JUNE 18-22, 2023
CVPR
VANCOUVER, CANADA

Center Focusing Network for Real-Time LiDAR Panoptic Segmentation

Xiaoyan Li, Gang Zhang, Boyue Wang, Yongli Hu, Baocai Yin
Beijing Institute of Artificial Intelligence, Faculty of Information Technology,
Beijing University of Technology, Beijing 100124, China

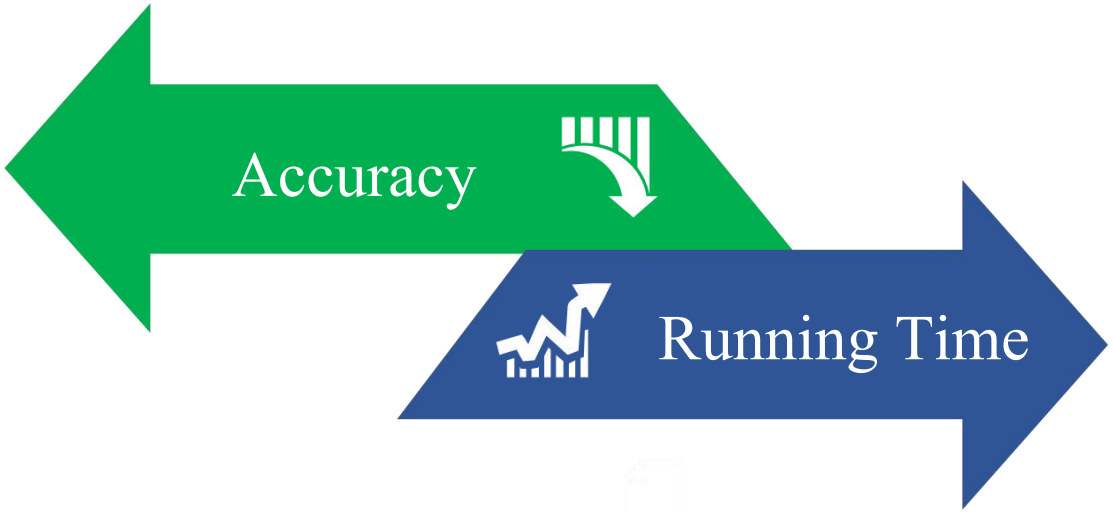
Paper ID: 2616
Poster ID: WED-PM-102

Preview

Key problems that limit the **accuracy** and **speed** of **proposal-free** LIDAR panoptic segmentation approaches

Center Encoding

Difficulty of encoding non-existent instance centers results in the errors of LiDAR instance segmentation

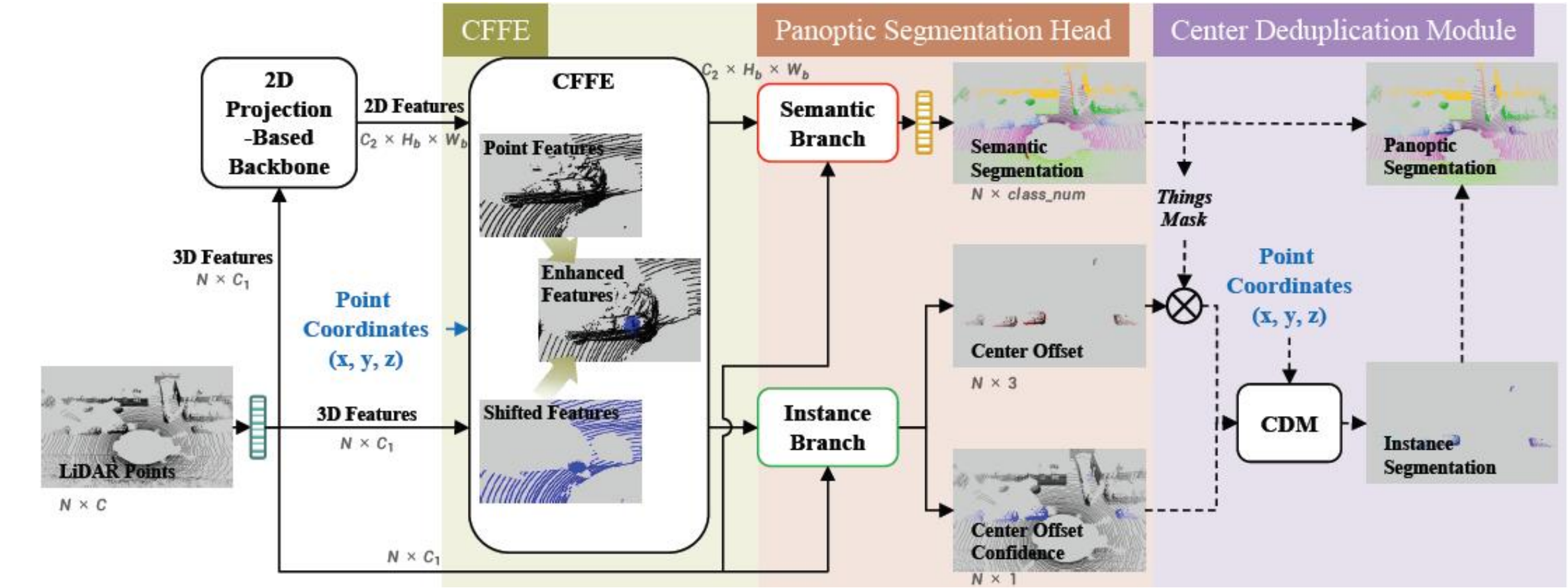


Center Exploiting

The costly center-based clustering algorithms slow down the overall inference speed

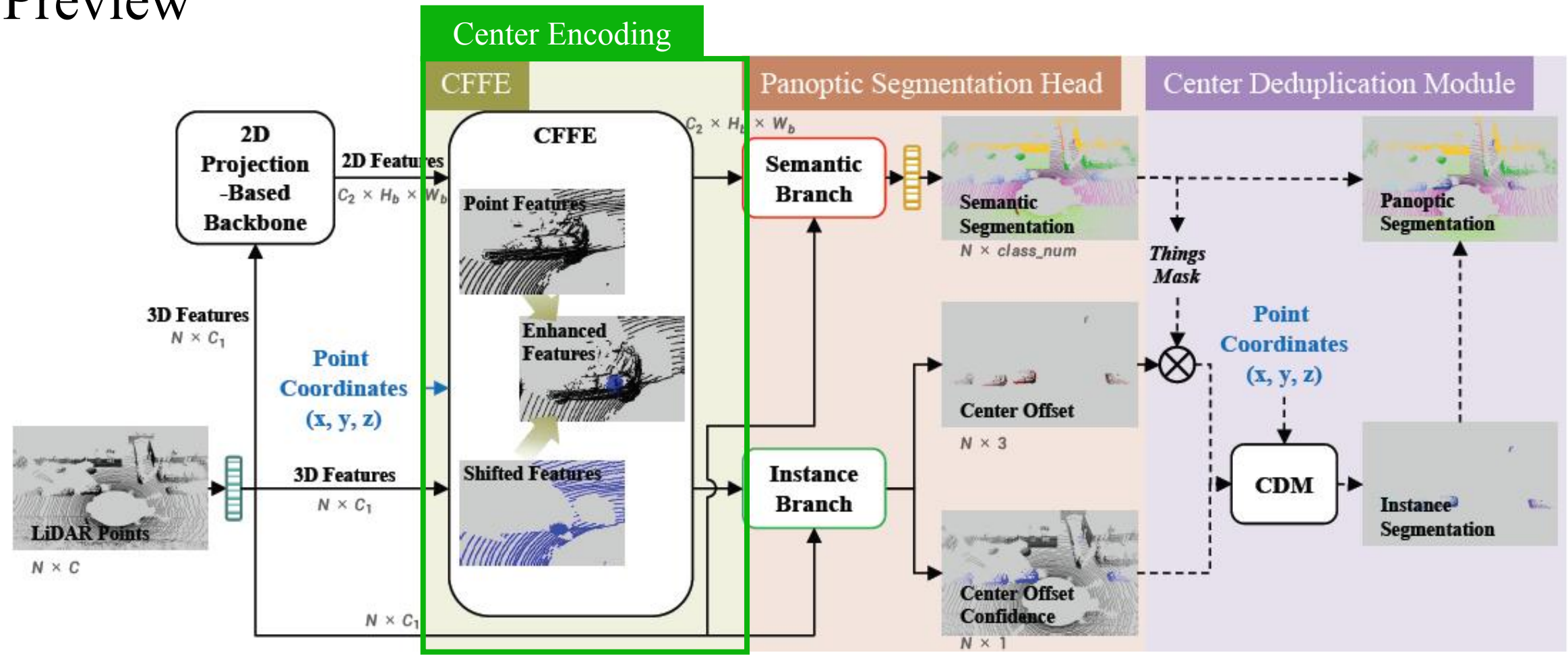
JUNE 18-22, 2023
CVPR
VANCOUVER, CANADA

Preview



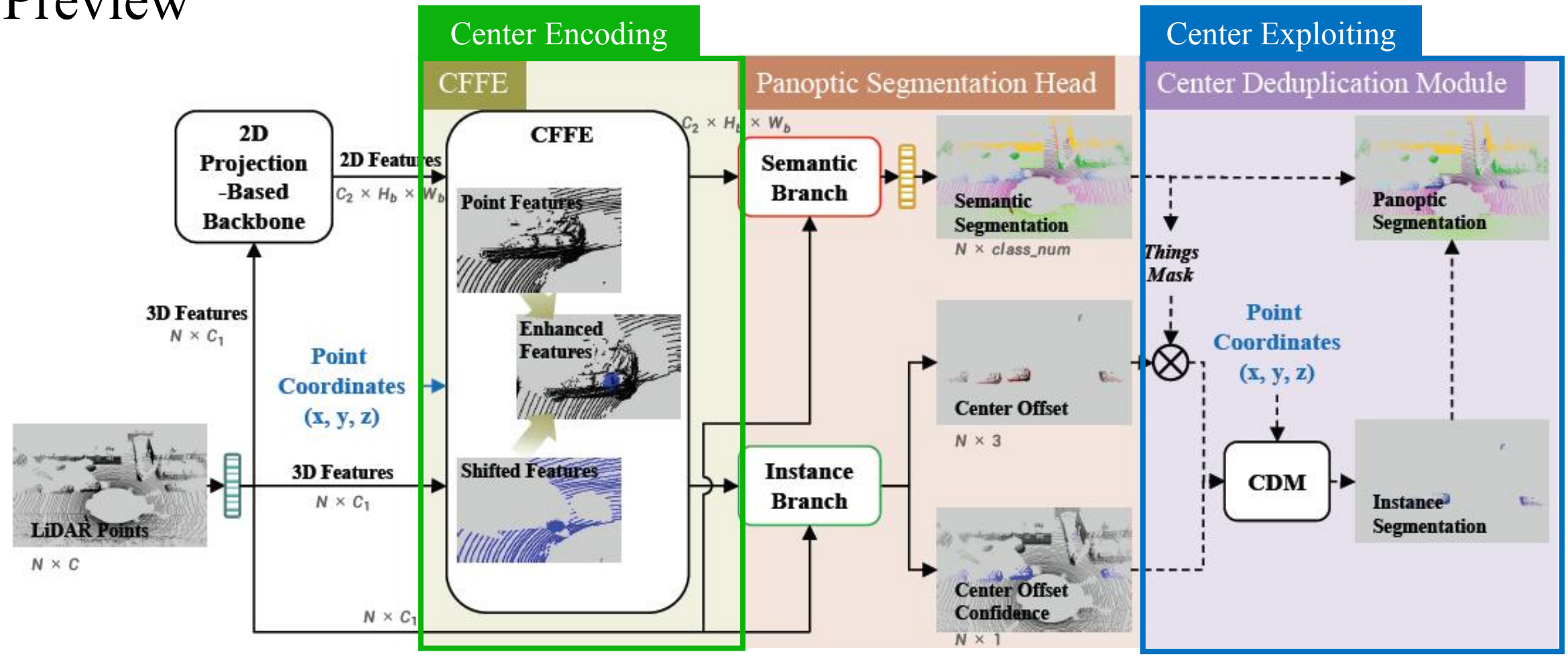
JUNE 18-22, 2023
CVPR
VANCOUVER, CANADA

Preview



JUNE 18-22, 2023
CVPR
VANCOUVER, CANADA

Preview

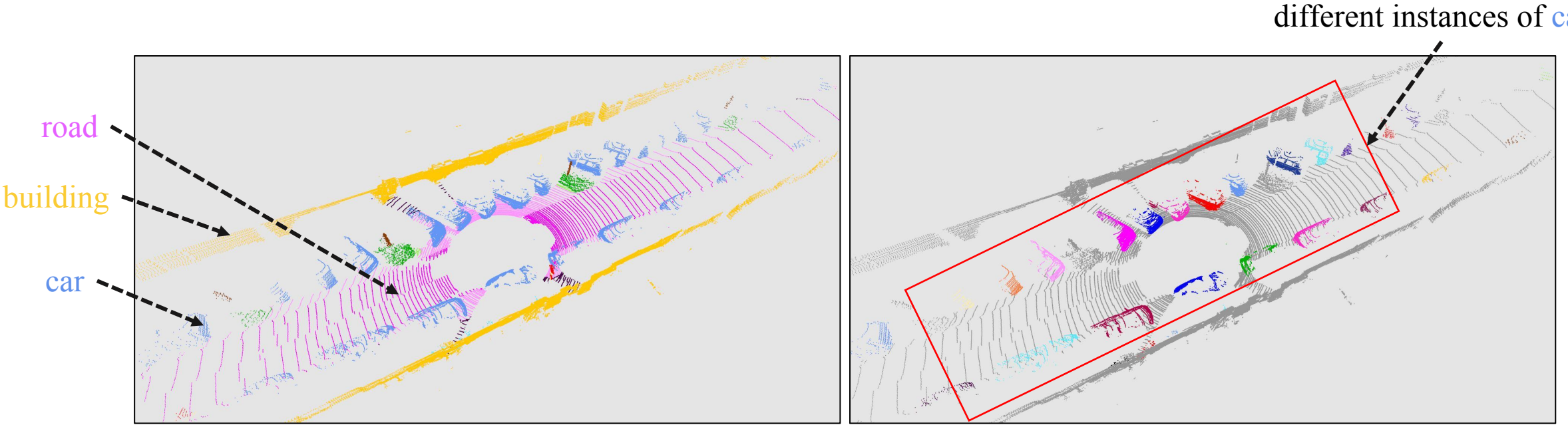


JUNE 18-22, 2023
CVPR
VANCOUVER, CANADA

LiDAR Panoptic Segmentation

1. **LiDAR Semantic Segmentation**: for both ***Stuff*** *(e.g. road, building)* and ***Things*** *(e.g. car, pedestrian)* classes

2. **LiDAR Instance Segmentation**: only for ***Things*** *(e.g. car, pedestrian)* classes



(a) LiDAR Semantic Segmentation

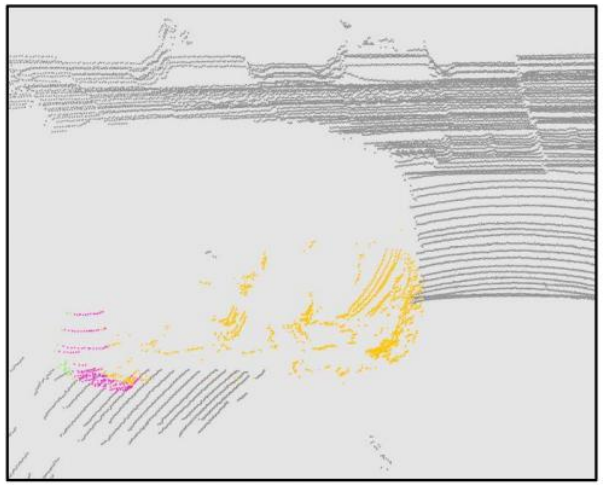
(b) LiDAR Instance Segmentation

JUNE 18-22, 2023
CVPR
VANCOUVER, CANADA

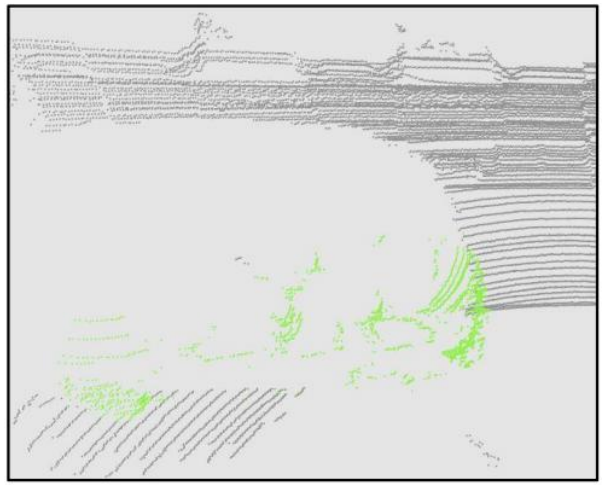
Accurate and Real-Time LiDAR Panoptic Segmentation

1. Accurate: The difficulty of modeling **non-existent instance centers** results in the errors of LiDAR instance segmentation

- LiDAR scans are surface-aggregated



(a) CFNet without CFFE



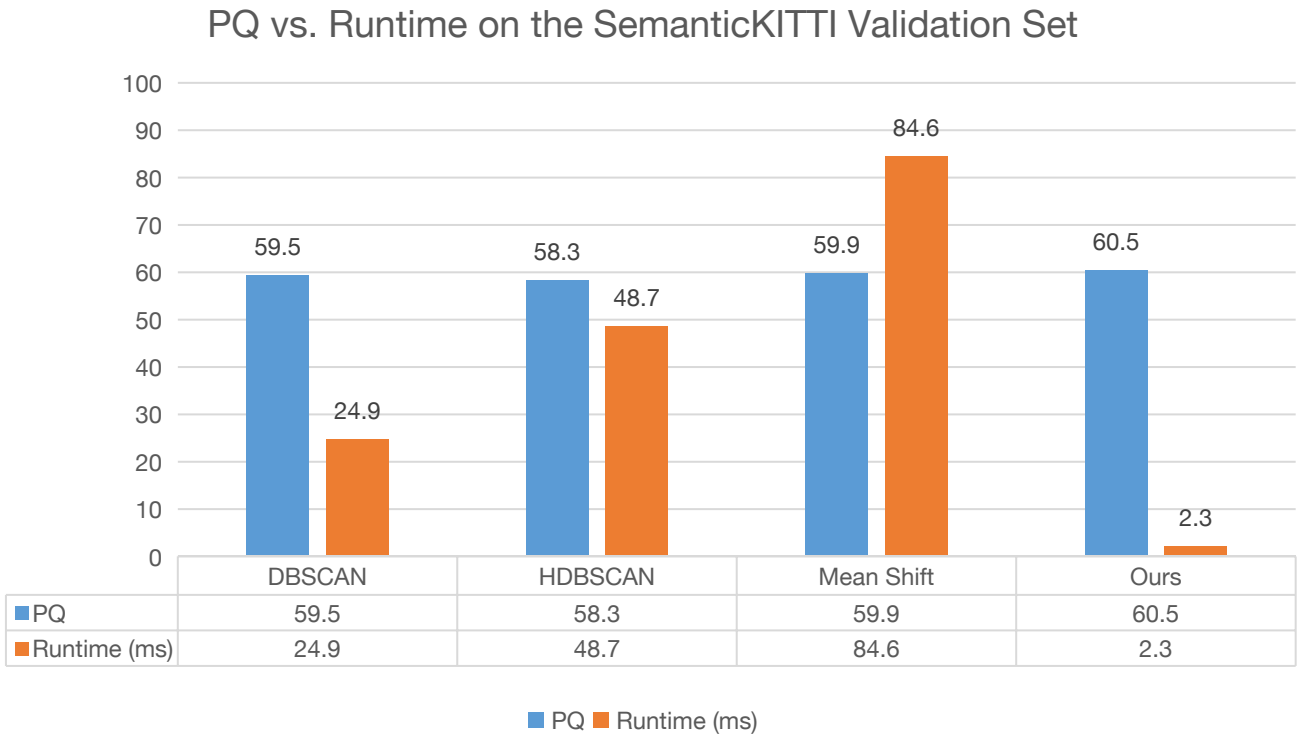
(b) CFNet with CFFE

JUNE 18-22, 2023
CVPR
VANCOUVER, CANADA

Accurate and Real-Time LiDAR Panoptic Segmentation

2. Real-Time: The costly center-based clustering algorithms slow down the overall inference speed

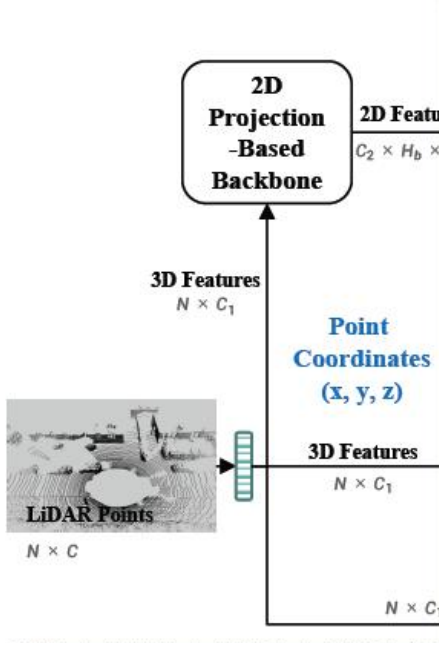
- These clustering algorithms follow a complicated iterative process



JUNE 18-22, 2023
CVPR
VANCOUVER, CANADA

Our Method—Center Focusing Network (CFNet)

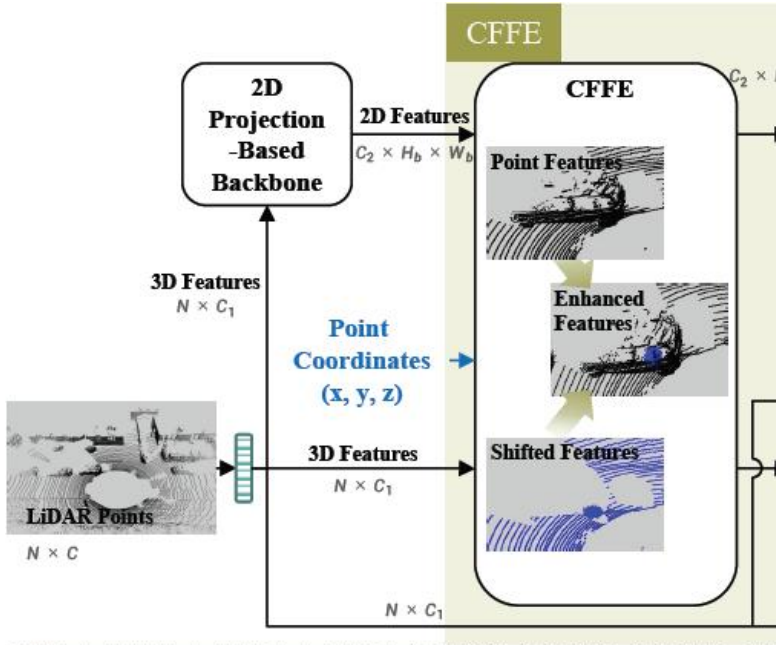
Our Method—Center Focusing Network (CFNet)



Four Step:

1. 2D projection-based backbone

Our Method—Center Focusing Network (CFNet)

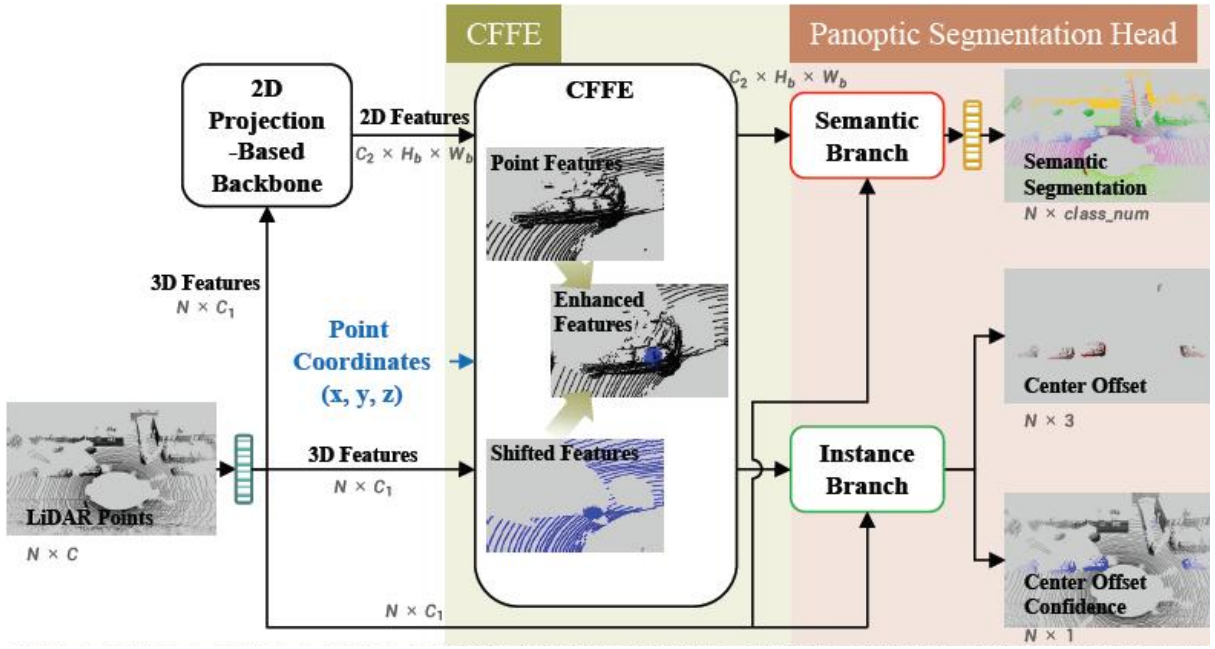


Four Step:

1. 2D projection-based backbone
2. **Center focusing feature encoding (CFFE)**

JUNE 18-22, 2023
CVPR
VANCOUVER, CANADA

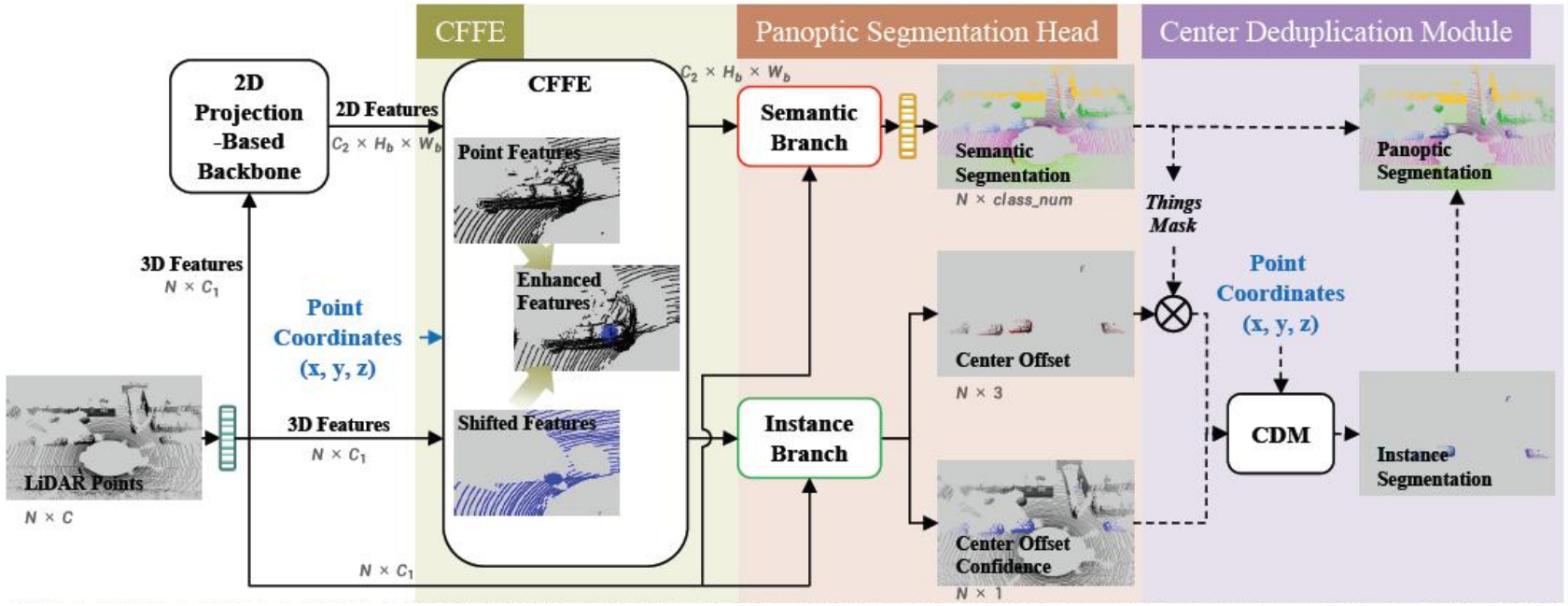
Our Method—Center Focusing Network (CFNet)



Four Step:

1. 2D projection-based backbone
2. **Center focusing feature encoding (CFFE)**
3. Panoptic segmentation head

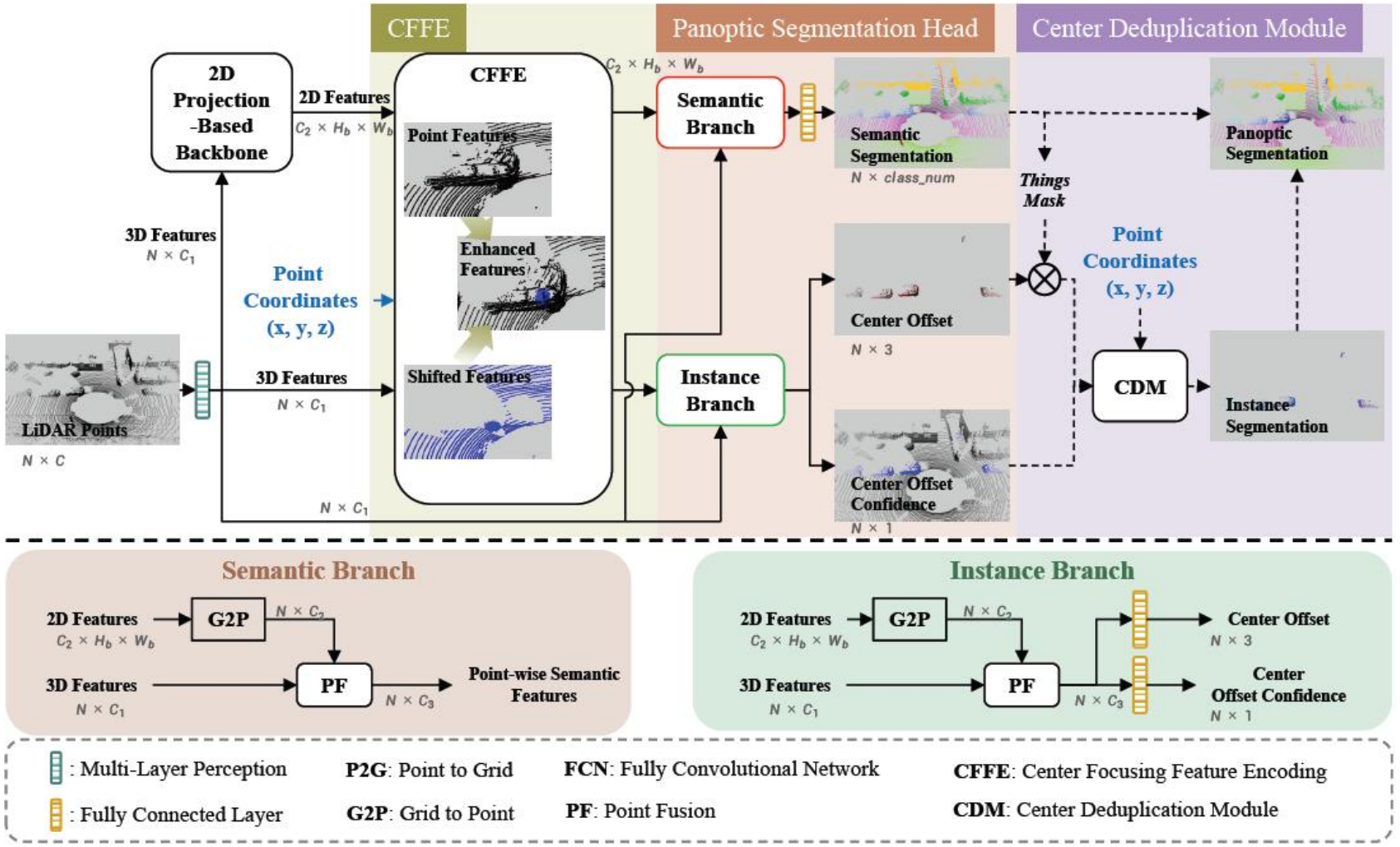
Our Method—Center Focusing Network (CFNet)



Four Step:

1. 2D projection-based backbone
2. **Center focusing feature encoding (CFFE)**
3. Panoptic segmentation head
4. **Center deduplication module (CDM)**

Our Method—Center Focusing Network (CFNet)



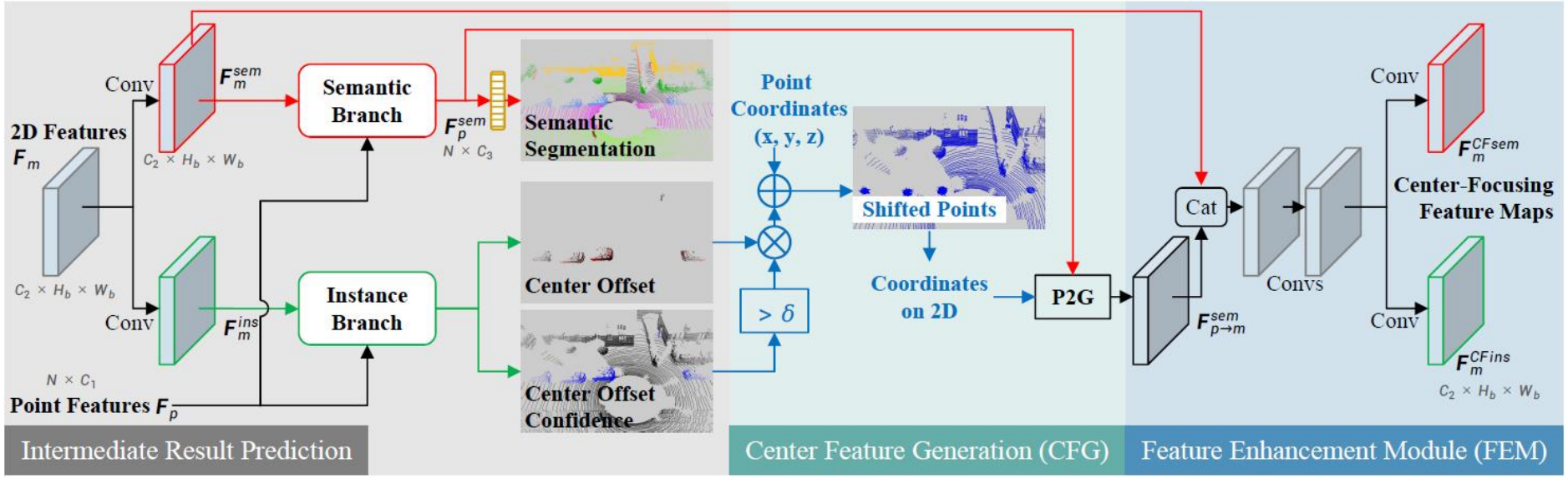
Four Step:

1. 2D projection-based backbone
2. **Center focusing feature encoding (CFFE)**
3. Panoptic segmentation head
4. **Center deduplication module (CDM)**

Our Method—Center Encoding

1. CFFE assists the model in understanding the relationships between the LiDAR points and **instance centers**

- Using the intermediate result predictions to **shift the LiDAR points and fill in the instance center points**
- Three modules: **Intermediate Result Prediction, Center Feature Generation, Feature Enhancement Module**

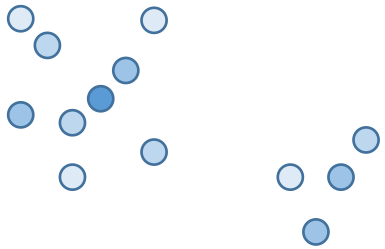


Framework of the proposed center focusing feature encoding (CFFE)

Our Method—Center Deduplication

2. Center deduplication module (CDM) efficiently **selects only one center for each instance**

- Suppressing the candidate instance centers with lower scores within a Euclidean distance threshold
- Being implemented efficiently in CUDA without iterative processing
- **Instance segmentation** is achieved by assigning the shifted ***Things*** points to the closest **deduplicated instance centers**

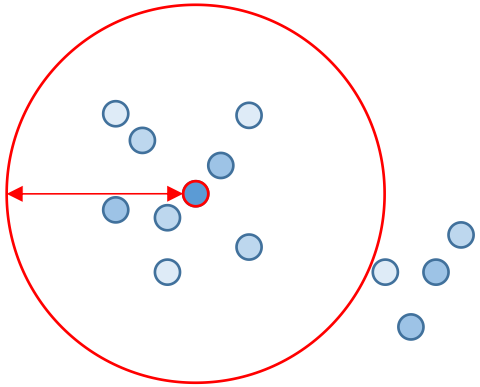


shifted *Things* points

Our Method—Center Deduplication

2. Center deduplication module (CDM) efficiently **selects only one center for each instance**

- Suppressing the candidate instance centers with lower scores within a Euclidean distance threshold
- Being implemented efficiently in CUDA without iterative processing
- **Instance segmentation** is achieved by assigning the shifted ***Things*** points to the closest **deduplicated instance centers**



shifted *Things* points

Our Method—Center Deduplication

2. Center deduplication module (CDM) efficiently **selects only one center for each instance**

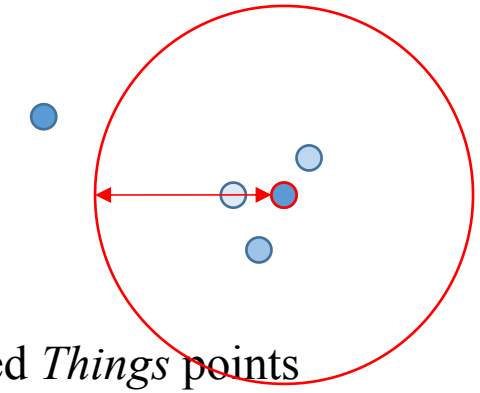
- Suppressing the candidate instance centers with lower scores within a Euclidean distance threshold
- Being implemented efficiently in CUDA without iterative processing
- **Instance segmentation** is achieved by assigning the shifted ***Things*** points to the closest **deduplicated instance centers**

shifted *Things* points

Our Method—Center Deduplication

2. Center deduplication module (CDM) efficiently **selects only one center for each instance**

- Suppressing the candidate instance centers with lower scores within a Euclidean distance threshold
- Being implemented efficiently in CUDA without iterative processing
- **Instance segmentation** is achieved by assigning the shifted ***Things*** points to the closest **deduplicated instance centers**



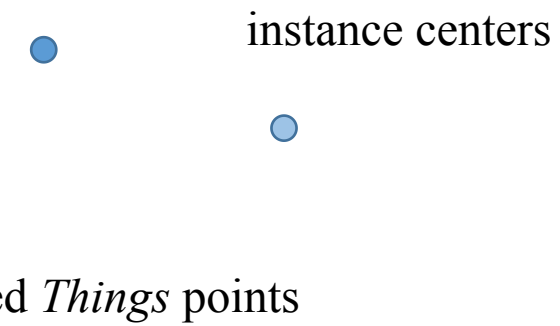
shifted *Things* points

JUNE 18-22, 2023
CVPR
VANCOUVER, CANADA

Our Method—Center Deduplication

2. Center deduplication module (CDM) efficiently **selects only one center for each instance**

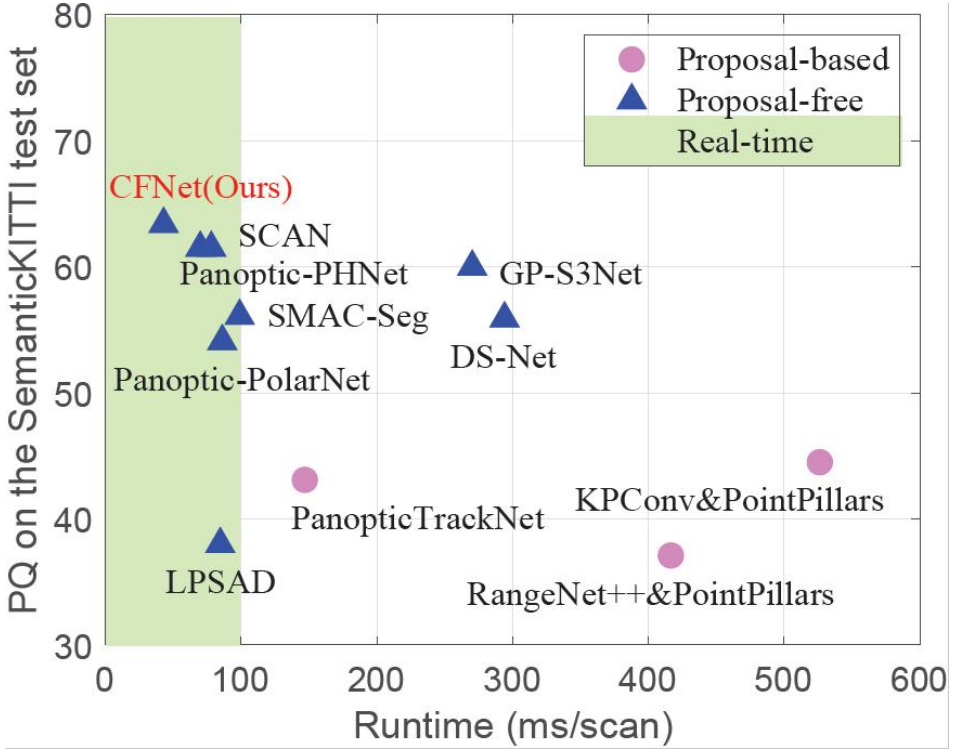
- Suppressing the candidate instance centers with lower scores within a Euclidean distance threshold
- Being implemented efficiently in CUDA without iterative processing
- **Instance segmentation** is achieved by assigning the shifted ***Things*** points to the closest **deduplicated instance centers**



Experiment

1. Our CFNet achieves the **SOTA results with a real-time inference speed** on the SemanticKITTI leaderboard

   - Our CFNet achieves the **highest PQ of 63.4**
   - Our CFNet is **1.6 times** faster than the the current most efficient method **Panoptic-PHNet**



Experiment

2. Our CFNet achieves the **SOTA results** on the nuScenes validation set

- Our CFNet achieves the **highest PQ of 75.1**

Methods	PQ	PQ†	SQ	RQ	PQ^{Th}	SQ^{Th}	RQ^{Th}	PQ^{St}	SQ^{St}	RQ^{St}	mIoU
Proposal-based Methods											
Cylinder3D [40]&PointPillars [19]	36.0	44.5	83.3	43.0	23.3	83.7	27.0	57.2	82.7	69.6	52.3
Cylinder3D [40]&SECOND [36]	40.1	48.4	84.2	47.3	29.0	84.4	33.6	58.5	83.7	70.1	58.5
PanopticTrackNet [17]	51.4	56.2	80.2	63.3	45.8	81.4	55.9	60.4	78.3	75.5	58.0
EfficientLPS [31]	62.0	65.6	83.4	73.9	56.8	83.2	68.0	70.6	83.8	83.6	65.6
Proposal-free Methods											
LPSAD [24]	50.4	57.7	79.4	62.4	43.2	80.2	53.2	57.5	78.5	71.7	62.5
DS-Net [15]	42.5	51.0	83.6	50.3	32.5	83.1	38.3	59.2	84.4	70.3	70.7
GP-S3Net [29]	61.0	67.5	84.1	72.0	56.0	85.3	65.2	66.0	82.9	78.7	75.8
SMAC-Seg [20]	67.0	71.8	85.0	78.2	65.2	87.1	74.2	68.8	82.9	82.2	72.2
Panoptic-PolarNet [39]	63.4	67.2	83.9	75.3	59.2	84.1	70.3	70.4	83.6	83.5	66.9
SCAN [35]	65.1	68.9	85.7	75.3	60.6	85.7	70.2	72.5	85.7	83.8	77.4
Panoptic-PHNet [22]	74.7	77.7	88.2	84.2	74.0	89.0	82.5	75.9	86.8	86.9	**79.7**
CFNet [Ours] w/CPGNet [23]	**75.1**	**78.0**	**88.8**	**84.6**	**74.8**	**89.8**	**82.9**	**76.6**	**87.1**	**87.3**	79.3

Experiment—Ablation Studies

3. Effects of the CFFE

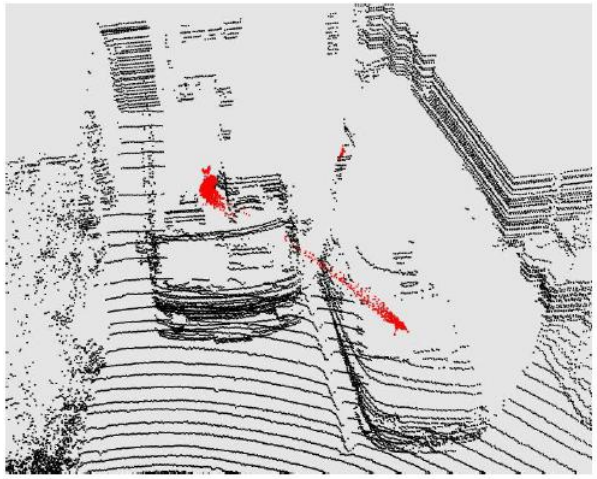
- The CFFE can improve the PQ for the backbones of PolarNet (+**1.3 PQ**) and CPGNet (+**2.2 PQ**)
- The CFFE brings acceptable latency for the backbones of PolarNet (+**6.6 ms**) and CPGNet (+**8.5 ms**)

	Backbone	CFFE		Deduplication	PQ	PQ†	PQ^{Th}	PQ^{St}	mIoU	RT(ms)
		CFG	FEM							
a	PolarNet [38]			BEV Center Heatmap [39]	59.1	64.1	65.7	54.3	64.5	65.3 + 20.9
b			✓	BEV Center Heatmap [39]	59.5	64.3	66.2	54.4	64.6	69.7 + 20.9
c		✓	✓	BEV Center Heatmap [39]	60.4	65.6	67.2	55.1	65.2	71.9 + 20.9
d		✓	✓	CDM [Ours]	**60.6**	**65.7**	**67.8**	**55.2**	**65.4**	**71.9 + 3.6**
e	CPGNet [23]			DBSCAN [11]	59.5	64.1	63.8	56.3	64.6	31.6+24.9
f				HDBSCAN [6]	58.3	62.9	60.9	56.3	64.9	31.6+48.7
g				MeanShift [9]	59.9	64.5	64.7	56.3	64.9	31.6+84.6
h				CDM [Ours]	60.5	65.6	66.1	56.5	66.3	**32.7+2.3**
i			✓	CDM [Ours]	60.8	65.9	66.3	56.9	66.5	37.4+2.3
j		✓	✓	CDM [Ours]	**62.7**	**67.5**	**70.0**	**57.3**	**67.4**	41.2+2.3

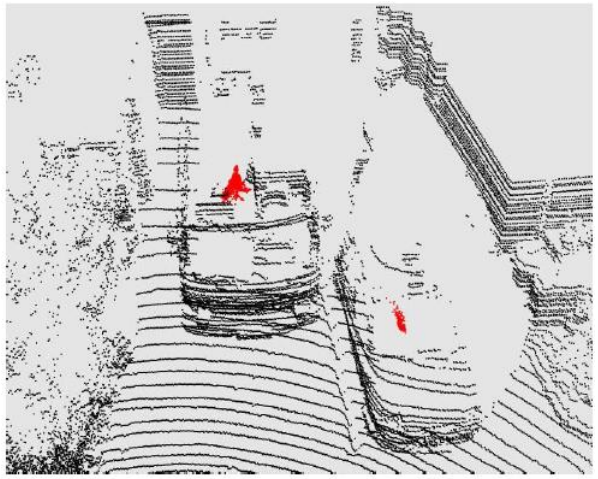
Experiment—Ablation Studies

3. Effects of the CFFE

- The CFFE makes the predicted instance centers more compact



(a) Intermediate Predictions

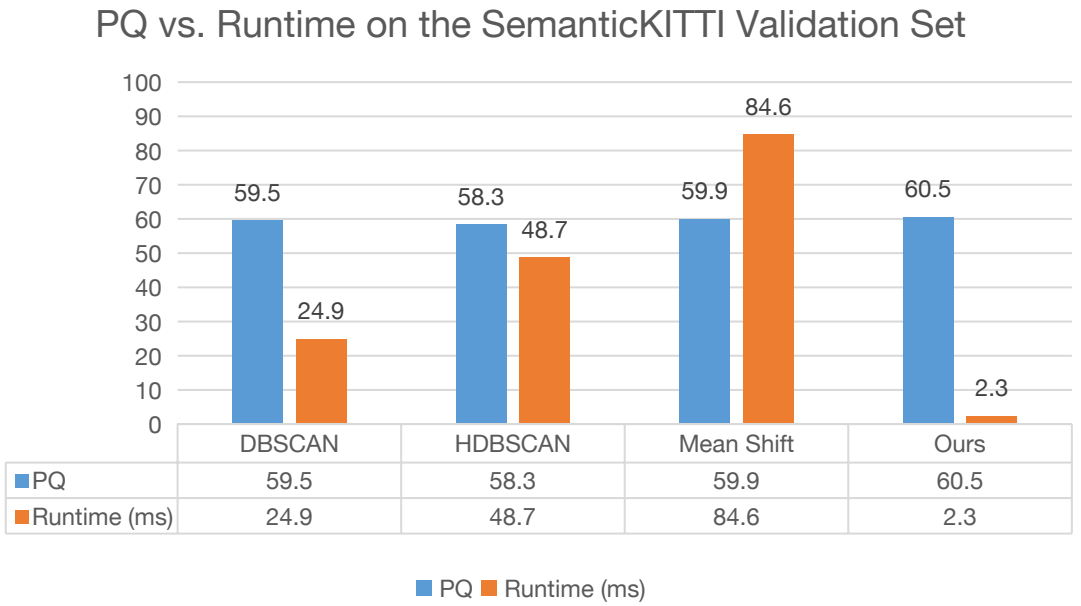


(b) Final Predictions

Experiment—Ablation Studies

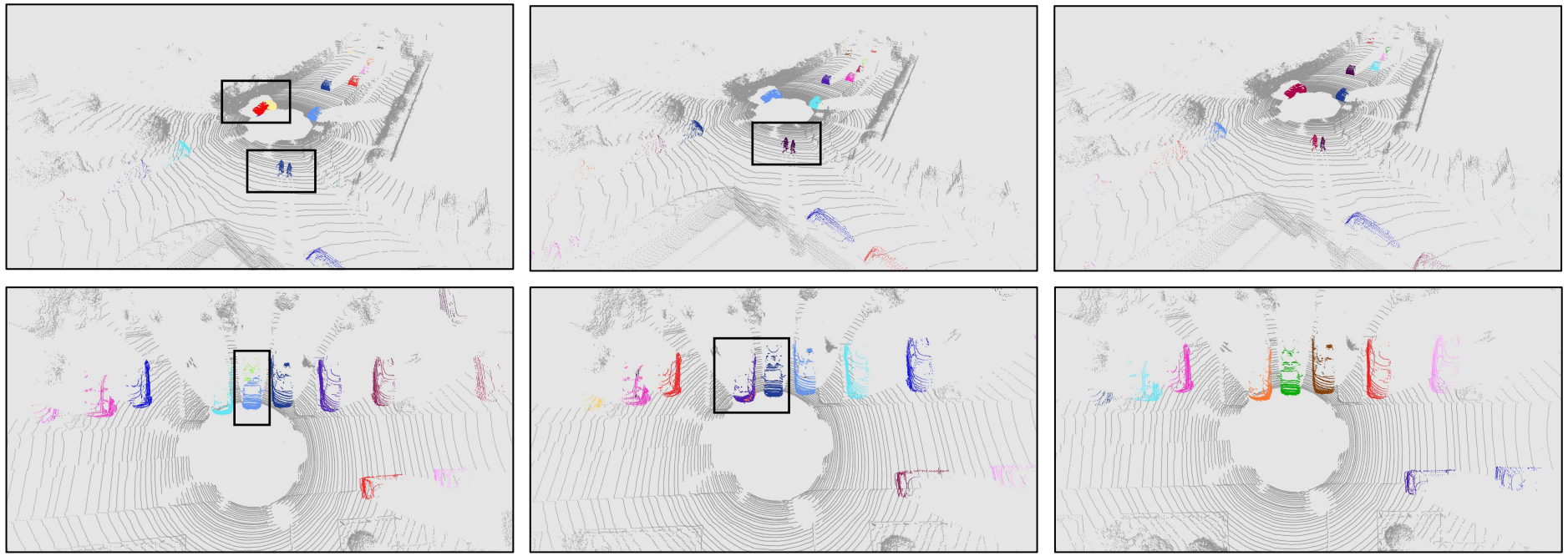
4. Effects of the CDM

- The CDM achieves the **best performance (+0.6 PQ~+2.2 PQ)**
- The CDM **significantly accelerates the inference speed**, since it is implemented efficiently in CUDA **without iterative processing**



Experiment—Ablation Studies

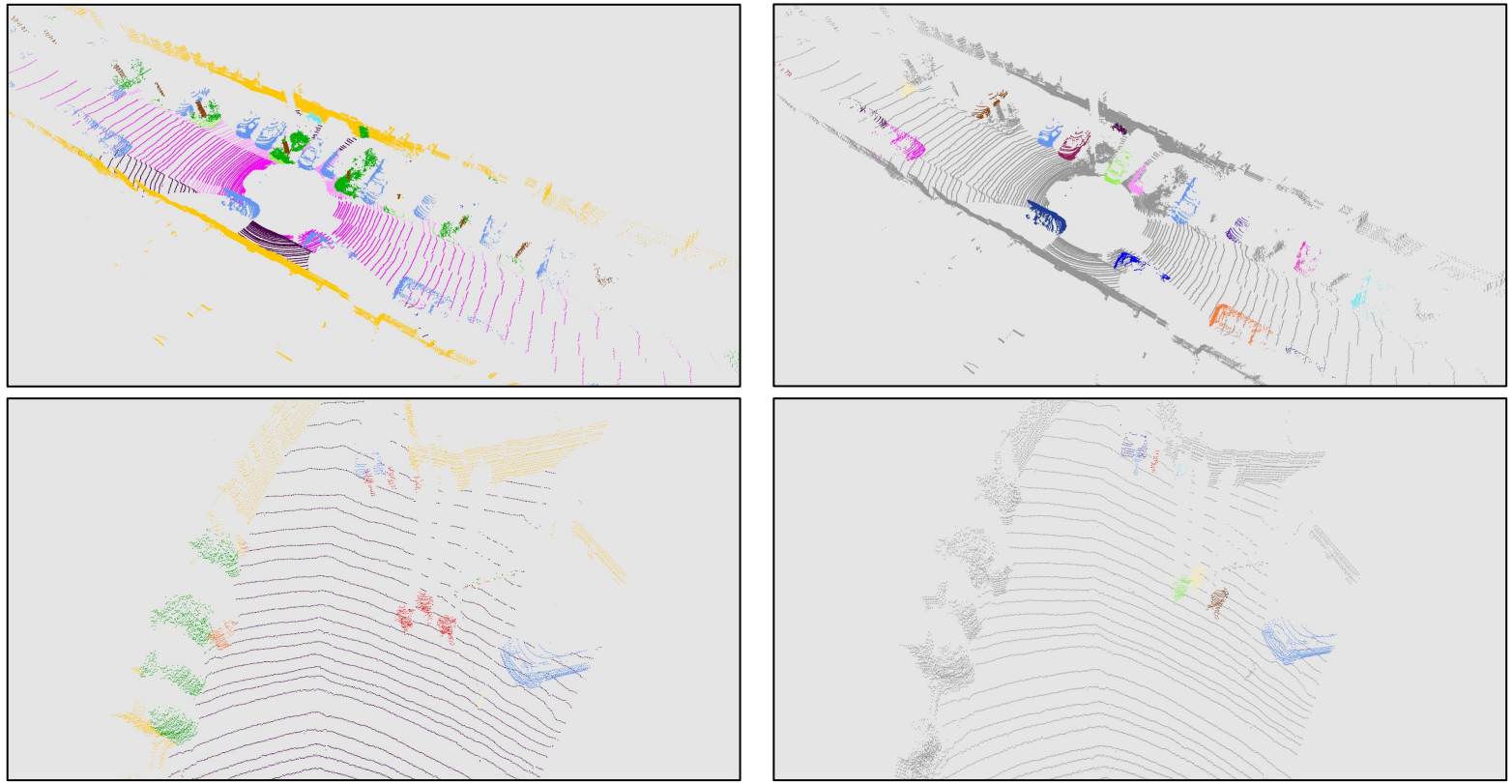
5. Comparison visualizations of instance segmentation from the Panoptic-PolarNet, DS-Net, and our CFNet



Panoptic-PolarNet DS-Net CFNet (Ours)

Experiment—Ablation Studies

6. Visualization results of our CFNet on the SemanticKITTI test set



semantic segmentation

instance segmentation

Thanks!