Large-scale Medical Semantic Segmentation Pretraining from Automatically Generated Masks
Fachbebiet Digital Health – Machine Learning

Project Supervisors

benjamin.bergner@hpi.de

christoph.lippert@hpi.de

Large-scale Medical Semantic Segmentation Pretraining from Automatically Generated Masks

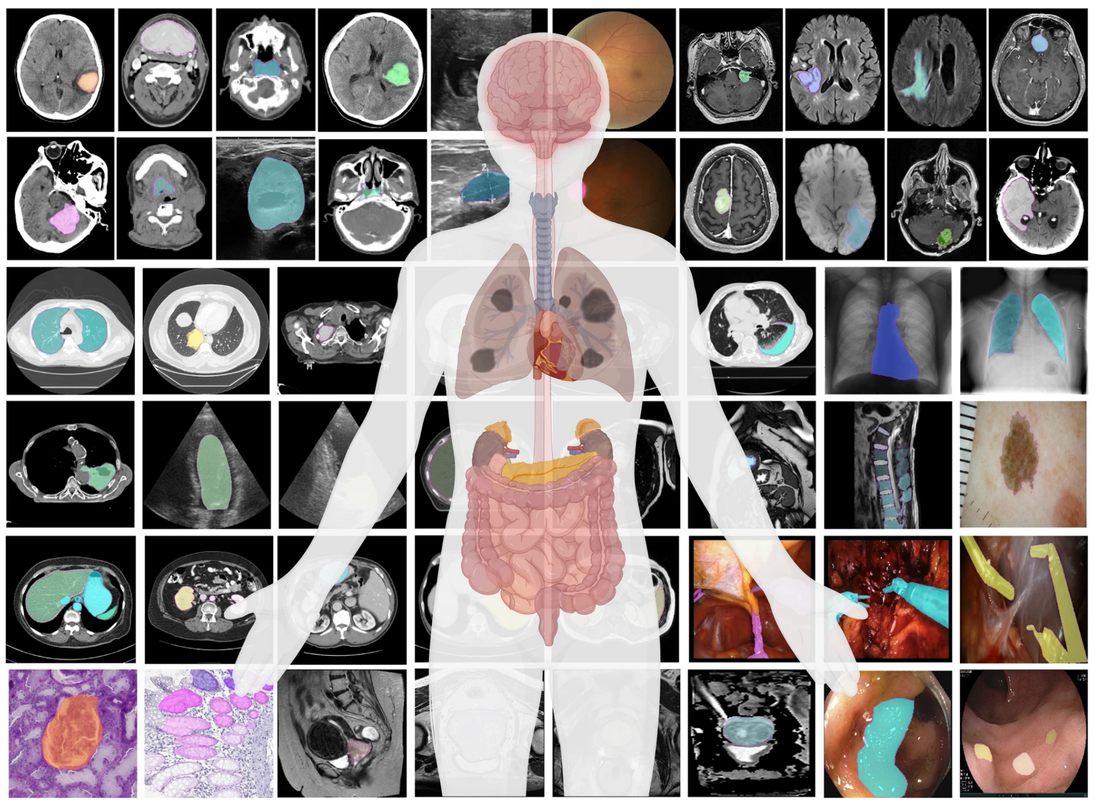
Fachgebiet Digital Health - Machine Learning

Project Overview:

This project aims to leverage recent advancements in foundation models for semantic segmentation to address the scarcity of annotated data in medical imaging. By automatically generating segmentation masks and using these masks for large-scale pretraining of vision transformers, we intend to create a powerful model that can be fine-tuned for various medical imaging tasks. The ultimate goal is to enhance performance on small labeled datasets, where models trained from scratch often struggle.

Introduction and idea:

Annotated data in medical imaging is scarce, as it requires the expertise of medical professionals who are often time-constrained. Recently, foundation models for semantic segmentation, such as SAM and MedSAM, have demonstrated the ability to generate valid segmentation masks from prompts like text, bounding boxes, and points. For instance, using a dense grid of points to prompt these models can produce multiple plausible segmentation masks. This project aims to use these automatically generated masks as labels to pretrain vision transformers on a large-scale dataset like the UK Biobank, which includes various imaging modalities (e.g., T1/T2 MRI) and body parts (e.g., chest, brain, liver). The pretrained model is expected to perform significantly better on small labeled datasets compared to a model trained from scratch.



Hasso Plattner Institut
Digital Engineering · Universität Potsdam

Project Objectives:

1. **Dataset Collection**: Gather a large and diverse dataset for segmentation pretraining.
2. **Mask Generation**: Automatically generate masks using foundation models such as the Segment Anything Model (SAM).
3. **Pretraining**: Conduct large-scale pretraining of vision transformers using the generated masks.
4. **Fine-tuning**: Finetune the pretrained model on various downstream tasks.
5. **Evaluation**: Evaluate your models and compare them to various baselines.

Expected Outcomes:

By the end of this project, students will have developed powerful pretrained segmentation models of various sizes that significantly improve performance on various medical downstream tasks. Besides training and evaluation of these models, we intend to share the automatically generated segmentation masks to the research community.

Learning Outcomes for Students:

- **Transformers:** Gain a deep understanding of transformer architectures and their applications.
- **Programming:** Acquire hands-on experience with PyTorch.
- **Model dev:** Develop, debug, pretrain and fine-tune deep learning models on our clusters.
- **Application:** Learn more about the intriguing field of medical imaging.
- **Research and Writing:** Enhance your skills in academic research and paper writing.

Project Phases:

1. **Identify Segmentation Foundation Models:** Research and select suitable foundation models for medical image segmentation.
2. **Dataset Collection:** Collect and preprocess data to create a large-scale medical imaging dataset.
3. **Automatic Mask Generation:** Automatically generate segmentation masks using the selected foundation models and dataset.
4. **Pretraining:** Pretrain vision transformers on the generated masks.
5. **Evaluation:** Evaluate the pretrained models on various downstream tasks and compare to several baselines.
6. **Report Writing:** Document your findings in a research report.