X-PDNet: Accurate Joint Plane Instance Segmentation and Monocular Depth Estimation by Cross-Task Distillation and Boundary Correction

Cao Dinh Duc[1, 3], Jongwoo Lim[2]

[1] Dept of Computer Science, Hanyang University, Seoul, South Korea
[2] Dept of Mechanical Engineering, Seoul National University, Seoul, South Korea
[3] AI Research Lab, MAXST Co., Ltd, Seoul, South Korea

Paper ID: 326

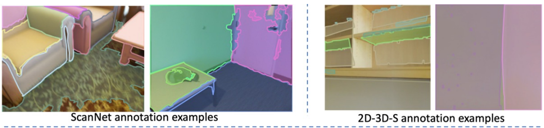
1 Overview

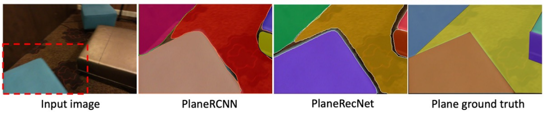
- Design X-PDNet, a multitask learning framework for joint plane instance segmentation and monocular depth estimation.
- Propose a novel Depth-Guided Boundary Preserving Loss that uses depth information to precise the plane instance segmentation results at boundary related regions.
- Contribute a manually annotated test set as a standard dataset for the plane instance segmentation problem.

Acknowledgement

This work was partially supported by the Technology Innovation Program (No. 20018110, "Development of a wireless tele-operable relief robot for detecting searching and responding in narrow space") funded by the Ministry of Trade, Industry & Energy (MOTIE, Korea) and National Research Foundation of Korea (NRF) grant funded by the Korea government (MSIT) (NRF-2021R1A2C2010245).

2 Related Works & Limitations

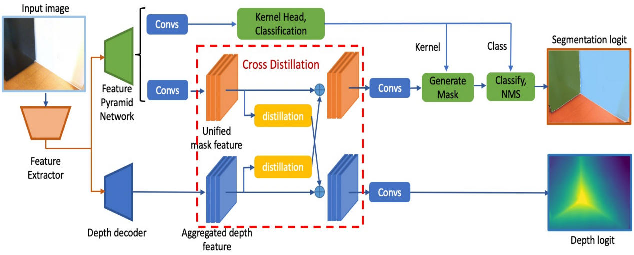


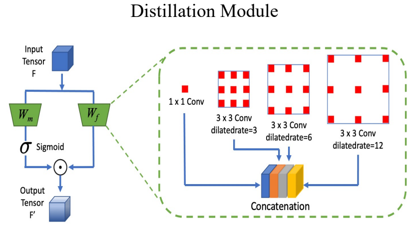


- Poor quality of plane instance ground truth in existing datasets
- Lack of boundary preserving loss => incorrect prediction at boundary regions
- Traditional boundary regression loss is vulnerable with incorrect GT boundary

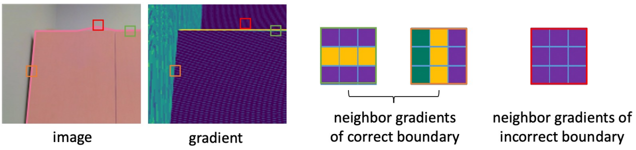
3 Methods

Propose the X-PDNet, with the Cross-Task feature distillation design, which promotes early information sharing between cross for the specific task optimization



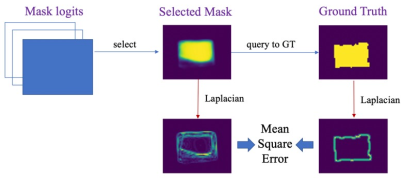


Propose a novel Depth Guided Boundary Preserving Loss, which employes depth information to combats with incorrect ground truth boundary, pricise predicted plane instance at boundary related areas



Gradient: $G_{gt} = abs(G_x) + abs(G_y)$ with $G_x = Sobel_x(D_{gt}),\ G_y = Sobel_y(D_{gt})$.

For each GT boundary point, measure the standard deviation of a set of points constructed from its gradient and that of its neighbors.



Then utilize set of std values to reweight the boundary regression loss

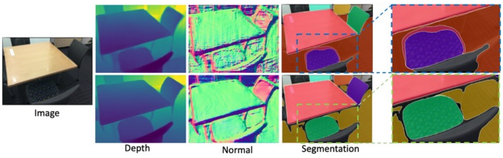
4 Experimental results

Methods	Dataset	Segmentation Metrics AP_m	AP_m^{50}	AP_m^{75}	AP_b	AP_b^{50}	AP_b^{75}	Depth Metrics $rel \downarrow$	$log_{10} \downarrow$	$RMS \downarrow$	$\delta 1$	$\delta 2$	$\delta 3$
PlaneAE [illegible]	ScanNet	5.92	14.72	4.00	7.86	17.83	6.25	0.111	0.049	0.409	0.864	0.967	0.991
PlaneRCNN [illegible]	ScanNet	14.23	28.23	12.88	17.51	33.00	16.00	0.124	0.050	0.265	0.865	0.972	0.994
PlaneRecNet [illegible]	ScanNet	16.61	31.59	15.56	21.05	36.45	20.29	0.076	0.032	0.180	0.950	0.992	0.998
X-PDNet	ScanNet	**17.62**	**33.05**	**16.60**	**22.23**	**37.53**	**21.91**	**0.069**	**0.029**	**0.175**	**0.955**	**0.993**	**0.999**
PlaneRecNet [illegible]	2D-3D-S	24.10	38.99	24.39	27.13	41.14	27.23	0.062	0.027	**0.294**	**0.966**	0.990	**0.996**
X-PDNet	2D-3D-S	**25.20**	**39.63**	**25.79**	**28.62**	**41.80**	**29.15**	**0.061**	**0.026**	**0.294**	**0.966**	**0.991**	**0.996**

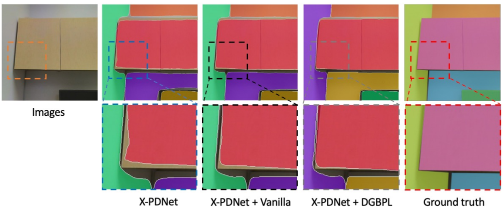
Table 1: Evaluation of plane instance segmentation and depth estimation on **ScanNet** and **2D-3D-S** datasets. **X-PDNet** outperforms existing methods in most metrics.

Methods	Eval set	Boundary IoU	Segmentation Metrics AP_m	AP_m^{50}	AP_m^{75}	AP_b	AP_b^{50}	AP_b^{75}
X-PDNet	Provided by [21]	-	25.20	39.63	25.79	28.62	41.80	29.15
X-PDNet+Vanilla	Provided by [21]	-	**26.49**	41.61	**27.09**	**30.23**	44.18	**30.7**
X-PDNet+DGBPL	Provided by [21]	-	25.86	**41.79**	26.34	29.94	**45.55**	29.98
X-PDNet	Manually annotated	13.36	24.09	36.84	25.08	25.80	37.08	26.72
X-PDNet+Vanilla	Manually annotated	14.82	25.27	38.24	26.59	27.08	38.93	**27.77**
X-PDNet+DGBPL	Manually annotated	**16.68**	**26.12**	**39.47**	**26.68**	**28.18**	**40.86**	27.46

Table 2: Evaluation of segmentation results on **2D-3D-S** annotation provided by [21] and human labelling evaluation datasets.



Example produced by the baseline (above) and X-PDNet (below)



Planes predicted by X-PDNet, with Vanilla, and with DGBPL

한양 HANYANG UNIVERSITY 1939

서울대학교 SEOUL NATIONAL UNIVERSITY

MAXST

BMVC 2023