3D Medical Image Reconstruction from Limited Data

Technical Abstract

The project aims to develop robust methods for **3D medical image reconstruction from limited data**, addressing key challenges in data scarcity, reconstruction accuracy, and interpretability for clinical applications. With applications in diagnostics and treatment planning, achieving high-quality 3D reconstructions from limited 2D or sparse input data is critical. Our work is structured around three primary objectives:

Objective 1: Data-Efficient Reconstruction Using Generative and Self-Supervised Models

To mitigate the issue of limited data, we propose implementing **Generative Adversarial Networks (GANs)** and **Variational Autoencoders (VAEs)** that incorporate self-supervised learning strategies. By training the model to generate realistic 3D representations from sparse 2D images or incomplete 3D scans, we aim to improve reconstruction quality while reducing reliance on large datasets. Our model will leverage **VAE-GAN hybrids** that allow for robust learning of 3D structures even in cases of partial or noisy input data, drawing on foundational work by Goodfellow et al. (2014) for GANs and Kingma & Welling (2013) for VAEs.

Objective 2: Enhancing Reconstruction Accuracy through Cross-Domain Transfer Learning

Leveraging **cross-modal and cross-domain transfer learning**, our model will utilize knowledge from synthetic and multi-modal medical images to improve reconstruction accuracy. This approach enables adaptation to diverse data modalities, such as MRI, CT, and ultrasound, enhancing the model's versatility and accuracy. For this purpose, we will use pre-trained networks like **ResNet** and **UNet**, fine-tuning them for 3D reconstruction tasks. This technique will be informed by research such as the transfer learning methodologies discussed in Tajbakhsh et al. (2016) and cross-domain applications by Zhou et al. (2021).

Objective 3: Ensuring Clinical Interpretability with Explainable AI (XAI) Techniques

To facilitate clinical adoption, interpretability is essential. We will integrate **Explainable AI (XAI)** techniques, such as **Layer-wise Relevance Propagation (LRP)** and **Class Activation Mapping (CAM)**, which provide insight into model decision-making processes. By highlighting the model's focal points during reconstruction, these methods can assure clinicians of the model's reliability, ultimately aiding in its acceptance and deployment. XAI techniques will be guided by prior work from Montavon et al. (2017) on LRP and Selvaraju et al. (2017) on CAM.

Implementation Strategy

Our pipeline will be implemented in a staged manner:

1. **Pre-Training Phase**: GANs and VAEs will be trained on synthetic 3D datasets with partial supervision to enhance self-learning capabilities.
2. **Transfer Learning and Cross-Domain Adaptation**: We will integrate pre-trained networks (ResNet, UNet) and fine-tune them for specific medical imaging tasks using limited and noisy medical datasets.
3. **Explainability Layer Integration**: LRP and CAM methods will be implemented to generate interpretable visualizations during the reconstruction process.

References to Foundational Research

- Goodfellow et al. (2014). "Generative Adversarial Networks" – foundational for GAN architecture.
- Kingma & Welling (2013). "Auto-Encoding Variational Bayes" – essential for understanding VAEs.
- Tajbakhsh et al. (2016). "Convolutional Neural Networks for Medical Image Analysis" – foundational work in transfer learning in medical imaging.
- Montavon et al. (2017). "Explaining Nonlinear Classification Decisions with Deep Taylor Decomposition" – essential for interpretability using LRP.
- Selvaraju et al. (2017). "Grad-CAM: Visual Explanations from Deep Networks via Gradient-based Localization" – critical for explainability in medical imaging applications.