CVPR
JUNE 17-21, 2024
SEATTLE, WA

Depth-Aware Concealed Crop Detection in Dense Agricultural Scenes

Liqiong Wang[1†], Jinyu Yang[2,3†], Yanfu Zhang[4], Fangyi Wang[1*], Feng Zheng[2*]

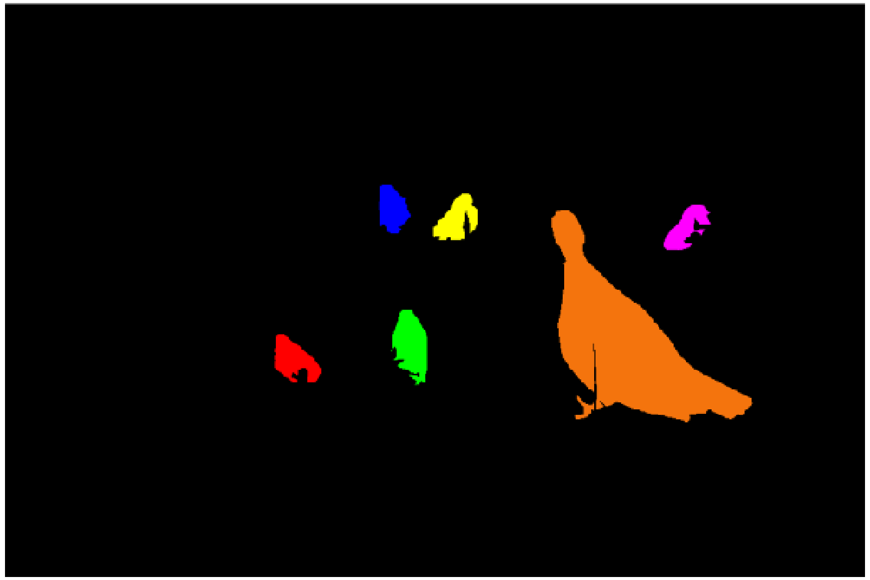
Poster: THU-PM-252

[1]China Three Gorges University, [2]Southern University of Science and Technology, [3]University of Birmingham, [4]College of William and Mary

三峡大学
CHINA THREE GORGES UNIVERSITY
南方科技大学
SOUTHERN UNIVERSITY OF SCIENCE AND TECHNOLOGY
UNIVERSITY OF BIRMINGHAM
WILLIAM & MARY
CHARTERED 1693

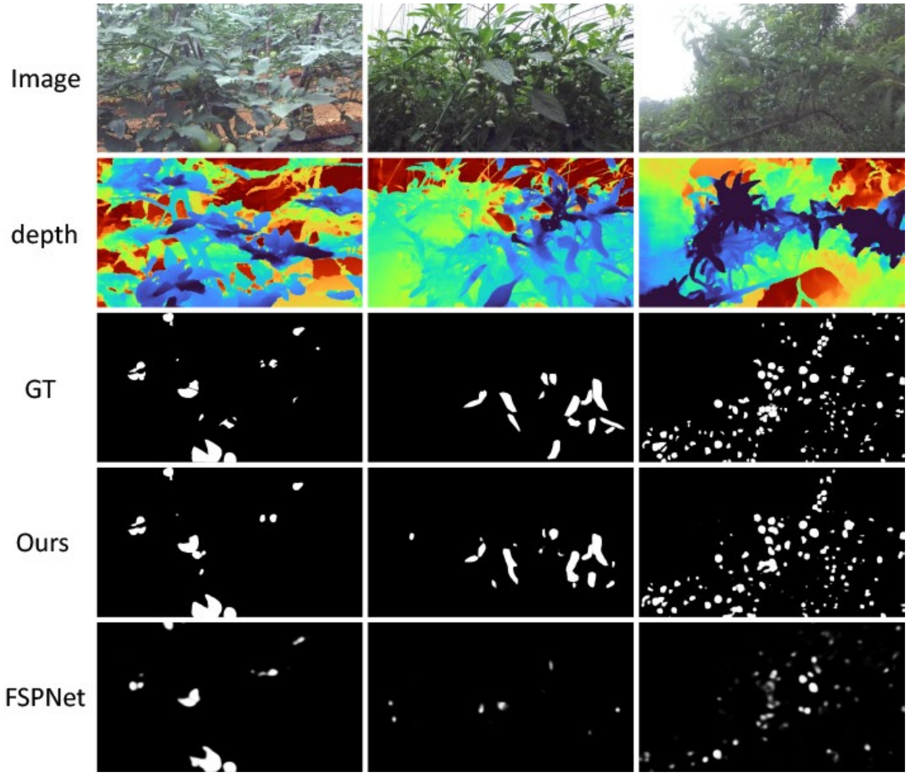
Background

Concealed Object Detection (COD) aims to identify objects that are visually embedded in their background.



Reference:
[1] D.-P. Fan, G.-P. Ji, M.-M. Cheng, and L. Shao, "Concealed object detection," IEEE TPAMI, 2022.

Limitation



We mainly focus on concealed objects in dense agricultural scenes, which pose unique challenges:

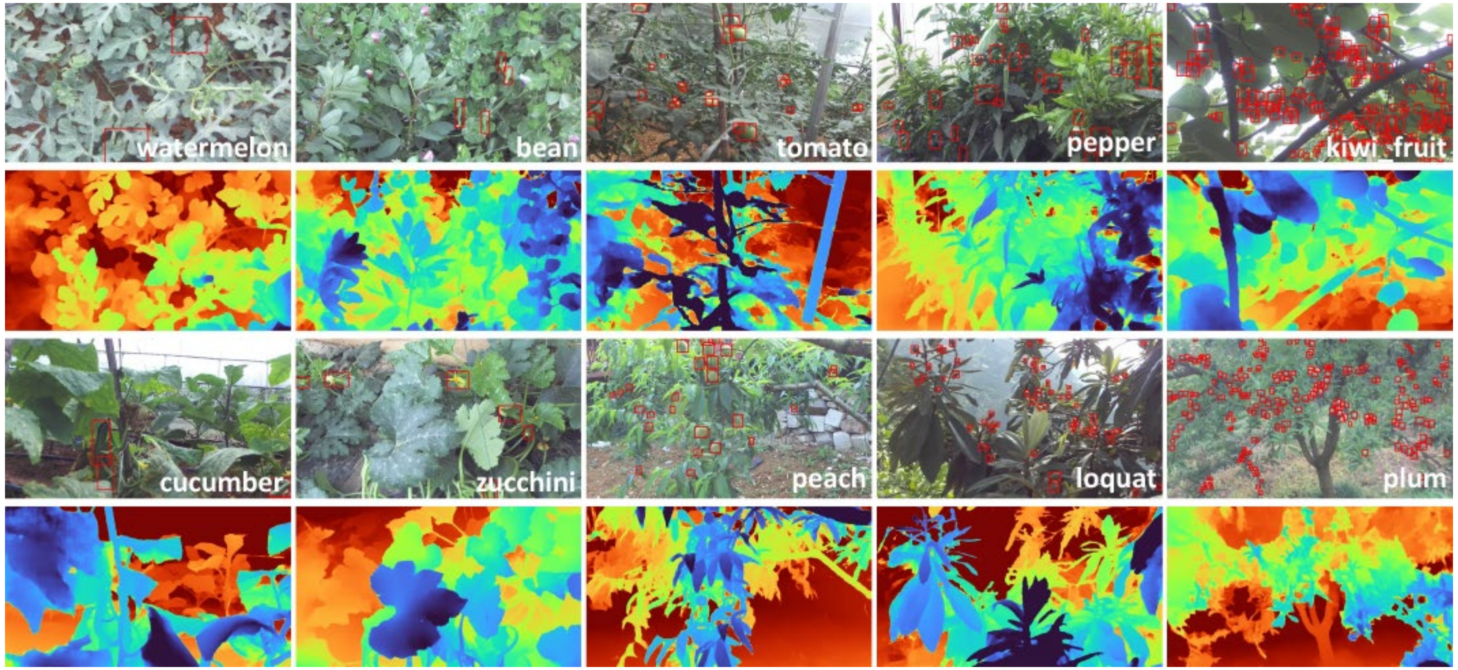
- ➢ Dense objects
- ➢ Intricate backgrounds
- ➢ Severe occlusion
- ➢ Small objects

Introduction

Dataset	Year	Img	Avg.Res.	Free View	Mul.	Object Statistics Total	Min	Avg	Max	Link
CHAMELEON[50]	2018	76	742 × 981	✗	✗	79	1	1	3	N/A
CAMO[34]	2019	1250	509 × 653	✓	✗	1368	1	1	7	Link
COD10K[13]	2020	5066	737 × 964	✓	✗	5899	1	1	8	Link
NC4K[42]	2021	4121	530 × 709	✓	✗	4584	1	1	8	Link
ACOD-12K(Ours)	2023	6092	1080 × 1920	✓	✓	71417	1	11	412	Link

- This paper aims to address the overlooked challenge of detecting small, concealed crops amidst dense plant populations and occlusions in the agricultural domain.
- Existing COD methods primarily focus on animals or humans, and this paper aims to bridge this gap by introducing Concealed Crop Detection (CCD) and developing novel techniques to help detection in this area.

Dataset



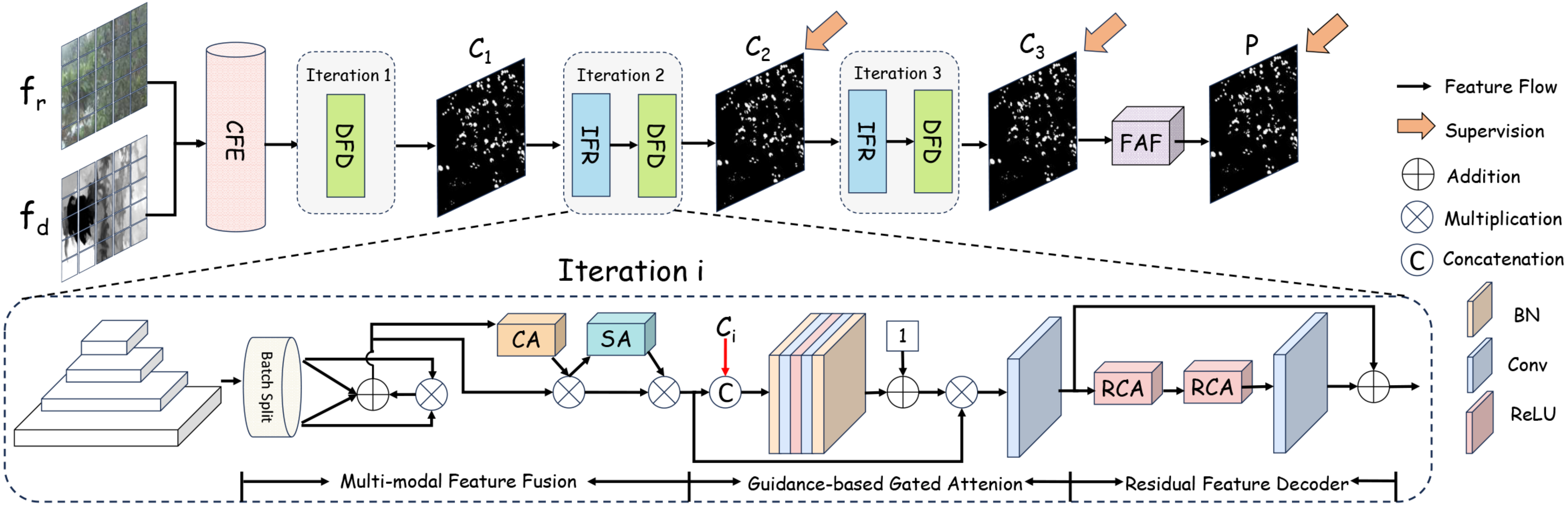
To facilitate research on CCD, we propose our ACOD-12K.

- The sole existing multi-modal COD dataset.
- The largest-scale COD dataset.
- Higher object density.
- Focus on the distinctive challenges presented by concealed objects in dense agricultural scenes.

Contribution

- Introduce **Concealed Crop Detection** (CCD), extending classic COD to agriculture.
- Collect a new large-scale RGB-D dataset **ACOD-12K**, which is the first multi-modal dataset on COD tasks.
- Develop a new baseline **RISNet**, which fuses depth features for CCD and achieves SOTA on COD and CCD.

The proposed framework



Comparison with SoTA COD methods

Quantitative comparisons on CCD

Model	Publications	ACOD-12K		
		$S_\alpha \uparrow$	$F_\beta^\omega \uparrow$	$E_\theta \uparrow$
Concealed Object Detection				
SINet[13]	CVPR20	0.745	0.474	0.826
MGL[67]	CVPR21	0.808	0.685	0.872
PFNet[45]	CVPR21	0.805	0.685	0.942
UGTR[65]	ICCV21	0.798	0.632	0.858
SINet-V2[14]	TPAMI22	0.804	0.691	0.947
C2FNet[7]	TCSVT22	0.833	0.746	0.947
PreyNet[69]	MM22	0.832	0.760	0.937
SegMaR[31]	CVPR22	0.799	0.677	0.930
ZoomNet[49]	CVPR22	0.832	0.747	0.934
DaCOD[57]	MM23	0.803	0.705	0.910
PopNet[61]	ICCV23	0.844	0.778	0.955
HitNet[26]	AAAI23	0.853	0.787	0.955
FSPNet[28]	CVPR23	0.719	0.526	0.819
RGB-D Salient Object Detection				
CLNet[68]	ICCV21	0.826	0.747	0.936
SPNet[74]	ICCV21	0.818	0.731	0.949
DCMF[54]	TIP22	0.779	0.631	0.872
HINet[6]	PR22	0.776	0.651	0.853
SPSN[35]	ECCV22	0.834	0.739	0.930
CIRNet[9]	TIP22	0.794	0.675	0.865
HIDANet[60]	TIP23	0.822	0.734	0.950
XMSNet[62]	MM23	0.844	0.754	0.961
Ours		0.866	0.803	0.967

Quantitative comparisons on COD

Model	Publications	CAMO				COD10K				NC4K			
		$S_\alpha \uparrow$	$F_\beta^\omega \uparrow$	$E_\theta \uparrow$	$M \downarrow$	$S_\alpha \uparrow$	$F_\beta^\omega \uparrow$	$E_\theta \uparrow$	$M \downarrow$	$S_\alpha \uparrow$	$F_\beta^\omega \uparrow$	$E_\theta \uparrow$	$M \downarrow$
SINet[13]	CVPR20	0.745	0.644	0.804	0.092	0.776	0.631	0.864	0.043	0.808	0.723	0.871	0.058
LSR[42]	CVPR21	0.787	0.696	0.838	0.080	0.804	0.673	0.880	0.037	0.840	0.766	0.895	0.048
R-MGL[67]	CVPR21	0.775	0.673	0.812	0.088	0.814	0.666	0.852	0.035	0.833	0.740	0.867	0.052
JSCOD[36]	CVPR21	0.800	0.728	0.859	0.073	0.809	0.684	0.884	0.035	0.842	0.771	0.898	0.047
PFNet[45]	CVPR21	0.782	0.695	0.841	0.085	0.800	0.660	0.877	0.040	0.829	0.745	0.887	0.053
ZoomNet[49]	CVPR22	0.820	0.752	0.877	0.066	0.838	0.729	0.888	0.029	0.853	0.784	0.896	0.043
FDNet[73]	CVPR22	0.841	0.775	0.895	0.063	0.840	0.729	0.919	0.030	0.834	0.750	0.893	0.052
SegMaR[31]	CVPR22	0.815	0.753	0.874	0.071	0.833	0.724	0.899	0.034	0.841	0.781	0.896	0.046
DGNet[29]	MIR23	0.839	0.769	0.901	0.057	0.822	0.693	0.896	0.033	0.857	0.784	0.911	0.042
PopNet[61]	ICCV23	0.808	0.744	0.859	0.077	0.851	0.757	0.910	0.028	0.861	0.802	0.910	0.042
DaCOD[57]	MM23	0.855	0.796	0.905	0.051	0.840	0.729	0.907	0.028	0.874	0.814	0.924	0.035
HitNet[26]	AAAI23	0.844	0.801	0.902	0.057	0.868	0.798	0.932	0.024	0.870	0.825	0.921	0.039
FEDER[22]	CVPR23	0.822	0.738	0.886	0.067	0.851	0.716	0.917	0.028	0.863	0.789	0.917	0.042
FSPNet[28]	CVPR23	0.856	0.799	0.899	0.050	0.851	0.735	0.895	0.026	0.879	0.816	0.915	0.035
Ours		0.870	0.827	0.922	0.050	0.873	0.799	0.931	0.025	0.882	0.834	0.925	0.037

Qualitative results on ACOD-12K

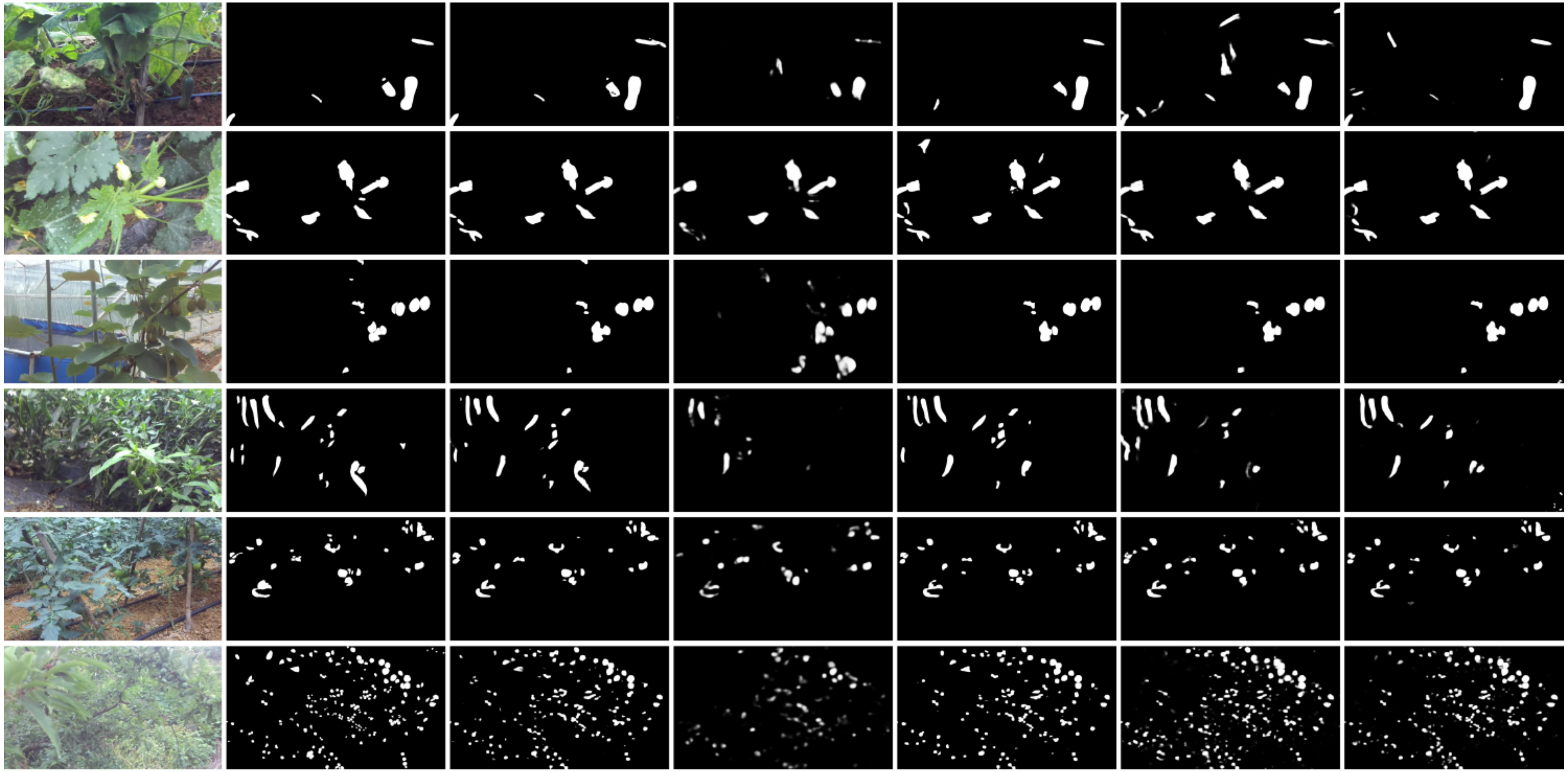
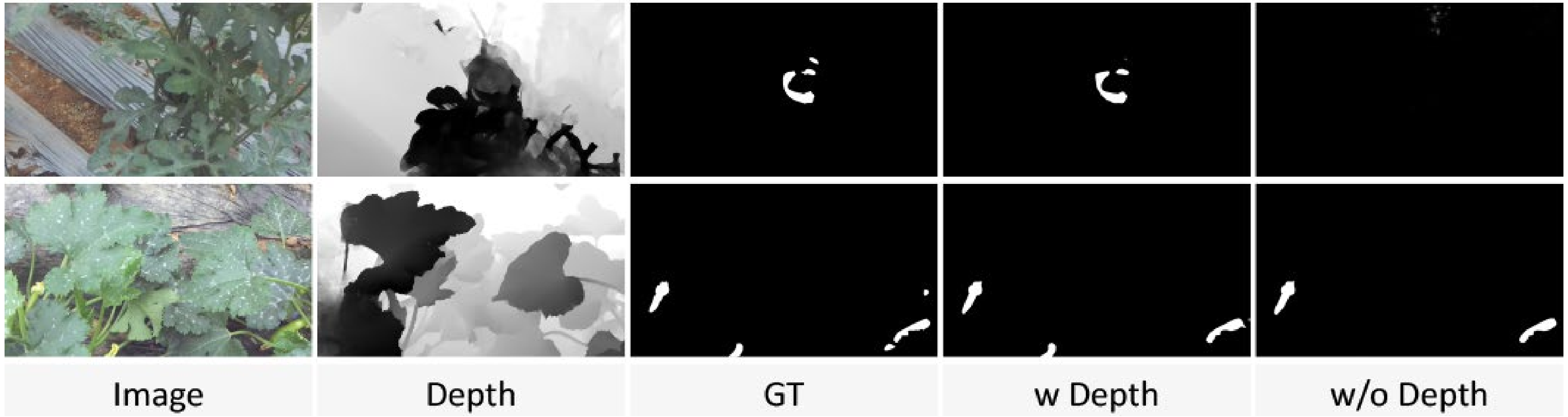
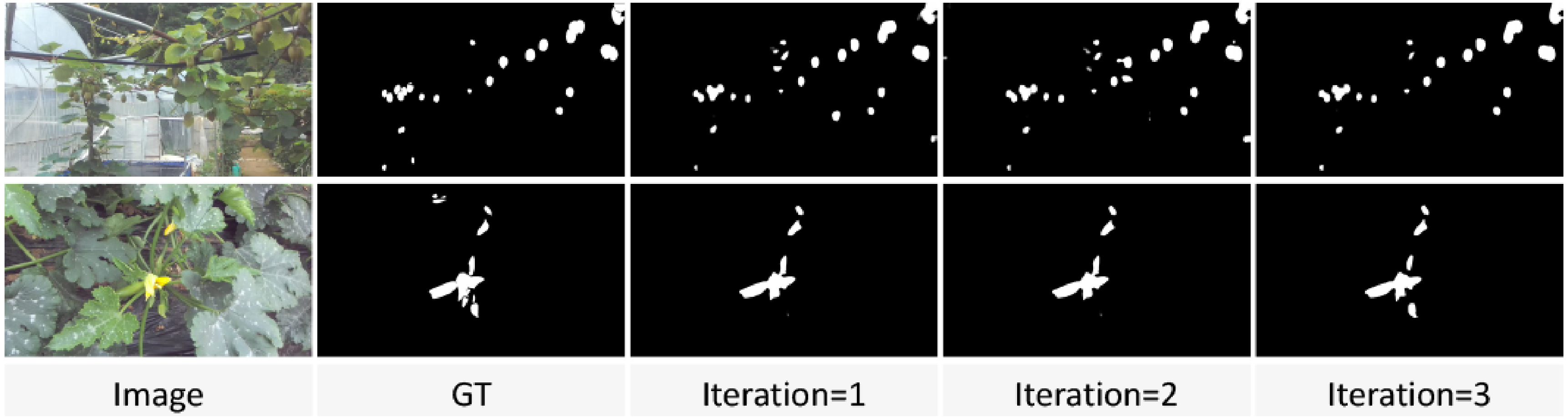


Image	GT	Ours	FSPNet	HitNet	XMSNet	PopNet

Ablation Study





Thanks!