Zero-shot Segmentation of Biomarkers in Optical Coherence Tomography Using Foundation Model Based on Bounding Box Prompting

Jungrae Cho[1], Byungeun Shon[1,2], Dongwon Woo[1], Hosang Yu[1] , Moonkyu Kim[3,4], Dong Ho Park[5], and Sungmoon Jeong[1,2*]

[1]Research Center for AI in Medicine, Kyungpook National University Hospital, Daegu, Korea
[2]Department of Medical Informatics, Kyungpook National University, Daegu, Korea
[3]Hair Transplantation Center, Kyungpook National University Hospital, Daegu, Korea
[4]Department of Immunology, Kyungpook National University School of Medicine, Daegu, Korea
[5]Department of Ophthalmology, School of Medicine, Kyungpook National University,
Kyungpook National University Hospital, Daegu, Korea

{zzemb6,jeongsm00}@gmail.com

Introduction

- Manual segmentation for deep learning in medical imaging is time-consuming [1,2] but precise, whereas bounding box annotation keeps local information and simpler but has coarse boundary of object. Segment Anything by Meta AI [3] has achieved impressive performance on semantic segmentation for unseen task with its promptable interface. In this paper, we investigate the potential and effect of leveraging SAM for zero-shot segmentation of biomarkers in optical coherence tomography (OCT) B-scans from age-related macular degeneration (AMD) patients.

Method

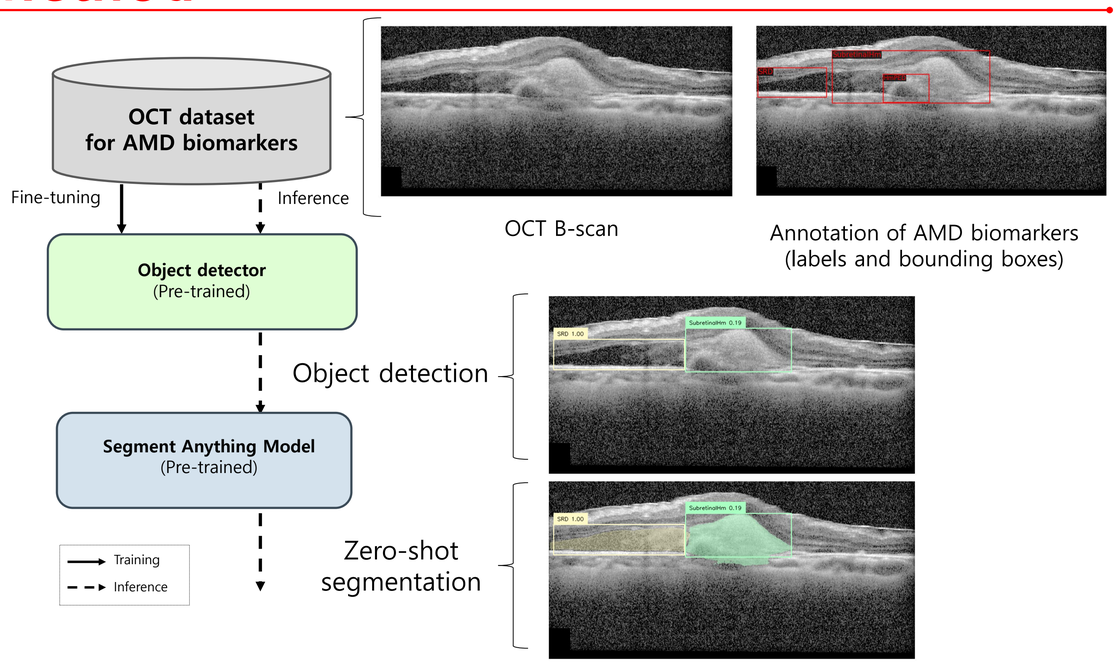


Figure 1. Overview of zero-shot segmentation process in this work

- In this work, we propose a method for zero-shot segmentation with only bounding box annotations by using Segment Anything Model (SAM), which has been released by Meta [3]. The trained object detectors on medical images with bounding box annotations produce labels and bounding boxes of medical objects. These outputs are fed into Segment Anything Model (SAM) as prompts, and the foundation model segments the foreground in each bounding box.
- On implementation details, a pretrained Faster R-CNN with ResNet50(Caffe)-FPN trained by MMDetection [4] framework was fine-tuned for medical image dataset with object annotations. The model was trained with 12 epochs with default settings of the framework in a single GPU, Quadro RTX 6000. For the foundation model, the pre-trained SAM with ViT-H as image encoder was utilized for only inference of semantic segmentation.

Dataset

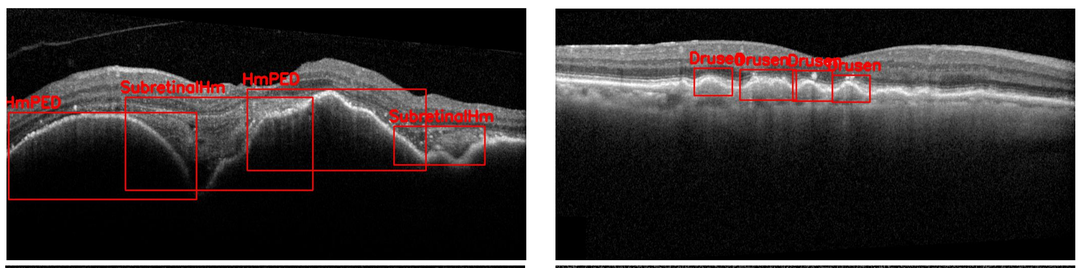


Figure 2. Examples of OCT B-scan images with biomarkers of age-related macular degeneration (AMD). (Red: ground-truth annotation of AMD biomarkers)

- We have collected OCT volumes from 96 and 238 eyes from dry and wet AMD patients, respectively in Kyungpook National University Hospital. Data were approved by IRB and anonymized. From each OCT volume, OCT B-scans close to macular were annotated by drawing bounding box and labeling the name of biomarker type, e.g. drusen, geometric atrophy for AMD biomarker localization as shown in table 1. For each OCT scan, annotators also labeled type of AMD, i.e. normal, dry AMD, and wet AMD for classification. Finally, 1,661 OCT scans have been collected and annotated. Ophthalmologists participated in annotation of OCT scans. Fig. 2 shows examples of different types of OCT scans. The number of dataset are described in table 2.

Table 1. Description of object classes

Object class	Description
CME	Cystoid macular edema
Retinal edema	Retinal edema
SRD	Subretinal detachment
SubretinalHm	Subretinal hemorrhage
SerousPED	Serous pigment epithelial detachment
FvPED	Fibrovascular pigment epithelial detachment
HmPED	Hemorrhagic pigment epithelial detachment
Polyp	Polyp
HRF	Hyperreflective foci
Retinal exudate	Retinal exudates
Drusen	Drusen
SDD	Subretinal drusenoid deposits
GA	Geographic atrophy
RPE atrophy	Retinal pigment epithelium atrophy
DisciformScar	Disciform scar

Table 2. Summary of OCT AMD dataset

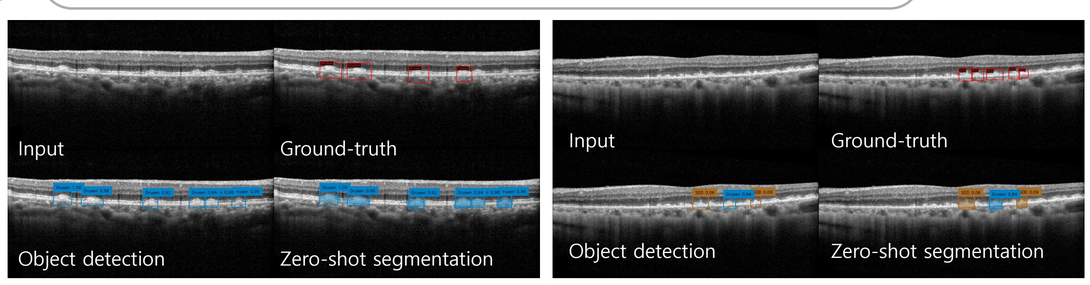
Dataset	OCT B-scan
Train+Val. Set	1,384
Test Set	287
Total	1,661
Object classes	15 AMD biomarkers

Results

1 Performance of Medical Object Detectors

Dataset	mAP@0.5:0.95	mAP@0.5	mAP@0.75
OCT B-scan (AMD biomarker)	0.164	0.343	0.152

2 Zero-shot Segmentation



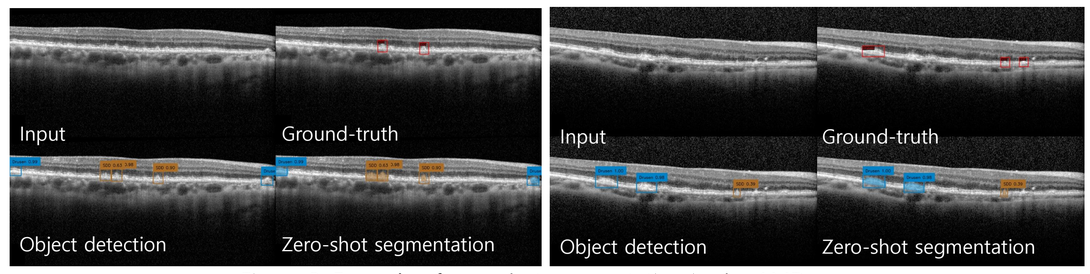


Figure 3. Example of zero-shot segmentation in dry AMD cases

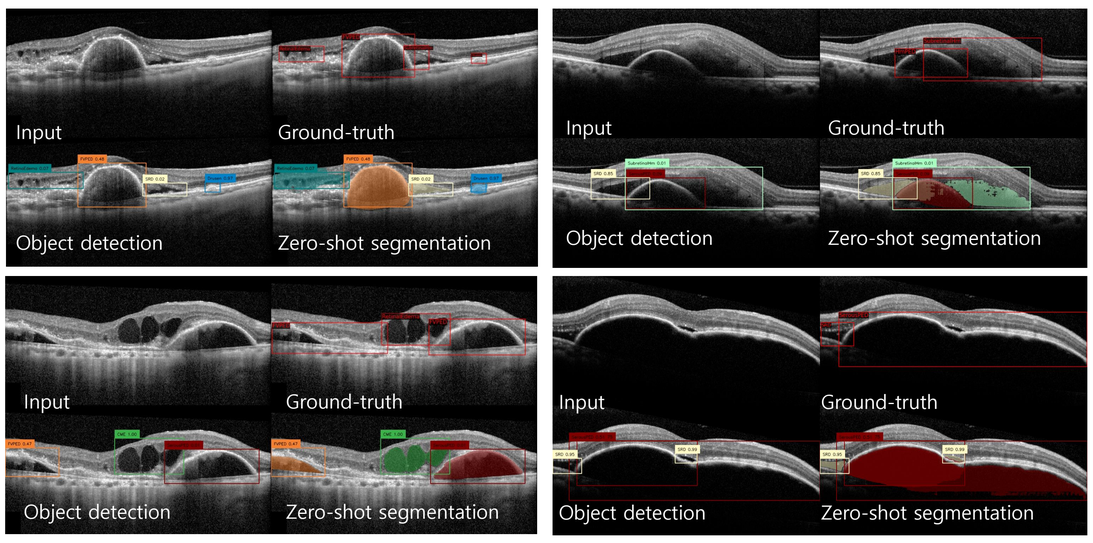


Figure 4. Examples of zero-shot segmentation in wet AMD cases

Discussion

- Our experimental results demonstrate that the foundation model of semantic segmentation trained with large dataset can segment AMD biomarkers with only annotated with bounding box and class label in a zero-shot manner.
- This can be applied to fine-grained quantification of biomarkers with relatively cheaper annotation, such as amount of intraretinal fluid or hemorrhage in an OCT B-scan image.
- Based on our results, the quality of semantic segmentation may depend on the performance of object detectors and the extent to which objects are contained within their bounding boxes. Variation of SAM may be one of solutions to handle the quality of zero-shot segmentation of AMD biomarkers in OCT.

Acknowledgement

This work was partly supported by the MSIT(Ministry of Science and ICT), Korea, under the ITRC(Information Technology Research Center) support program(IITP-2023-2020-0-01808) supervised by the IITP(Institute of Information & Communications Technology Planning & Evaluation) and National IT Industry Promotion Agency(NIPA) grant funded by the Korea government(MSIT) (S0252-21-1001, Development of AI Precision Medical Solution(Doctor Answer 2.0)).

References

[1] Lin, T. Y., Maire, M., Belongie, S., Hays, J., Perona, P., Ramanan, D., . . . , Zitnick, C. L. (2014). Microsoft coco: Common objects in context. In Computer Vision–ECCV 2014: 13th European Conference, Zurich, Switzerland, September 6-12, 2014, Proceedings, Part V 13 (pp. 740–755). Springer International Publishing.https://www.overleaf.com/project/6483f883c610a58da8024f33
[2] Hsu, H. Y., Chou, Y. B., Jheng, Y. C., Kao, Z. K., Huang, H. Y., Chen, H. R., . . . , Wu, Y. T. (2022). Automatic segmentation of retinal fluid and photoreceptor layer from optical coherence tomography images of diabetic macular edema patients using deep learning and associations with visual acuity. Biomedicines, 10(6), 1269.
[3] Kirillov, A., Mintun, E., Ravi, N., Mao, H., Rolland, C., Gustafson, L., Xiao, T., Whitehead, S., Berg, A.C., Lo, W.Y. and others: "Segment anything." arXiv preprint arXiv:2304.02643 (2023)
[4] Chen, K., Wang, J., Pang, J., Cao, Y., Xiong, Y., Li, X., ... & Lin, D. (2019). MMDetection: Open mmlab detection toolbox and benchmark. arXiv preprint arXiv:1906.07155.
https://github.com/open-mmlab/mmdetection