USC Viterbi
School of Engineering

Multimodal Medical Image Synthesis for Underrepresented Populations

EE 641 – FINAL PROJECT

Neelakandan Abirami, Nissanth
Devavaram Jebaraj, Sam

Fall 2024

PROJECT SUMMARY

MOTIVATION

Medical imaging datasets, particularly for underrepresented populations (e.g., Sub-Saharan Africa), are insufficient to train robust diagnostic models. Existing diagnostic tools trained on biased datasets lack generalizability across diverse demographics, leading to poor performance in real-world settings.

PROBLEMS

Glioma MRI Scans: Limited availability of multi-modal glioma MRI datasets (T1c, T1n, T2, T2-FLAIR) for underrepresented populations.

Challenges: Ensuring spatial and contextual coherence during synthetic image generation.

PROPOSED SOLUTION

Use different versions of **StyleGAN 2** with ADA(*Adaptive Discriminator Augmentation*) to incorporates dynamic style conditioning.

LITERATURE SURVEY

StyleGAN (Karras et al., 2019):
Introduced a style-based generator using AdaIN, enabling fine-grained control over generated image details (e.g., texture, colors) via a style-focused latent space.

Conditional GAN (Mirza & Osindero, 2014):
Extended GANs with conditional inputs like labels to enable class-specific, targeted generation through a domain-guided conditioning mechanism.

BraTS Challenge (Menze et al., 2015):
Provided a benchmark dataset of multi-parametric MRI for glioma segmentation (LGG, HGG), establishing a standard for evaluating models in medical imaging.

3D GANs (Wu et al., 2016):

Pioneered 3D object generation by extending GANs to synthesize volumetric voxel grids, modeling probabilistic latent spaces for realistic 3D shapes.

TIMELINE

- **November 11:** Project Proposal
- **November 20:** Made the first custom StyleGAN without ADA to test on MNIST dataset - Worked
- **November 26:** Trained the same on a fixed axial slice of T2W modality of all training subjects, with a lower resolution of (32x32x1) - Worked
- **December 2:** Upgraded the architecture to accommodate a 3D MRI scan - Trained with a lower resolution (32x32x32) - Result wasn't convincing
- **December 4:** Reverted back to 2D slices but increased the resolution to 128x128x1 and used A100 for training.
- **December 9-11:** Tested different version including replacement of Adaptive Instance Normalization with Weight modulation. Use of Feature quantization.
- **December 10:** Prepare project documentation
- **December 13:** Submit the completed project.

DATASET

BraTS-Africa 2024

Source: The Cancer Imaging Archive (TCIA) - BraTS-Africa Collection

- **Modalities**: Multi-parametric MRI scans:
  - T1-weighted (T1)
  - Post-contrast T1-weighted (T1Gd)
  - T2-weighted (T2)
  - T2-FLAIR (Fluid Attenuated Inversion Recovery)

- Includes annotated scans for **Low-Grade Gliomas (LGG)** and **High-Grade Gliomas (HGG)**
- Useful for training models for **segmentation, classification**, and **synthetic data generation**.

DATA PREPROCESSING & MODELING

PREPROCESSING - 2D StyleGAN

- **Normalization:**

  Intensity values of slices are rescaled to [0,1][0, 1][0,1] using matplotlib for visualization.

- **Padding:**

  Slices are padded to a uniform size of 256x256×256, adding an 8-pixel border for consistency.

- **Filtering:**

  Slices are selected if more than 25% of T1C pixels have an intensity > 50, ensuring relevant data is retained and saved in grayscale.

PREPROCESSING - 3D StyleGAN

- **Normalization:**

  Rescaled intensities to [-1, 1] for model compatibility.

- **Volume Resampling:**

  Standardized spatial resolution to 128×128×128 for uniform input.

- **Data Augmentation:**

  Applied transformations with torchio to enhance variability:

  - Random Affine: Scaling and rotation.
  - Elastic Deformation: Mimics realistic distortions.
  - Random Noise: Adds Gaussian noise for robustness.

MODEL - 3D StyleGAN

- **Input**:

  Random noise vector z (latent space)

- **Latent Mapping Network**:

  Combines z and labels to produce the **style vector** w.

- **Generator3D**:

  Key Components:

  - **Latent Space Mapping**: Maps input noise z and clinical labels into a **style vector**.
  - **Adaptive Instance Normalization (AdaIN)**:
    - Dynamically modulates feature maps using the style vector.
    - Controls style at multiple convolutional layers for fine-grained details.
  - **Residual Blocks (3D)**: Maintain spatial consistency across volumetric data.
  - **Noise Injection**: Adds stochastic variation for texture diversity.

- **Discriminator3D**:

  Composed of **3D convolutional layers** with **Leaky ReLU activations**.

MODEL - 2D StyleGAN

- Same as the precious architecture in 2D, but with improved layers and usage of CUDA (adapted from stylegan2-ada-pytorch/_pytorch.py at master · Di-Is/stylegan2-ada-pytorch · GitHub), this includes:
    - Adaptive Discriminator Augmentation
    - FID Calculation: Periodically calculates Fréchet Inception Distance for evaluation.
    - Style Mixing: Supports style mixing regularization.

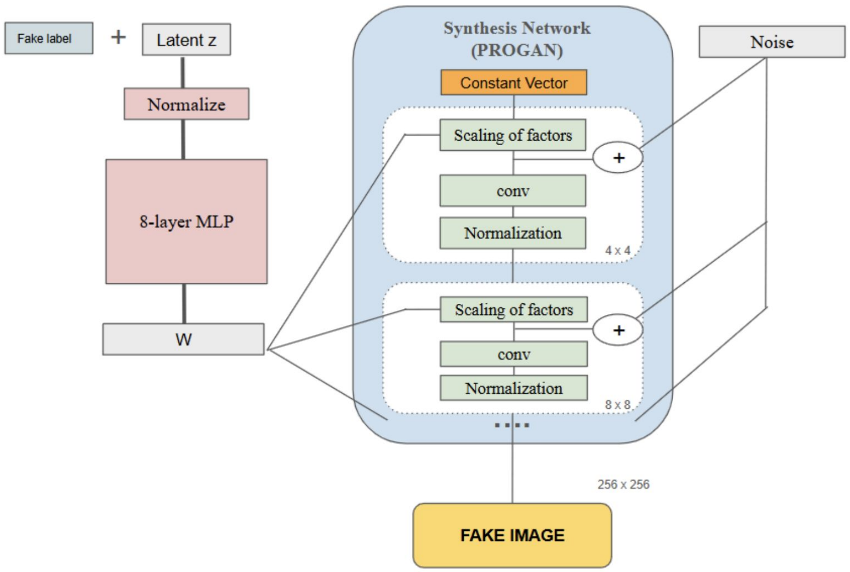


Figure 1: StyleGAN Generator

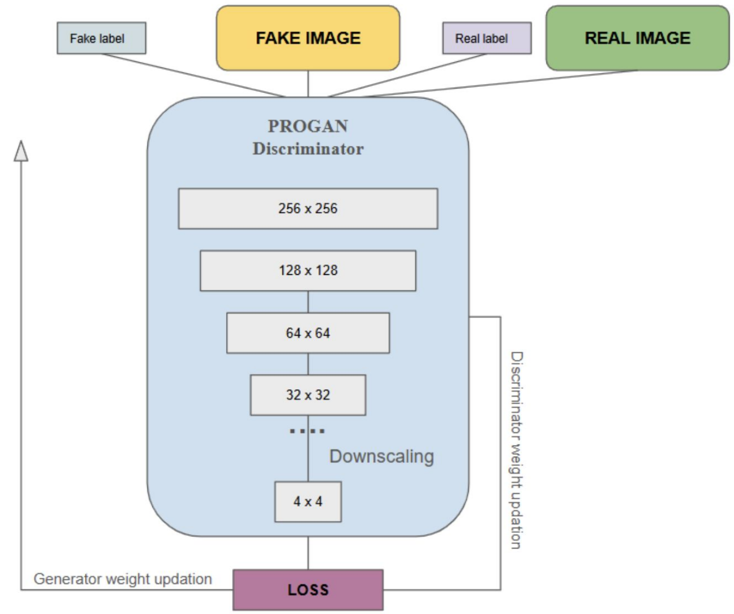


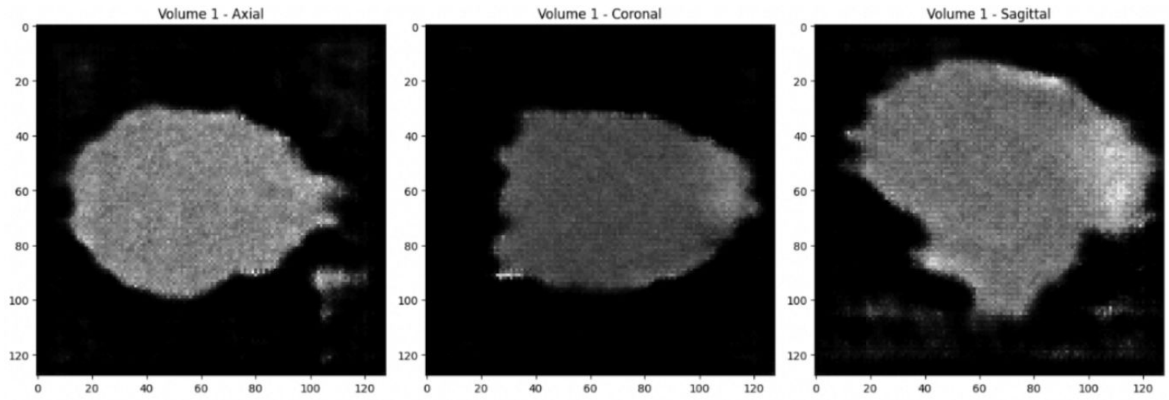
Figure 2: StyleGAN Discriminator

Discriminator Loss: $\mathcal{L}_D = -\mathbb{E}_{x \sim P_{\text{real}}}[\log D(x)] - \mathbb{E}_{z \sim P_z}[\log(1 - D(G(z)))]$

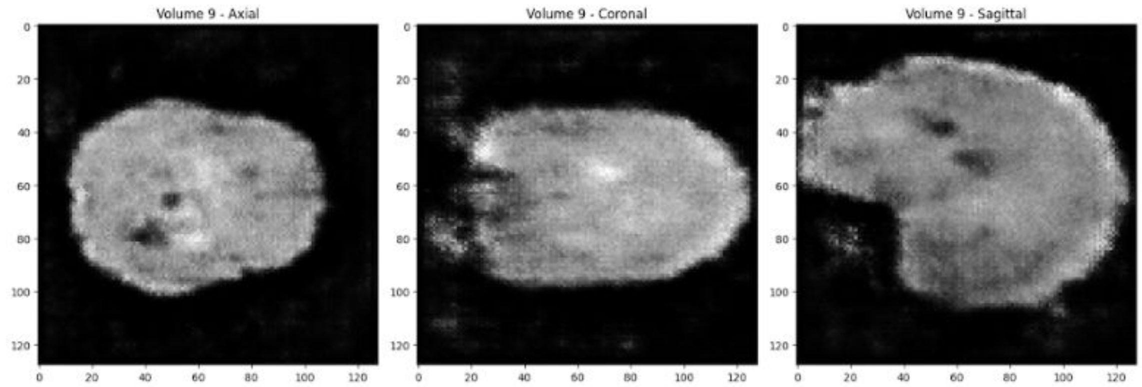
Generator Loss: $\mathcal{L}_G = -\mathbb{E}_{z \sim P_z}[\log D(G(z))]$

RESULTS

USC Viterbi
School of Engineering



StyleGAN 3D - Results
(As of Status report)



StyleGAN 3D - Best results

StyleGAN 2 - Different versions

Attention Mechanism:

Enhances the generator and discriminator by focusing on relevant parts of the image, helping to extract more meaningful features from fewer images, which improves image quality.

ADA:

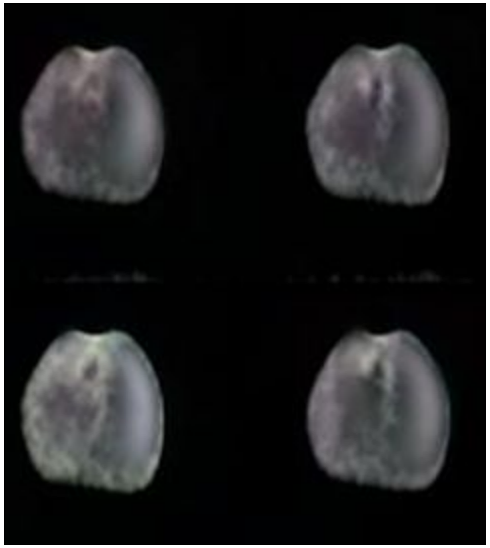
(Adaptive Discriminator Augmentation) is a technique used to improve training efficiency, especially when there is a limited dataset. ADA allows StyleGAN2 to train effectively with as few as 1,000-2,000 high-quality images by applying dynamic data augmentations to the discriminator during training.

Top-k Training for Generator:

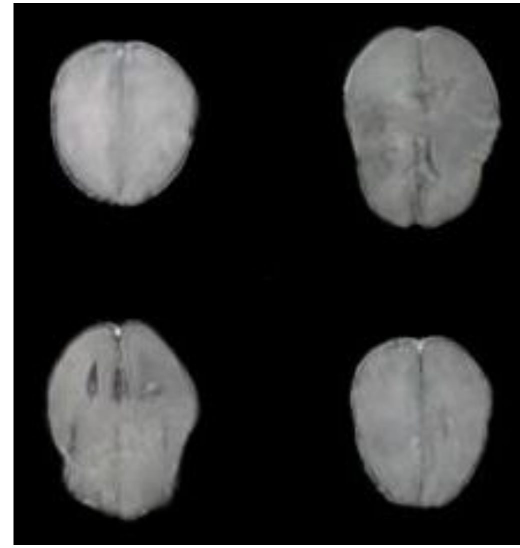
Focuses the generator on high-quality images by zeroing out gradients from fake images that the discriminator identifies, accelerating the generator's improvement, especially when data is limited.

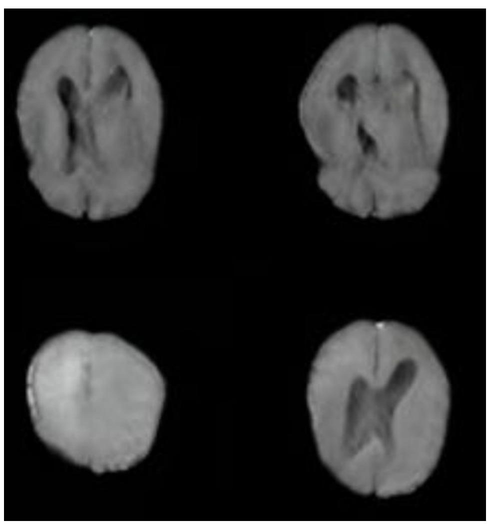
Feature Quantization:

Involves quantizing the discriminator's intermediate features, which can stabilize and regularize the model, aiding generalization in limited data scenarios.

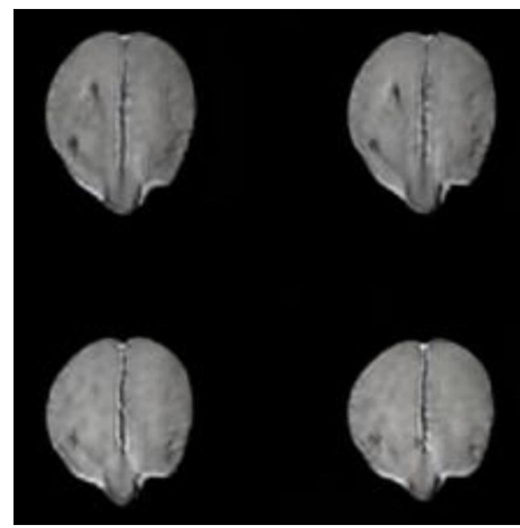


Epoch 1000 - Epoch 10000:
StyleGAN 2D

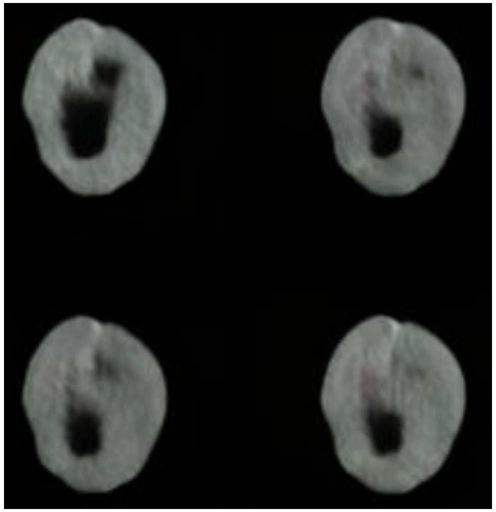




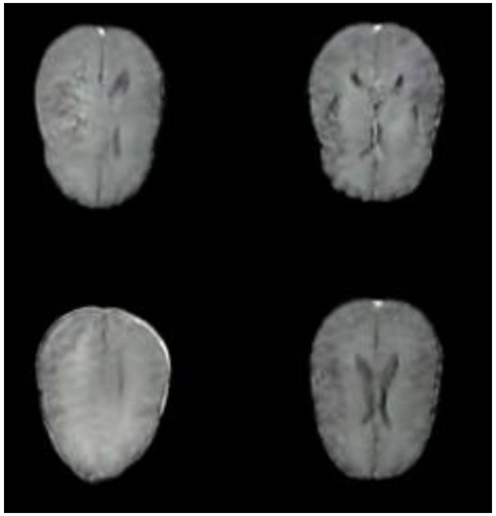
Best evaluation: StyleGAN 2D
with ADA



Best evaluation: StyleGAN 2D
with Attention

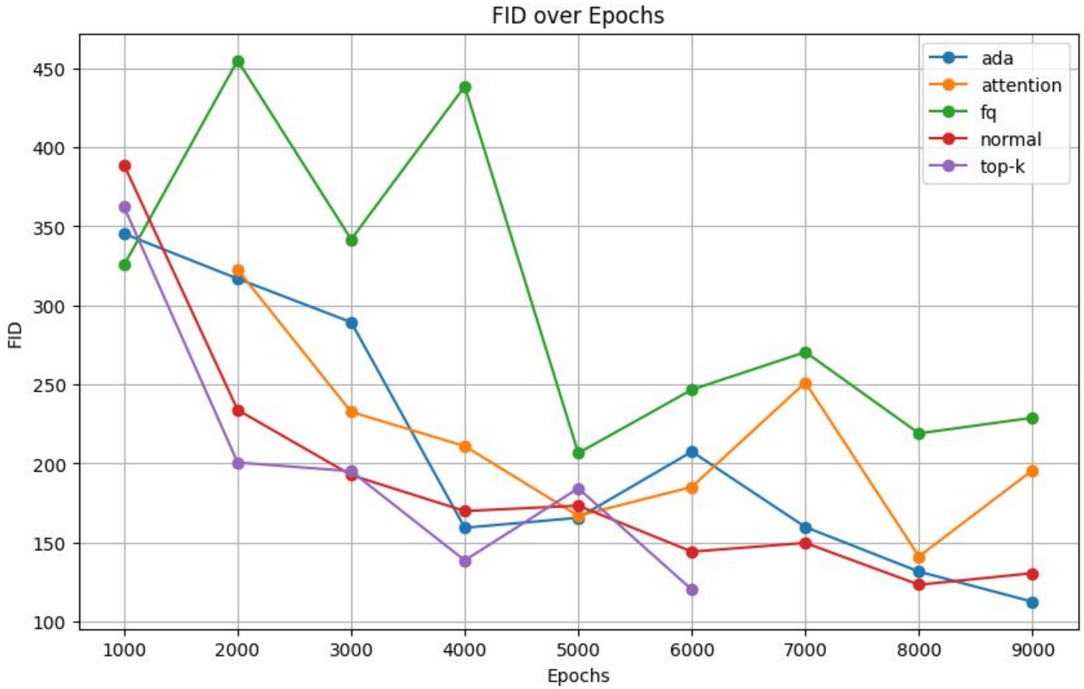


Best evaluation: StyleGAN 2D with Feature Quantization



Best evaluation: StyleGAN 2D with Top-K

FID Scores



The Fréchet Inception Distance (FID) score is the metric we used to evaluate the quality of images.

$$\text{FID} = \|\mu_{\text{real}} - \mu_{\text{gen}}\|^2 + \text{Tr}(\Sigma_{\text{real}} + \Sigma_{\text{gen}} - 2(\Sigma_{\text{real}}\Sigma_{\text{gen}})^{\frac{1}{2}})$$

OUTCOMES AND RESULTS

Key Contributions:

1. Developed a **StyleGAN** to generate realistic and spatially coherent MRI images.
2. Addressed the data scarcity of glioma MRI scans from **underrepresented populations** (Sub-Saharan Africa).
3. Integrated **Adaptive Instance Normalization (AdaIN) and further** for dynamic style control based on clinical metadata.
4. Improved image fidelity and diversity with techniques like **noise injection** and **residual blocks**.

Outcomes:

- Improved **data diversity**, potentially aiding in better diagnostic model training and generalizability through GANs.
    -

INCOMPLETE AND DESCOPED FEATURES

Feature Not Implemented:

- **3D StyleGAN:**

  Training a 3D StyleGAN required substantial computational resources. Also due to time constraints, we couldn’t reach the necessary goal as training was performed on smaller batch sizes and limited epochs.

Impact:

- The specific slice (256x256x1) with more than 75% brain matter will be generated rather than a a complete 3D volume. (128x128x128)