Text-Guided 3D MRI Segmentation with the Segment Anything Model

Scientific Context

[CREATIS](#) is a multidisciplinary laboratory with broad expertise in medical imaging and a leading role in health technologies in Lyon. Within this framework, the [Myriad team](#) develops innovative methodologies in image processing, biomechanical modeling, and image simulation for medical imaging. A key strength lies in bringing scientists and clinicians together around targeted medical applications, with a focus on cardiovascular, pulmonary, and neurological imaging. The team currently includes 30 permanent members (researchers, clinicians, and engineers), supervising about 30 PhD students and 3 postdocs.

Recent advances in foundation models, such as the *Segment Anything Model* [1], have enabled broadly generalizable, promptable segmentation. While medical and 3D adaptations of SAM now exist (e.g., MedSam, SAM3D, adapter-based 3D variants) [2–6], effective use in **MRI stroke** still requires **task- and modality-specific adaptation**. Most current segmentation models primarily rely on image features. However, in complex brain MRI, where lesions and normal tissues may exhibit very similar signal intensities or appearances, models may fall short and struggle to distinguish between them. In contrast, image interpretation is rarely performed in isolation in clinical practice: radiologists complement imaging data with prior anatomical knowledge and additional clinical information (e.g., symptoms, medical history, examination findings, or prior discussions with colleagues).

Such contextual knowledge helps to localize, characterize, and prioritize findings beyond what is visible in the images alone. Building on this idea, the proposed work will **translate available clinical information into SAM-compatible prompts**—and integrate them into a SAM pipeline to guide segmentation, with the aim of improving accuracy and robustness.

This project proposes to bridge this gap by leveraging clinical text as a source of guidance for SAM, aiming to enhance its accuracy in 3D MRI segmentation.

Cohorts and Data

The evaluation will be carried out on two complementary multi-modal datasets:

- **HIBISCUS cohort** [8]: A stroke imaging dataset including multimodal MRIs (notably diffusion-weighted MRI) together with clinical variables such as age, sex, lesion laterality, and clinical scores (mRS, NIHSS, ASPECTS). Expert lesion segmentations are available, making HIBISCUS the primary dataset for model development and multimodal integration.

- **ETIS dataset**: A multi-center ischemic stroke dataset that also combines diffusion MRI with similar clinical variables. However, not all cases are segmented. ETIS will therefore be used mainly as a testbed to evaluate the generalization of models trained on HIBISCUS across different centers and acquisition protocols.

This design allows studying both the feasibility of multimodal integration (HIBISCUS) and the robustness of the approach in a broader, heterogeneous, multi-institutional setting (ETIS).

Keywords : Medical image segmentation, Foundation models / SAM, Multimodal learning, Stroke imaging (MRI)
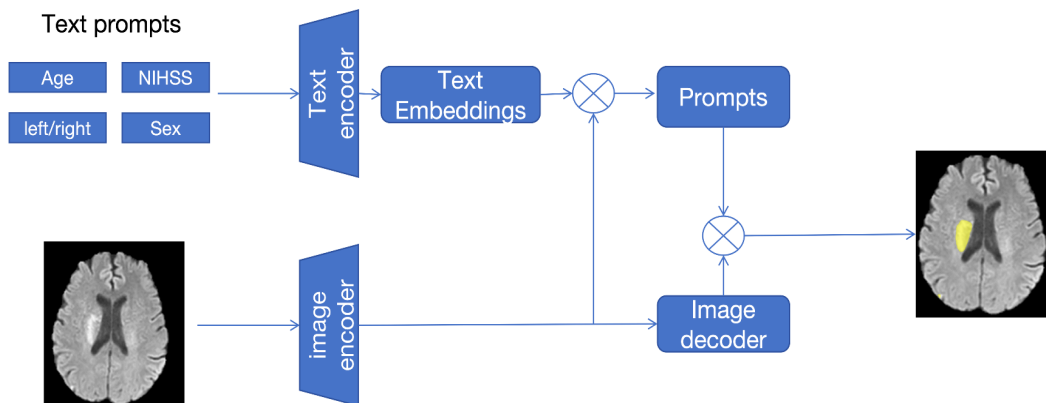

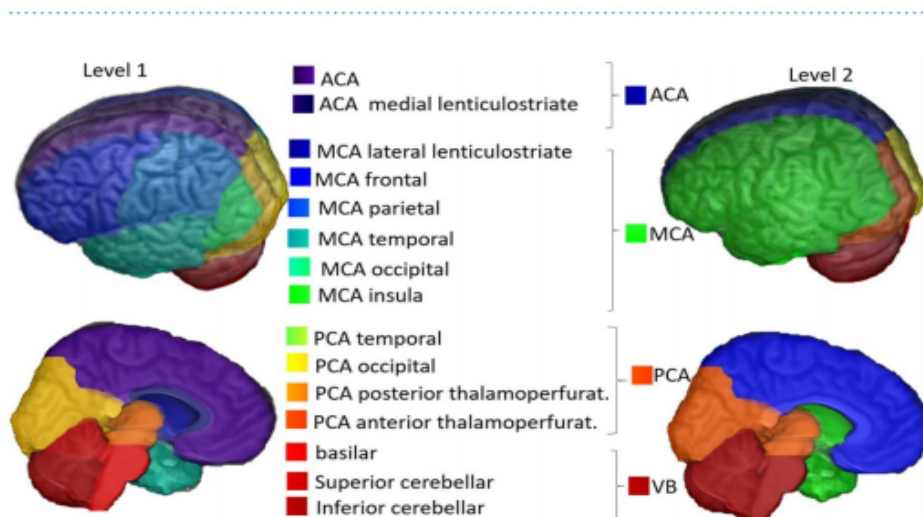Figure 1: The proposed approach for text-guide segmentation in stroke MRI

Figure 2: Digital 3D Brain MRI Arterial Territories Atlas [7]

Objectives

Develop a clinically guided prompting pipeline for SAM-based stroke lesion segmentation in diffusion MRI. The internship will (i) convert structured clinical descriptors—and, when available, short free-text phrases[6]—into SAM-compatible geometric prompts (3D boxes, positive/negative points, weak masks) via an atlas-based[7] probabilistic prior, and (ii) integrate these prompts into both MedSAM[2] (slice-wise baseline) and a selected SAM-3D variant, aiming to improve accuracy, focus, and robustness.

- Standardize inputs (side, vascular territory, NIHSS, ASPECTS), build an atlas-informed spatial prior $P(x)$[9], leveraging prior anatomical knowledge and arterial territories atlas segmentation of specific occlusion arterial to guide SAM.
- Run a prompt-guided MedSAM baseline or a SAM variant (e.g., SAM3D/adapter-based) with the same prompt interface; ensure volumetric consistency and reasonable memory/runtime.
- Handle missing/ambiguous clinical info via multi-hypothesis prompts (e.g., left/right, territory candidates) and coarse-to-fine prompting; stress-test sensitivity to prompt noise/placement.
- Train/validate on HIBISCUS; test out-of-distribution generalization on ETIS (multicenter, partial labels). Metrics: Dice, HD95, lesion-wise F1, volumetric coherence, and inter-center robustness. Compare to no-prompt SAM baselines and standard CNN/Transformer models.

Deliverables

Reusable prompt-generation module + configs, prompt-guided MedSAM or SAM-3D pipelines, ablation studies (prompt types/quantities), and a proof-of-concept demonstrating measurable gains from clinically guided prompting. (Optional: small text-encoder variant for direct language guidance if time permits[6].)

References

1. Kirillov, Alexander, et al. "*Segment anything*" Proceedings of the IEEE/CVF international conference on computer vision. 2023.
2. Ma, J., He, Y., Li, F. *et al.* Segment anything in medical images. *Nature Communication* 15, 654 (2024).
3. Bui, Nhat-Tan, et al. "*SAM3D: Segment anything model in volumetric medical images*" 2024 IEEE International Symposium on Biomedical Imaging (ISBI). IEEE, 2024.
4. Gong, Shizhan, et al. "*3DSAM-adapter: Holistic adaptation of SAM from 2D to 3D for promptable tumor segmentation*" Medical Image Analysis 98 (2024): 103324.
5. Chen, Cheng, et al. "*Ma-SAM: Modality-agnostic SAM adaptation for 3D medical image segmentation*" Medical Image Analysis 98 (2024): 103310.
6. Zhao, Z., Zhang, Y., Wu, C. *et al.* Large-vocabulary segmentation for medical images with text prompts. *npj Digit. Med.* 8, 566 (2025).
7. Liu, CF., Hsu, J., Xu, X. *et al.* Digital 3D Brain MRI Arterial Territories Atlas. *Sci Data* 10, 74 (2023). https://doi.org/10.1038/s41597-022-01923-0
8. Moreau, Juliette, et al. "*Contrast quality control for segmentation task based on deep learning models—Application to stroke lesion in CT imaging*" Frontiers in Neurology 16 (2025): 1434334.
9. Bloch, Isabelle. "Fuzzy sets for image processing and understanding." *Fuzzy sets and systems* 281 (2015): 280-291

Contact : carole.frindel@creatis.insa-lyon.fr