AI in Health Technologies (ELEC-E8739): Final Project

Knee cartilage Segmentation via soft label regression

Prof. Aleksei Tiulpin, PhD

Fall 2025

1 Project Overview

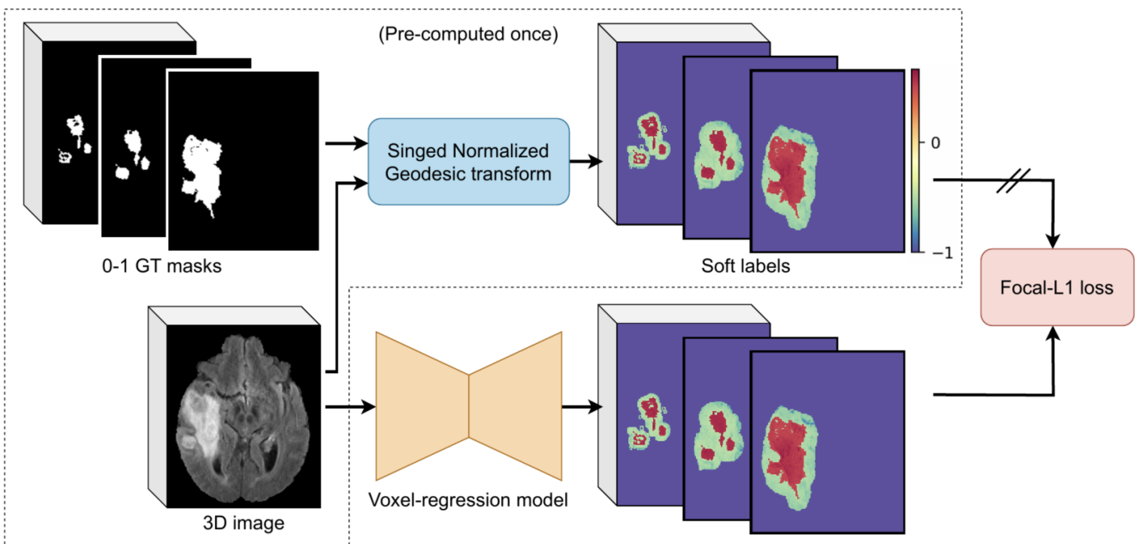

Figure 1: Example of converting voxel-wise classification based segmentation into regression. In this project, you will need to the same, but for the segmentation of cartilage and meniscus (Assignment 3).

The final project for our course challenges you to synthesize and extend the concepts from the course beyond what is readily available. The core task is to adapt the uncertainty-aware segmentation methods from the SAUNA [1] and [2] papers to the cartilage segmentation problem from **Assignment 3** that re-implements works by Panfilov et al [3, 4].

Instead of treating segmentation as a binary classification problem (pixel is 0 or 1), you will reframe it as a **regression problem**. This involves creating ”soft” ground truth labels that encode spatial uncertainty and training your implicit neural network to predict these continuous values.

This project will be completed in **groups of three**. Although assignment 3 is an individual task, this project requires you to collaborate to build a more complex and robust system.

2 Core technical challenge

2.1 Ground Truth Transformation

Assignment 3 uses “hard” masks (0, 1, 2 ...) as the ground truth. You need implement a transformation to convert these ”hard” labels into ”soft” regression targets that capture uncertainty, similar to the methods in the provided papers.

A robust and highly recommended approach is to convert your binary masks into Signed Distance Fields (SDFs). An SDF maps each coordinate (x, y) to its distance to the *nearest* boundary, with the sign indicating whether it is inside (e.g., negative) or outside (e.g., positive) the object.

This directly provides a continuous label that is rich in geometric information and inherently encodes the boundary, which is the primary source of uncertainty.

2.2 Model and loss

For simplicity, all segmentation and analysis can be performed in **2D**. You still need to make two modifications to your pipeline:

- **Model output:** The implicit neural field (INR) from Assignment 3 must be adapted. Its final layer should output a single continuous value (the predicted distance) instead of a binary logit.
- **Loss function:** You must replace classification losses (like Binary Cross-Entropy or Dice) with a **regression loss**. A simple **L1 loss** ($|y_{\text{true}} - y_{\text{pred}}|$) is a strong starting point. You are encouraged to explore and report on other regression losses, such as L2 (MSE) or the Focal-L1 loss mentioned in the papers.

3 Deliverables

Your group should submit a single ZIP file containing the following:

1. **Google Colab Notebook:** A single, runnable '.ipynb' file. The code must be well-commented, correct, and execute from start to finish on Google Colab, including data loading, preprocessing, training, and evaluation.
2. **README File:** A 'README.md' file explaining exactly how to run your notebook, including how to arrange the data (e.g., "place the 'data.zip' file in the root directory") and any required setup.
3. **Project Report:** A 4-page, 2-column report in the IEEE template style (emulated by this document).

4 Grading Criteria

Your project will be graded based on the following criteria:

1. **Code Correctness & Quality (30%):** The code runs on Colab without errors. It is clean, well-commented, and logically structured.
2. **Report Quality (30%):** The report is well-written, clear, and has a logical flow. It contains insights, not just a description of work. **Must include plots** of training and validation loss curves.
3. **Data Splitting & Cross-Validation (10%):** You must correctly implement a patient-aware train/test split. No data from the same patient can be in both the training and test sets. You must go beyond the split in Assignment 3. *Hint: Use OAI metadata you have access to.*
4. **Baseline & Improvement (10%):** You must define a "trivial baseline" (e.g., the original Assignment 3 binary model, or a standard U-Net with binary loss) and demonstrate that your new regression-based method provides an improvement.
5. **Statistical Testing (10%):** You must formally compare your model's performance against the baseline using an appropriate statistical test (e.g., a paired t-test or Wilcoxon signed-rank test on the cross-validation fold metrics).
6. **Ablation Studies (10%):** Your report must include experiments that analyze and justify your design choices. (e.g., Why L1 loss vs. L2? How does the SDF representation compare to the binary mask model? What was the effect of a specific hyperparameter?)

5 Report and Submission

5.1 Report Guidelines

Your report PDF must follow the 4-page, 2-column IEEE format. It must include the following sections:

- **Introduction:** Describe the problem, the limitations of binary segmentation, and your proposed regression-based approach.
- **Methods:** Detail your data preprocessing (especially the soft-label transformation), your baseline model, your adapted INR model architecture, and the loss functions used.
- **Experiments:** Describe your training setup, evaluation metrics, data splitting strategy, and present your results. This section must include:
  - Plots of training and validation loss.
  - Comparison tables against your baseline.
  - The results of your statistical tests.
  - Your ablation studies.
- **Conclusion:** Summarize your findings and provide insights.
- **Member Contributions:** A brief, clear description of the responsibilities and contributions of each group member.

References

[1] T. D. Dang, H. H. Nguyen, and A. Tiulpin, "Image-level regression for uncertainty-aware retinal image segmentation," in *2025 IEEE/CVF Winter Conference on Applications of Computer Vision (WACV)*, pp. 4194–4204, IEEE, 2025.

[2] T. Dang, H. H. Nguyen, and A. Tiulpin, "Singr: Brain tumor segmentation via signed normalized geodesic transform regression," in *International Conference on Medical Image Computing and Computer-Assisted Intervention*, pp. 593–603, Springer, 2024.

[3] E. Panfilov, A. Tiulpin, S. Klein, M. T. Nieminen, and S. Saarakkala, "Improving robustness of deep learning based knee mri segmentation: Mixup and adversarial domain adaptation," in *Proceedings of the IEEE/CVF international conference on computer vision workshops*, pp. 0–0, 2019.

[4] E. Panfilov, A. Tiulpin, M. T. Nieminen, S. Saarakkala, and V. Casula, "Deep learning-based segmentation of knee mri for fully automatic subregional morphological assessment of cartilage tissues: data from the osteoarthritis initiative," *Journal of Orthopaedic Research®*, vol. 40, no. 5, pp. 1113–1124, 2022.