Project 4
3D Liver Image Segmentation under Limited Annotation

Project Overview

Accurate liver and liver tumor segmentation from abdominal CT images is a critical task in medical image analysis, supporting applications such as surgical planning, tumor burden assessment, and treatment response evaluation. Despite recent advances in deep learning–based segmentation models, high-quality voxel-wise annotations for liver imaging remain costly, time-consuming, and dependent on expert radiologists, limiting scalability in real-world clinical settings.

This project investigates how liver segmentation performance degrades under limited annotation availability, and how different segmentation architectures cope with varying proportions of labeled training data. By systematically reducing the percentage of labeled samples from the Liver Tumor Segmentation (LiTS) dataset, this study aims to benchmark classical and advanced segmentation models and establish a clear performance–annotation trade-off for liver CT segmentation. The project will first target an international conference and then be extended toward a Q2 journal publication with more advanced learning paradigms.

Research Objectives

- Quantify liver segmentation performance under different proportions of labeled data (e.g., 1%, 5%, 10%, 25%, 50%, 100%).
- Benchmark segmentation models including U-Net, U-Net++, DeepLabV3, TransUNet, and nnU-Net.
- Study model generalization as annotation availability decreases.
- Provide qualitative and quantitative analysis of failure cases under sparse supervision.
- Produce a conference paper and an extended Q2 journal submission.

Methodology

1. Dataset Preparation:
- Use the LiTS dataset with liver and tumor annotations.
- Convert volumes into 2D or 2.5D representations.
- Create multiple training subsets with different annotation ratios.

2. Limited Annotation Simulation:
- Systematically vary labeled data proportions.
- Maintain consistent preprocessing and splits.

3. Model Benchmarking:
- Train CNN- and Transformer-based segmentation models.
- Include nnU-Net as a strong automated baseline.

4. Evaluation Metrics:
- Dice, IoU, Hausdorff Distance, Average Surface Distance.
5. Visualization and Analysis:
- Qualitative overlays.
- Failure case analysis.
- Architecture sensitivity comparison.

Timeline (6 Months)

Phase 1 (Months 1–3): Conference Target
- 1 dataset (LiTS)
- Fully supervised training with limited annotations
- Conference submission (ICISN, CITA, MAPR)
- Compute: Google Colab, Kaggle

Phase 2 (Months 4–6): Journal Target
- Extend to 2 datasets or multiple LiTS splits
- Apply semi-supervised, self-supervised, and few-shot learning
- Q2 journal submission
- Compute: AIVN GPUs, Private servers