Poster Session
THU-PM-313

Highlight

JUNE 18-22, 2023
CVPR
VANCOUVER, CANADA

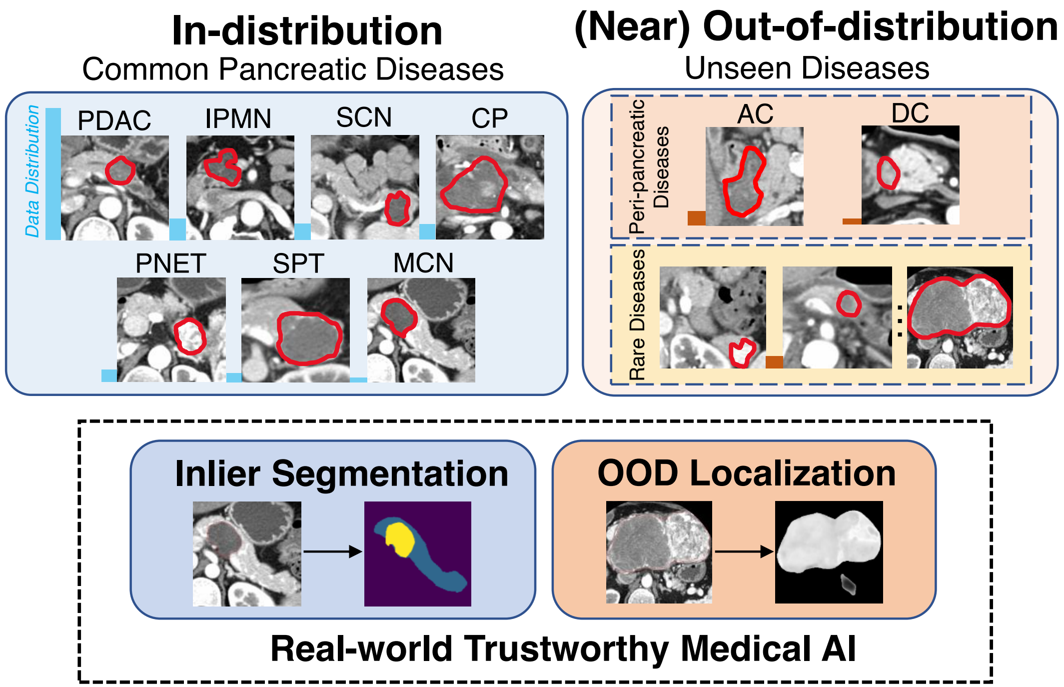
Devil is in the Queries: Advancing Mask Transformers for Real-world Medical Image Segmentation and Out-of-Distribution Localization

Mingze Yuan[1,2], Yingda Xia[1], Hexin Dong[1,2], Zifan Chen[2], Jiawen Yao[1], Mingyan Qiu[1], Ke Yan[1], Xiaoli Yin[4], Yu Shi[4], Xin Chen[3], Zaiyi Liu[3], Bin Dong[1,5], Jingren Zhou[1], Le Lu[1], Ling Zhang[1], Li Zhang[2]

[1]Alibaba Group [2]Peking University [3]Guangdong Province People's Hospital [4]Shengjing Hospital
[5]Peking University Changsha Institute for Computing and Digital Economy

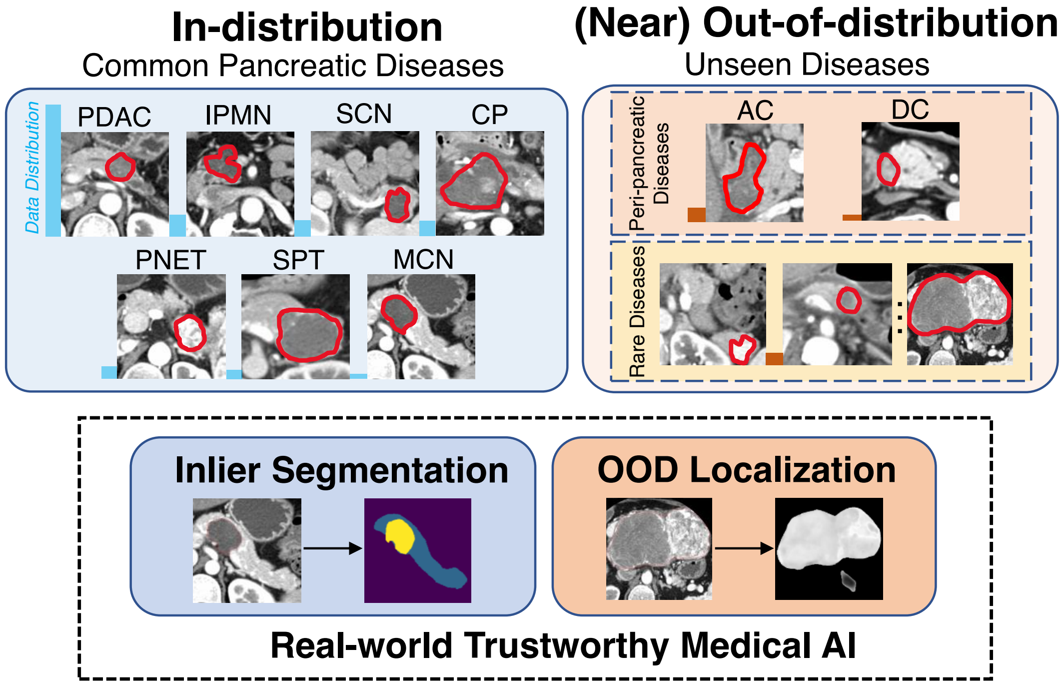
ALIBABA DAMO ACADEMY

MaxQuery: Quick Preview



- Real-world medical image segmentation
  - Tail conditions: rare disease
- MaxQuery for out-of-distribution (OOD) localization
- Query distribution loss
- Validated on real-world pancreatic & liver tumor datasets

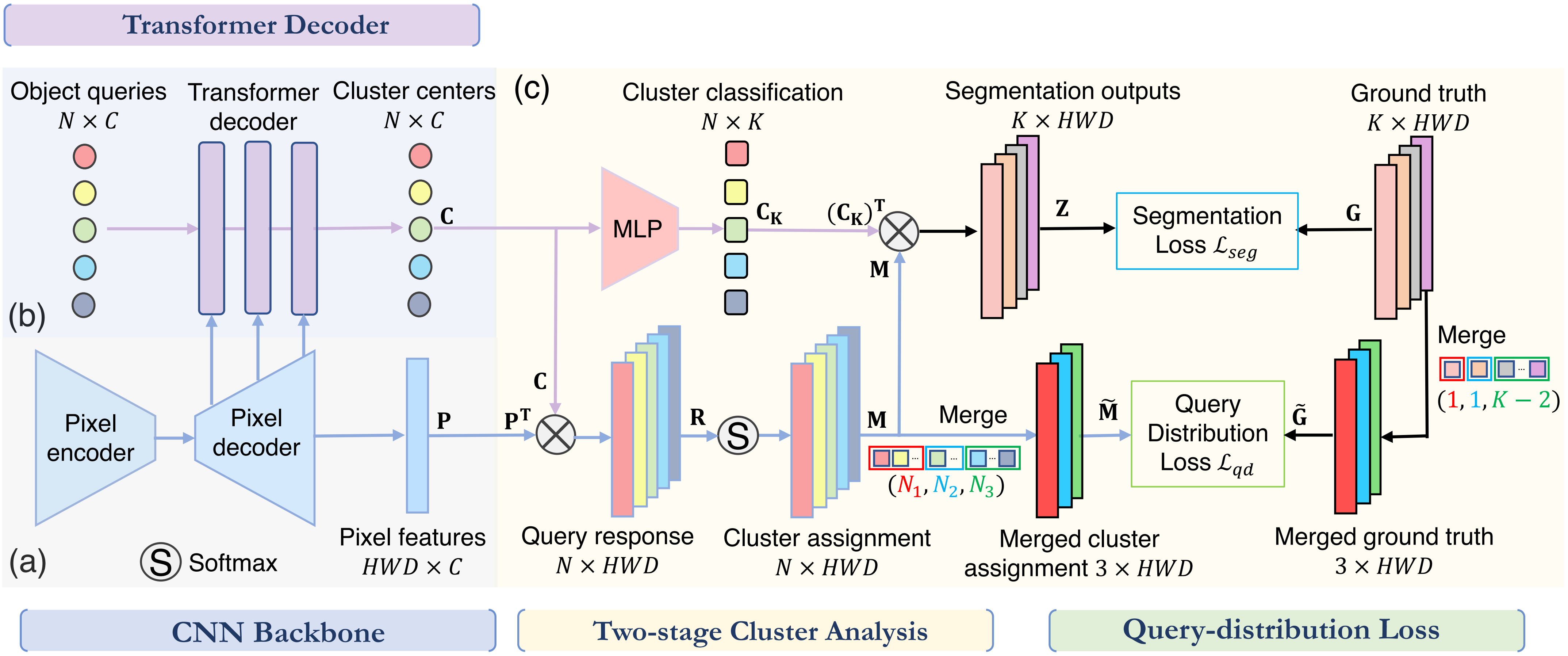
Real-world Medical Image Segmentation



- Long-tailed medical conditions
- Near-OOD Problem
  - Outliers: unseen/rare tumors
  - Inliers: labeled lesions
  - Subtle Differences
- Address both inlier segmentation & OOD localization

[1] Roy et al., *Does your dermatology classifier know what it doesn't know? detecting the long-tail of unseen conditions*, *Medical Image Analysis* 75 (2022): 102274.

Framework Overview

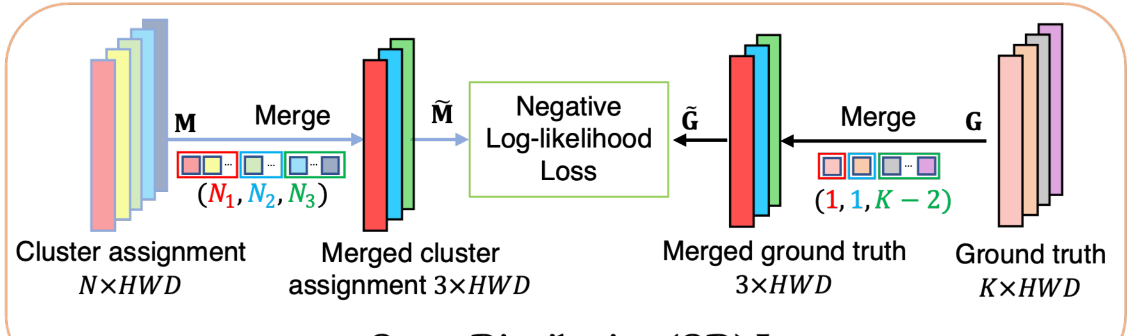


[2] Cheng et al., *Per-pixel classification is not all you need for semantic segmentation, NeurIPS 2021.*
[3] Yu et al., *Cmt-deeplab: Clustering mask transformers for panoptic segmentation, CVPR 2022.*
[4] Yu et al., *k-means Mask Transformer., ECCV 2022.*

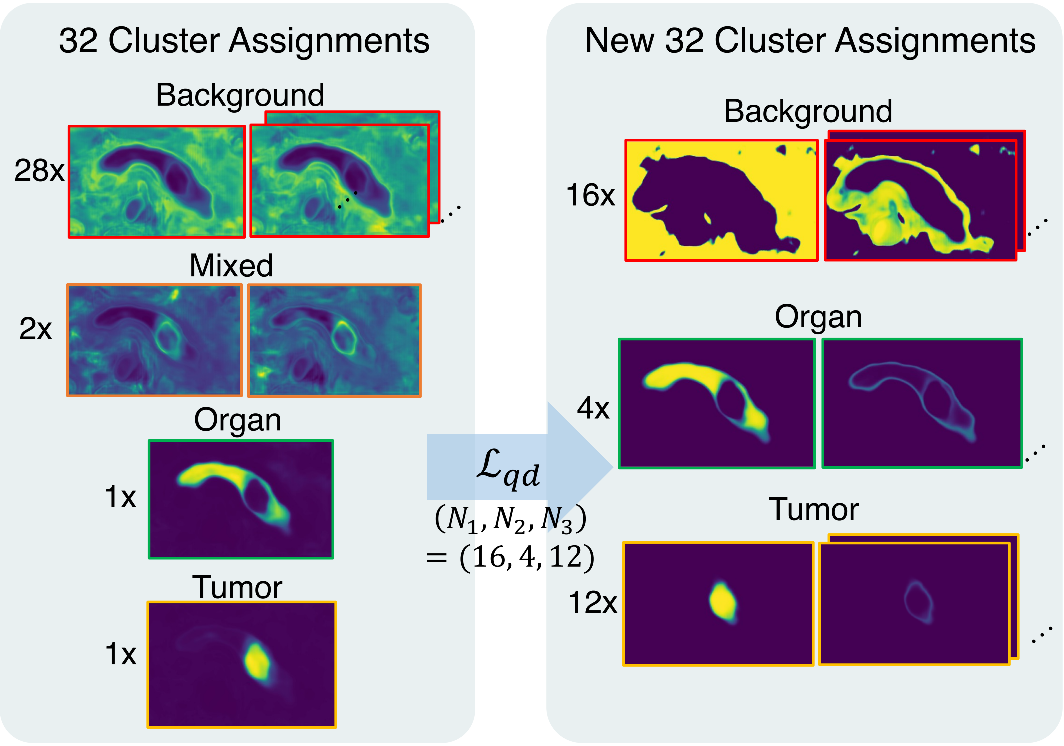
Managing Cluster with QD Loss

JUNE 18-22, 2023
CVPR
VANCOUVER, CANADA

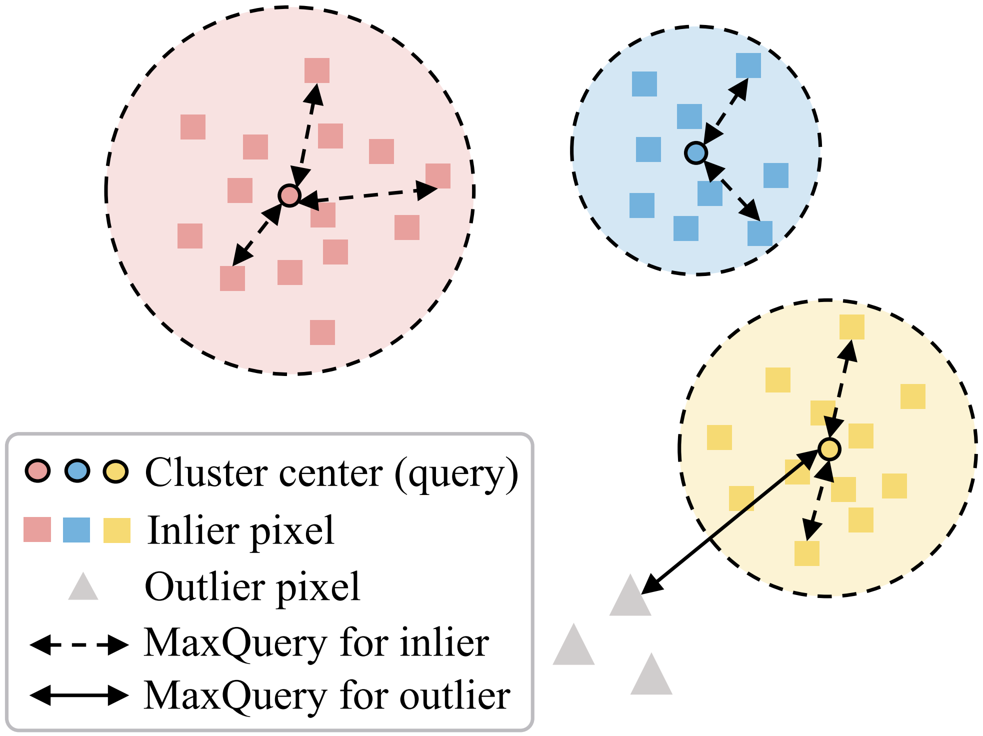
- Manipulate object queries to focus on tumors
- Enforce query-level boundaries
- Benefits both two tasks



Query Distribution (QD) Loss

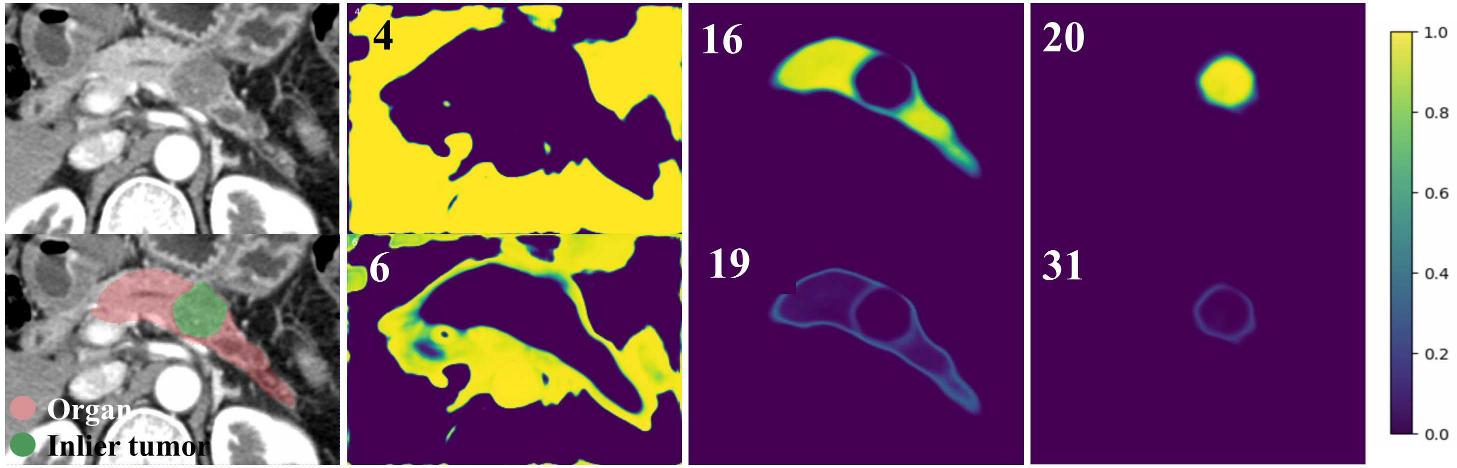


Localizing OOD Regions with MaxQuery

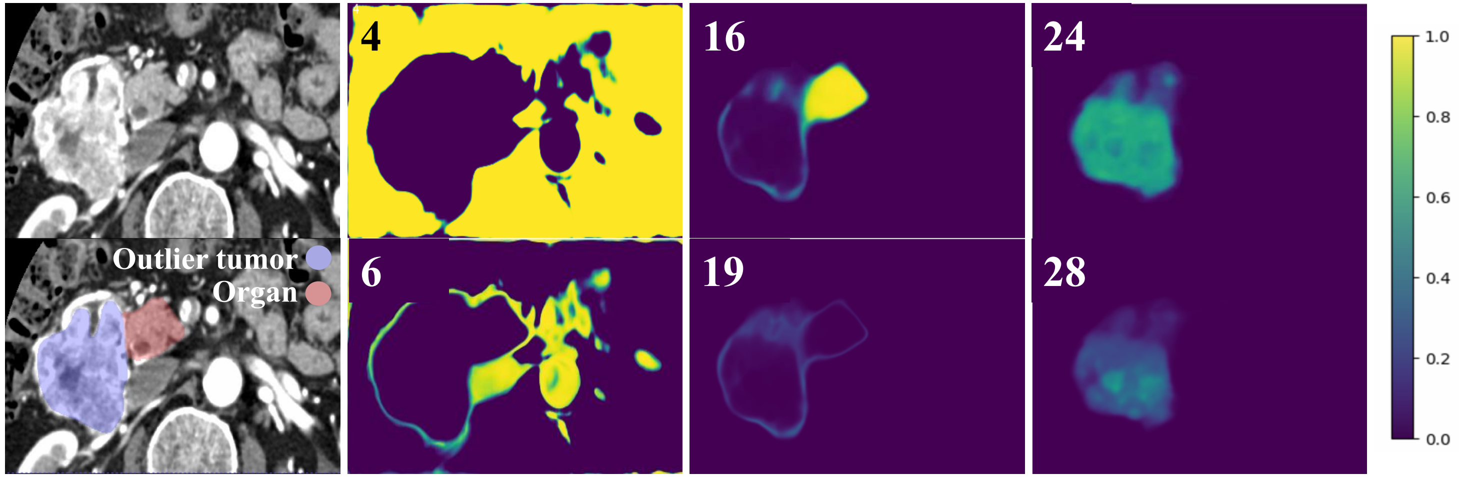


- Motivation: cluster analysis
  - OOD pixels far from inlier centers
- MaxQuery: the negative of maximal query response
  - Reflects the distance/similarity of a pixel and its assigned center
- Intuitively, MaxQuery of inlier pixels is usually smaller than outliers

An illustrative Example



Cluster assignments for an in-distribution example



Cluster assignments for an OOD example

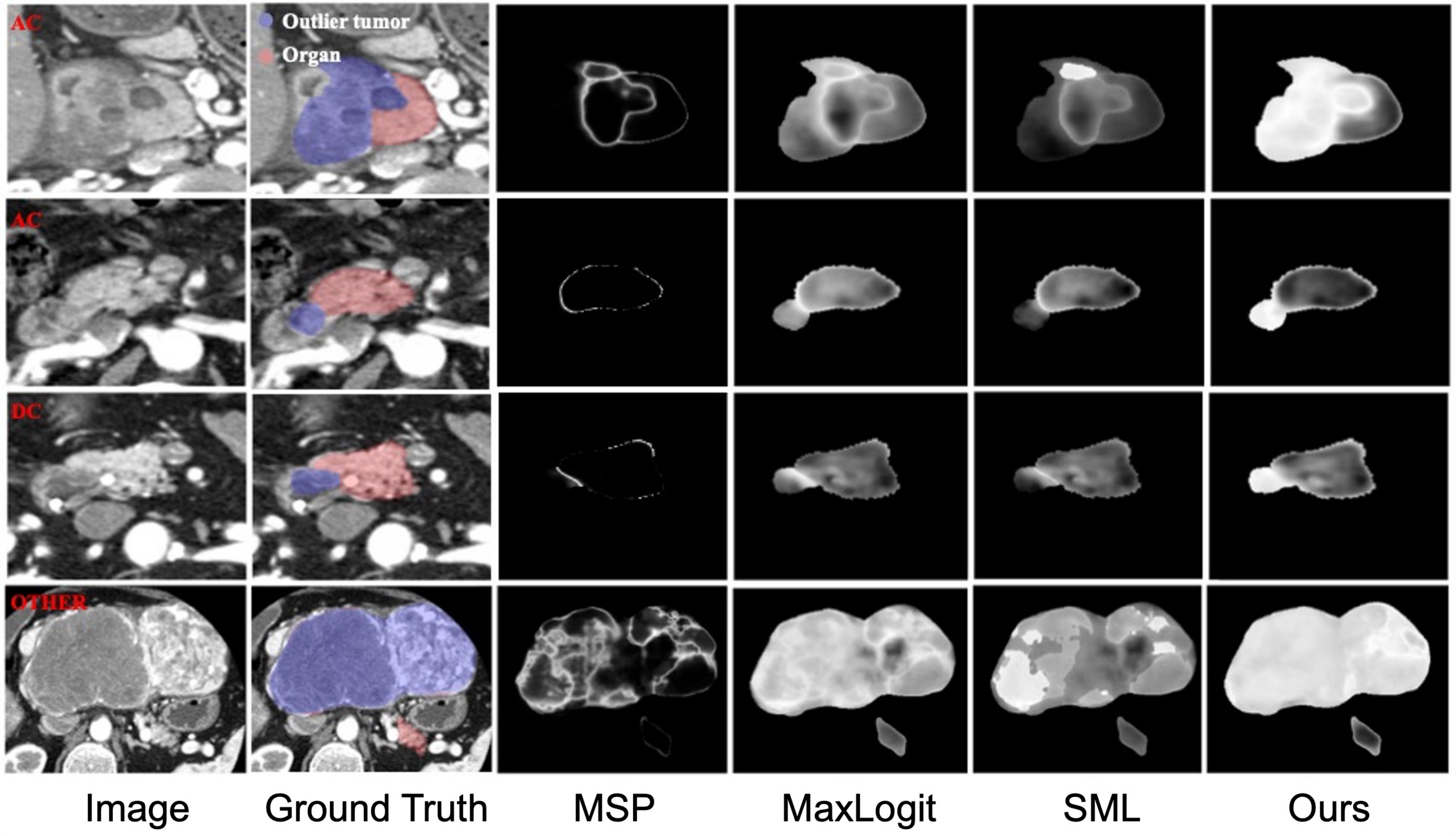
- Background & organ: activated confidently by one/few queries
- Inlier tumor: concentrate at a single query
- Outlier tumor: split into multiple centers with lower response

Main Results

- Two curated real-world tumor segmentation datasets, pancreatic & liver
- State-of-the-art OOD detection performance at both pixel & case level

Methods	Pancreatic %				Liver %			
	OOD Localization			OOD_{case}	OOD Localization			OOD_{case}
	AUROC↑	AUPR↑	FPR_{95} ↓	AUC↑	AUROC↑	AUPR↑	FPR_{95} ↓	AUC↑
MC Dropout [27]	49.08	11.47	84.60	72.91	39.61	16.05	91.13	34.05
MSP [21]	53.81	13.44	86.44	73.38	75.14	25.27	70.04	66.76
MaxLogit [20]	58.46	21.93	83.68	73.42	78.60	35.47	48.73	65.68
SynthCP [52]	69.86	26.50	66.65	68.43	74.93	34.03	57.91	63.34
SML [26]	56.10	30.44	77.81	62.26	86.64	44.59	31.04	63.85
Ours (w/o $\mathcal{L}_{qd}$)	63.54	25.25	67.09	74.87	74.95	42.31	53.52	65.91
Ours	**82.52**	**55.60**	**46.19**	**77.97**	**88.75**	**48.80**	**23.93**	**69.04**

Qualitative Results



Results & Ablation Studies

- Outperforms strong baseline (e.g., nnUNet) for inlier segmentation

	Pancreatic %								Liver %					
Methods	PDAC	IPMN	PNET	SCN	CP	SPT	MCN	Avg.	HCC	ICC	Meta.	Heman.	Cyst	Avg.
UNet [7]	63.96	21.07	21.72	30.70	17.88	33.96	18.10	29.62	61.59	28.76	43.77	65.01	37.39	47.30
UNet++ [12]	63.43	22.85	14.52	25.09	15.02	21.36	10.07	24.62	56.51	29.13	36.88	56.74	46.60	45.17
TransUNet [1]	64.91	31.18	26.78	38.96	22.39	29.87	30.27	34.91	52.26	25.50	42.31	70.90	47.52	47.70
nnUNet [5]	65.65	27.60	32.59	36.46	23.33	31.73	30.96	35.47	57.22	28.16	52.81	77.55	46.49	52.45
Ours	67.91	46.92	32.07	42.51	31.36	42.67	28.97	**41.77**	67.61	30.78	60.40	77.07	47.61	**56.69**

- Robust to query distribution selection

Query Dist. (N_1, N_2, N_3)	AUROC↑	AUPR↑	FPR↓	$\mathrm{DSC}_{\mathrm{inlier}}$ ↑
SML [26]	56.10	30.44	77.81	35.47
(8, 4, 20)	84.44	51.32	42.10	36.50
(8, 20, 4)	83.73	49.76	43.32	39.79
(16, 4, 12)	82.52	**55.60**	46.90	**41.77**
(20, 4, 8)	85.66	55.17	37.24	38.19
(24, 4, 4)	**86.41**	52.58	**33.70**	39.43

- Query-level vs. category-level score

Level	Softmax	AUROC↑	AUPR↑	FPR95↓
Category	post	58.14	16.28	79.29
	pre	52.70	24.59	88.40
Query	post	76.88	33.98	55.82
	pre	**82.52**	**55.60**	**46.19**

Conclusion

- Processing a large collection of medical imaging data with long-tailed distribution is challenging.
- Two curated real-world datasets
- Interpreting segmentation as query/cluster assignment
- Novel MaxQuery & QD loss are evidently helpful.
- Great potential to further boost the adoption of medical image segmentation in designing various clinical applications