Lightweight Dual-Task Framework for Semi-Supervised Lesion Segmentation with Knowledge Distillation from SAM

Xuan-Loc Huynh, Huy-Thach Pham, Anh Mai Vu, Thanh-Minh Nguyen, Tran Quang Khai Bui, Tat-Bach Nguyen, Quan Nguyen, Minh Huu Nhat Le, and Phat K. Huynh

Boston University & North Carolina A&T State University

Background & Related Work

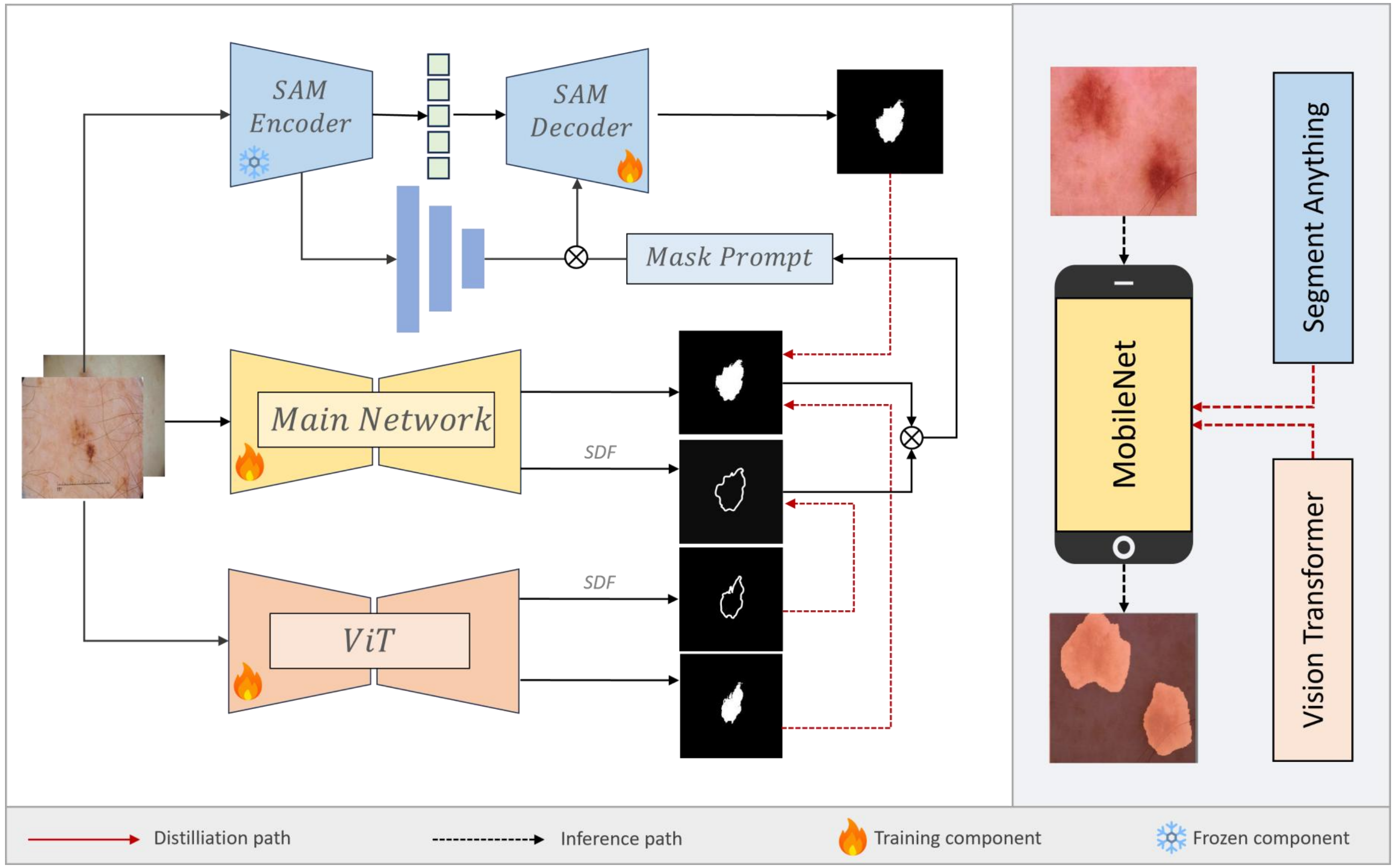
- Semi-supervised segmentation: consistency regularization, pseudo-labeling (FixMatch, CPS, PseudoSeg).
- Multi-task & dual-branch models → enhance boundary accuracy.
- Knowledge distillation from foundation models (SAM, LPS, SKD).
- Need for boundary-aware prompts + efficient students

Contributions

- Lightweight Co-Training: MobileNet main network + ViT/SAM teacher for efficiency vs representation.
- Fused Mask Prompt: merge coarse mask & SDF map for boundary-aware SAM guidance.
- SAM-Guided KD: refine SAM outputs as pseudo-labels for the student.
- Achieves accurate, boundary-aligned predictions under low labels

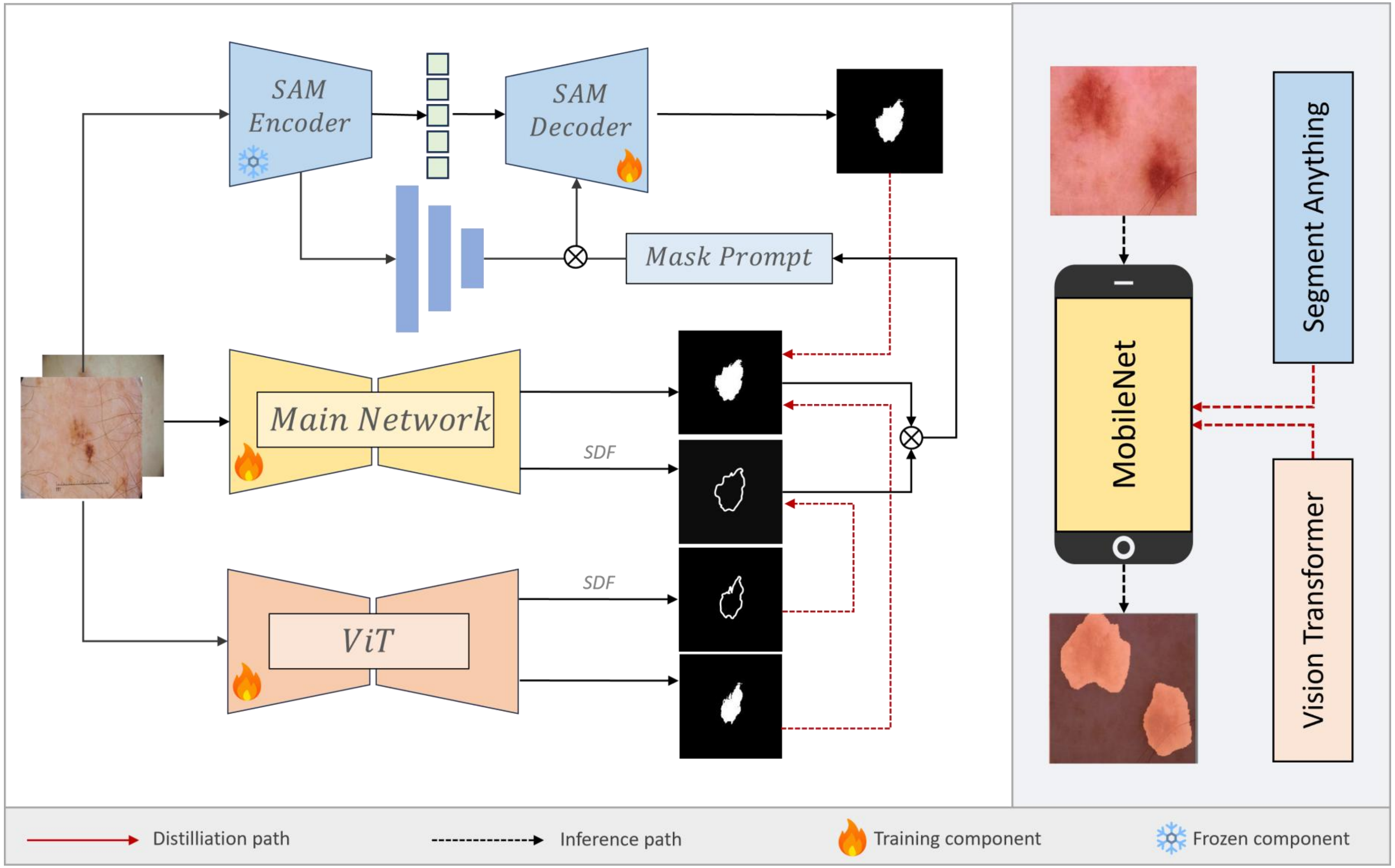
Method Overview

- Main Network predicts:Binary mask (semantic).Signed Distance Function (boundary).
- Outputs fused → mask prompt → SAM decoder → refined masks.
- SAM & Teacher supply pseudo-labels; only Main net used at inference



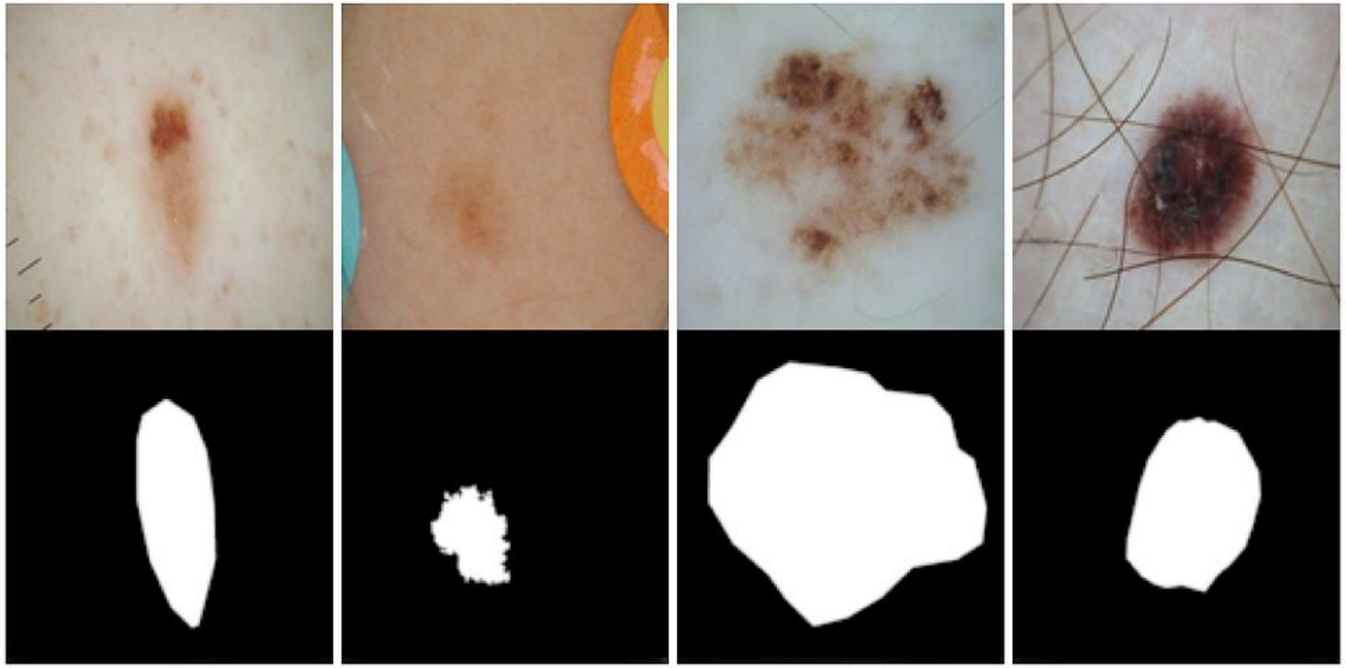
Key Modules

- Dual-Output Head: simultaneous mask + SDF prediction → spatial & structural cues.
- Prompt Generation: learnable decoder + fused mask/SDF tokens for SAM.
- Knowledge Distillation: KL between SAM logits & student predictions (soft targets)



Experimental Setup

- Datasets: ISIC-2018 & HAM10000 (2%, 4%, 8% labeled).
- Baselines: PseudoSeg, CCT, CPS, GTA-Seg, Unimatch, DME-FD, SemiSAM.
- Metrics: Dice, IoU, Sensitivity, Specificity.
- Model size & inference speed also reported



Experimental Results

Table 1. Segmentation performance on ISIC-2018 under 2% and 4% labeled data settings. The SupOnly row reports results using fully supervised training.

Method	Data (%)		Metrics			
	Label	Unlabel	Dice (%) ↑	IoU (%) ↑	Sensitivity (%) ↑	Specificity (%) ↑
SupOnly	2%	0%	74.65 ±2.92	60.81 ±2.99	76.16 ±2.44	93.38 ±3.01
	4%	0%	77.23 ±0.48	65.35 ±0.56	79.53 ±1.37	95.10 ±0.59
	8%	0%	82.28 ±0.61	70.66 ±0.85	81.35 ±0.42	96.12 ±0.21
	100%	0%	87.66 ±0.93	78.49 ±1.38	87.11 ±1.05	96.90 ±0.35
PseudoSeg	2%	98%	79.76 ±2.11	67.16 ±2.77	76.65 ±3.72	96.26 ±0.83
CCT			78.66 ±2.02	65.80 ±2.63	77.17 ±4.15	95.56 ±0.69
CPS			79.61 ±1.66	67.04 ±2.28	78.43 ±4.64	95.52 ±1.04
GTA-Seg			77.33 ±2.20	64.21 ±2.59	80.04 ±3.87	93.37 ±2.00
Unimatch			80.03 ±2.04	67.55 ±2.71	78.46 ±4.74	95.84 ±1.50
DME-FD			80.07 ±1.75	67.62 ±2.37	78.97 ±3.51	95.69 ±0.58
Ours			**83.44 ±1.91**	**68.97 ±2.50**	**80.67 ±3.79**	**97.01 ±0.72**
PseudoSeg	4%	96%	81.77 ±0.66	71.18 ±1.03	81.98 ±3.13	96.37 ±0.85
CCT			80.96 ±1.11	68.95 ±1.41	79.75 ±1.68	95.93 ±0.28
CPS			80.89 ±0.91	70.31 ±1.07	82.08 ±2.35	95.67 ±0.96
GTA-Seg			80.83 ±0.80	70.03 ±1.07	82.54 ±2.35	94.64 ±1.46
Unimatch			81.41 ±1.22	69.46 ±1.58	79.50 ±1.76	96.34 ±0.56
DME-FD			82.06 ±0.69	71.54 ±1.04	82.87 ±1.58	96.23 ±0.36
Ours			**85.33 ±0.72**	**73.61 ±0.98**	**84.01 ±1.21**	**96.91 ±0.60**

Experimental Results

Table 2. Segmentation performance on HAM10000 under 2% and 4% labeled data settings. The SupOnly row reports results using fully supervised training.

Method	Data (%) Label	Unlabel	Metrics Dice (%) ↑	IoU (%) ↑	Sensitivity (%) ↑	Specificity (%) ↑
SupOnly	2%	0%	88.15 ±0.21	78.90 ±0.31	88.37 ±0.72	95.72 ±0.27
	4%	0%	89.59 ±0.07	81.24 ±0.12	88.58 ±0.97	96.80 ±0.40
	8%	0%	91.46 ±0.22	84.33 ±0.37	91.01 ±0.29	97.17 ±0.06
	100%	0%	93.54 ±0.25	87.92 ±0.42	93.30 ±0.07	97.80 ±0.22
PseudoSeg	2%	98%	90.02 ±0.17	81.94 ±0.28	88.18 ±1.32	97.29 ±0.51
CCT			89.93 ±0.10	81.79 ±0.15	88.54 ±0.86	97.09 ±0.30
CPS			89.94 ±0.14	81.81 ±0.23	87.95 ±0.48	97.35 ±0.28
GTA-Seg			89.55 ±0.32	81.17 ±0.54	88.89 ±0.70	96.60 ±0.09
Unimatch			89.66 ±0.15	81.35 ±0.26	87.89 ±0.38	97.15 ±0.31
DME-FD			90.45 ±0.17	82.65 ±0.27	88.74 ±0.83	97.39 ±0.44
Ours			**91.02 ±0.21**	**83.02 ±0.32**	**89.00 ±0.67**	**97.91 ±0.35**
PseudoSeg	4%	96%	90.97 ±0.39	83.21 ±0.64	89.11 ±0.77	97.49 ±0.45
CCT			90.64 ±0.53	82.97 ±0.86	89.08 ±0.94	97.39 ±0.15
CPS			90.76 ±0.51	83.17 ±0.84	89.20 ±0.54	97.44 ±0.19
GTA-Seg			90.86 ±0.19	83.34 ±0.31	89.74 ±0.69	97.24 ±0.32
Unimatch			90.32 ±0.44	82.43 ±0.73	88.93 ±1.29	97.20 ±0.61
DME-FD			91.13 ±0.30	83.79 ±0.50	**90.05 ±0.43**	97.33 ±0.08
Ours			**91.56 ±0.37**	**83.81 ±0.52**	90.02 ±0.60	**97.82 ±0.12**

Experimental Results

Table 3. Comparison with SemiSAM methods on ISIC-2018, where SemiSAM denotes approaches that incorporate SAM for semi-supervised segmentation.

Method	Data (%)		Metrics		Params (M)	Speed (s)
	Label	Unlabel	Dice ↑	IoU ↑		
Unet (SupOnly)	100%	0%	0.7723 ± 0.0048	0.6535 ± 0.0056	7.8	0.8
DME-FD	4%	96%	0.8206 ± 0.0069	0.7154 ± 0.0104	41.4	1.2
SemiSAM			0.8412 ± 0.0110	0.7213 ± 0.0149	7.8	0.8
Ours			$\mathbf{0.8533 \pm 0.0072}$	$\mathbf{0.7361 \pm 0.0098}$	**2.9**	**0.02**

Conclusion

- Proposed a lightweight, SAM-guided SSL framework for lesion segmentation.
- Fused mask prompt + SDF improves boundaries; KD enables learning with few labels.
- Meets accuracy–efficiency trade-off for clinical deployment.
- Future works: extend to other modalities, explore adaptive prompts & uncertainty filtering