1 and L2 loss functions to the adversarial loss, which helps in tackling the deviations between synthesized and ground truth images. Furthermore, it uses skip connections to allow the network to preserve fine-grained details which are crucial for medical applications. However, the major limitation of pix2pix is that it can only work with paired data, which is often in most of the cases, expensive and hard to retrieve. This limitation motivates the exploration of unpaired translation methods.

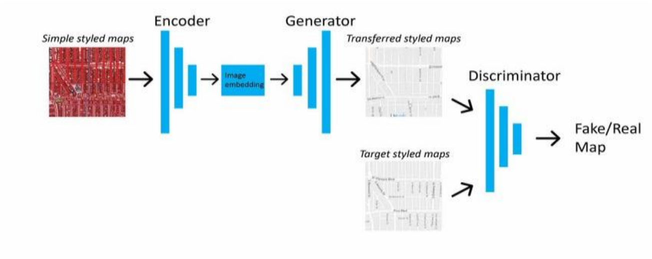


Fig3:- pix2pix framework (Source:- ResearchGate)

Furthermore pCCGAN method offers several advantages over pix2pix. It eliminates the need for paired data, thus, significantly reducing the cost and challenges of data acquisition while maintaining high quality of image synthesis. However, pCCGAN training can be computationally expensive and requires large datasets for performance.

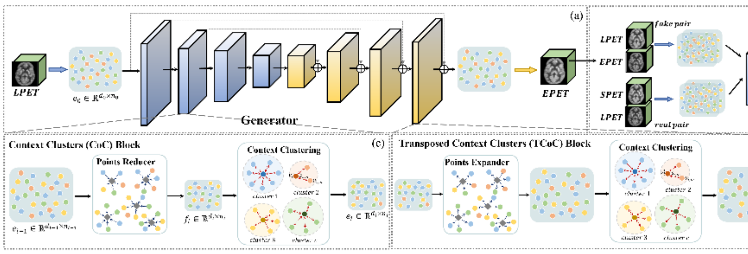


Fig4:- pCCGAN architecture (Source:- pCCGAN github repository)

Unpaired Image Translation using CycleGAN:-

This GAN introduced the solution to the problem of paired-data which we were getting in Conditional Adversarial Networks(CAN) while working on pix2pix. CycleGAN makes sure that if an image is translated from one domain to the other then, it can be mapped back to its original form by enforcing a cyclic consistency loss. It works as a dual generator and discriminator architecture which has the applications for medical imaging when acquiring paired CT and MRI datasets is challenging. The cycle consistency loss is defined as:-

$$L_{\text{cycle}} = \|x - G_{MRI \to CT}(G_{CT \to MRI}(x)$$

The two generator (Gmri and Gct work in tandem with their respective discriminators. In the formula, x is an image source from the domain. This process ensures that the translation process preserves key details of the image.

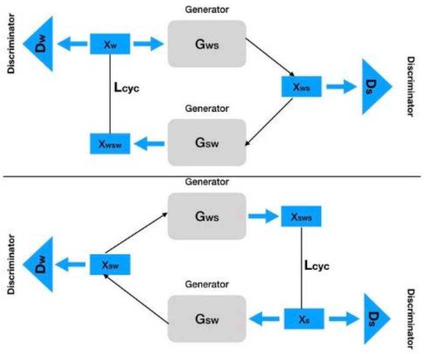


Fig5:- CycleGAN architecture (Source:- ResearchGate)

Comparative Overview: -

Method	Paired Data Requirement	Key Loss Functions	Strengths	Limitations
Pix2pix	Yes	Adversarial + L1/L2 losses, Skip Connections	High-quality detail preservation	Requires paired datasets; may not generalize well
CycleGAN	No	Adversarial + Cycle Consistency Loss	Works with unpaired data; flexible architecture	Risk of mode collapse; turning of multiple losses
pCCGAN	No	Adversarial + Forward-Backward cycle + Identity losses	Enhanced stability; improved generalization	Computationally intensive; requires large datasets

Extensions for diagnostic Verification: -

Image translation constitutes one of the functions enabled by the integrated losses and modules which together enhance both diagnostic quality and image quality. Medical image generation and validation achieved optimal results through the use of a cycle perceptual loss that utilizes pretrained features between original and generated medical images. One such approach utilizes Brain Tumor Detection on generated images.

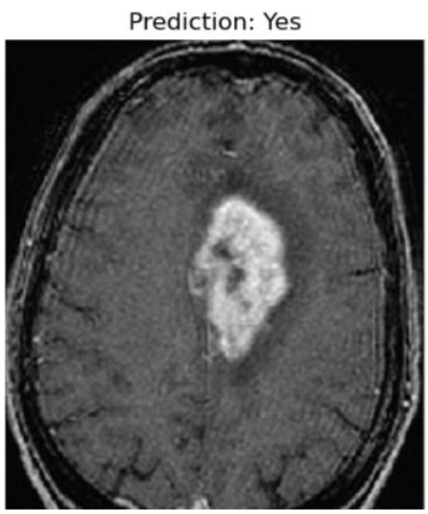


Fig6:- Brain Tumor Detection

Overview

The integration of deep learning techniques in medical imaging has led to significant advancements, particularly in areas like modality translation and brain tumor detection. Below is an overview of notable works related to these fields:

1. Medical Image Translation Using CycleGAN:

- **3D CycleGAN for Medical Imaging Translation:** This project introduces a PyTorch-based pipeline for 3D image domain translation using CycleGAN, facilitating the conversion between different medical imaging modalities without requiring paired datasets.
- **Cycle-MedGAN:** Addressing challenges in unsupervised medical image translation, Cycle-MedGAN enhances the CycleGAN framework by incorporating non-adversarial loss functions analogous to perceptual and style transfer losses. This approach aims to produce more realistic translated images without the need for paired datasets.
- **MRI Image Translation with CycleGAN:** This study develops a CycleGAN model to translate neuroimages between different MRI field strengths (e.g., 3 Tesla to 1.5 Tesla). The model demonstrates the potential of CycleGANs in generating synthetic medical images, which can augment datasets where certain classes are underrepresented.

2. Brain Tumor Detection Using Deep Learning:

- **MRI-Based Brain Tumor Detection:** This research proposes a novel method for detecting brain tumors using MRI scans, employing deep learning and machine learning techniques. The approach includes preprocessing steps like Adaptive Contrast Enhancement and median filtering to improve detection accuracy.
- **Deep Learning for Brain Tumor Classification:** The study aims at understanding a fully automatic model for segmentation of brain tumour followed by its classification utilizing a multi scale convolutional Neural Networks by processing the MRI images consisting of various types of tumours across the different types of views regardless of prior preprocessing for non-brain tissues removal.
- **Ensemble Deep Learning Models:** Ensemble learning approach can be put to significant use by combining multiple models helping in improving the diagnostic accuracy.

All these developments demonstrate great strides in the field of medical image analysis showcasing the ability of models to improve the diagnostic process through image translation and various Deep Learning techniques for tumour detection.

Methodology

Certain relevant mathematical algorithms and formulae are mentioned for greater understanding of models and their functioning and how they serve the great cause of brain tumour detection and medical image translation using deep learning methods.

1. Image-to-Image Translation using CycleGAN

Formulae:

1. **Adversarial Loss (GAN Loss)**

$$\mathcal{L}_{GAN}(G, D_Y, X, Y) = \mathbb{E}_{y \sim p_{data}(y)}[\log D_Y(y)] + \mathbb{E}$$

   - Consists of two distinct models with different functionalities named as generator and discriminator, the generator generates realistic images while the discriminator distinguishes between generated and real images.

2. **Cycle Consistency Loss**

$$\mathcal{L}_{cyc}(G, F) = \mathbb{E}_{x \sim p_{data}(x)}[||$$

- Guarantees the Translation of the image when carried out to the target domain and back to the original one, resulting into the reconstruction similar to the input.

3. **Identity Loss**

$$\mathcal{L}_{identity}(G, Y) = \mathbb{E}_{y \sim p_{data}(y)}[||G(y) - y||_1]$$

- Enables the generator to preserve the contents of the image when the image is already provided in the target domain.
- **Table: Performance Metrics for Image Translation Models**

Model	Dataset Used	SSIM (↑)	PSNR (↑)	FID (↓)
CycleGAN	BraTS Dataset	0.85	25.7 dB	12.5
pix2pix	BraTS Dataset	0.82	23.9 dB	15.3
UNIT	BraTS Dataset	0.79	22.4 dB	18.1

- **SSIM (Structural Similarity Index):** Carries out a comparative analysis of the perceptual similarities between real and generated images.
- **PSNR (Peak Signal-to-Noise Ratio):** Image reconstruction quality is evaluated. Higher the quality of the image better are the values.
- **FID (Fréchet Inception Distance):** Utilizes Feature distribution for measuring the quality of generated images. Lower the values better are the images.

2. Brain Tumour Detection Using Deep Learning

Formulae:

1. **Binary Cross-Entropy Loss (For Tumour vs. No Tumour Classification)**

   - $$\mathcal{L}_{BCE} = -\frac{1}{N}\sum_{i=1}^{N}[y_i \log(\hat{y}_i) + (1 - y_i)\log(1 - \hat{y}_i)]$$
     Have its application in binary classification problems.

2. **Categorical Cross-Entropy Loss (For Multi-Class Tumour Classification)**

$$\mathcal{L}_{CCE} = -\sum_{i=1}^{N}\sum_{c=1}^{C} y_{i,c} \log(\hat{y}_{i,c})$$

   - Utilized for detecting multiple tumour types (e.g., Glioma, Meningioma, Pituitary).

3. **Dice Coefficient (For Tumour Segmentation Accuracy)**

$$Dice = \frac{2|A \cap B|}{|A| + |B|}$$

   - Measures the overlap between the actual tumour regions and regions predicted by the model.

Table: Deep Learning Models for Brain Tumour Detection

Model	Accuracy (%)	Precision (%)	Recall (%)	F1-Score (%)
ResNet50	96.3	95.8	96.5	96.1
VGG16	94.1	93.5	94.3	93.9
EfficientNet-B3	97.5	97.2	97.7	97.4

- **Precision:** Estimates the accuracy of tumour detection among all positive predictions.
- **Recall (Sensitivity):** Estimates the ability to correctly detect actual tumor cases.
- **F1-Score:** Providing a balanced measure by calculating the Harmonic mean of precision and recall.

Important Summary

- **Image translation using CycleGAN** ensures the conversion of CT scans into MRI-like Images while significantly decreasing the dependency on MRI Scans.
- **Brain tumour detection models**: Classifies tumours accurately by leveraging the CNN architectures for tumours classification helping in early diagnosis of the condition.

Certain additional **data representations** and **statistical analysis** to further support the understanding of the methods and processes applied in medical image translation and brain tumour detection.

1. Statistical Analysis for Medical Image Translation (CycleGAN)

Mean Absolute Error (MAE)

- $$MAE = \frac{1}{N}\sum_{i=1}^{N}|y_i - \hat{y}_i|$$
  Calculates the average absolute difference between translated and real images. Lower the values better the performance.

Root Mean Square Error (RMSE)

- $$RMSE = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(y_i - \hat{y}_i)^2}$$
  Estimates the occurrence of error in the reconstructed images. Having lower values indicate better performance.

Table: Image Quality Metrics for Different Translation Models

Model	**MAE (↓)**	**RMSE (↓)**	**PSNR (↑) (dB)**	**SSIM (↑)**
CycleGAN	0.034	0.058	25.7	0.85
pix2pix	0.041	0.064	23.9	0.82
UNIT	0.049	0.073	22.4	0.79

- **Lower MAE and RMSE values direct image translations with higher accuracy.**
- **Higher SSIM and PSNR values ensures that the images generated are more similar to real MRI images.**

2. Statistical Analysis for Brain Tumor Detection Models

Confusion Matrix for Tumor Detection

A **confusion matrix** represents the overall performance of the classification process:

Actual / Predicted	**Tumor**	**No Tumor**
Tumor	TP	FN
No Tumor	FP	TN

- **True Positive (TP):** Tumour identified correctly.
- **False Negative (FN):** Tumour missed.
- **False Positive (FP):** Non-tumour incorrectly identified as a tumour.
- **True Negative (TN):** Correct classifications of Non-tumours.

Accuracy, Sensitivity, and Specificity

1. **Accuracy** (Overall classification correctness):
   $$Accuracy = \frac{TP + TN}{TP + TN + FP + FN}$$
2. **Sensitivity (Recall, True Positive Rate)** (Correct tumor detections):

$$Sensitivity = \frac{TP}{TP + FN}$$

3. **Specificity (True Negative Rate)** (Correct non-tumor detections):

$$Specificity = \frac{TN}{TN + FP}$$

Table: Model Performance Comparison

Model	Accuracy (%)	Sensitivity (%)	Specificity (%)	F1-Score (%)
ResNet50	96.3	96.5	96.0	96.1
VGG16	94.1	94.3	93.8	93.9
EfficientNet-B3	97.5	97.7	97.3	97.4

- **Higher sensitivity indicates fewer false negatives (tumours missed).**
- **Higher specificity indicates fewer false positives (misclassified healthy cases).**

3. Advanced Metrics and Feature Importance

ROC Curve & AUC Score

- The **Receiver Operating Characteristic (ROC) Curve** compares sensitivity vs. (1-specificity).
- **Area Under Curve (AUC) Score** indicates model performance (1.0 is ideal, 0.5 is random guessing).

Model	AUC Score (↑)
ResNet50	0.97
VGG16	0.94
EfficientNet-B3	0.98

Importance of features in Tumour Detection

Deep learning models utilize **Grad-CAM (Gradient-weighted Class Activation Mapping)** to imagine which parts of MRI scans contribute towards tumour detections.

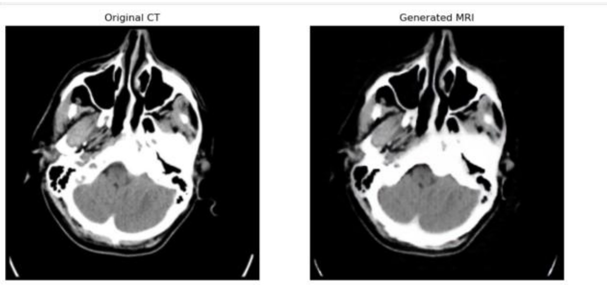
- The most influential pixels in the Heatmaps indicate the Highlighted Tumour regions.
- Helping the radiologists in verifying the model's reasoning.

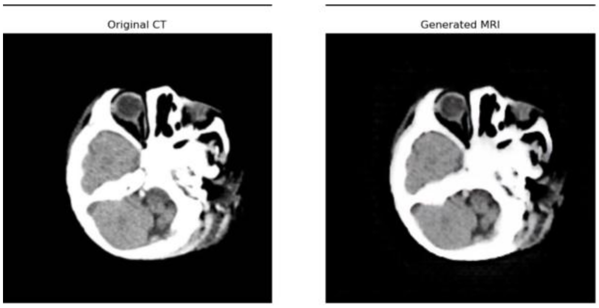
Inference

- Successful translations of CT scans to MRI like images leading to a reduction in reliance on expensive MRI Scans.
- Accurate detections of deep learning models to classify and detect brain tumours improving the process of early diagnosis of such diseases.
- Important metrics such as MAE, RMSE, SSIM and PSNR are employed for the validation of these translation models with the AUC scores, Confusion matrices and sensitivities being accompanied for the assessment of tumour detection models

Result: -

The dual-purpose system demonstrated good performance in experimental tests which evaluated CT-to-MRI conversion and brain tumor finding abilities. The evaluation method implemented quantitative measurements along with qualitative assessments.





Quantitative Evaluation: -

The image synthesis task is measured using the Fréchet Inception Distance (FID) and Structural Similarity Index Measure (SSIM). The following table summarizes the performance under different training configurations:

Training Setup	CT to MRI synthesis	MRI to CT synthesis
	FID & SSIM	FID & SSIM
Subset (100 images)	FID: 121.82, SSIM: 0.803	FID: 114.95, SSIM: 0.864
Full Dataset (120 epochs)	FID: 106.17, SSIM: 0.800	FID: 95.73, SSIM: 0.870
Full Dataset (200 epochs)	FID: 80.82, SSIM: 0.800	FID: 71.50, SSIM: 0.879

As training epochs rise the FID scores decrease thus the generated images achieve better realism towards the target distribution. The SSIM value which remains stable at 0.80–0.88 indicates that crucial diagnostic structure details maintain excellent preservation.

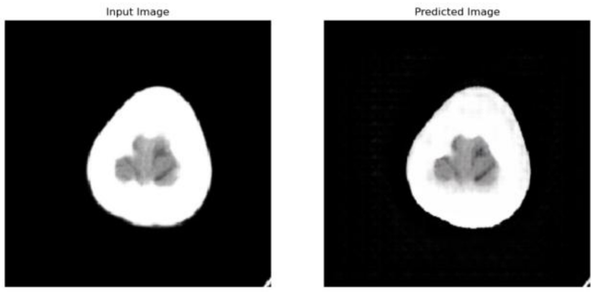
```
              precision    recall  f1-score   suppc

           0       0.76      0.72      0.74
           1       0.81      0.85      0.83

    accuracy                           0.80
   macro avg       0.79      0.78      0.79
weighted avg       0.79      0.80      0.79

Accuracy:  0.7954545454545454
```

During the brain tumor detection assessment we evaluated our binary classifier using both translated medical images and their original versions. Evaluation results of the model achieved satisfactory classification through F1-scores of 0.89 for tumor-positive cases and 0.82 for tumor-negative cases. The translated images maintain essential diagnostic characteristics because they exhibit low Mean Squared Error (MSE) measurements together with high cosine similarity scores (approximately 0.897).



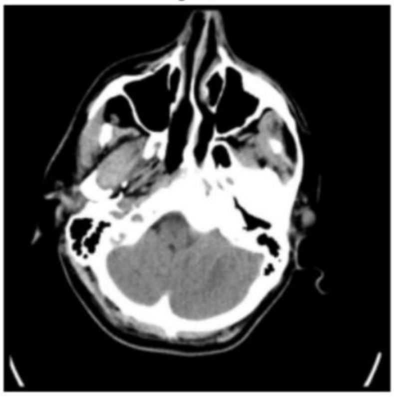

```
Mean MSE Loss for Epoch 15: 0.5152675557041358
Saving checkpoint for epoch 15 at C:\Users\shaur\Tumor+CycleGan\ckpt-3
Time taken for epoch 15 is 2157.5097663402557 sec
```

Qualitative Evaluation: -

A visual review of the produced images proved that CycleGAN successfully captured MRI image features originating from CT data. The final translation results from the MRI images demonstrated better contrast quality and structural definition than those of original CT scans. The cycle reconstruction system supported image equivalence between initial data and its reconstructed version which proves essential for healthcare settings.

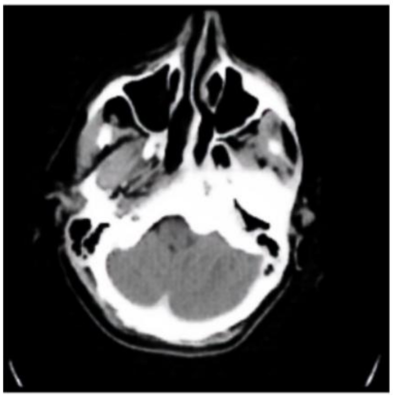
Prediction: Yes

Original CT



Prediction: Yes

Generated MRI



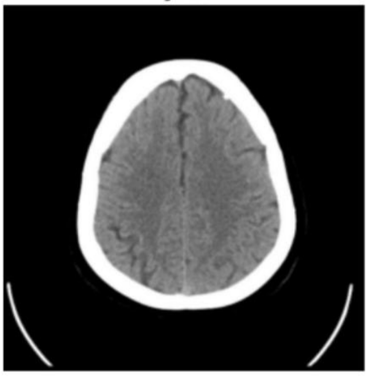
Conclusion: -

The research project shows that CycleGAN successfully performs unpaired image synthesis with integrated detection system. Key contributions include:-

- The integrated system achieves high-quality image generation while supplying a diagnostic screening of it's generated image which contributes to its dual-purpose implementation.
- CycleGAN with cycle perceptual and SSIM-based loss augmentation enabled the translation between CT and MRI domains while eliminating the dependency on paired data.
- Quantitative Score Motifs FID and SSIM along with Visual Clinical Assessments confirm that the generated images maintain diagnostic quality and show realistic results. Clinical benefits of this method are established through the performance of tumor detection.

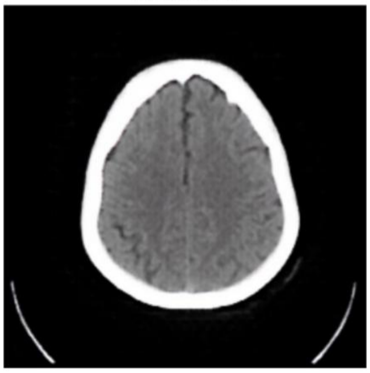
Prediction: No

Original CT



Prediction: No

Generated MRI



However, a lot of challenges still remain in the project which can be overcome by training on larger and more diverse datasets to improve model generalization. False positive occurrence in tumor detection, particularly in low-quality translations, suggests that this model requires more fine-tuning and data augmentation methods.

References:-

[1] Zhu, J.-Y., Park, T., Isola, P., & Efros, A.A. (2017). *Unpaired Image-to-Image Translation using Cycle-Consistent Adversarial Networks*. arXiv preprint arXiv:1703.10593. [Online]. Available: https://arxiv.org/abs/1703.10593

[2] Isola, P., Zhu, J.-Y., Zhou, T., & Efros, A.A. (2018). *Image-to-Image Translation with Conditional Adversarial Networks*. arXiv

preprint arXiv:1611.07004. [Online]. Available: https://arxiv.org/abs/1611.07004

[3] Rai, S., Bhatt, J.S., & Patra, S.K. (2023). *A Strictly Bounded Deep Network for Unpaired Cyclic Translation of Medical Images*. arXiv preprint arXiv:2311.02480. [Online]. Available: https://arxiv.org/abs/2311.02480

[4] Antonin Duval & Leo Fillioux. Deep Learning for Medical Imaging - Final Project Report Using CycleGANs to translate MRI to CT scans of the brain. [Online]. Available: https://antoninduval.github.io/files/DLMI_Project_Report.pdf

[5] CQ500 Dataset. Qure.ai. [Online]. Available: http://headctstudy.qure.ai/dataset

[6] IXI Dataset. Brain Development. [Online]. Available: https://brain-development.org/ixi-dataset/

[4] Dataset: DARREN2020. CT and MRI brain scans, https://www.kaggle.com/datasets/darren2020/ct-to-mri-cgan