清華大學
Tsinghua University

Focused and Collaborative Feedback Integration for Interactive Image Segmentation

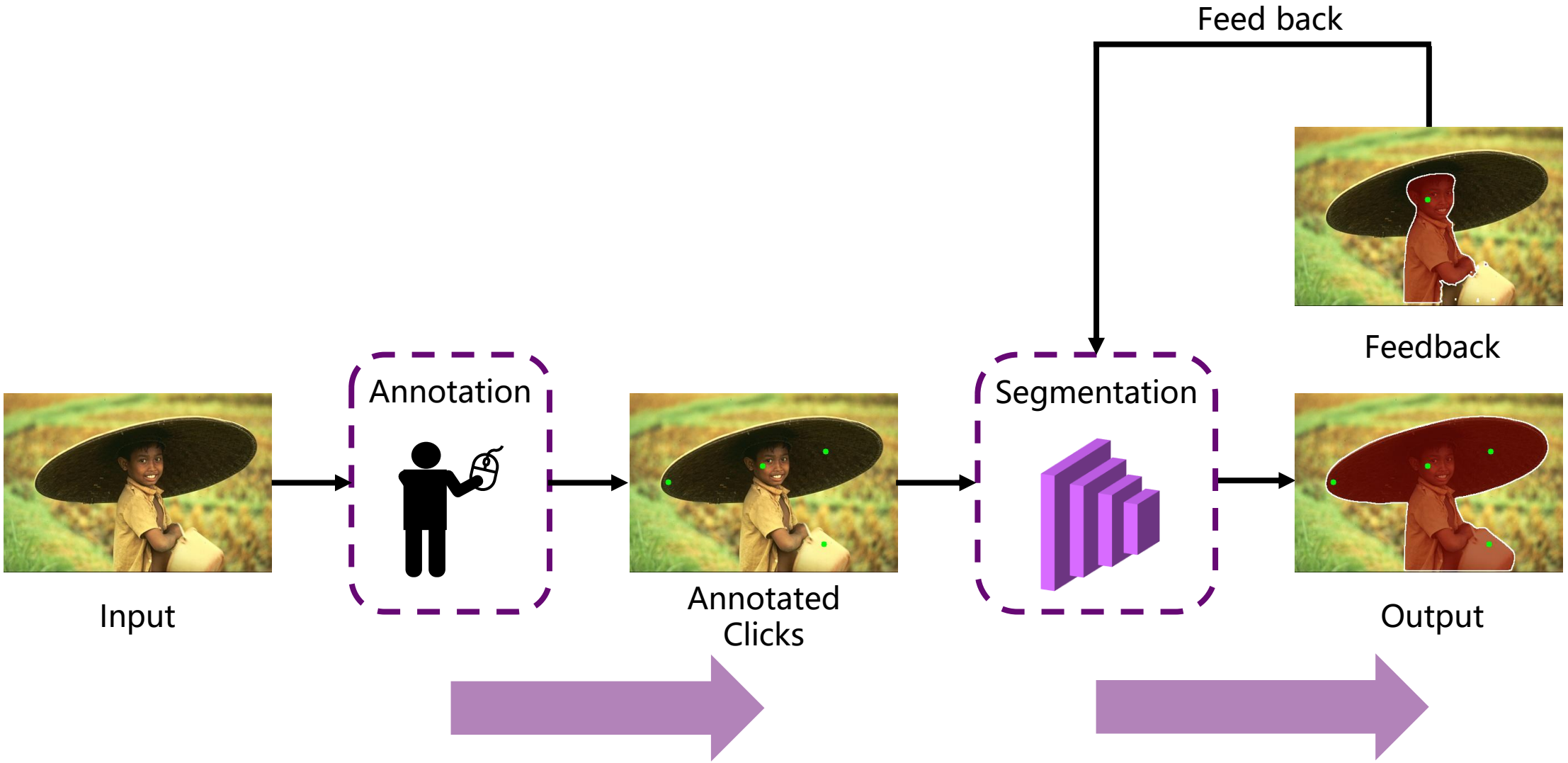
Qiaoqiao Wei, Hui Zhang, Jun-Hai Yong

Tsinghua University

Paper Tag: THU-AM-207

Takeaways

Goal


Feed back
Feedback
Annotation
Segmentation
Input
Annotated
Clicks
Output

Contributions

FCFI

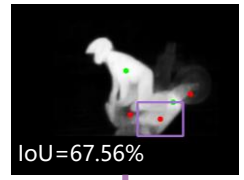
Focused and Collaborative Feedback Integration

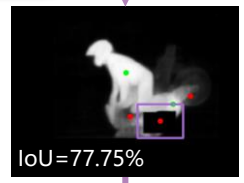
Focused Feedback Correction Module

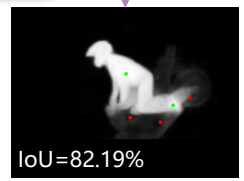
- Corrects feedback from a local perspective based on similarities in the feature space.
- Parameter-free and fast.

Collaborative Feedback Fusion Module

- Collaboratively updates the feedback and deep features and fuses them together.
- Effective and efficient.







Method

Motivation

- Overlook the importance of feedback and do not use the prior information provided by feedback [5, 12, 15, 19, 24, 26]
- Simply concatenate the feedback with the original input, easily leading to an information dilution problem [6, 23, 25, 34, 35]

Feedback M_{t-1}

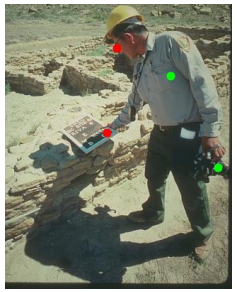
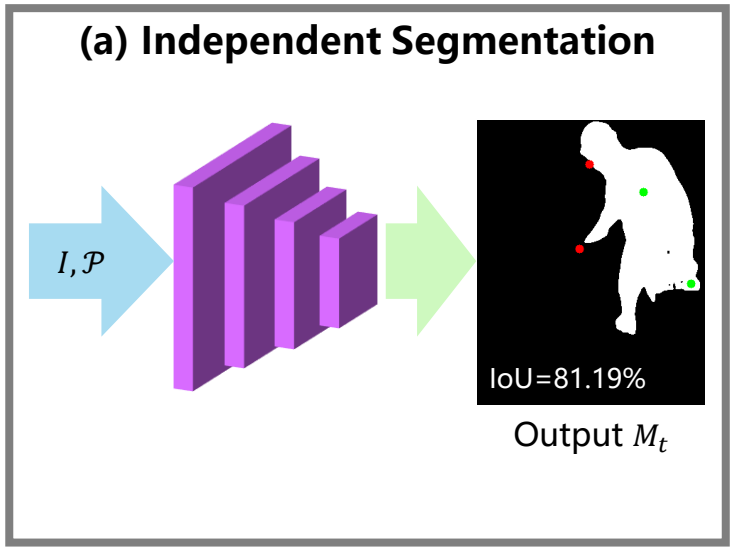
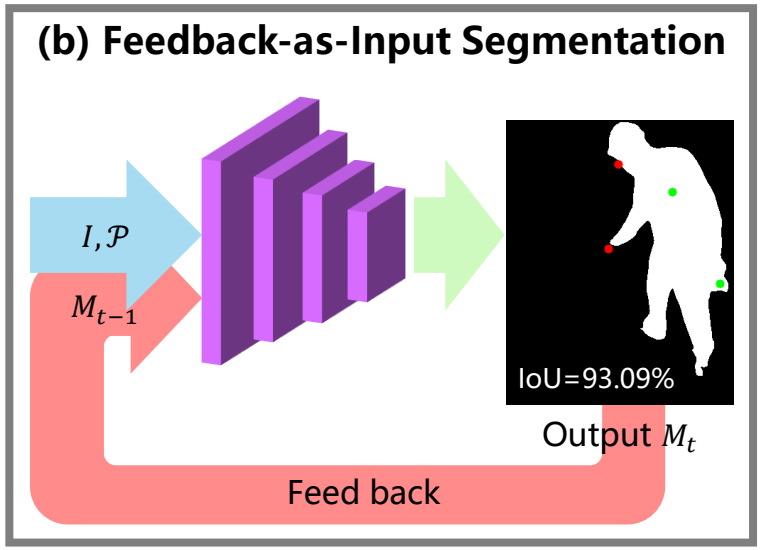
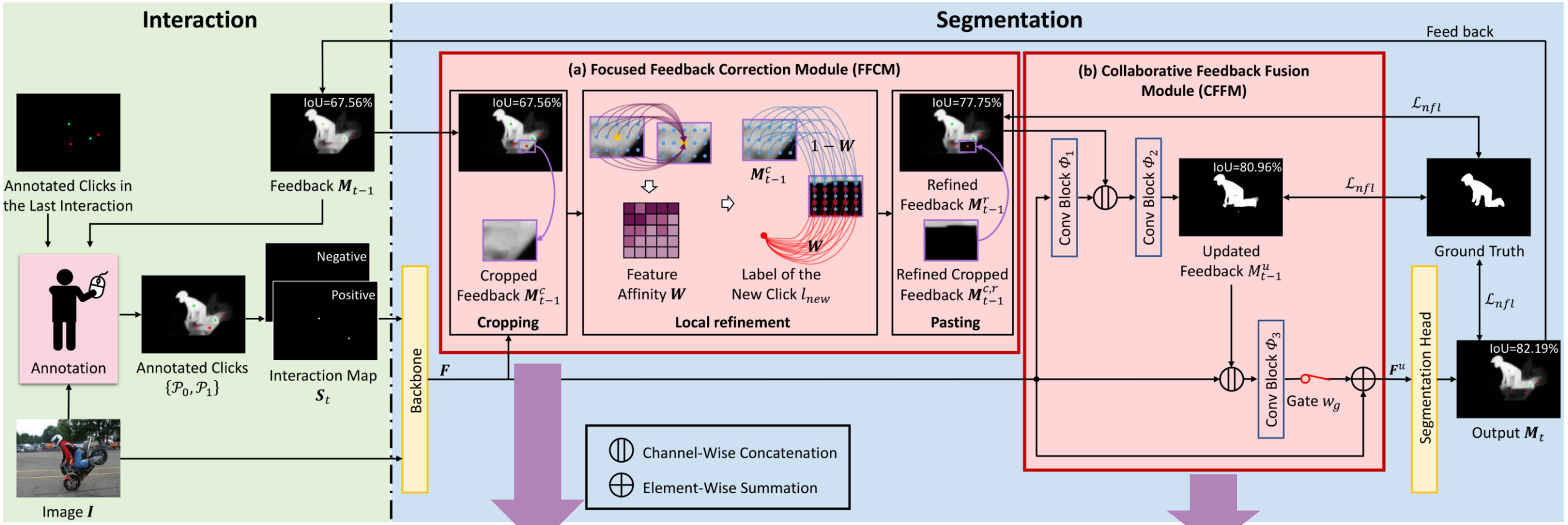


Image I and Clicks $\mathcal{P}$





Pipeline

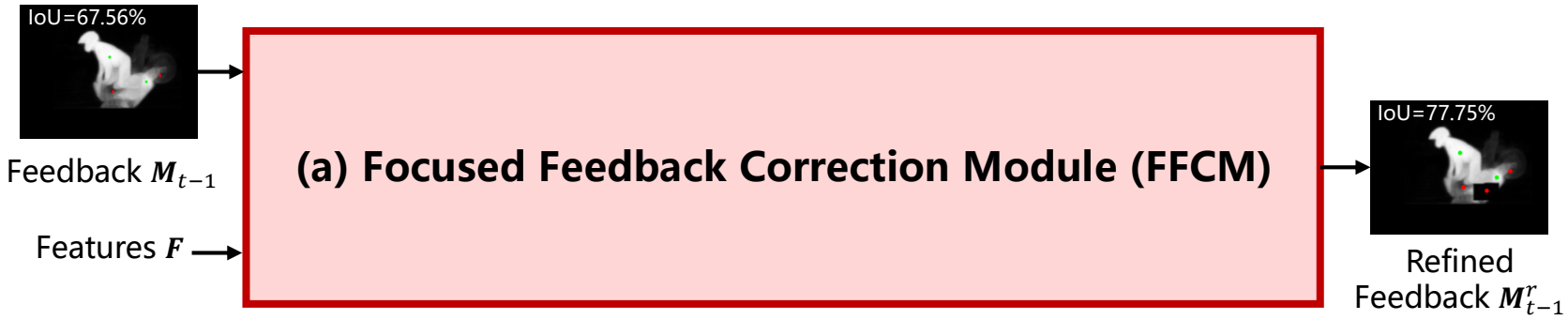


Corrects the feedback from a local perspective

Integrates the corrected feedback into deep features

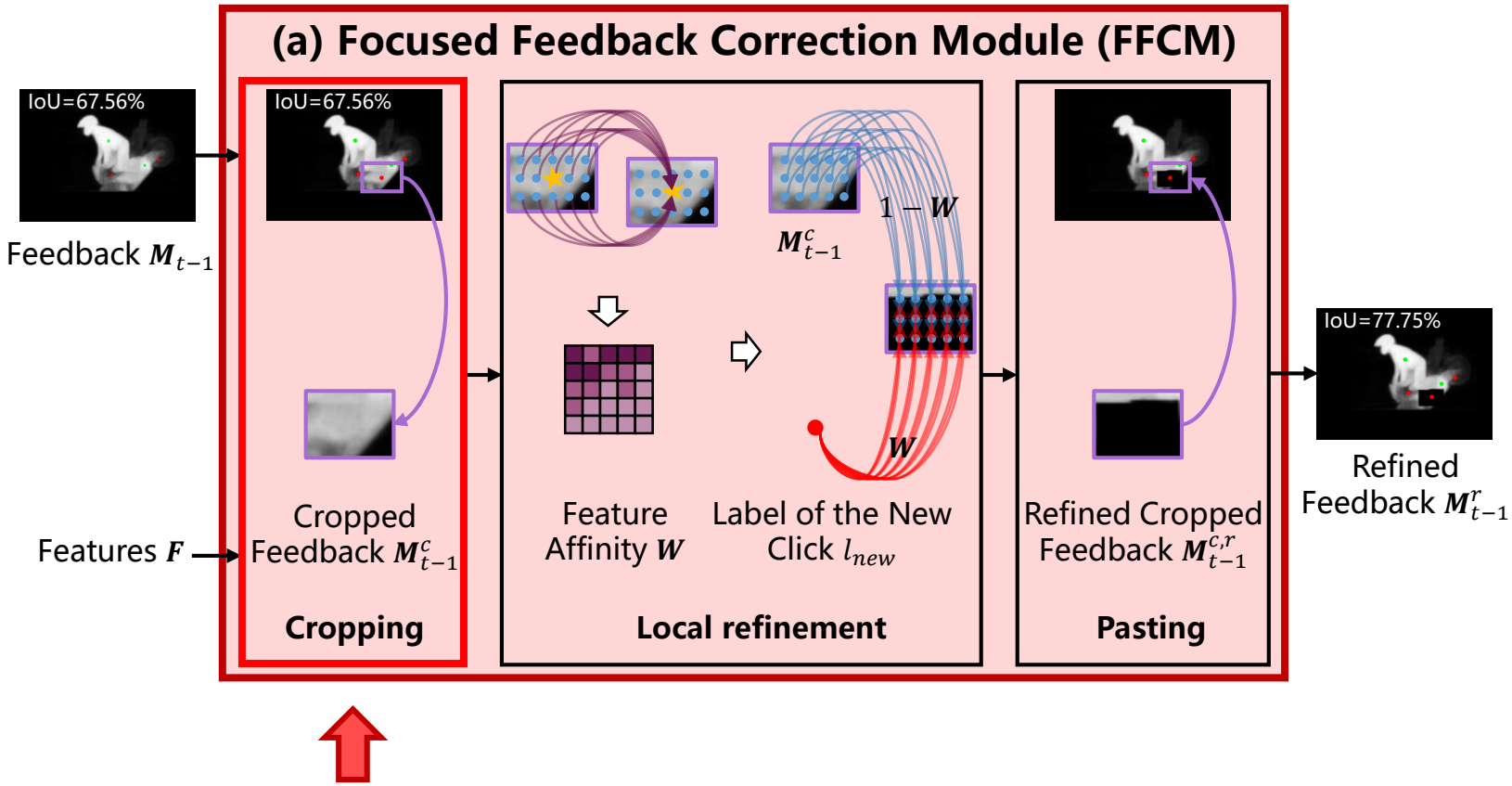
Focused Feedback Correction Module

- Corrects the segmentation feedback from a local perspective



Focused Feedback Correction Module

- Cropping
  - Crops a rectangle patch to exclude irrelevant pixels



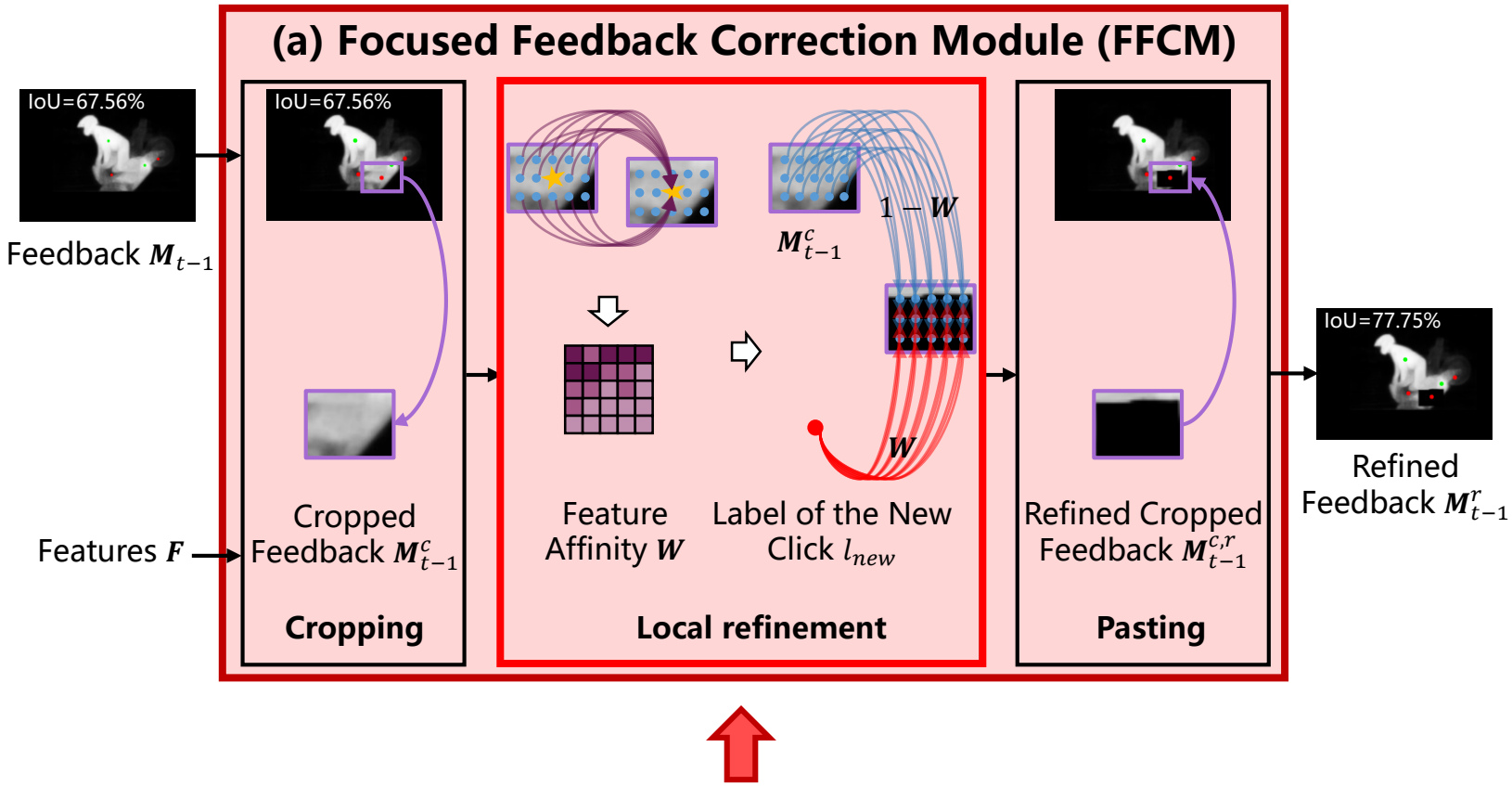
Focused Feedback Correction Module

- Local refinement
  - Calculates similarities

$$\boldsymbol{W}(i,j) = \frac{\boldsymbol{F}^c(i,j)\boldsymbol{F}^c(x_{new}, y_{new})}{\|\boldsymbol{F}^c(i,j)\|_2 \|\boldsymbol{F}^c(x_{new}, y_{new})\|_2}$$

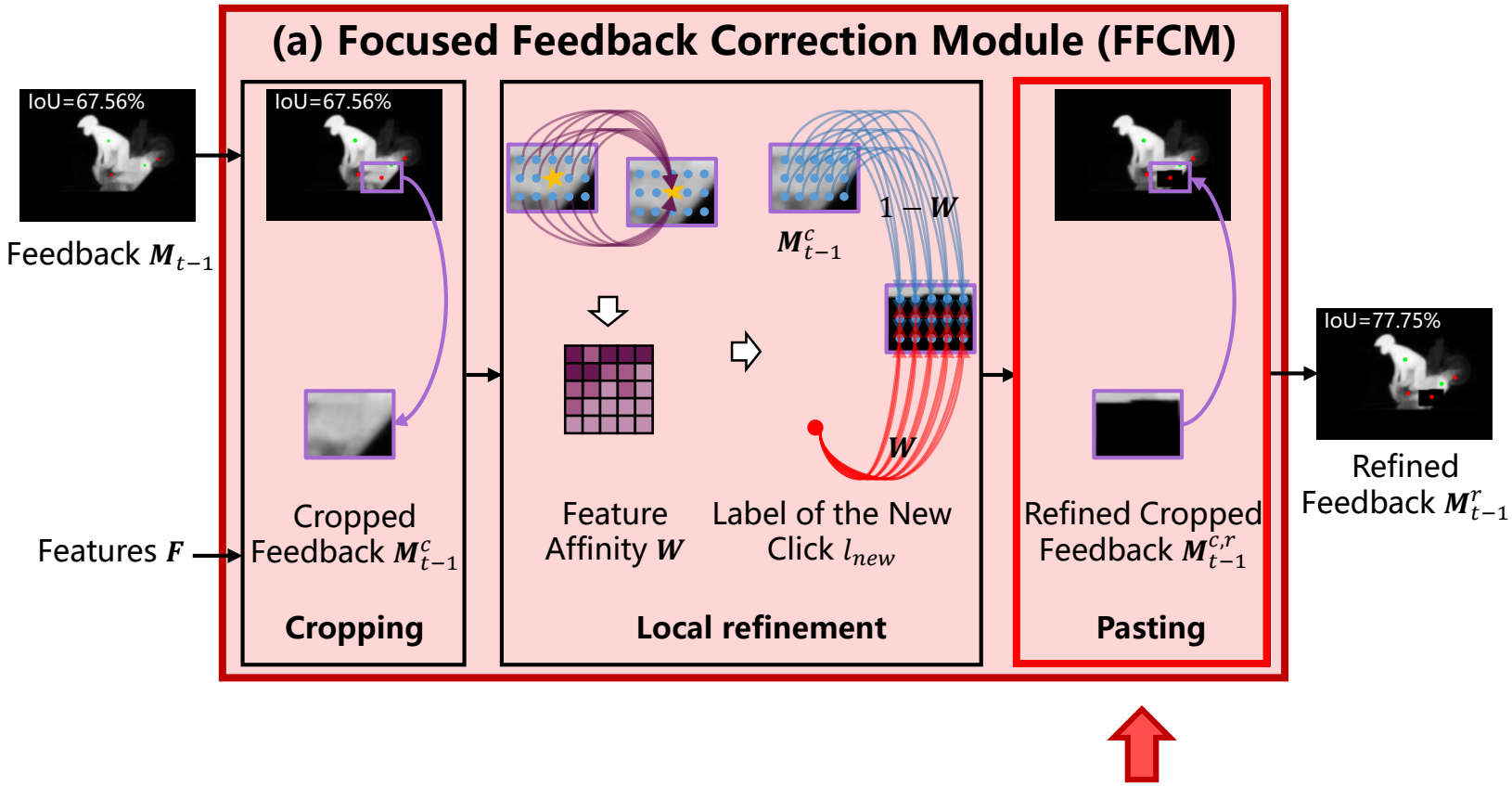
  - Corrects the feedback

$$M_{t-1}^{c,r} = l_{new} \cdot W + (1 - W) \odot M_{t-1}^{c}$$



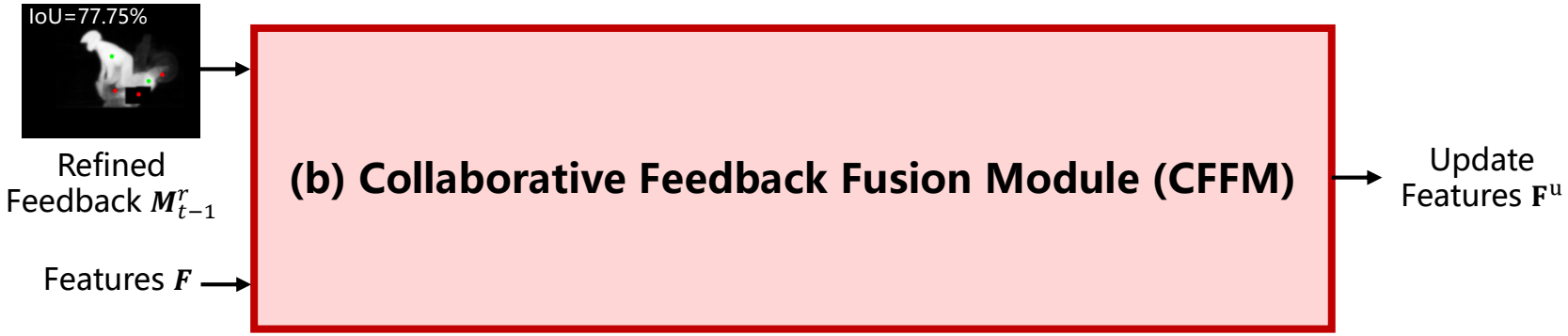
Focused Feedback Correction Module

- Pasting
  - Pastes back to the original position



Collaborative Feedback Fusion Module

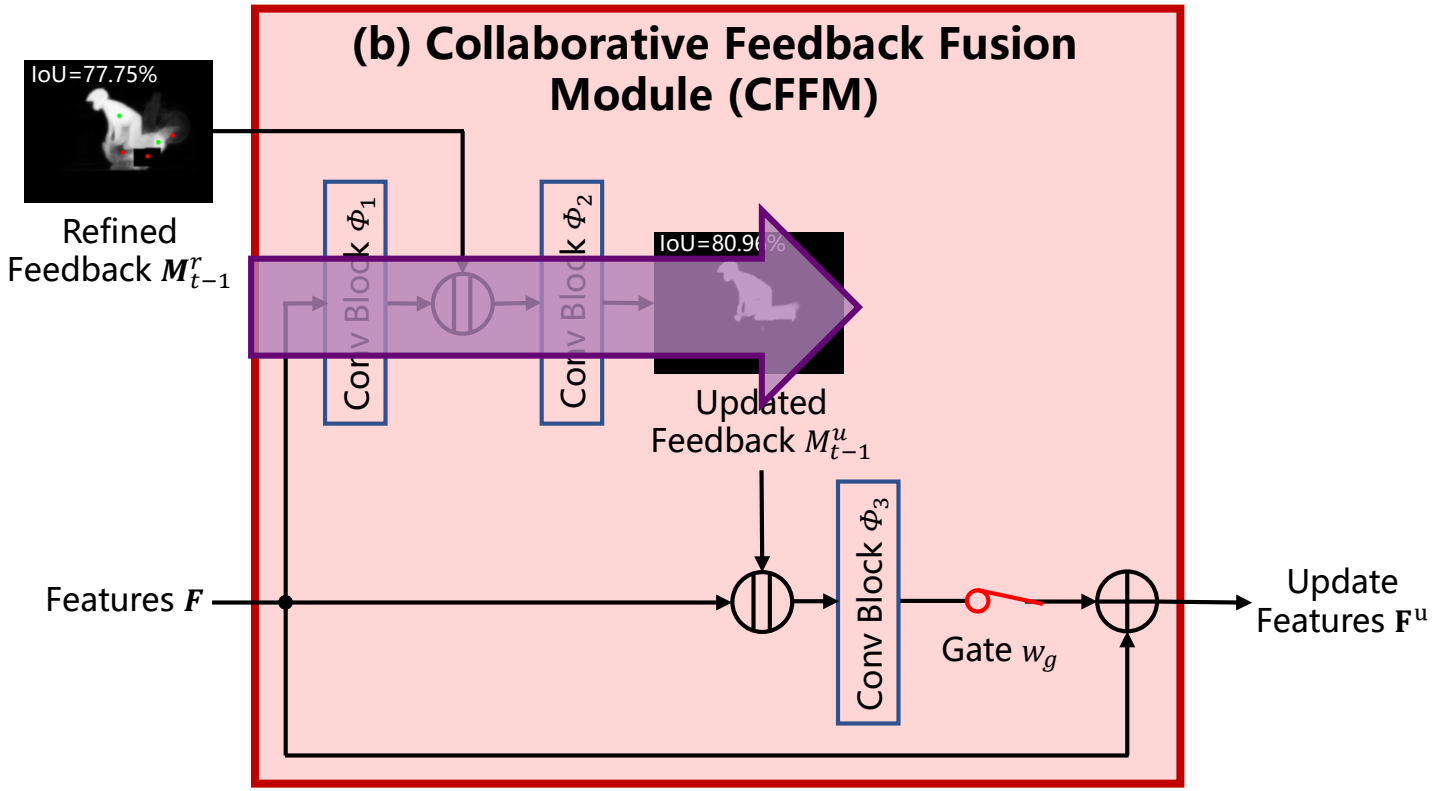
- Refines the segmentation feedback from a global view
- Integrates the feedback into deep features



Collaborative Feedback Fusion Module

- Feedback pathway
  - Globally refines the feedback with the aid of deep features

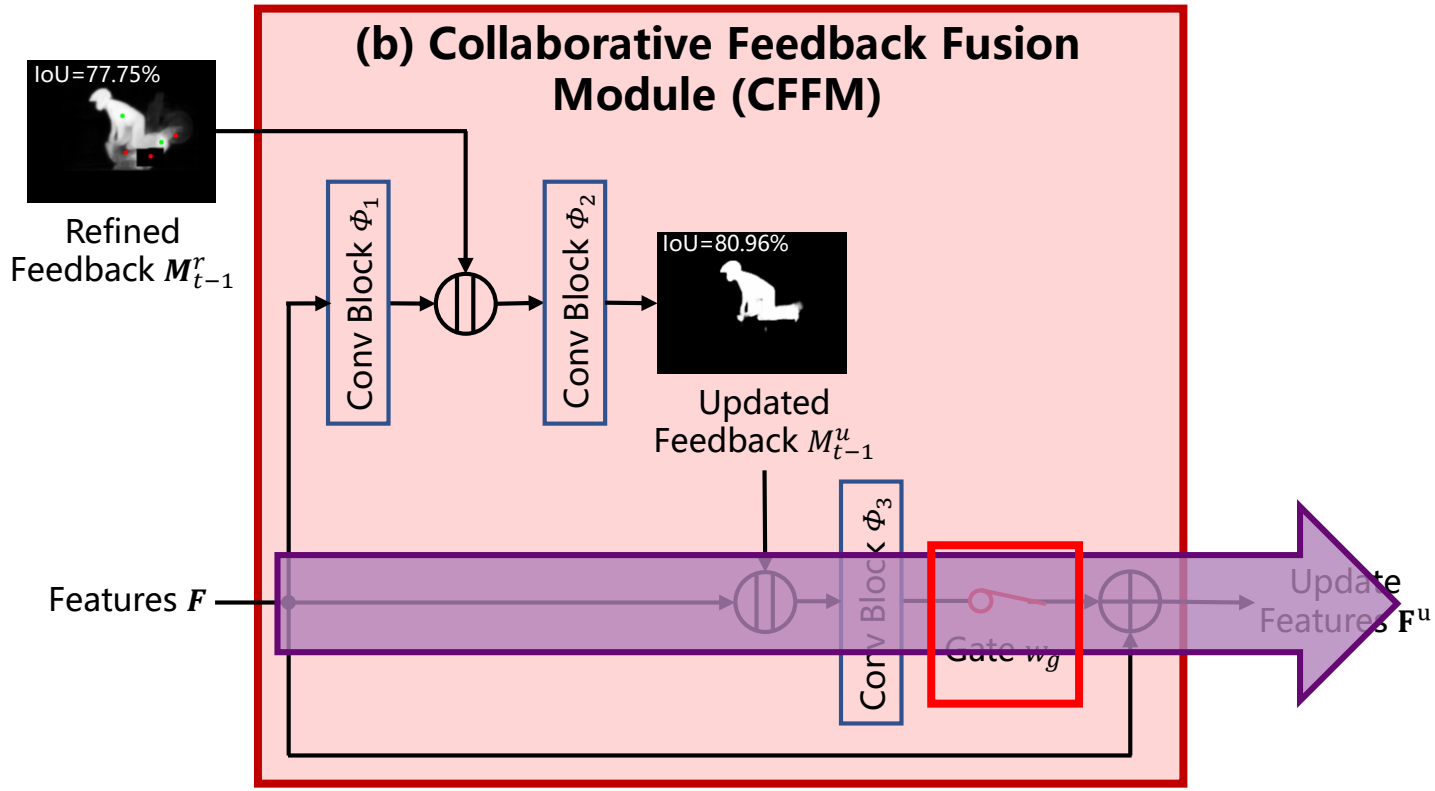
$$\boldsymbol{M}_{t-1}^{u} = \Phi_2(\text{concat}(\Phi_1(\boldsymbol{F};\theta_1), \boldsymbol{M}_{t-1}^{r});\theta_2)$$



Collaborative Feedback Fusion Module

- Feature pathway
  - Updates the features by integrating the feedback

$$\boldsymbol{F}^{u} = w_g \cdot \Phi_3(\text{concat}(\boldsymbol{F}, \boldsymbol{M}^{u}_{t-1}); \theta_3) + \boldsymbol{F}$$



Results

Evaluation

Method	Backbone	Train set	GrabCut		Berkeley	SBD		DAVIS	
			NoC@85%	NoC@90%	NoC@90%	NoC@85%	NoC@90%	NoC@85%	NoC@90%
DOS w/o GC [39]	FCN-8s	SBD	8.02	12.59	-	14.30	16.79	12.52	17.11
DOS w/ GC [39]	FCN-8s	SBD	5.08	6.08	-	9.22	12.80	9.03	12.58
RIS-Net [20]	VGG-16	Pascal VOC	-	5.00	-	6.03	-	-	-
LD [19]	VGG-19	SBD	3.20	4.79	-	7.41	-	5.95	9.57
CAG [26]	FCN-8s	Augmented SBD	-	3.58	5.60	-	-	-	-
BRS [15]	DenseNet	SBD	2.60	3.60	5.08	6.59	9.78	5.58	8.24
f-BRS-B [34]	ResNet-101	SBD	2.30	2.72	4.57	4.81	7.73	5.04	7.41
FCA-Net [24]	ResNet-101	Augmented SBD	1.88	2.14	4.19	-	-	5.38	7.90
IA+SA [16]	ResNet-101	Augmented SBD	-	3.07	4.94	-	-	5.16	-
CDNet [5]	ResNet-101	SBD	2.42	2.76	3.65	4.73	7.66	5.33	6.97
FocusCut [23]	ResNet-101	SBD	**1.46**	**1.64**	3.01	3.40	**5.31**	4.85	**6.22**
Ours	ResNet-101	SBD	1.64	1.80	**2.84**	**3.26**	5.35	**4.75**	6.48
RITM [35]	HRNet-18s	COCO+LVIS	1.54	1.68	2.60	4.26	6.86	4.79	6.00
FocalClick-S1 [6]	HRNet-18s	COCO+LVIS	1.72	1.94	3.40	4.75	7.22	5.19	7.95
FocalClick-S2 [6]	HRNet-18s	COCO+LVIS	1.52	1.66	2.41	4.37	6.59	4.20	5.49
Ours	HRNet-18s	COCO+LVIS	**1.50**	**1.56**	**2.05**	**3.88**	**6.24**	**3.70**	**5.16**
EdgeFlow [12]	HRNet-18	COCO+LVIS	1.60	1.72	2.40	-	-	4.54	5.77
RITM [35]	HRNet-18	COCO+LVIS	1.42	1.54	2.26	3.80	6.06	4.36	5.74
Ours	HRNet-18	COCO+LVIS	**1.38**	**1.46**	**1.96**	**3.63**	**5.83**	**3.97**	**5.16**

Ablation Study

Method	Backbone	Berkeley		DAVIS	
		NoC @90%↓	NoF_{20} @90%↓	NoC @90%↓	NoF_{20} @90%↓
Baseline	ResNet-101	4.31	6	7.56	75
+ FFC	ResNet-101	3.75	4	6.75	65
+ CFF	ResNet-101	3.07	**1**	6.62	66
+ FFC + CFF	ResNet-101	**2.84**	3	**6.48**	**59**
Baseline	HRNet-18	2.60	1	5.73	54
+ FFC	HRNet-18	2.11	1	5.52	54
+ CFF	HRNet-18	2.05	**0**	5.32	52
+ FFC + CFF	HRNet-18	**1.96**	**0**	**5.16**	**51**

Demo

Demo

FCFI Interactive Interface

Project Page

Choose click type: Background

Submit Clear Save

Click size: 5

Conclusion

Conclusion

- We propose **Focused and Collaborative Feedback Integration (FCFI)** to exploit segmentation feedback for interactive image segmentation.
- The proposed method consists of a **Focused Feedback Correction Module** for locally correcting segmentation feedback and a **Collaborative Feedback Fusion Module** for integrating feedback into deep features.
- The experimental results have demonstrated that FCFI exhibits a strong generalization ability and achieves outstanding performance.

References

- Xi Chen, Zhiyan Zhao, Feiwu Yu, Yilei Zhang, and Manni Duan. Conditional diffusion for interactive segmentation. In ICCV, pages 7345–7354, 2021. 2, 3, 5, 6
- Xi Chen, Zhiyan Zhao, Yilei Zhang, Manni Duan, Donglian Qi, and Hengshuang Zhao. Focalclick: Towards practical interactive image segmentation. In CVPR, pages 1300–1309, 2022. 1, 2, 3, 4, 5, 6
- Yuying Hao, Yi Liu, Zewu Wu, Lin Han, Yizhou Chen, Guowei Chen, Lutao Chu, Shiyu Tang, Zhiliang Yu, Zeyu Chen, and Baohua Lai. Edgeflow: Achieving practical interactive segmentation with edge-guided flow. In ICCVW, pages 1551–1560, 2021. 2, 4, 5
- Won-Dong Jang and Chang-Su Kim. Interactive image segmentation via backpropagating refinement scheme. In CVPR, pages 5297–5306, 2019. 1, 2, 3, 5, 6
- Zheng Lin, Zheng-Peng Duan, Zhao Zhang, Chun-Le Guo, and Ming-Ming Cheng. Focuscut: Diving into a focus view in interactive segmentation. In CVPR, pages 2637–2646, 2022. 1, 2, 5, 6
- Zheng Lin, Zhao Zhang, Lin-Zhuo Chen, Ming-Ming Cheng, and Shao-Ping Lu. Interactive image segmentation with first click attention. In CVPR, pages 13339–13348, 2020. 1, 2, 5, 6
- Konstantin Sofiiuk, Ilia Petrov, Olga Barinova, and Anton Konushin. f-brs: Rethinking backpropagating refinement for interactive segmentation. In CVPR, pages 8623–8632, 2020. 1, 2, 3, 4, 5, 6
- Konstantin Sofiiuk, Ilia A Petrov, and Anton Konushin. Reviving iterative training with mask guidance for interactive segmentation. arXiv preprint arXiv:2102.06583, 2021. 1, 2, 3, 4, 5, 6

清華大學
Tsinghua University

JUNE 18-22, 2023
CVPR
VANCOUVER, CANADA

Thanks for watching!

Focused and Collaborative Feedback Integration for Interactive Image Segmentation