

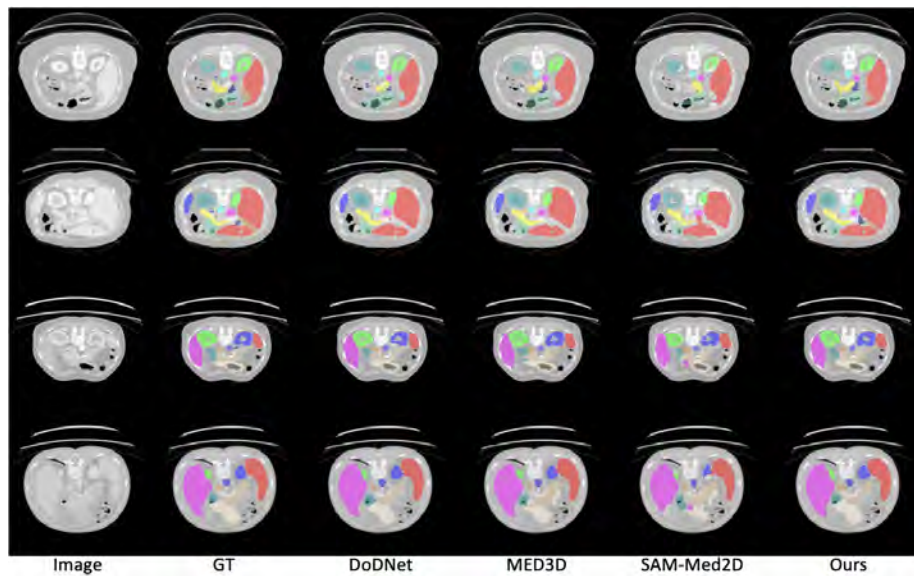
# M<sup>4</sup>oE: Foundation Model for Medical Multimodal Image Segmentation with Mixture of Experts Supplementary Material

Yufeng Jiang<sup>1</sup>[0009-0004-4987-2683] and Yiqing Shen<sup>2</sup>() [0000-0001-7866-3339]

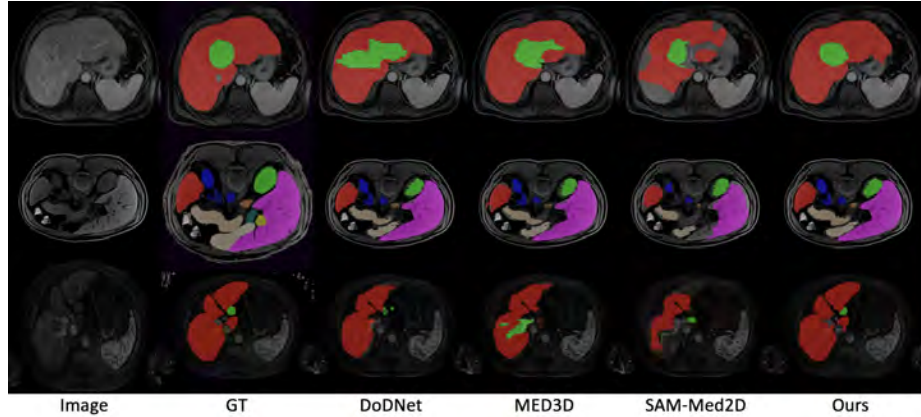
<sup>1</sup> Department of Computer Science, Hong Kong Baptist University, Hong Kong SAR, China

<sup>2</sup> Department of Computer Science, Johns Hopkins University, Baltimore, USA  
yshen92@jhu.edu

## 1 Results



**Fig. 1.** Visualization of segmentation results from various methods. The first and second rows depict FLARE22 with CT images, and the third and fourth rows showcase AMOS22 with CT images. The six columns from left to right correspond to the original image, the ground truth (GT), the DoDNet results, the MED3D results, the SAM-MED2D results, and our M<sup>4</sup>oE results.



**Fig. 2.** Visualization of segmentation results from various methods. The first and third rows depict ATLAS2023 with CE-MRI images, and the second row showcases AMOS22 with MR images. The six columns from left to right correspond to the original image, the ground truth (GT), the DoDNet results, the MED3D results, the SAM-MED2D results, and our  $M^4oE$  results.

**Table 1.** Comparison of mIoU (%) across multimodal segmentation methods. Our  $M^4oE$  model outperforms others, achieving the highest mIoU in FALRE and AMOS-CT+MRI tasks. It shows adaptability and improvement (\*) upon new modality introduction, contrary to DoDNet and MED3D which see performance drops. Our model’s robustness is attributed to training on three varied datasets, and SAM-Med2D’s performance is evaluated using a single data point. Results are highlighted in red for best and blue for second best.

Dataset	Method	Liver	Left Kidney	Right Kidney	Gallbladder	Pancreas	Spleen	Stomach	Mean
FLARE22	DoDNet	87.25	87.23	88.59	68.28	71.23	83.76	84.10	81.49
	MED3D	80.37	79.21	77.82	63.12	66.57	79.31	80.21	75.23
	SAM-Med2D	75.37	70.35	73.11	43.24	46.56	75.91	70.03	64.86
	Ours	94.69	89.72	89.71	75.53	76.99	91.89	81.58	85.73
AMOS-CT	DoDNet	92.14	83.93	82.48	70.73	63.69	90.58	85.14	81.24
	MED3D	82.78	81.19	79.45	63.89	54.18	79.37	75.41	73.75
	SAM-Med2D	60.34	65.31	63.16	53.86	43.14	69.87	67.50	60.45
	Ours	91.32	87.08	84.76	72.12	58.46	88.66	79.07	80.21
AMOS-CT+MRI	DoDNet	83.25	80.90	74.33	68.44	64.87	83.25	79.32	76.34
	MED3D	76.31	77.96	80.23	70.90	54.88	89.35	71.29	74.42
	SAM-Med2D	45.15	43.86	54.12	43.68	43.92	53.20	52.38	48.04
	Ours	92.55	87.13	85.95	73.74	65.33	89.07	83.45	82.46