Associatively Segmenting Instances and Semantics in Point Clouds

Xinlong Wang, Shu Liu, Xiaoyong Shen, Chunhua Shen and Jiaya Jia

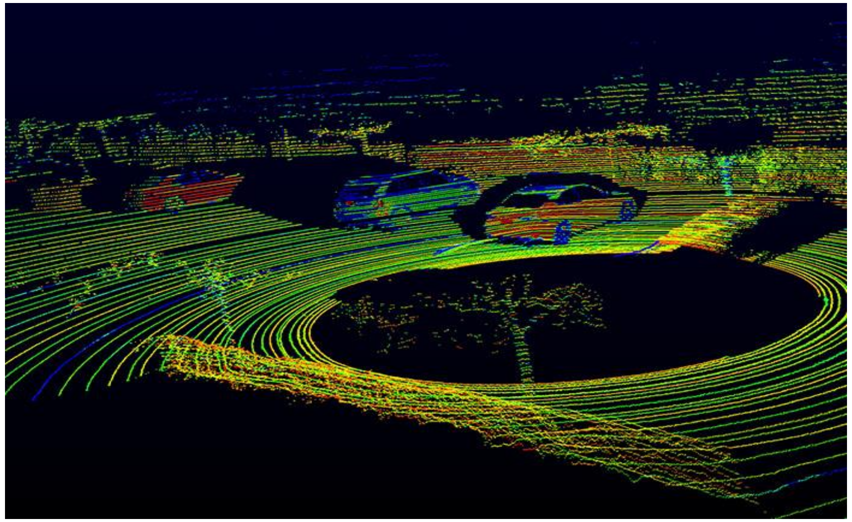
CVPR 2019

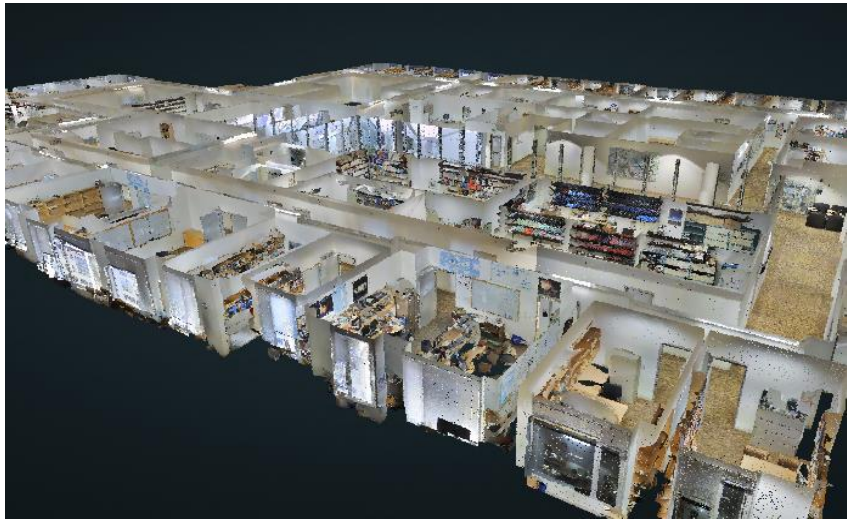
THE UNIVERSITY of ADELAIDE

香港中文大學
The Chinese University of Hong Kong

腾讯优图

Potential Applications



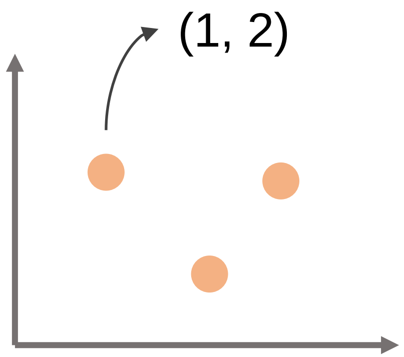


[1] Narita et al, PanopticFusion, arXiv, 2019

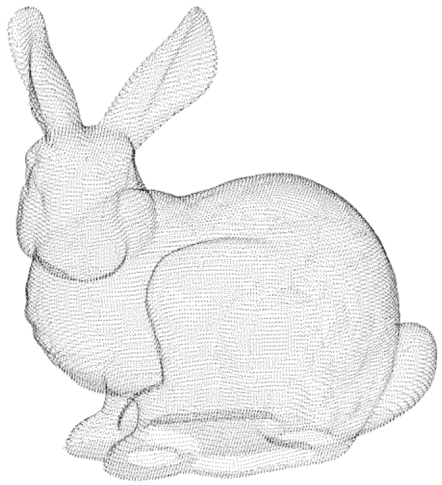
Preliminaries

What is Point Cloud?

- A set of data points:



In 2D space

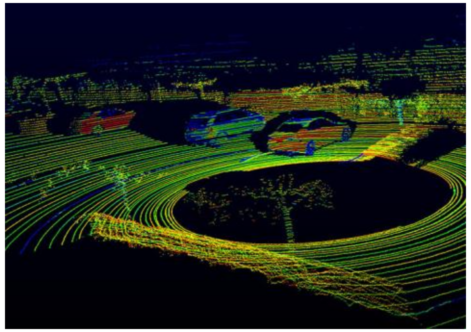


In 3D space

In 3D space (with color)

Preliminaries

How to Get Point Cloud?

Sensor	Example	Data format
 LiDAR		N x 4 [x, y, z, reflectance]
 3D Camera		N x 6 [x, y, z, r, g, b]

Preliminaries

Deep Learning on Point Clouds

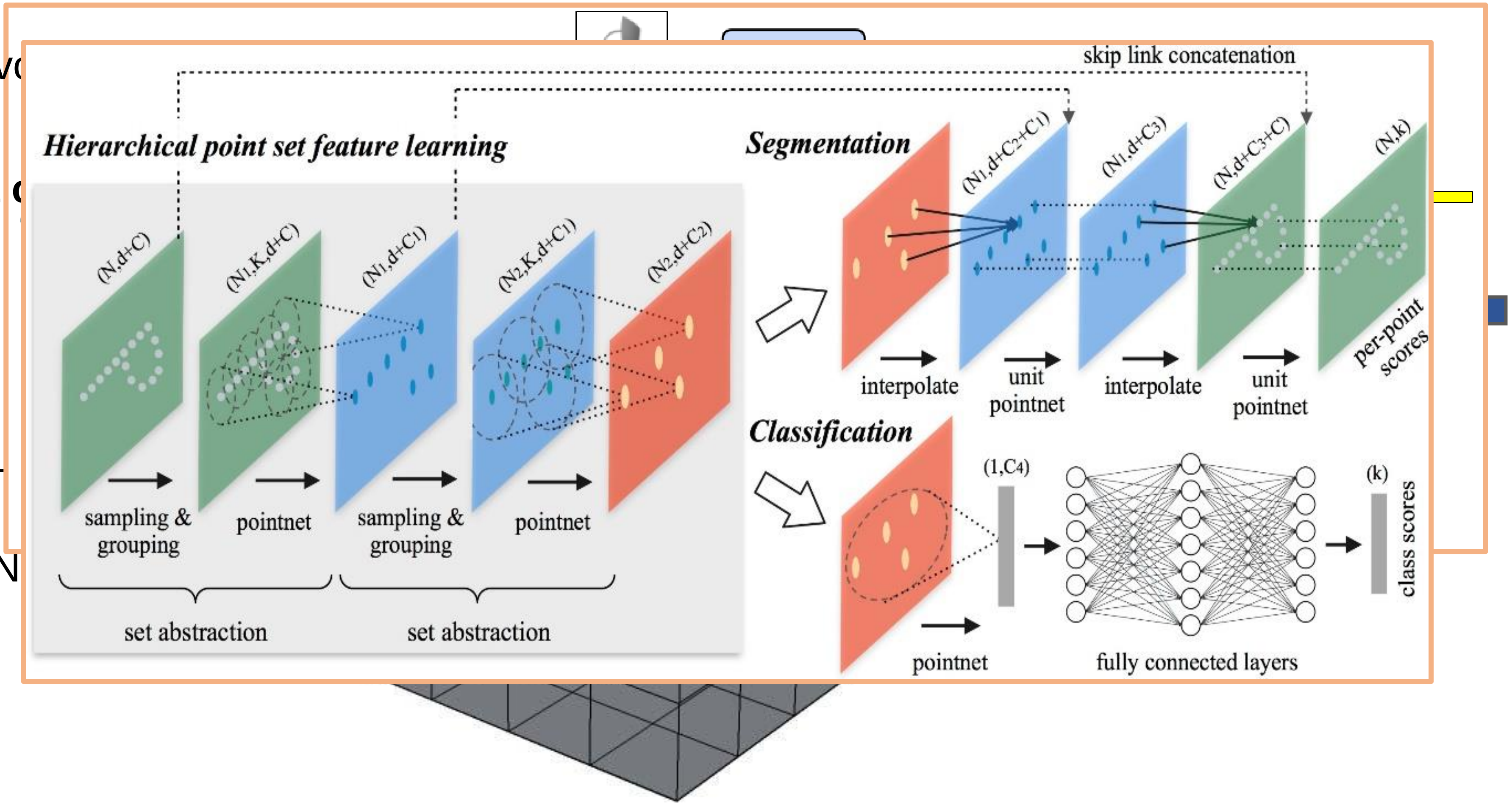
- Multiview rendering images + 2D CNNs: loss of contextual information

- Voxelized v

- **Raw point c**

PointNet

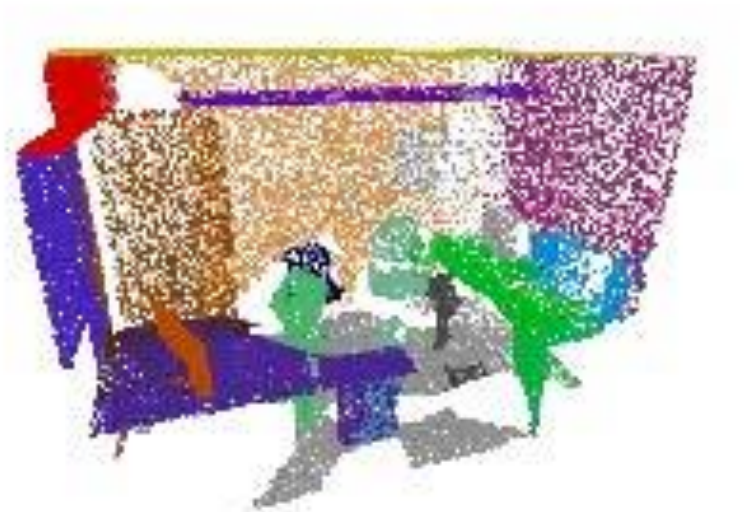
PointNet-
DGCNN
PointCNN
…



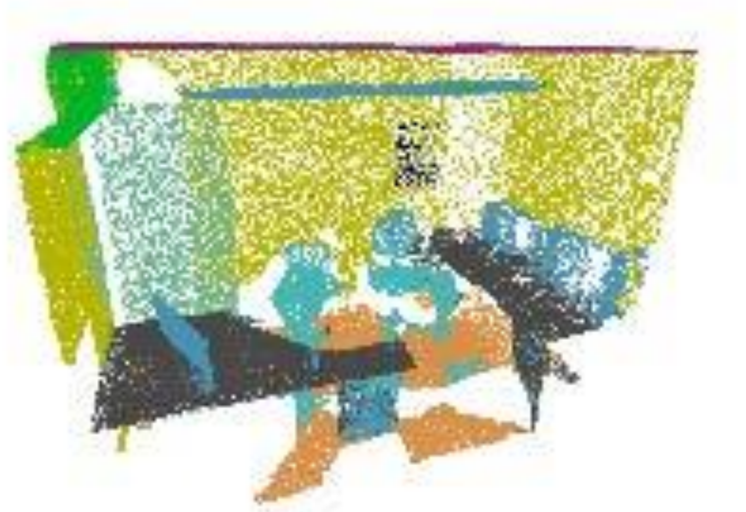
Task

Input

Per-point Predictions

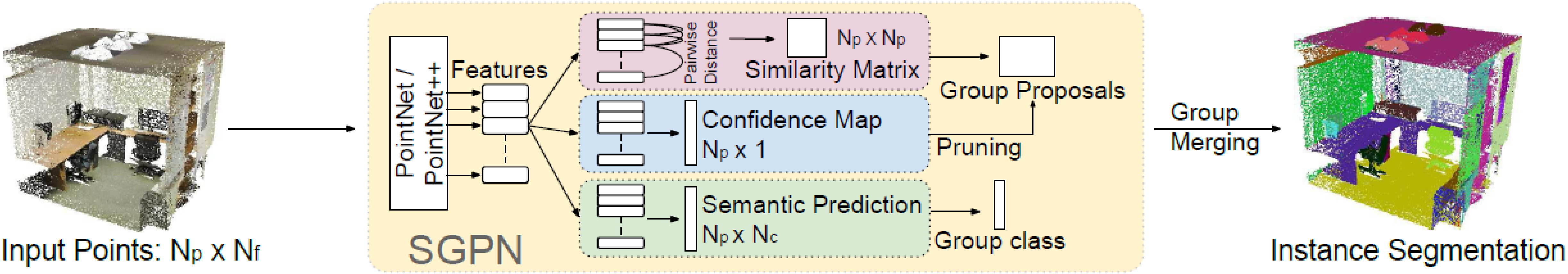


Instance Labels



Semantic Labels

Previous Method



- To learn the similarity matrix of a point cloud to get instance proposals.

Wang et al. SGPN: Similarity group proposal network for 3d point cloud instance segmentation, CVPR 2018

Associatively Segmenting Instances and Semantics in Point Clouds

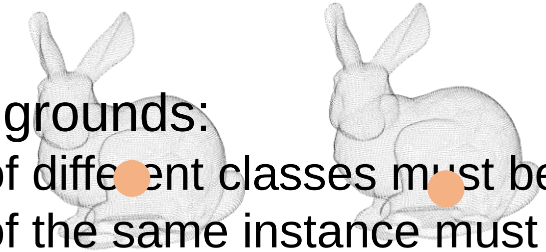
Motivation

Q: how could instance and semantic segmentation networks reinforce each other to make more accurate predictions?

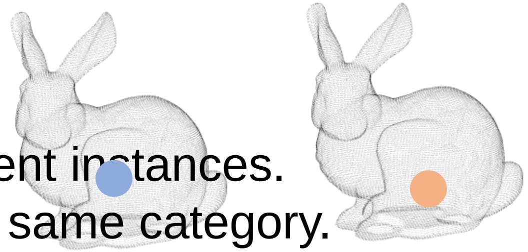
In fact, the two tasks conflict with each other in some respects.

Common grounds:

- Points of different classes must belong to different instances.
- Points of the same instance must belong to the same category.



Semantic Seg.



Instance Seg.

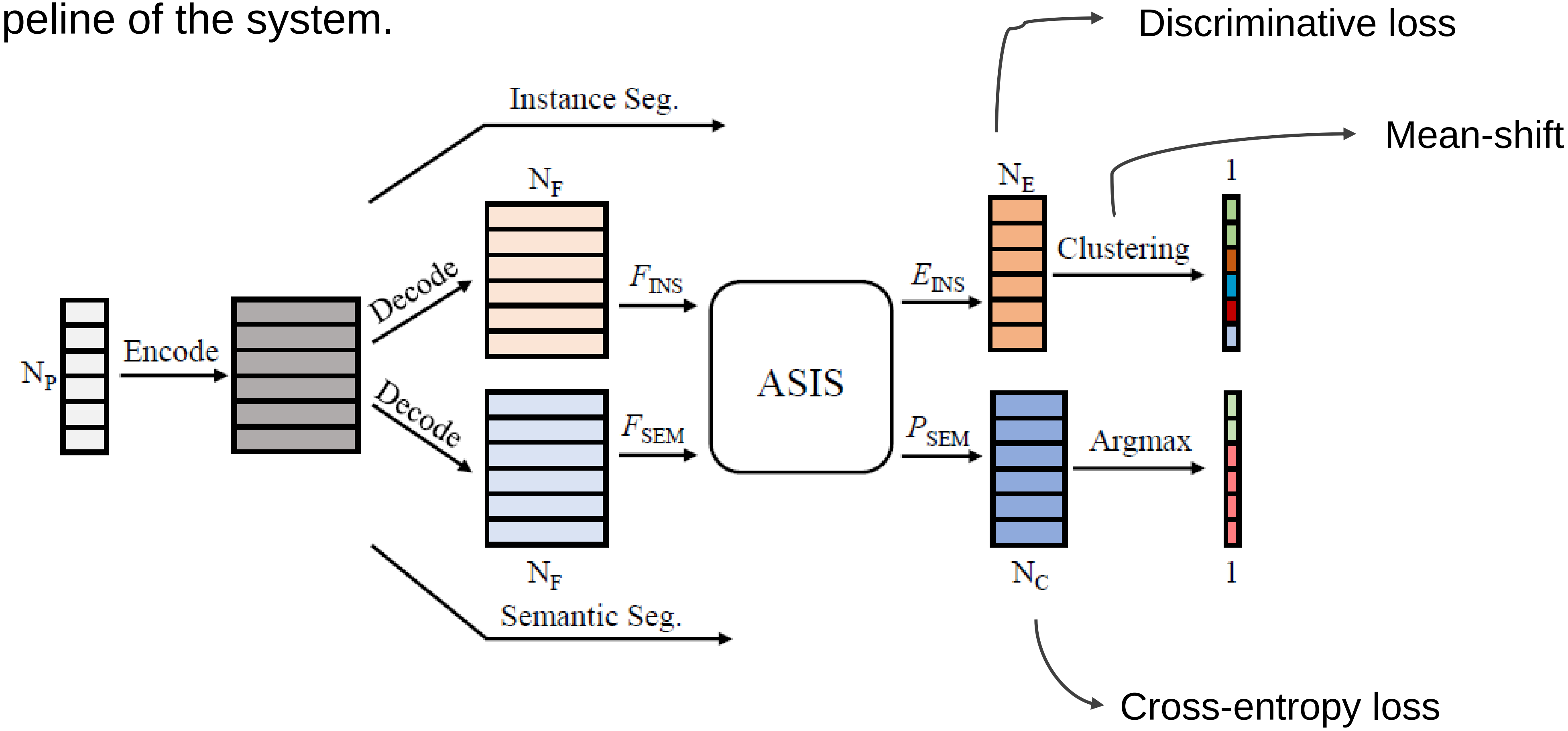
Two straightforward approaches:

- Given the semantic labels, we could run instance segmentation independently. on every semantic class.
- Given the instance labels, one could classify each instance and assign the predicted class label to each point of this instance.

Associatively Segmenting Instances and Semantics in Point Clouds

Method

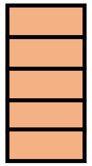
- Full pipeline of the system.



Associatively Segmenting Instances and Semantics in Point Clouds

Method

- Use Dscriminative Loss [1] to train the instance embeddings.



$$L = L_{var} + L_{dist} + \alpha \cdot L_{reg}, \quad (1)$$

$$L_{var} = \frac{1}{I}\sum_{i=1}^{I}\frac{1}{N_i}\sum_{j=1}^{N_i}\left[\|\mu_i - e_j\|_1 - \delta_{\mathrm{v}}\right]_+^2, \quad (2)$$

$$L_{dist} = \frac{1}{I(I-1)}\sum_{\substack{i_A=1 \\ i_A \neq i_B}}^{I}\sum_{i_B=1}^{I}\left[2\delta_{\mathrm{d}} - \|\mu_{i_A} - \mu_{i_B}\|_1\right]_+^2, \quad (3)$$

$$L_{reg} = \frac{1}{I}\sum_{i=1}^{I}\|\mu_i\|_1, \quad (4)$$

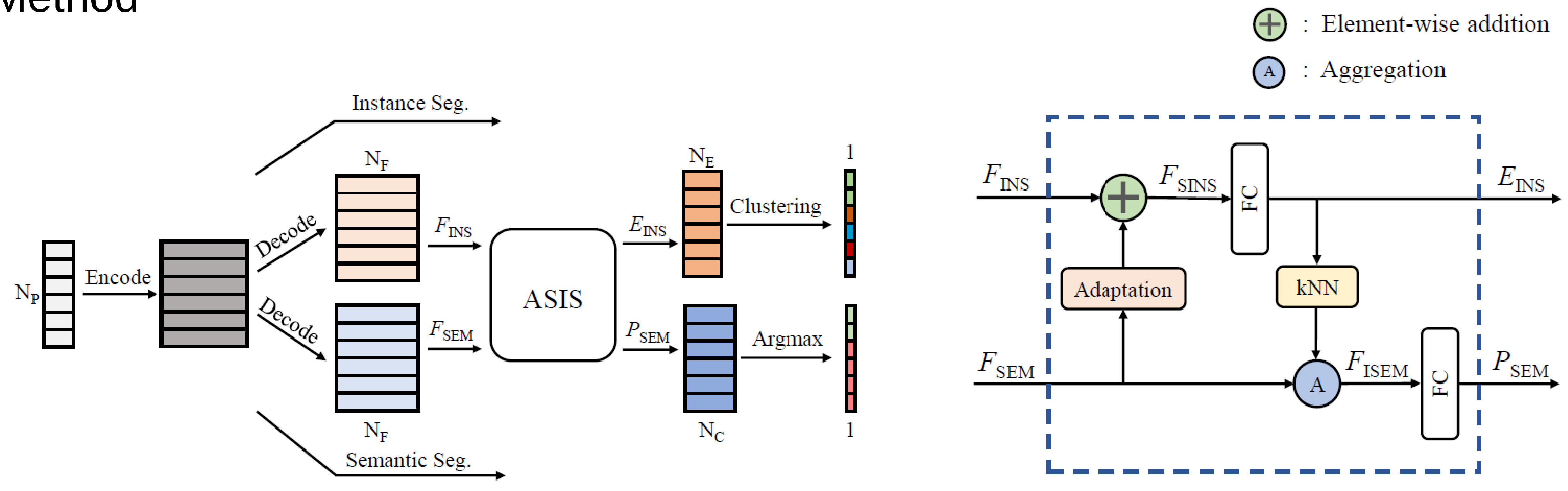
Differences:

- We adopt the class-agnostic instance embedding learning strategy while the loss used in [1] is class-specific.
- We use L1 distance in our loss terms based on our practice on 3D case.

[1] Brabandere et al. Semantic instance segmentation with a discriminative loss function, arXiv, 2017

Associatively Segmenting Instances and Semantics in Point Clouds

Method



(a) (b)

- Semantic-aware Instance Segmentation
- Instance-fused Semantic Segmentation

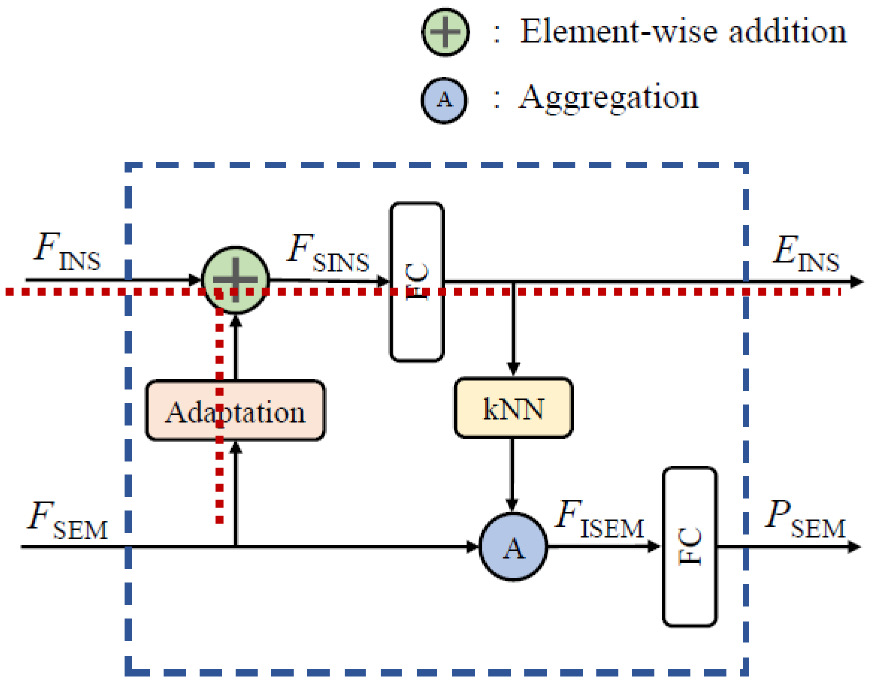
Method

- Semantic-aware Instance Segmentation

In semantic feature space, points are naturally positioned according to their categories.

$$F_{\text{SINS}} = F_{\text{INS}} + FC(F_{\text{SEM}}). \qquad (5)$$

In this soft and learnable way,
points belonging to different category instances are further repelled in instance feature space



(b)

Associatively Segmenting Instances and Semantics in Point Clouds

Method

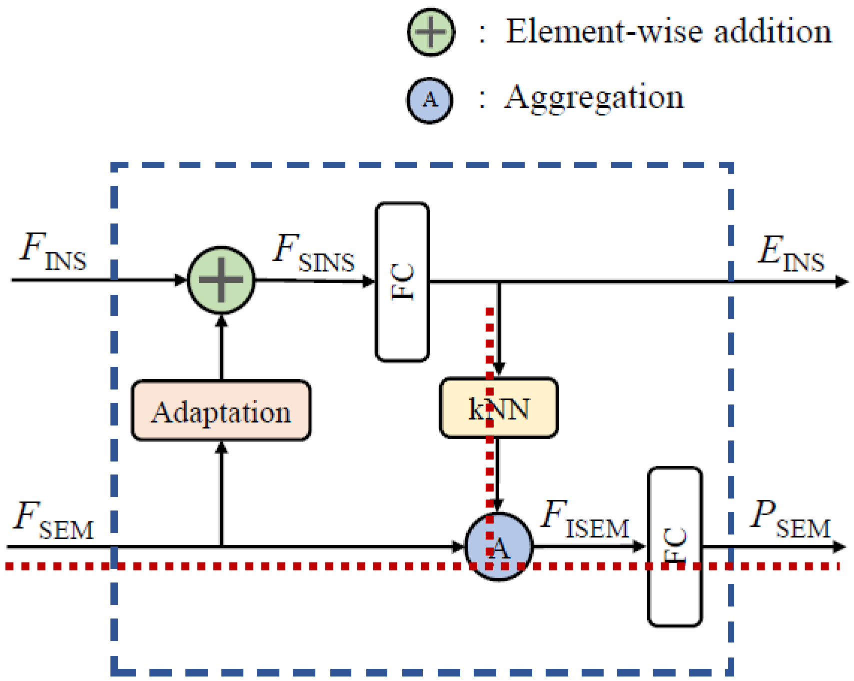
- Instance-fused Semantic Segmentation

A point being assigned to one of the categories is because the instance containing that point belongs to that category.

F_{SEM} $\qquad\qquad\qquad\qquad$ F_{ISEM}

$$N_P \times N_F \rightarrow N_P \times K \times N_F \rightarrow N_P \times N_F$$

channel-wise max aggregation



(b)

Associatively Segmenting Instances and Semantics in Point Clouds

Results

- Evaluation on S3DIS Dataset

Backbone	Method	mCov	mWCov	mPrec	mRec
Test on Area 5					
PN	SGPN [35]	32.7	35.5	36.0	28.7
	ASIS (*vanilla*)	38.0	40.6	42.3	34.9
	ASIS	**40.4**	**43.3**	**44.5**	**37.4**
PN++	ASIS (*vanilla*)	42.6	45.7	53.4	40.6
	ASIS	**44.6**	**47.8**	**55.3**	**42.4**
Test on 6-fold CV					
PN	SGPN [35]	37.9	40.8	38.2	31.2
	ASIS (*vanilla*)	43.0	46.3	50.6	39.2
	ASIS	**44.7**	**48.2**	**53.2**	**40.7**
PN++	ASIS (*vanilla*)	49.6	53.4	62.7	45.8
	ASIS	**51.2**	**55.1**	**63.6**	**47.5**

Table 1: Instance segmentation results on S3DIS dataset.

Backbone	Method	mAcc	mIoU	oAcc
Test on Area 5				
PN	PN (*RePr*)	52.1	43.4	83.5
	ASIS (*vanilla*)	52.9	44.7	83.7
	ASIS	**55.7**	**46.4**	**84.5**
PN++	ASIS (*vanilla*)	58.3	50.8	86.7
	ASIS	**60.9**	**53.4**	**86.9**
Test on 6-fold CV				
PN	PN [26]	-	47.7	78.6
	PN (*RePr*)	60.3	48.9	80.3
	ASIS (*vanilla*)	60.7	49.5	80.4
	ASIS	**62.3**	**51.1**	**81.7**
PN++	ASIS (*vanilla*)	69.0	58.2	85.9
	ASIS	**70.1**	**59.3**	**86.2**

Table 2: Semantic segmentation results on S3DIS dataset.

Associatively Segmenting Instances and Semantics in Point Clouds

Results

- Evaluation on S3DIS Dataset

Method	Inference Time (ms)			mWCov
	Overall	Network	Grouping	
SGPN	726	18	708	35.5
ASIS (*vanilla*)	212	11	201	41.4
ASIS	205	20	185	43.6
ASIS (*vanilla*.PN++)	150	35	115	45.7
ASIS (PN++)	179	54	125	47.8

Table 5: Comparisons of computation speed and performance. Inference time is estimated and averaged on Area 5, which is the time to process a point cloud with size 4096×9. The instance segmentation results on Area 5 are reported.

Associatively Segmenting Instances and Semantics in Point Clouds

Results

(a) Input (b) Ins. (c) Ins. GT (d) Sem. (e) Sem. GT

Figure 6: Qualitative results of ASIS on the S3DIS test fold.

Associatively Segmenting Instances and Semantics in Point Clouds

Results

- Evaluation on ShapeNet Dataset

Method	mIoU
PointNet [26]	83.7
PointNet (*RePr*)	83.4
PointNet++ [28]*	84.3
ASIS (PN)	84.0
ASIS (PN++)	85.0

Table 6: Semantic segmentation results on ShapeNet datasets. *RePr* is our reproduced PointNet. PointNet++* denotes the PointNet++ trained by us without extra normal information.


(a) Ins. (b) Ins. GT (c) Sem. (d) Sem. GT

Figure 8: Qualitative results of ASIS on ShapeNet test split. (a) Instance segmentation results of ASIS. (b) Generated ground truth for instance segmentation. (c) Semantic segmentation results of ASIS. (d) Semantic segmentation ground truth.

Associatively Segmenting Instances and Semantics in Point Clouds

Analysis

- Ablative Analysis

Method	+IF	+SA	mIoU	mWCov
Baseline			49.5	46.3
	✓		50.0	47.0
		✓	49.8	47.4
	✓	✓	**51.1**	**48.2**

Table 3: Ablation study on the S3DIS dataset. IF refers to instance fusion; SA refers to semantic awareness.

- Category-based Analysis

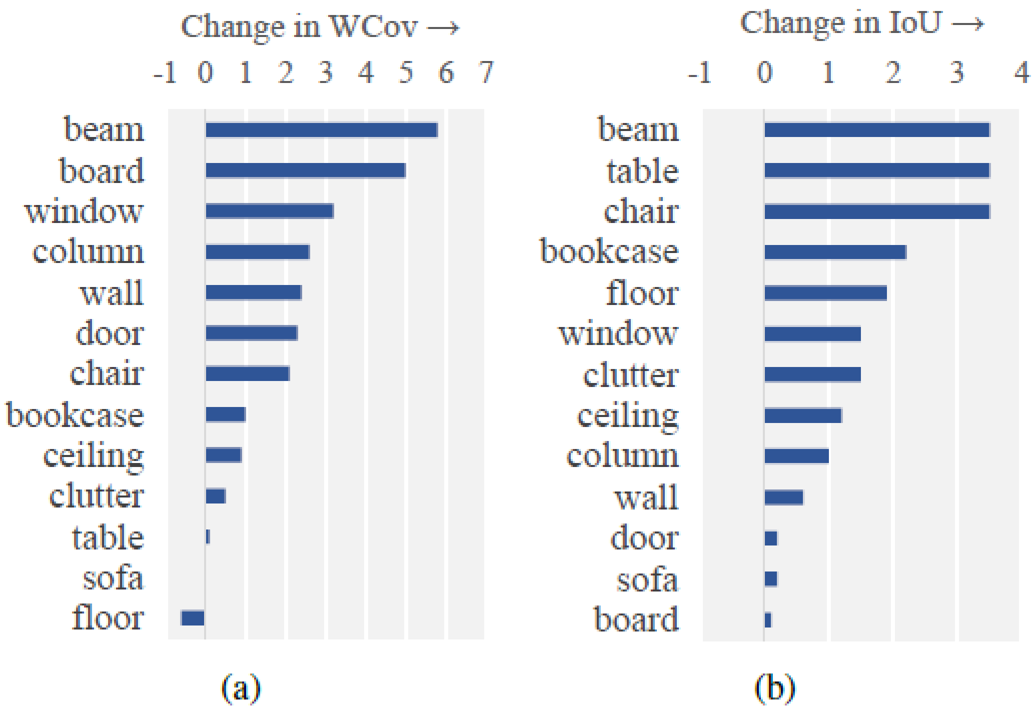
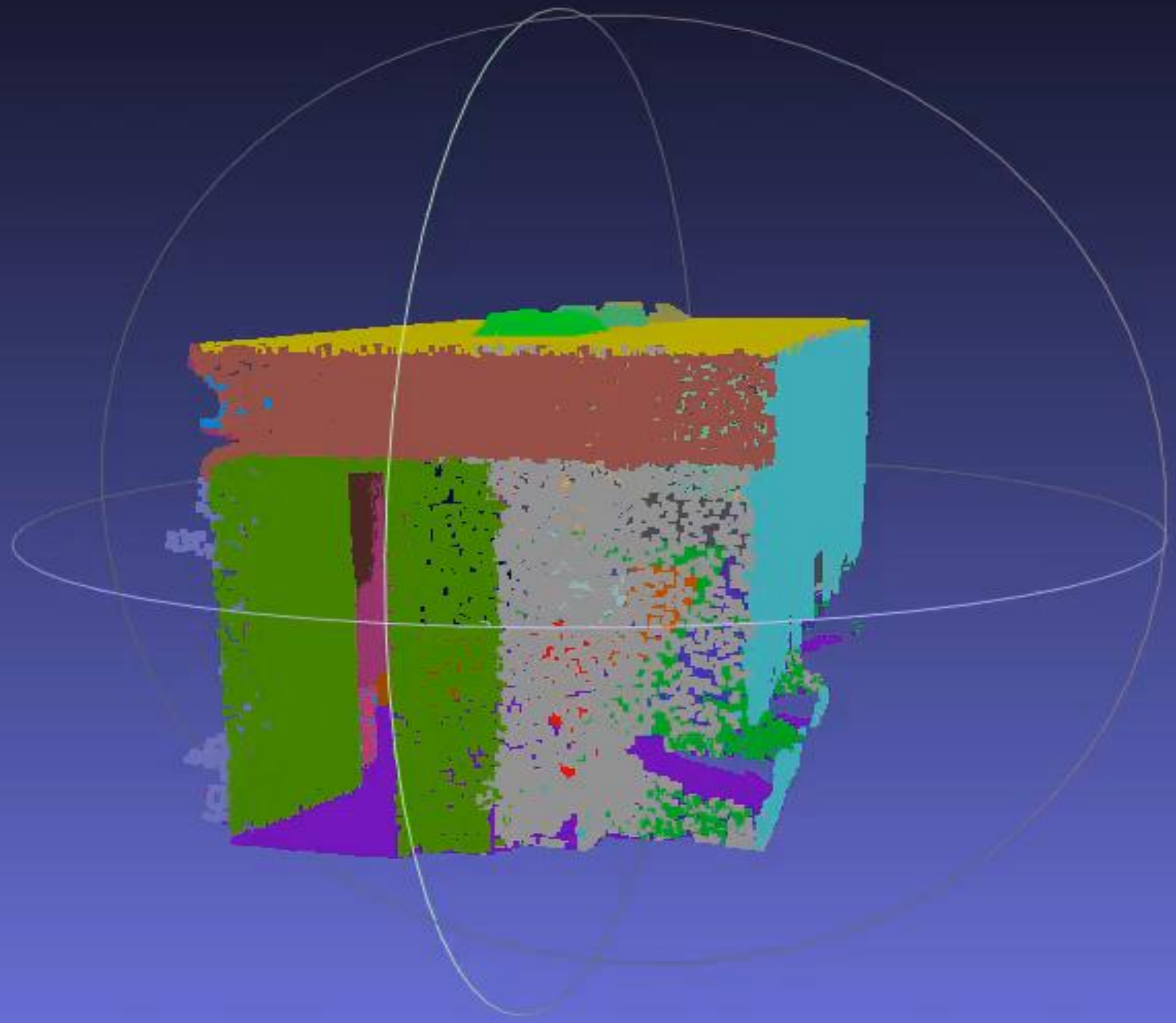


Figure 7: Per class performance changes. (a) Changes of instance segmentation performance compared to our baseline method. (b) Changes of semantic segmentation performance compared to our baseline method.

Contributions

- We propose a fast and efficient simple baseline for simultaneous instance segmentation and semantic segmentation on 3D point clouds.
- We propose a new framework, termed ASIS, to associate instance segmentation and semantic segmentation closely together.
- Our method largely outperforms the state-of-the-art method in 3D instance segmentation along with a significant improvement in 3D semantic segmentation.

Demo



Associatively Segmenting Instances and Semantics in Point Clouds

Xinlong Wang, Shu Liu, Xiaoyong Shen, Chunhua Shen and Jiaya Jia

CVPR 2019

Code: https://github.com/wxinlong/asis

Contact: 1452405wxl@tongji.edu.cn