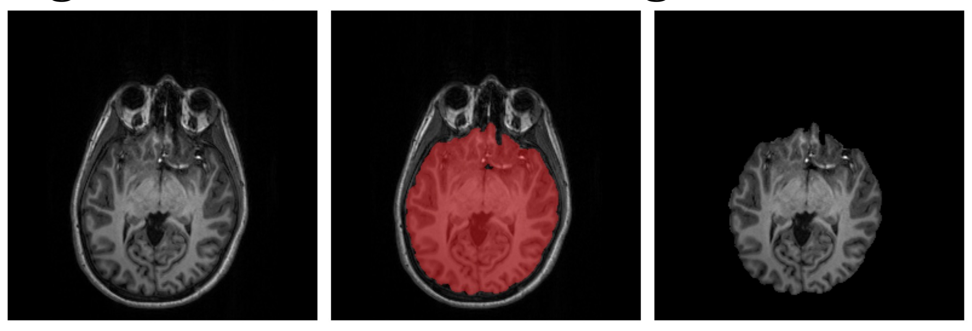
Retrieval-Augmented Few-shot Skull Stripping: Implementation and Evaluation

Kim Dohoon[1,2], Jong-Min Lee[2]
{ kdhyu, ljm } @ hanyang.ac.kr
Department of Data Science[1]
Department of Biomedical Engineering[2]

GitHub

INTRODUCTION

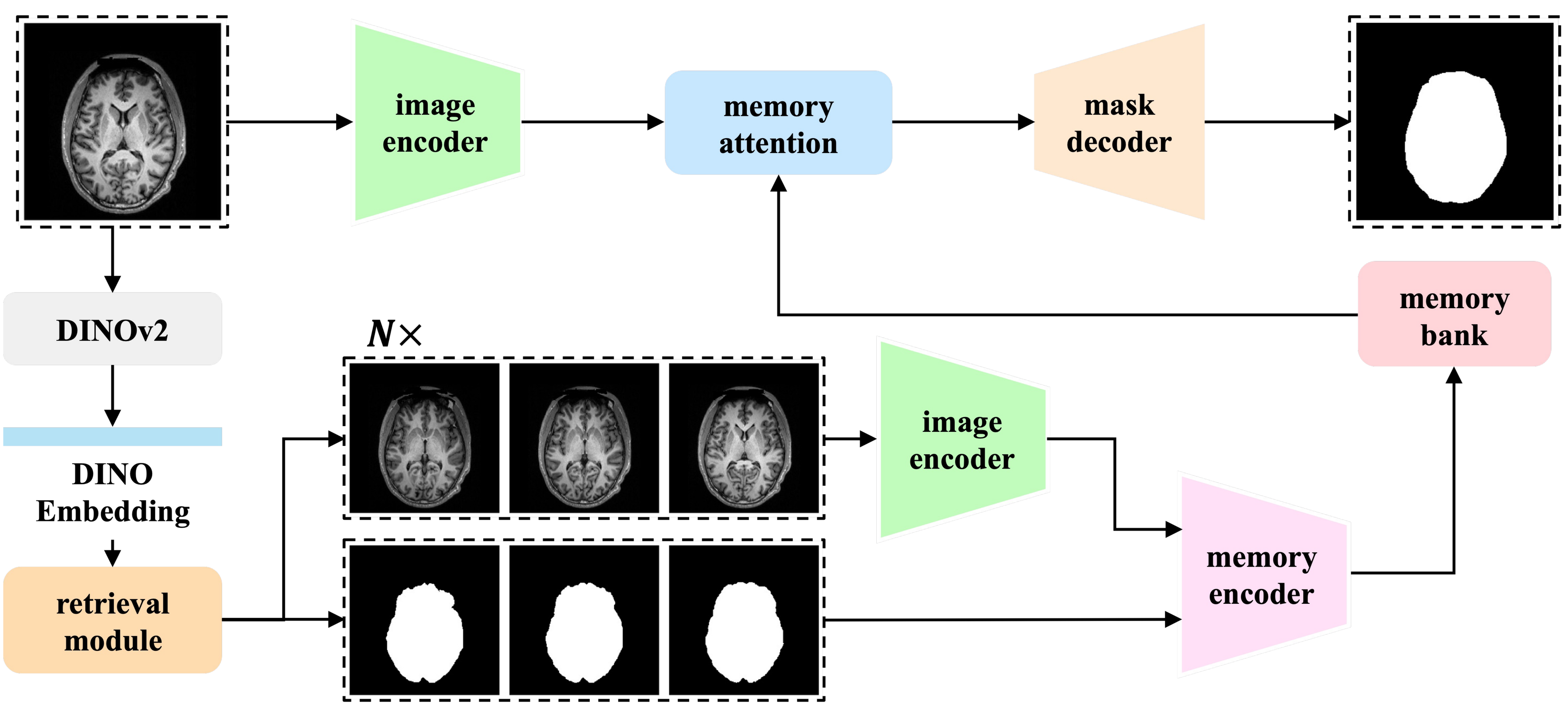
- Skull Stripping is a crucial preprocessing step in brain image analysis, as it removes non-brain tissues (e.g., skull, scalp) from MRI scans. This process improves the accuracy of downstream tasks such as brain tissue segmentation, registration, and disease diagnosis.



- Recent visual foundation models such as DINOv2 [2] and SAM 2 [3] show strong zero-shot generalization across various domain shifts, including the medical images.
- Furthermore, recent study[1] has demonstrated that Retrieval-Augmentation (RA) significantly enhances performance in medical image segmentation tasks.
- In this work, we implement the Retrieval-Augmented Few-shot Medical Image Segmentation (RAFS) framework and evaluate its effectiveness on the skull stripping task. Note that, this method is "training-free approach", which means it does not require any additional training.
- All codes are available at https://github.com/kdh-yu/BME_Capstone1.

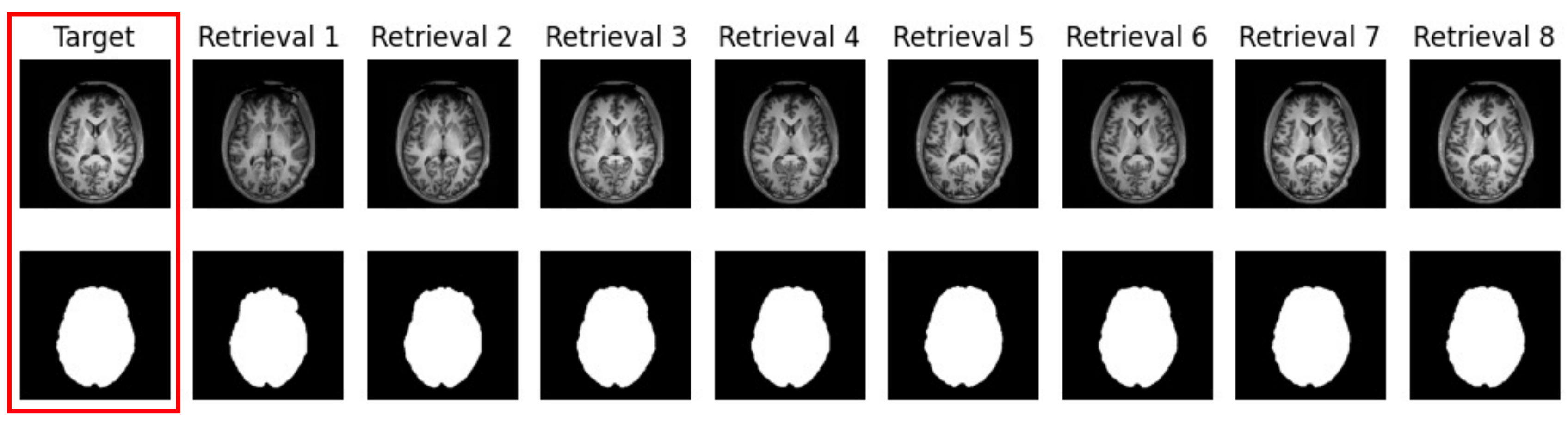
METHOD

- **Pipeline Overview**



- **First Stage: FAISS Index Construction**
  - DINOv2 Vision Transformer extracts normalized [CLS] embedding from brain MRI scans.
  - 784 axial slices from 3 annotated volumetric data are added to a FAISS index

- **Second Stage: Retrieval-Augmented Segmentation**
  - Retrieval: The FAISS Index is queried to retrieve N similar images, which are then treated as "Past Frames".
  - Segmentation: SAM 2 Video Segmentation is performed to predict the current image's brain mask.



RESULTS

- **Implementation Details**
  - FAISS Index: Configurated with Euclidean Distance (L2)
    - Vectors are normalized during both constructing and querying
    - Model retrieves 8 slices to predict current mask. (N=8)
  - SAM 2: Hiera-Large Model
  - DINOv2: ViT-S/14 with registers
- **Dataset**
  - NFBS Skull Stripped Repository [4]
    - Anonymized T1w MRI Dataset
    - Includes 125 volumetric data with size (256, 256, 192)
- **Quantitative Analysis (Dice Score)**

	Slice-wise	Volume-wise
RAFS (ours)	**0.8969**	**0.9594**
U-Net (3 Subj)	0.9250	0.9747
U-Net	0.9677	0.9918
Swin U-Net (3 Subj)	0.8279	0.9204
Swin U-Net	0.9671	0.9881

- Our implemented method demonstrates performance on par with fully-supervised U-Net and Swin U-Net
- Notably, it also excels when applied to volumetric data, suggesting it's a promising solution for practical skull stripping tasks.

CONCLUSION

- Our work demonstrates that the Retrieval-Augmented Few-shot Medical Image Segmentation (RAFS) framework achieves robust performance, not only on 2D slices but also on 3D volumetric data.
- One key advantage of RAFS is its training-free nature, eliminating the need for extensive fine-tuning on new datasets, which makes RAFS a highly efficient and adaptable solution for skull stripping, with the potential to be expanded for diverse tasks such as tissue segmentation.

LIMITATIONS

- One limitation is its reduced ability to produce highly intricate or geometrically complex masks. We hypothesize this stems from two main factors:
  - A potential loss of fine-grained detail when integrating retrieved images and masks into memory
  - Performance bottlenecks exacerbated by significant domain shift

ACKNOWLEDGEMENT

- This project was conducted as part of the Biomedical Engineering Capstone Design 1 course. I would like to sincerely thank Prof. Jong-Min Lee and Sangwoo Ra from CNA Lab for their guidance and support.

REFERENCES

[1] Zhao, Lin, et al. "Retrieval-augmented Few-shot Medical Image Segmentation with Foundation Models." arXiv preprint arXiv:2408.08813 (2024).
[2] Oquab, Maxime, et al. "DINOv2: Learning robust visual features without supervision." Transactions on Machine Learning Research (2024).
[3] Ravi, Nikhila, et al. "Sam 2: Segment anything in images and videos." arXiv preprint arXiv:2408.00714 (2024).
[4] Puccio, B., et al. "The preprocessed connectomes project repository of manually corrected skull-stripped T1-weighted anatomical MRI data." Gigascience 5.1 (2016)