Learning to Model Pixel-Embedded Affinity for Homogeneous Instance Segmentation

Wei Huang[1], *Shiyu Deng*[1], *Chang Chen*[1], *Xueyang Fu*[1], and *Zhiwei Xiong*[1,2]✉

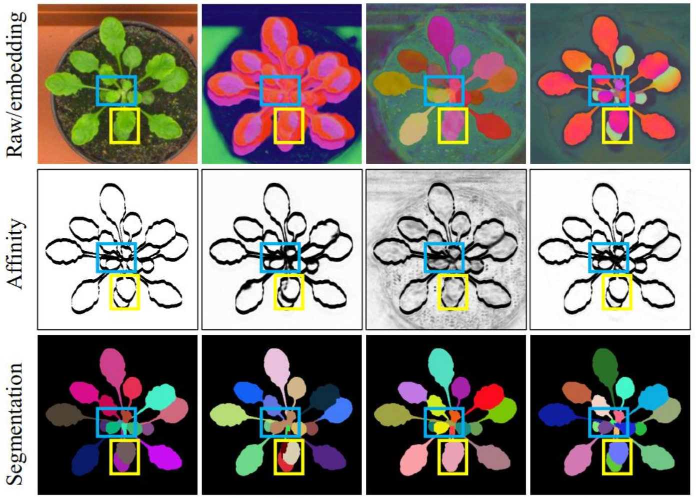
[1]University of Science and Technology of China, [2]Institute of Artificial Intelligence, Hefei Comprehensive National Science Center

Introduction

Homogeneous Instance Segmentation

- Homogeneous instance segmentation focuses on the identification of instances belonging to the same category in an image, which is desired in many practical applications, especially for biomedical image analysis.

Motivation



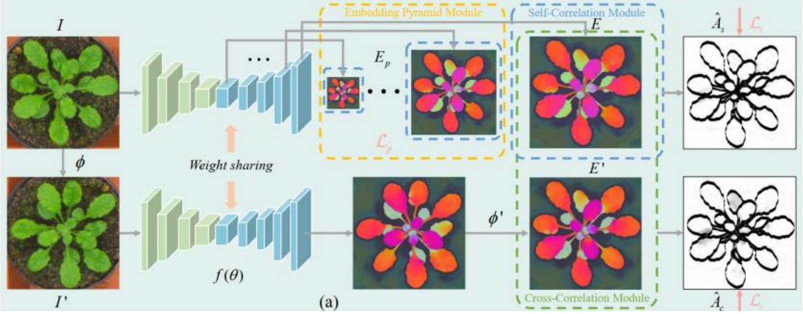
(a) Groundtruth (b) AL-Gao (c) ML-De (d) Ours

- Affinity learning (AL-Gao) straightforwardly generates affinity but suffers from the absence of the semantic instance information (b).
- Metric learning (ML-De) aims to push all instances away from each other but ignores the spatial information between instances (c).
- Our method better preserves the semantic information of instances and pays more attention to the distinguishability of adjacent instances, producing superior segmentation results (d).

Method

Overall Pipeline

- **Self-Correlation Module** explicitly models the pairwise relationship between pixels to preserve the semantic instance information.
- **Cross-Correlation Module** mutually estimates the pairwise relationships under different views and appearances of the input image to improve the distinguishability of adjacent instances.
- **Embedding Pyramid Module** models affinity on different scales to integrate the global instance information.



Training & Inference

- In the training phase, we combine these three modules together and train the network to generate affinities in an end-to-end way.
- In the inference phase, we only use the self-correlation affinity as the final affinity prediction, and adopt the Mutex algorithm as post-processing to obtain instance masks from the predicted affinity. In addition, we merge too small instances to further refine the final segmentation result.

Evaluation

Evaluation on CVPPP (Plant leaves)

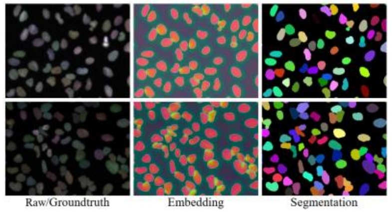
Methods	Param.	SBD	$\|DiC\|$
MSU (Scharr et al. 2016)	-	66.7	2.3
Nottingham (Scharr et al. 2016)	-	68.3	3.8
Wageningen (Yin et al. 2014)	-	71.1	2.2
IPK (Pape and Klukas 2014)	-	74.4	2.6
Coloring (Kulikov et al. 2018)	30.2M	80.4	2.0
ML-De (De Brabandere et al. 2017)	23.1M	84.2	1.0
Recurrent (Ren and Zemel 2017)	-	84.9	**0.8**
Aug. (Kuznichov et al. 2019)	-	88.7	5.3
Harmonic (Kulikov et al. 2020)	43.1M	89.0	3.0
Synthesis (Ward et al. 2018)	105.7M	90.0	-
PFFNet (Liu et al. 2021)	105.7M	91.1	-
Ours w/ ResNet-50	15.3M	91.7	1.5
Ours w/ ResNet-101	34.3M	91.9	1.4
Ours w/ ResUNet	**4.7M**	**92.3**	2.4

Dataset	Methods	Clustering	SBD	$\|DiC\|$
A1	AL-Gao	Mutex	87.1	1.25
	ML-De	Mean-shift	87.3	1.45
	ML-De	Mutex	88.5	1.10
	Ours	Mutex	**89.1**	**0.85**
A2	AL-Gao	Mutex	71.1	2.61
	ML-De	Mean-shift	71.2	2.52
	ML-De	Mutex	73.4	2.00
	Ours	Mutex	**76.3**	**1.71**

When trained on A1 and tested on A2, our method significantly outperforms the competitors.

Evaluation on BBBC039V1 (Nuclei of U2OS cells)

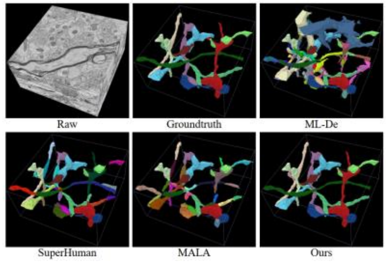
Methods	AJI	$Dice$	PQ
Mask R-CNN (He et al. 2017)	0.7983	0.9277	0.7773
Cell R-CNN (Zhang et al. 2018)	0.8070	0.9290	0.7959
UPSNet (Xiong et al. 2019)	0.8128	0.9274	0.7857
JSISNet (De Geus et al. 2018)	0.8134	0.9316	0.7913
PanFPN (Kirillov et al. 2019)	0.8193	0.9320	0.7960
OANet (Liu et al. 2019b)	0.8198	0.9372	0.8085
AUNet (Li et al. 2019)	0.8252	0.9377	0.8090
Cell R-CNN v2 (Liu et al. 2019a)	0.8260	0.9336	0.8010
PFFNet (Liu et al. 2021)	0.8477	0.9478	0.8330
Ours	**0.8674**	**0.9673**	**0.8420**



Evaluation on AC3/AC4 (Neurons of mouse brain)

Methods	VOI_S	VOI_M	VOI	$ARAND$
ML-De	1.5752	0.6151	2.1903	0.1964
SuperHuman	1.1445	0.2630	1.4075	0.1220
MALA	1.3039	0.2423	1.5462	0.1203
Ours	**0.8522**	**0.2322**	**1.0844**	**0.0938**

We also demonstrate the effectiveness of our method on the 3D instance segmentation task.



Ablation study

SCM	CCM	EPM	SBD	$\|DiC\|$
✓			87.7	1.15
✓		✓	88.1	1.00
✓	✓		88.5	0.95
✓	✓	✓	**89.1**	**0.85**

The effectiveness of modules.