EE367 Project Proposal: Accelerated Longitudinal MRI using Conditional Latent Diffusion Models

Yonatan Urman – yurman, Zachary Shah – zshah9

1 Problem Statement and Motivation

Brain Magnetic Resonance Imaging (MRI) studies remain a prevalent medical procedure, and many patients will undergo several MRIs during the course of their life [1]. However, Brain MRI can suffer from long-scan times, motion-related artifacts resulting in sub-optimal image quality, and limited accessibility with long wait times [2]. Consequently, there has been a push to accelerate scanning time with minimal compromise in image quality to ultimately improve patient throughput and accessibility [3]. Interestingly, information from prior MR scans is not utilized during acquisition or reconstruction for subsequent scans. We here propose to include prior patient MR information as part of a conditional latent diffusion model to further accelerate MR scanning. We further aim to add a data consistency step to better incorporate the underlying MR physics into our model and improve reconstruction.

2 Previous and Related Works

Score-based diffusion models with physics-based constraints have recently grown in popularity for medical image reconstruction. Song et al. [4] first tackled the challenge of reconstructing medical images from incomplete measurements by integrating a data-consistency step into the sampling of a score-based generative model, which was trained specifically to model the prior distribution of medical images. In a subsequent work, Song et al. [5] extended this approach, embedding the data consistency step into a latent diffusion model through an intermediate optimization over the forward model within the reverse sampling process.

3 Methodology

The authors of this proposal have some predicate work demonstrating a proof-of-concept of data consistency for prior-informed reconstruction via a light-weight conditional latent diffusion model trained from scratch. However, this proof-of-concept suffers multiple limitations, which we aim to improve as major milestones to our project:

- The trained model cannot well-capture the a high-resolution prior distribution over MRI brain images due to the lightweight model size. As such, the authors propose building a better synthetic brain image generation model as a fine-tune off of a more modern architecture, Stable Diffusion [6], which was trained on the LAION-5B dataset [7] (containing over 5 billion captioned images).
- The trained model had limited ability to capture prior information. We propose using a Controlnet [8] image-to-image architecture to direct the diffusion process. Controlnet is an image-encoding architecture better suited for enforcing learned structural similarity in its samples with a prior image distribution.
- Data consistency as proposed by [5] requires integrating the latent encoder as a part of the data fidelity model to enforce data consistenty in the latent space. Solving this requires stochastic optimization per sample timepoint, as the latent encoding step is highly non-convex operation, requiring extensive compute times to solve latent data consistency. The authors plan to explore methods to accelerate data-consistent LDM sampling, whether it be with fewer or faster latent optimization steps, such as with Posterior Sampling as described by [9].

Dataset: We will utilize longitudinal information from 246 healthy adult volunteers from the OASIS-3 dataset released in 2022 [10] to fine-tune Stable Diffusion and train a Controlnet adapter. We will utilize longitudinal T1-weighted scans obtained at varying time intervals, specifically 2-3 scans for each volunteer. We will generate prior-sample data pairs for 2D image scans, of which we should be
able to generate 60,000+ sample pairs. We additionally plan to prepare a set of retrospective clinical data to test our models and demonstrate viability of this approach on real measurements.

4 Evaluation

To assess the effectiveness of our model’s image reconstruction, we plan to assess the Peak Signal to Noise Ratio (PSNR) and Structural Similarity Index Measure (SSIM) of our model’s reconstructions. However, these metrics tend to fare poorly in assessing the usefulness of a medical image on end-point diagnostic tasks. As such, we hope to explore more perceptual metrics, such as FID and Inception Scores from inception models trained on medical image datasets, as a better objective to our reconstructions.

We will create plots that illustrate the relationship between these metrics and various undersampling rates. Our aim is to demonstrate that the inclusion of longitudinal data as a condition to the LDM enables us to increase the undersampling rate while maintaining high-quality reconstruction results.

5 Other information

- The group consists of students in Prof. Kawin Setsompop’s lab, but we don’t receive any special support for this specific project.

References


