EE-641 - Deep Learning Systems

Group 4: Nissanth Neelakandan Abirami, Sam Devavaram Jebaraj

Title: Multimodal Medical Image Synthesis for Underrepresented Populations via Conditional StyleGAN

Project Objective

The project aims to develop a Conditional StyleGAN architecture to synthesize medical images from underrepresented populations, specifically glioma MRI scans from Sub-Saharan Africa. This approach addresses the scarcity of medical imaging data to enhance diagnostic model training and generalizability.

Project Timeline

1. **Week 1:** Data acquisition and preprocessing.
2. **Week 2:** Environment setup and Conditional StyleGAN implementation.
3. **Week 3:** Training and refining the Conditional StyleGAN model.
4. **Week 4:** Model evaluation, fine-tuning, and documentation.

Data Summary

Data Source Description

The data comes from the BraTS-Africa Challenge, comprising multi-parametric MRI scans, including:

- T1-weighted (T1)
- Post-contrast T1-weighted (T1Gd)
- T2-weighted (T2)
- T2 Fluid Attenuated Inversion Recovery (T2-FLAIR)

Data Size and Distribution

The dataset includes glioma images from adult patients, covering low-grade gliomas (LGG) and glioblastomas (HGG).

Data Cleaning and Preprocessing Steps

- Normalized image intensities to a range of $[-1,1]$.
- Adjusted volumetric dimensions for consistent input sizes.

- Preprocessed clinical labels for conditioning.

Feature Engineering Techniques

- One-hot encoding for MRI slice types.
- Integration of latent style vectors to conditionally synthesize images.

Model Development

Chosen Model Architecture

The initial proposed architecture used StyleGAN with enhancements for conditional generation based on clinical metadata. The current code utilizes:

- ➔ Generator3D: A volumetric extension of StyleGAN with AdaIN layers.
- ➔ Discriminator3D: A volumetric discriminator for authenticity validation.
- ➔ Mapping Network: A shared network to condition the latent space on clinical labels.

Training Parameters

- ➔ Batch size: 4
- ➔ Learning rate:
    - ◆ Generator: 4×10^−4
    - ◆ Discriminator: 1×10^−4
    - ◆ Optimizer: Adam
- ➔ Epochs: 100
- ➔ Loss Function: LSGAN loss for generator and discriminator.

Hyperparameter Tuning Process

The model architecture was fine-tuned with:

1. Latent space parameterization for smoother style conditioning.
2. Adjusted noise injection to enhance realism.
3. Modifications to the generator's intermediate layers to ensure volumetric data compatibility.

Changes in Proposed and Current Architecture

Proposed Architecture (Proposal Document)

- ➔ Utilized a 2D StyleGAN generator with conditional inputs based on clinical metadata.

- ➔ Generated slices independently for Axial, Sagittal, or Coronal views.
- ➔ Focused on integrating 2D data for grayscale images of size 256×256x256

Current Architecture (Code)

- ➔ Transitioned to a 3D volumetric StyleGAN generator.
- ➔ Conditioned volumetric generation using AdaIN layers for better spatial consistency.
- ➔ Included noise injection at each convolutional layer for fine-grained stochastic details.
- ➔ Adapted discriminator to evaluate 3D volumes instead of 2D slices.

Rationale for Changes

- The current approach improves the spatial and contextual coherence of MRI volumes by generating entire 3D data blocks instead of isolated 2D slices.
- AdaIN layers allow for more dynamic style control, aligning with the dataset's multimodal nature.
- Noise injection at multiple levels provides increased texture diversity, mimicking natural medical images

Evaluation Metrics

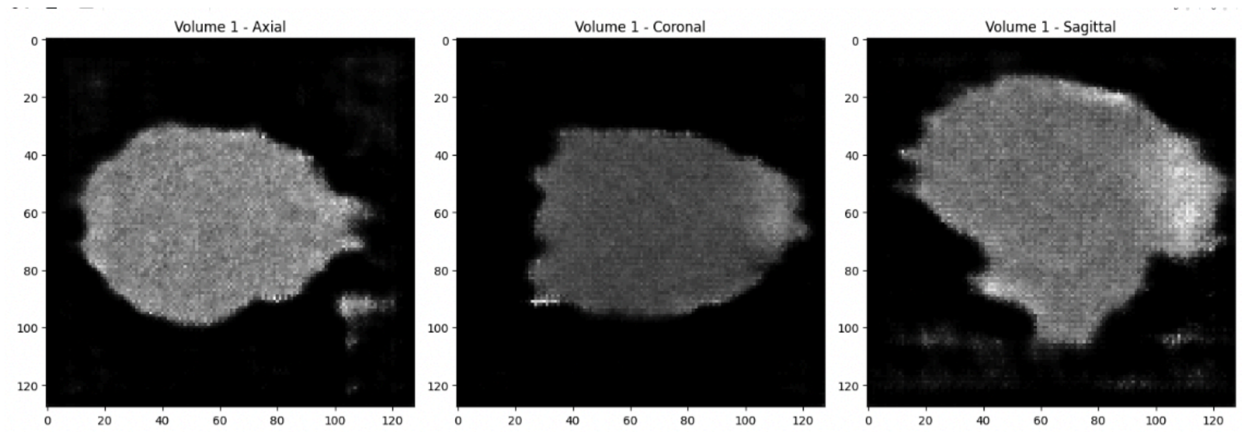
Primary Evaluation Metric(s):

- SSIM (Structural Similarity Index): To measure the structural similarity between real and generated 3D MRI volumes.
- MSE (Mean Squared Error): To evaluate pixel-level reconstruction accuracy.
- FID (Fréchet Inception Distance): To assess the quality and diversity of generated volumes compared to real volumes.

Key Accomplishments This Period

- ➔ **Milestone Achievements:**
    - ◆ Data preprocessing and normalization complete.
    - ◆ Implemented a 3D StyleGAN-based generator and discriminator.
    - ◆ Initial training and visualization of generated 3D MRI volumes completed.
- ➔ **Progress on Data Preparation:**
    - ◆ Successfully loaded and preprocessed the BraTS Africa dataset.
    - ◆ Applied MinMax scaling and resampled volumes to a uniform resolution of 128×128×128.
    - ◆ Visualized sample slices from axial, sagittal, and coronal planes to verify data integrity.
- ➔ **Progress on Model Training:**

- Trained the initial generator and discriminator for 200 epochs.
- Addressed shape mismatch issues and stabilized GAN training using rescaled outputs.
- Balanced generator and discriminator learning rates for improved convergence.



➔ **Progress on Evaluation:**
  - Periodically visualized generated MRI slices to assess training progression.
  - Evaluated generator outputs using SSIM, MSE, and FID metrics.

Next Steps:

- Fine-tune the generator and discriminator to further reduce mode collapse and improve image fidelity.
- Integrate additional regularization techniques (e.g., gradient penalty, spectral normalization).
- Conduct evaluation with segmentation models to verify anatomical realism.
- Improve SSIM and FID scores through additional training and architectural modifications.

Challenges and Risks

1. Data Quality Issues

1. **Challenge:** The BraTS-Africa dataset may have noise, inconsistencies, or missing metadata due to the diversity of medical imaging sources.
2. **Impact:** This could affect the accuracy and generalizability of the generated synthetic images.
3. **Mitigation:**
   a. Apply rigorous preprocessing steps like denoising, normalization, and intensity standardization.
   b. Use data augmentation techniques to enhance dataset variability.

2. Model Overfitting/Underfitting

1. **Challenge:** The generator may overfit to the limited dataset, producing unrealistic outputs, or underfit due to insufficient latent space representation.
2. **Impact:** Overfitting could lead to poor diversity in synthetic images, while underfitting might fail to capture critical tumor characteristics.
3. **Mitigation:**
    a. Implement regularization techniques such as dropout and L2-norm.
    b. Use early stopping based on validation metrics.
    c. Fine-tune hyperparameters, including learning rates and latent dimensions.

3. Computational Limitations

1. **Challenge:** Training a 3D volumetric GAN architecture demands substantial computational resources (e.g., GPU memory and processing power).
2. **Impact:** Limited hardware could result in prolonged training times or restricted model complexity.
3. **Mitigation:**
    a. Utilize gradient accumulation to manage large batch sizes within memory constraints.
    b. Optimize memory usage by reducing intermediate tensor sizes or using mixed-precision training.
    c. Leverage distributed training or cloud-based platforms (e.g., AWS or Colab Pro) for scalability.