UP-SAM: Uncertainty-Informed Adaptation of Segment Anything Model for Semi-Supervised Medical Image Segmentation

Reporter: Wenjing Lu

Co-corresponding Author: Yi Hong, Yang Yang

Department of Computer Science and Engineering, Shanghai Jiao Tong University, Shanghai, China

饮 水 思 源 • 爱 国 荣 校

Contents

01

Introduction

Motivation & Problem Definition & Contributions

Motivation

- **Medical Image Segmentation Challenges:**
  - Limited labeled data, high annotation costs

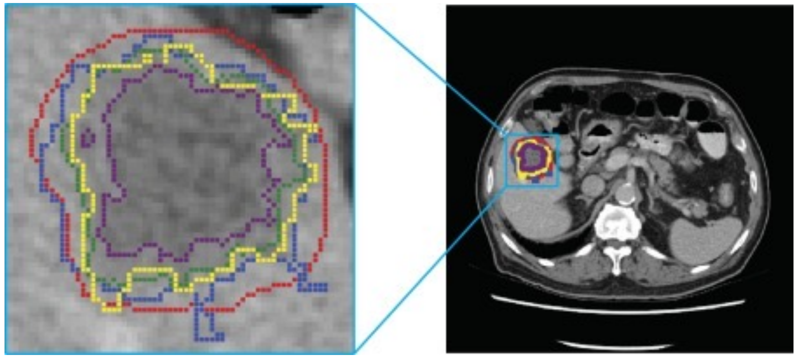


Fig. 1 Example of Expert Annotation

Manual Pixel-wise Annotation:

- Medical Expert
- 1 image ~ 30 mins

[1] S Kevin Zhou, Daniel Rueckert, and Gabor Fichtinger. Handbook of medical image computing and computer assisted intervention. Academic Press, 2019.

Motivation

Medical Image Segmentation Challenges:

- Limited labeled data, high annotation costs

Semi-Supervised Learning (SSL):

- Helps leverage unlabeled data to reduce annotation needs

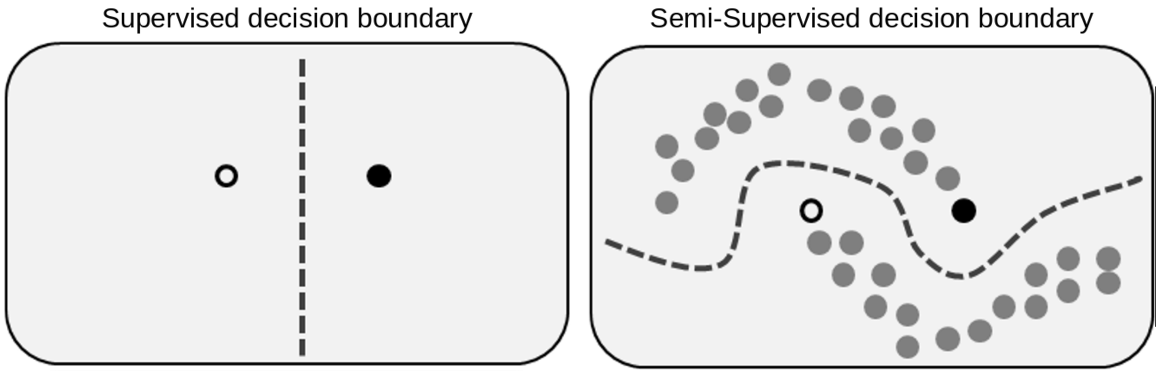


Fig. 2 An example of the influence of unlabeled data in semi-supervised learning.
(Image source: Wikipedia)

Motivation

- **Medical Image Segmentation Challenges:**
  - Limited labeled data, high annotation costs
- **Semi-Supervised Learning (SSL):**
  - Helps leverage unlabeled data to reduce annotation needs
- **Uncertainty in SSL:**
  - Need to handle Epistemic Uncertainty (EU) and Aleatoric Uncertainty (AU) effectively

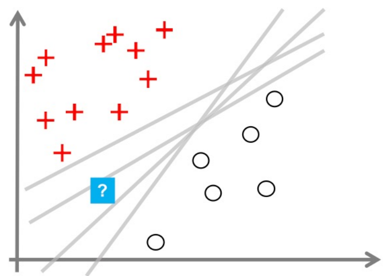


Fig.3 Epistemic Uncertainty

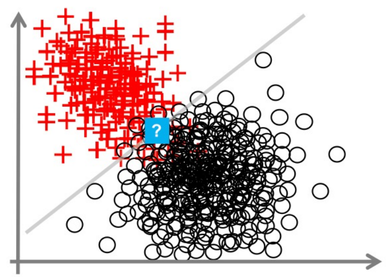


Fig.4 Aleatoric Uncertainty

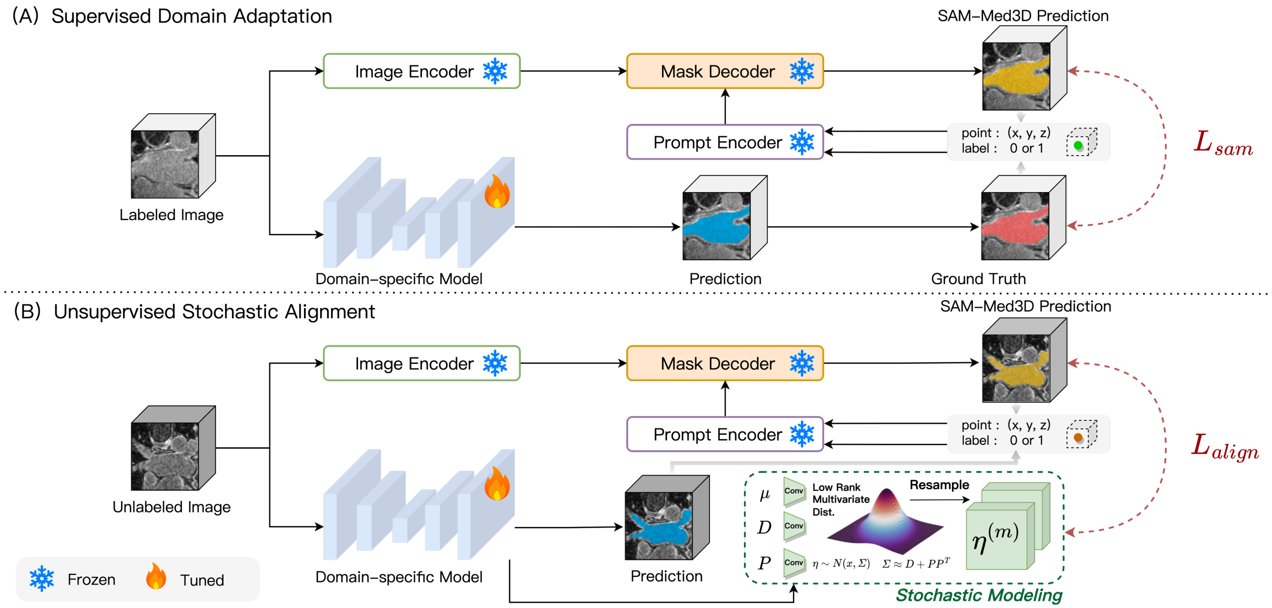
[2] Hüllermeier E, Waegeman W. Aleatoric and epistemic uncertainty in machine learning: An introduction to concepts and methods[J]. Machine learning, 2021, 110(3): 457-506.

Problem Definition

Objective: Improve segmentation performance in semi-supervised medical image analysis with **scarce labeled data**.

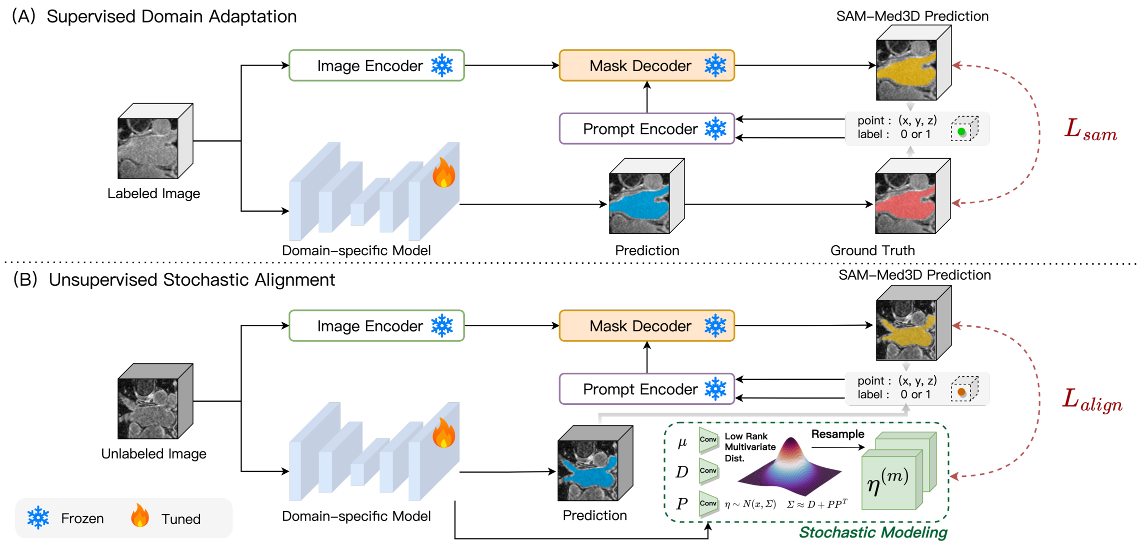
Challenge: Existing methods address EU but often **ignore AU**, which is crucial in medical contexts.

Proposed Framework - UP-SAM Overview



- **Framework Summary:** UP-SAM adapts the Segment Anything Model (SAM) with dual uncertainty assessment for medical image segmentation
- **Key Components:** Fine-tuning SAM to address EU, Stochastic modeling for AU on unlabeled data

Contributions



Overview of UP-SAM

- Dual uncertainty handling (EU & AU)
- Fine-tuning of SAM for domain-specific adaptation
- Stochastic alignment for AU using logit alignment
- Experimental validation showing superior performance

02

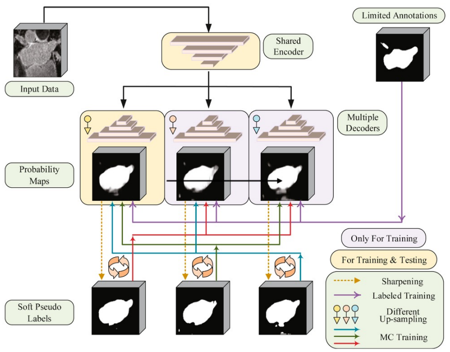
Related Work

Semi-supervised Learning & Foundation Models & Uncertainty Estimation

Related Work - Semi-Supervised Learning

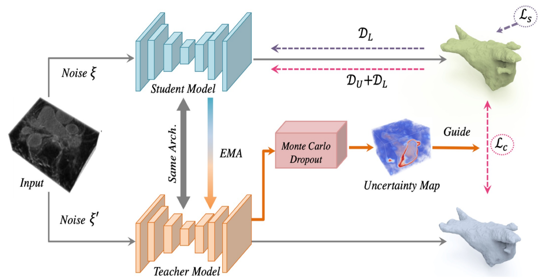
AC-MT (MIA’23)

Consistency Regularization



MC-Net+ (MIA’22)

Pseudo Labeling



UA-MT (MICCAI’19)

Uncertainty Estimation

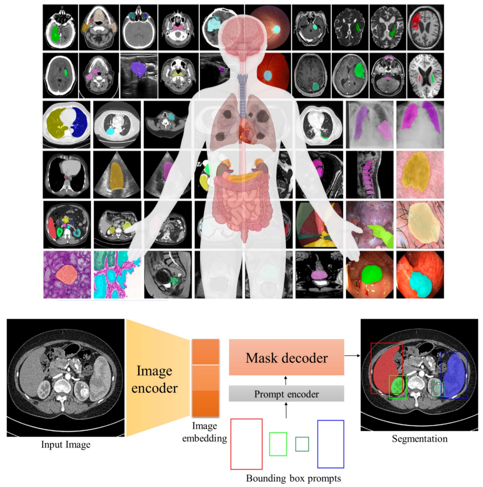
Limitations in Existing Methods:

<u>**Limited handling of structured uncertainty (AU),**</u> which affects segmentation reliability

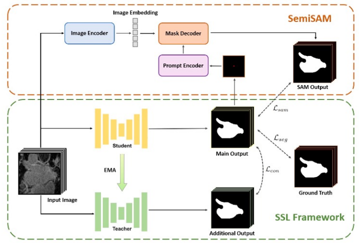
[3] Mean teachers are better role models: Weight-averaged consistency targets improve semi-supervised deep learning results. Advances in neural information processing systems
[4] “Mutual consistency learning for semi-supervised medical image segmentation,” Medical Image Analysis, vol. 81, p. 102530, 2022.
[5] Ambiguity-selective consistency regularization for mean-teacher semi-supervised medical image segmentation[J]. Medical Image Analysis, 2023, 88: 102880.

Related Work - Foundation Models in Medical Imaging

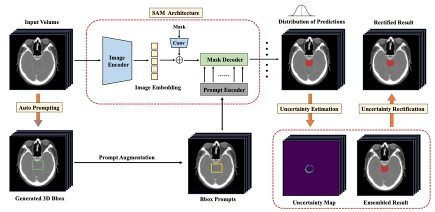
Segment Anything Model
(SAM)



Segment Anything in medical images
(MedSAM)



SemiSAM



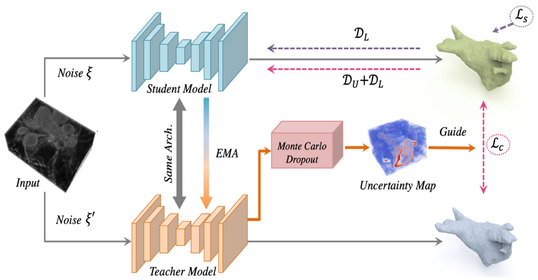
UR-SAM

SAM in Medical Imaging:

Strong results but often limited by dependence on **manual prompts and gaps in AU handling**

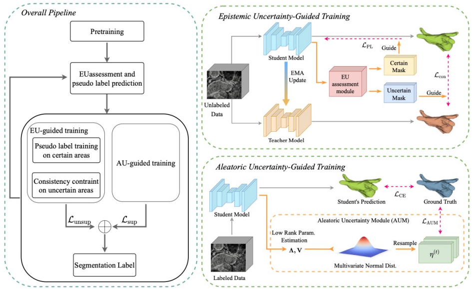
[6] Kirillov A, Mintun E, Ravi N, et al. Segment anything[C]//Proceedings of the IEEE/CVF International Conference on Computer Vision. 2023: 4015-4026.
[7] Zhang Y, Hu S, Jiang C, et al. Segment anything model with uncertainty rectification for auto-prompting medical image segmentation[J]. arXiv preprint arXiv:2311.10529, 2023.
[8] Zhang Y, Cheng Y, Qi Y. Semisam: Exploring sam for enhancing semi-supervised medical image segmentation with extremely limited annotations[J]. arXiv preprint arXiv:2312.06316, 2023.
[9] Ma J, He Y, Li F, et al. Segment anything in medical images[J]. Nature Communications, 2024, 15(1): 654.

Related Work - Uncertainty Estimation



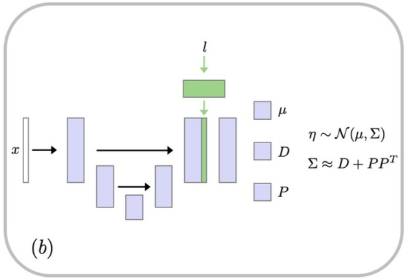
UA-MT (MICCAI'19)

Monte Carlo Dropout



FUSSNet (MICCAI'22)

Fuse AU and EU in SSL



c-SSNs (ICLR'23)

Conditional Modeling AU

Gap:

AU estimation remains underexplored, especially in SSL frameworks

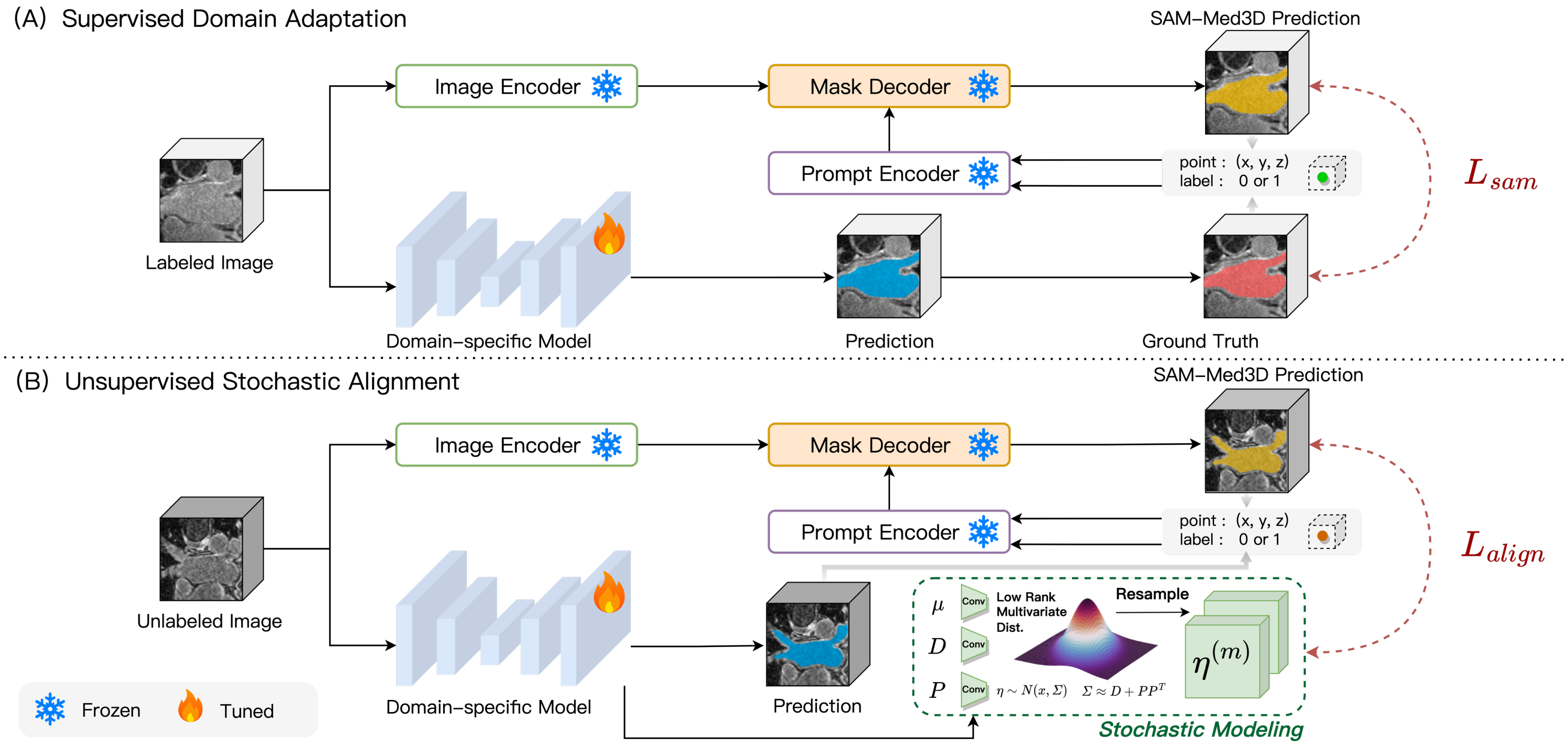
[10] Mean teachers are better role models: Weight-averaged consistency targets improve semi-supervised deep learning results. Advances in neural information processing systems 30 (2017).
[11] J. Xiang et al., "Fussnet: Fusing two sources of uncertainty for semi- supervised medical image segmentation," in MICCAI 2022. Springer, 2022, pp. 481–491.
[12] K. Zepf et al., "That label's got style: Handling label style bias for uncertain image segmentation," arXiv preprint arXiv:2303.15850, 2023.

03

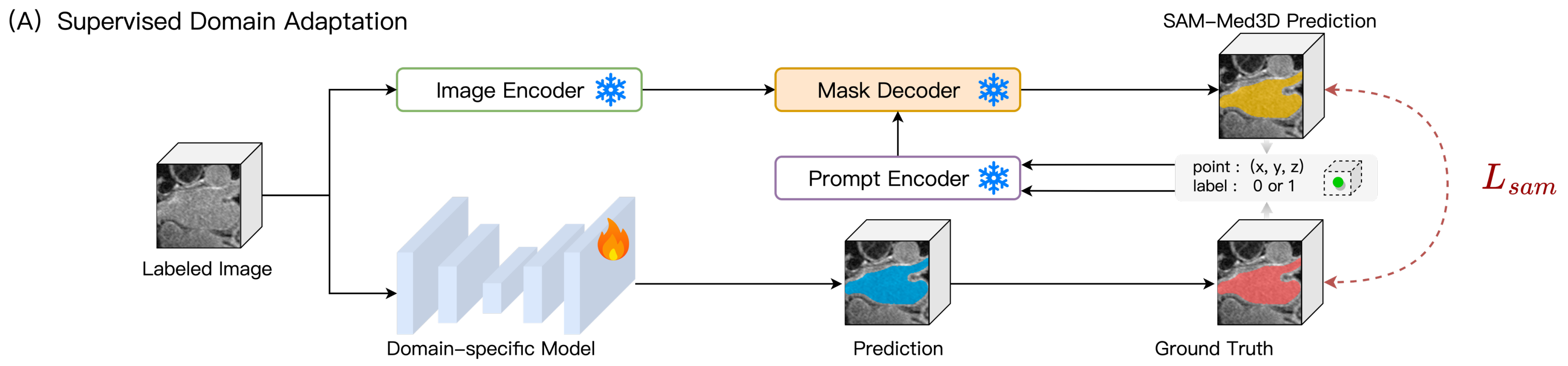
Methodology

Supervised Domain Adaptation & Unsupervised Stochastic Alignment

Methodology - Framework Structure



Methodology - Supervised Domain Adaptation



Domain-Specific Model Branch:

- Uses multiple classification heads for EU reduction

$$\mathrm{L}_{seg} = \sum_{h \in H} L_h\left(p_\phi^h(x_i^l), y_i^l\right)$$

$H = \{ce, dice, focal, iou\ loss\}$
p_ϕ^h domain-specific model predictions from EU head h
x_i^l, y_i^l labeled image and ground truth of sample i

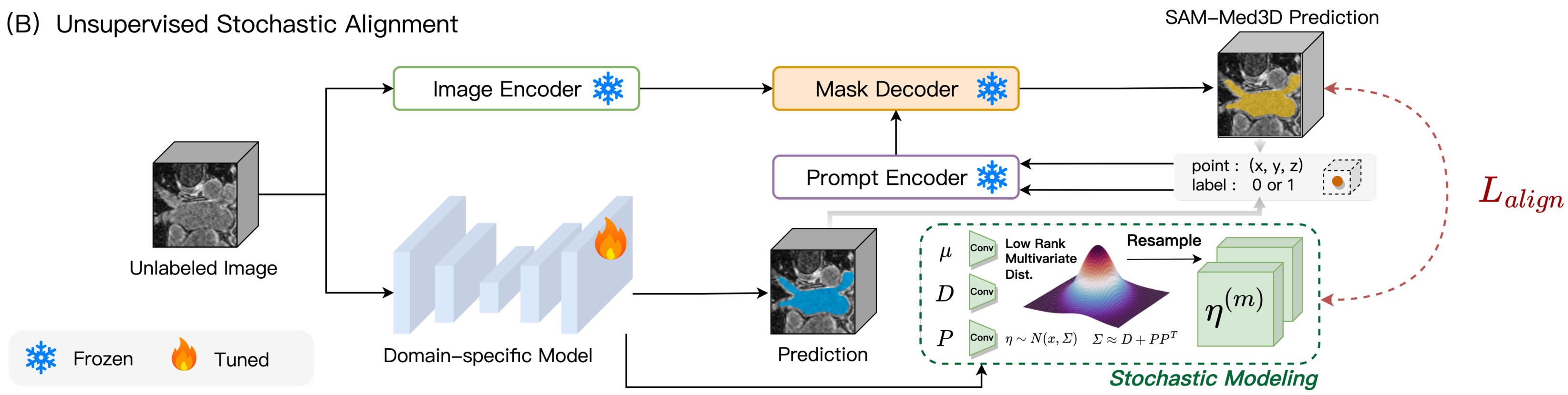
SAM-Med3D Branch:

- Fine-tuning of the mask decoder layers to maintain SAM's core knowledge

$$\mathrm{L}_{sam} = L_{ce+dice}\left(p_s(x_i^l), y_i^l\right)$$

p_s foundation model predictions

Methodology - Unsupervised Stochastic Alignment



Stochastic Modeling for AU:

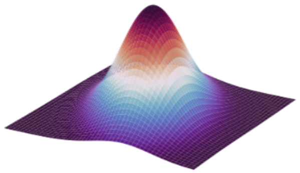
- Alignment of logit distributions for unlabeled data

Key Details:

- Multivariate normal distribution, Monte-Carlo sampling to refine predictions

Multivariate normal distribution

Multivariance Normal Distribution



$\eta | x \sim N(\mu(x), \Sigma(x))$

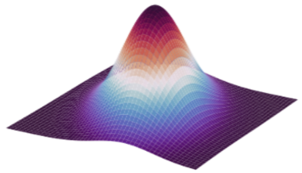
$S = \#\, pixels, C = \#\, classes$

$\mu(x)\ \epsilon\ R^{S\times C}$

$\Sigma(x)\ \epsilon\ R^{(S\times C)^2}$

Computational Cost High !!!

Low-Rank
Multivariance Normal Distribution



$\eta | x \sim N(\mu(x), \Sigma(x))$ $\quad \Sigma(x)=D(x)+P(x)P(x)^T$

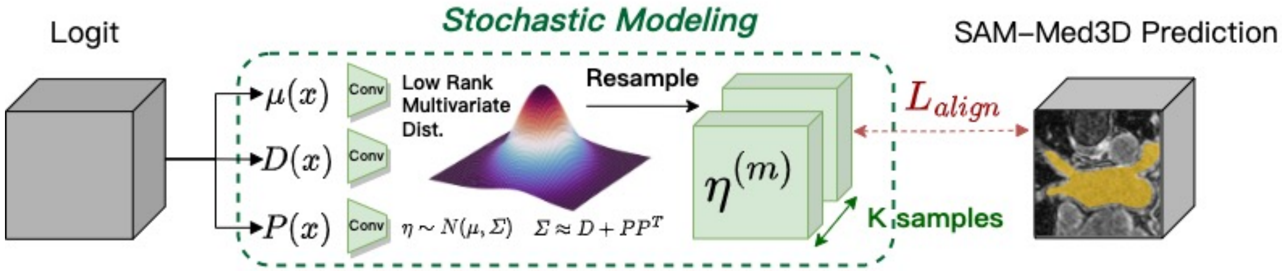
$S = \#\, pixels, C = \#\, classes, R = \#\, Rank$

$\mu(x)\ \epsilon\ R^{S\times C}$

$D(x)\ \epsilon R^{S\times C}$ $\quad P(x)\ \epsilon R^{(S\times C)\times R}$

Computational Efficiency Enhanced!!!

Monte-Carlo sampling to refine predictions



- **Estimating the conditional log-likelihood**
  - K Monte Carlo samples to refine μ(x), P(x), and D(x)

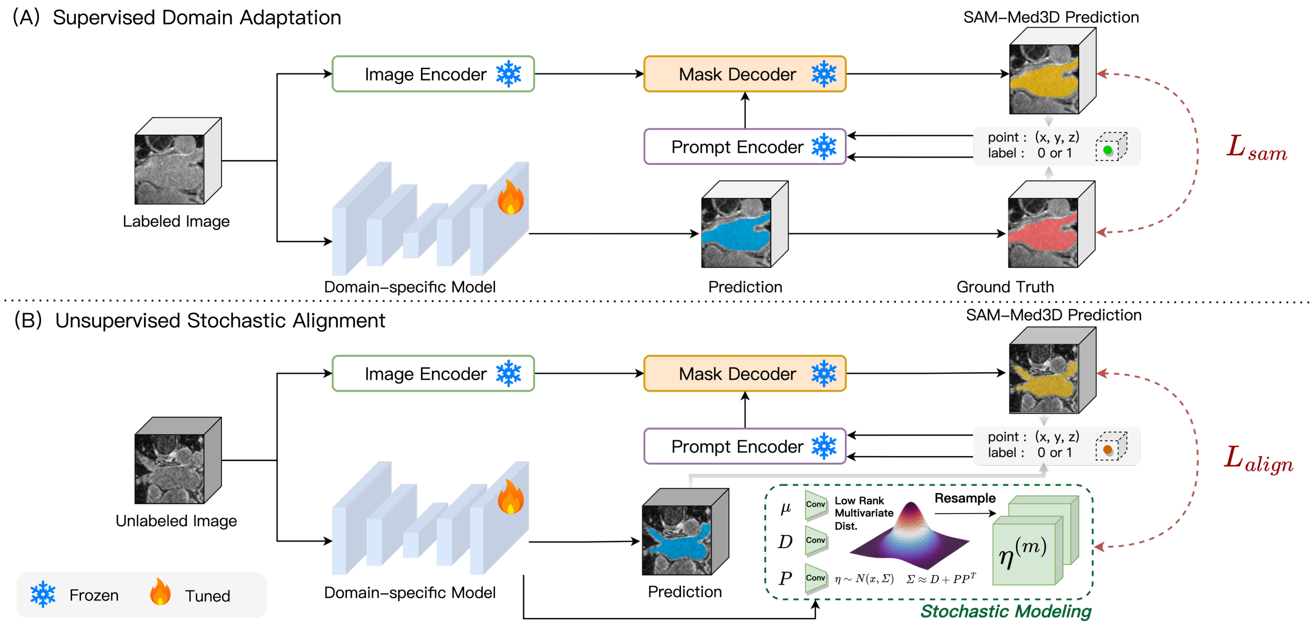
$$\log p(y|x) = \log \int p(y|\eta_\theta) p_\theta(\eta_\theta|x)\, d\eta \approx \log \frac{1}{K} \sum_{k=1}^{K} p(y|\eta_\phi^{(k)})$$

$p_\theta(\eta_\theta|x)$: *logit probability distribution conditioned on the image, parameterized by θ.*

- **Align logits to foundation model masks**

$$L_{align} = -\sum_{k=1}^{K} \sum_{i}^{H\times W\times D} \sum_{c}^{C} \hat{y}_{s,ic}^{u} \log(softmax(\eta_{\phi,ic}^{(k)}))$$

Methodology – Overall Loss



$$Loss = \alpha\left(L_{seg} + L_{sam}\right) + \lambda L_{align}$$

04

Experiments

Dataset and Experimental Setup & Experimental Results & Ablation Study

Datasets and Experimental Setup

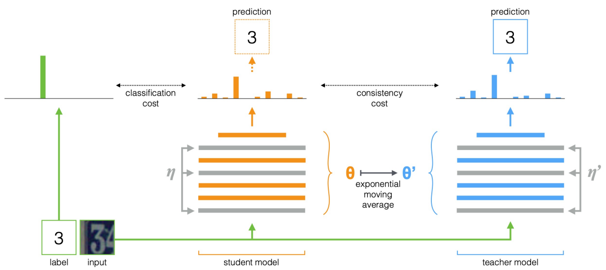
Datasets:

- Left Atrium (LA) : 100 scans
- Pancreas-CT datasets: 82 scans

Setup Details:

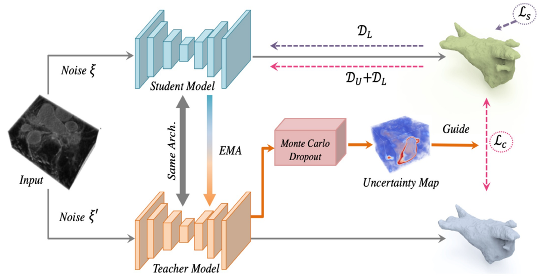
- Training iterations: pretraining 7.5k iterations, finetuning 7.5k
- Learning rates: SGD (lr = 10−2) AdamW (lr = 10−5)
- Batch sizes: 4 (2 labeled, 2 unlabeled) or 2 (1 labeled, 1 unlabeled)
- Performance metrics: Dice, Jac, 95HD, ASSD

Baseline Methods



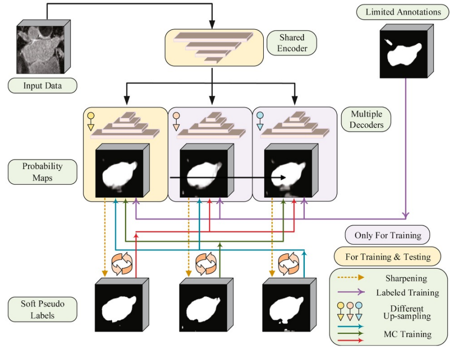
MT (NerulP'17)

Consistency Regularization



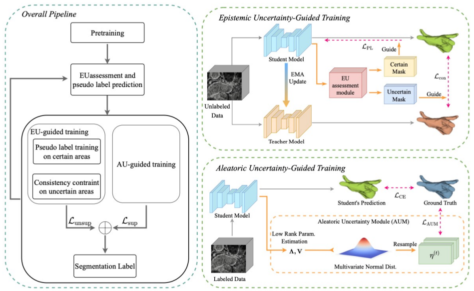
UA-MT (MICCAI'19)

Uncertainty Estimation



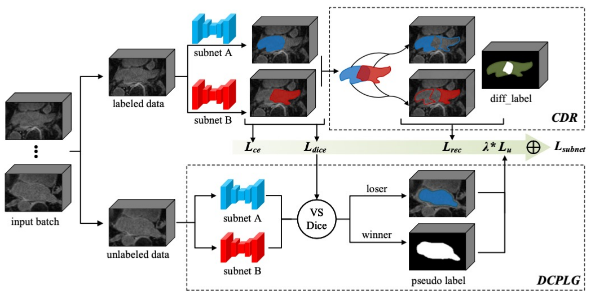
MC-Net+ (MIA'22)

Pseudo Labeling



FUSSNet (MICCAI'22)

Fuse AU and EU in SSL



MCF (CVPR'23)

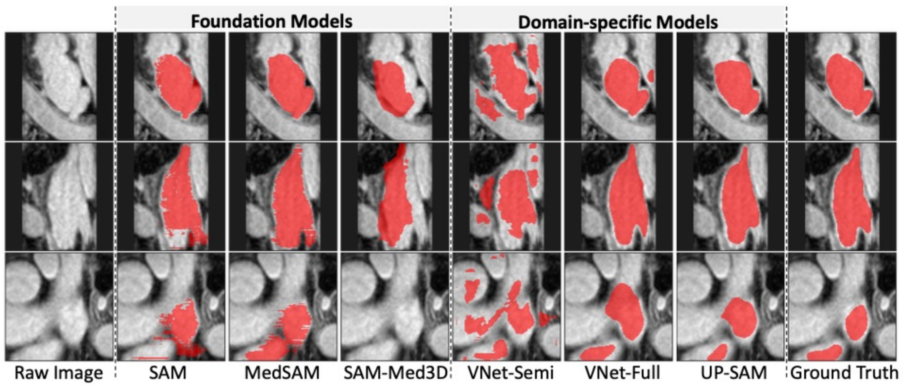
Dual-network Interactions

AC-MT (MIA'23)

Consistency Regularization

Experimental Results - Left Atrium Dataset

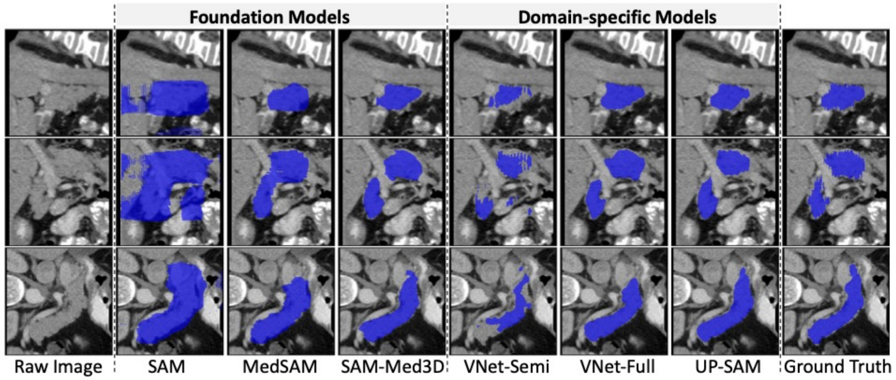
- UP-SAM shows **superior performance**, especially with **limited labeled scans**.
- UP-SAM provides **accurate and consistent segmentation boundaries**



Method	#Lb/#Unlb	Dice(%)	Jac(%)	95HD	ASD
Pretrained foundation models for Zero-shot Segmentation					
SAM	0 / 80	77.38	63.43	12.72	4.75
SAM-Med3D		77.96	64.78	13.03	3.73
Semi-Supervised Segmentation with Extremely Limited Labeled Data					
VNet	1 / 79	7.20	5.69	44.14	26.92
MT (NIPS'17)		44.16	29.66	47.02	**2.20**
UA-MT (MICCAI'19)		46.51	31.84	44.11	5.42
MC-Net+ (MIA'22)		4.09	2.92	37.26	9.98
FUSSNet (MICCAI'22)		65.88	50.23	40.69	13.92
MCF (CVPR'23)		5.33	4.42	32.41	19.88
AC-MT (MIA'23)		46.74	32.21	43.99	18.44
UP-SAM		**78.25**	**65.31**	**18.83**	4.39
VNet	2 / 78	46.74	32.60	39.02	6.24
MT (NIPS'17)		69.32	53.87	37.21	2.92
UA-MT (MICCAI'19)		73.38	58.82	31.94	2.69
MC-Net+ (MIA'22)		27.17	18.51	34.45	11.70
FUSSNet (MICCAI'22)		74.92	61.58	27.17	8.81
MCF (CVPR'23)		67.57	52.49	32.33	**2.60**
AC-MT (MIA'23)		81.33	68.87	16.50	5.16
UP-SAM		**82.84**	**71.25**	**16.03**	4.36
VNet	4 / 76	67.34	55.26	25.70	7.23
MT (NIPS'17)		74.64	60.77	34.14	2.72
UA-MT (MICCAI'19)		75.82	62.09	28.36	3.27
MC-Net+ (MIA'22)		78.66	65.88	22.27	6.04
FUSSNet (MICCAI'22)		81.97	70.46	20.54	5.94
MCF (CVPR'23)		82.73	71.32	16.59	2.28
AC-MT (MIA'23)		**86.04**	**76.03**	**9.23**	**2.50**
UP-SAM		84.06	72.90	13.78	3.14
Fully Supervised Segmentation					
VNet	80 / 0	91.42	84.27	5.15	1.50
MedSAM		81.34	68.73	10.47	3.71

Experimental Results - Pancreas-CT Dataset

- UP-SAM nearly **matches fully supervised models**, highlighting effectiveness with minimal annotations
- UP-SAM **captures complex shapes**, comparable to fully supervised models



Method	#Lb/#Unlb	Dice(%)	Jac(%)	95HD	ASD
Pretrained foundation models for Zero-shot Segmentation					
SAM*	0 / 62	35.95	22.06	37.41	17.04
SAM-Med3D*		79.60	66.36	5.34	1.49
Semi-Supervised Segmentation with Extremely Limited Labeled Data					
VNet	1 / 61	19.94	11.86	43.29	11.38
MT (NIPS'17)		19.22	10.84	67.01	2.38
UA-MT (MICCAI'19)		9.51	5.03	70.06	2.69
MC-Net+ (MIA'22)		10.42	5.54	65.35	33.52
FUSSNet (MICCAI'22)		30.67	18.61	43.86	19.86
MCF (CVPR'23)		11.45	6.10	67.45	**2.15**
AC-MT (MIA'23)		15.94	9.18	56.85	30.23
UP-SAM		**70.79**	**55.25**	**11.00**	3.41
VNet	2 / 60	37.20	24.73	41.18	5.55
MT (NIPS'17)		37.34	23.66	51.17	2.49
UA-MT (MICCAI'19)		21.44	12.33	61.12	3.63
MC-Net+ (MIA'22)		17.57	9.82	59.81	29.32
FUSSNet (MICCAI'22)		40.53	26.10	45.95	18.86
MCF (CVPR'23)		19.35	10.85	63.64	**1.71**
AC-MT (MIA'23)		21.63	13.04	65.81	32.84
UP-SAM		**71.12**	**55.61**	**11.39**	3.38
VNet	4 / 58	50.01	34.46	41.01	3.67
MT (NIPS'17)		35.67	22.71	57.53	3.51
UA-MT (MICCAI'19)		34.60	22.31	53.38	3.63
MC-Net+ (MIA'22)		33.78	21.11	53.33	25.13
FUSSNet (MICCAI'22)		58.81	43.37	34.44	12.51
MCF (CVPR'23)		42.28	28.70	53.56	27.53
AC-MT (MIA'23)		46.42	31.36	51.71	2.94
UP-SAM		**72.65**	**57.42**	**10.49**	**2.82**
Fully Supervised Segmentation					
VNet	62 / 0	82.53	70.63	6.01	1.88
MedSAM*		70.92	55.09	7.71	2.68

Ablation Study

- **Component Analysis:**
  - Impact of each component (EU Heads, SDA, USA) on Dice scores
- **Result:**
  - Dual EU & AU treatment improves performance significantly

Method		Backbone		Loss			Metrics	
		VNet	SAM-Med3D	L_{seg}	L_{sam}	L_{align}	Dice(%)	ASD
Sup-Only	Baseline	✓					46.74	6.24
	+ EU Heads	✓		✓			72.81	10.86
	+ SDA	✓	✓	✓	✓		77.78	7.10
Semi-Sup	+ USA	✓	✓	✓		✓	81.31	3.93
	UP-SAM	✓	✓	✓	✓	✓	82.84	4.36

05

Conclusion

Conclusion & Future Work

Conclusion

- **Efficient Mitigation of Label Sparsity**
  - Address label sparsity through epistemic and aleatoric uncertainty assessment within a dual-model architecture
- **Collaboration between foundation models and domain-specific models.**
  - Fine-tunes SAM-Med3D to address EU and logit space alignment to handle AU
- **Exceptional Segmentation Performance**
  - Outperforms existing models with minimal labeled data and prompt-free inference.

Future Work

- Extending this framework to multi-class segmentation tasks
- Further refining it to minimize redundant knowledge interference.

Thanks

饮 水 思 源 · 爱 国 荣 校