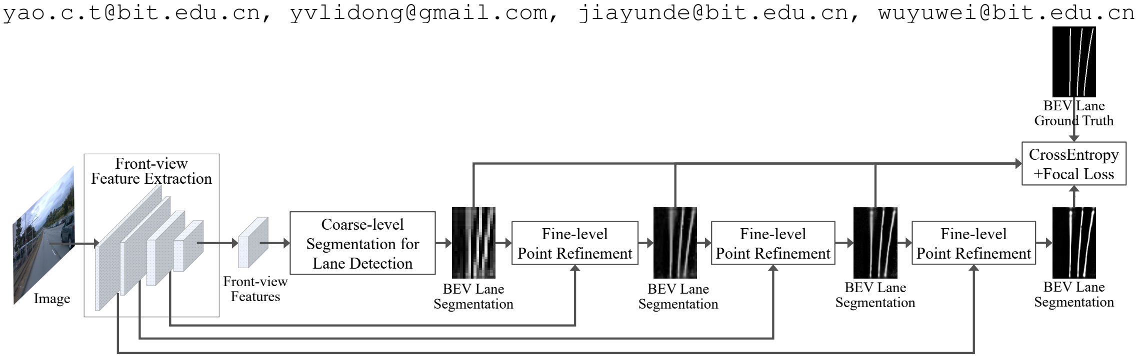
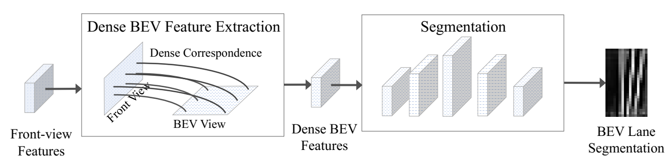
Sparse Point Guided 3D Lane Detection

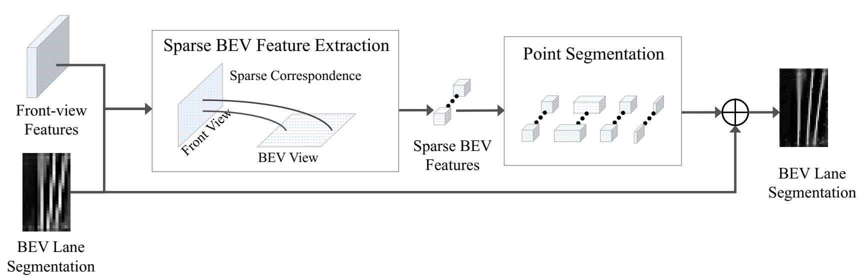
Chengtang Yao[1,3], Lidong Yu[3], Yuwei Wu[1]*, Yunde Jia[2]
[1]Beijing Key Laboratory of Intelligent Information Technology,
School of Computer Science & Technology, Beijing Institute of Technology, China
[2]Guangdong Laboratory of Machine Perception and Intelligent Computing,
Shenzhen MSU-BIT University, China
[3]Autonomous Driving Algorithm, DeepRoute
yao.c.t@bit.edu.cn, yvlidong@gmail.com, jiayunde@bit.edu.cn, wuyuwei@bit.edu.cn



(a) The overall pipeline of our model.



(b) The coarse-level segmentation for lane detection.



(c) The fine-level point refinement. $\oplus$ represents superposition.

Figure 1: The overall structure realization of our sparse-point guided lane detection for the segmentation-based approach is shown in (a). We decompose the 3D lane detection into the coarse-level segmentation for lane detection shown in (b) and the fine-level point refinement as shown in (c).

1. Realization for Segmentation-based Approach

In the supplementary material, we present a realization of our idea in the segmentation-based approach. As shown in Figure 1, the model takes an image as input to extract multi-scale front-view features. The lowest-resolution features are fed into a coarse-level segmentation to extract dense BEV features at the lowest resolution. We use the dense BEV features to predict a coarse BEV lane segmentation through an encoder-decoder-like network. We then use

*Corresponding author.

a series of fine-level point refinement to gradually improve the BEV lane segmentation from the previous scale. At each refinement, we fetch sparse features at high resolution for each point around lanes. The sparse BEV features are more informative and discriminative, providing better segmentation results for each point around lanes. The sparse segmentation is subsequentially superimposed on top of the dense segmentation from the previous scale as a dense result at the current scale.

2. Experiments

2.1. Influence of Window Size

We explore the influence of window size by changing the sampling number along the x-/z-dimension. As shown in Table 1, a smaller size along z-dim is better, while the best size along x-dim is around 3 and 4.

size	F1	size	F1	size	F1	size	F1
(3,3)	53.65	(4,3)	53.66	(5,3)	53.58	(5,5)	53.37
(3,4)	53.64	(4,4)	53.5	(3,5)	53.55		

Table 1: Ablation study on the window size (s_x, s_z) of candidate point sampling. s_x/s_z is the size along x-dim/z-dim.

2.2. Apollo 3D Lane

As PersFormer [1] did not open source the training code on Apollo, we reproduced the model and only got F1 89.6 on balanced scenes, while based on this model, our method got 90.7. This promotion is close to the results on ONCE-3DLanes [2], where the benchmark is too simple to demonstrate our priority.

2.3. Segmentation-based Approach

We analyze the performance of our method for segmentation-based approaches on the OpneLane dataset [1]. We take IOU as the evaluation metric and the BEV segmentation from the official code as the ground truth. We mainly compare our method with the segmentation branch of PersFormer by abandoning the other 3D lane branches. The IOU of PersFormer is 64.3, while our method achieves 66.1. The results show that our method achieves comparable results while reducing the memory cost by 80% and speeding up 2 times.

References

[1] Li Fei Chen, Chonghao Sima, Yang Li, Zehan Zheng, Jiajie Xu, Xiangwei Geng, Hongyang Li, Conghui He, Jianping Shi, Yu Qiao, and Junchi Yan. Persformer: 3d lane detection via perspective transformer and the openlane benchmark. In *Proceedings of the European Conference on Computer Vision (ECCV)*, 2022.

[2] Ruijin Liu, Dapeng Chen, Tie Liu, Zhiliang Xiong, and Zejian Yuan. Learning to predict 3d lane shape and camera pose from a single image via geometry constraints. In *Proceedings of the AAAI Conference on Artificial Intelligence (AAAI)*, pages 1765–1772, 2022.