Research Project in the Data Science Profile

Automated Muscle Segmentation in Medical Imaging Scans

Situation

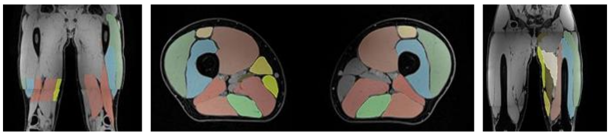

Figure 1: Example MR images of muscles of the upper thigh with sparse segmentation labels (own data)

Muscle-volume analysis is becoming essential for the planning of pre-surgical interventions, pre- and rehabilitation and recovery protocols (e.g., joint replacements), as well as for the design of training programs for performance enhancement or injury prevention. However, a tool for automated muscle volume analysis directly after MRI scanning is still missing. Although numerous deep learning-based auto-segmentation models for medical and muscle imaging have emerged, in practice these automatically generated labels often require manual refinement by experts. This process is particularly challenging because annotations in 3D are time-consuming and are often sparsely available, which limits the scalability and robustness of current approaches.

Goals / Methodology / Tasks

The goal of this project is to develop deep learning segmentation models that effectively exploit data sparsity to enable robust performance under varying forms of semi-annotated training. The specific research questions are:

- Which muscle types are most consistently represented in the datasets and provide a reliable basis for segmentation?
- Which evaluation metrics are best suited for sparsely annotated segmentation data?
- How can positively sparse annotations be utilized to improve segmentation performance without biasing error metrics?
- How does the self-trained model compare to auto-segmentation models (e.g., MuscleMap, TotalSegmentator, SAM/Med-SAM)?
- How do 2D, 2.5D, and 3D models compare in terms of training time, accuracy, and adaptability?

This project uses 3D MRI data from 50 athletes (approx. 704 × 508 × 640 voxels, water and fat channels) acquired at the Swiss Center for Musculoskeletal Imaging, Balgrist Campus. Ethical approval is in place, with expert annotations for 9-13 thigh and hip muscles and about 200 slices per scan. Additionally, public data from ~10 participants (Jouvencel et al., 2022) can be used, though it includes fewer labels and lower-resolution images. Further unlabeled data is available for research purposes from UK-Biobank platform. These datasets present different forms of annotation sparsity that can help develop more robust models. The Balgrist dataset includes annotations for one leg, sometimes incomplete or inconsistent across muscle groups, while the public dataset provides intentionally inner-contour annotations to address boundary uncertainty and artefacts along the third axis. The UK biobank data has no annotations at all.

Required Skills

Interest in medical image segmentation, data analysis (data wrangling, data exploration, visualization).

- Deep learning
- Python

Tasks for the Master Student

The student project can build on a previous student project (U-Net, MRI channel analysis, public dataset). We propose the following project phases, while gradually increasing dataset diversity and sample size:

P7: Establish a 2D segmentation baseline with defined muscle types, a training & evaluation pipeline for muscle/fat channel inputs

P8: Develop and test varying enhanced training strategies that leverage positively sparse annotations for improved segmentation

P9: Integrate all datasets, extend sparse training to 3D, and compare models with existing auto-segmentation tools

Full/Parttime: [x] Full time study [x] Part time study

Location: FHNW Brugg-Windisch

Advisor: Prof. Dr. Susanne Suter

Collaborators: Prof. Dr. Jörg Spörri (Balgrist), Prof. Dr. Daniel Nanz (Balgrist Campus), Dr. Stefan Sommer (Siemens Healthineers)