

# DeLS-3D: Deep Localization and Segmentation with a 3D Semantic Map – Supplementary Materials

- Video of labeled Zpark dataset, *i.e.* images, labeled 3D semantic maps, camera pose inside and semantic labels of each frame.
- Video of localization and parsing results over the Zpark and Drake dataset, *i.e.* rendered label map with simulated noisy pose, rectified pose after pose RNN and parsing results with the approach.

Due to the limitation of upload size, we can not upload full-size videos. Thus, we uploaded low resolution videos for reviewers to briefly check the results. Additionally, we upload the original videos of dataset and results to YouTube for better visualization. Here, we illustrate a snapshot for each of the videos to clarify their meaning.

## 1. Dataset

Youtube link <https://youtu.be/M6lhkzKFEhA>

Baidu Cloud <https://pan.baidu.com/s/1jHC1cmq>

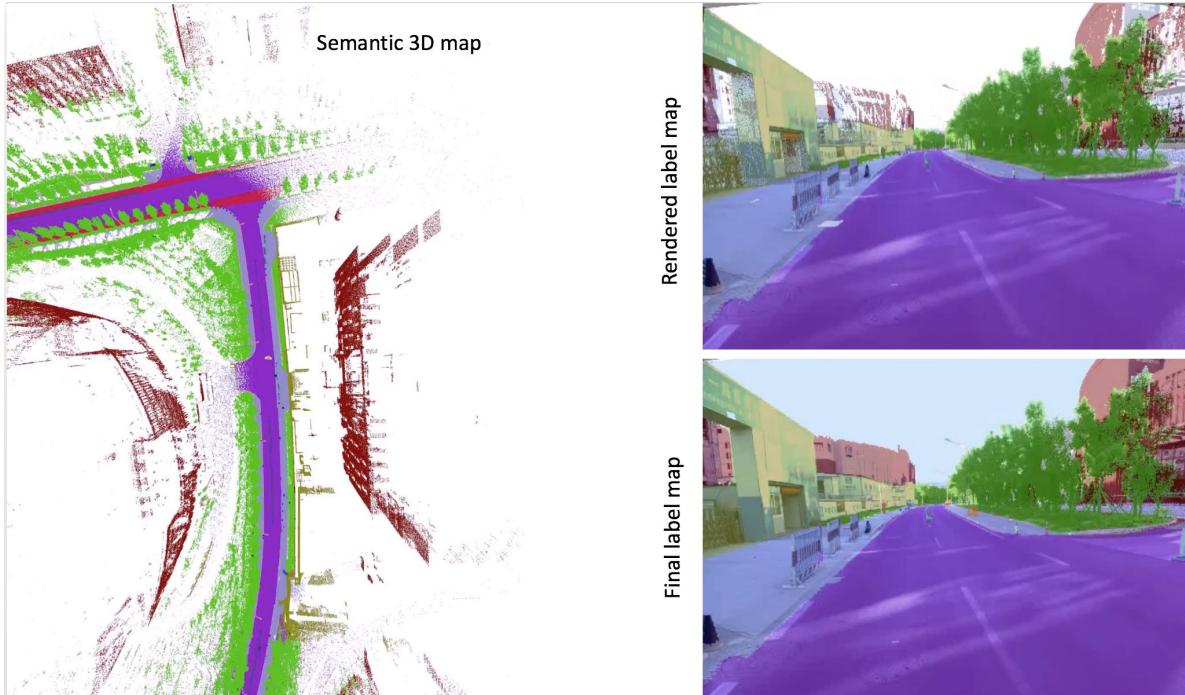


Figure 1: Zpark dataset overview. Left pane: the bird's eye view of our 3D map. Right top: the rendered label map from 3D with ground truth pose. Right bottom: the manually labeled map. Check video for the full content.

## 2. Results

Here we upload part of the results video online. Our full results will be released jointly with the datasets. Left top: the rendered label map with noisy poses. Left bottom: the rendered label map from rectified poses. Right top: the parsing results.

Right bottom: ground truth labeled map. Check the video online for the full content.

**Results for Zpark dataset.** Youtube link <https://youtu.be/HNPQVtgpjbE>

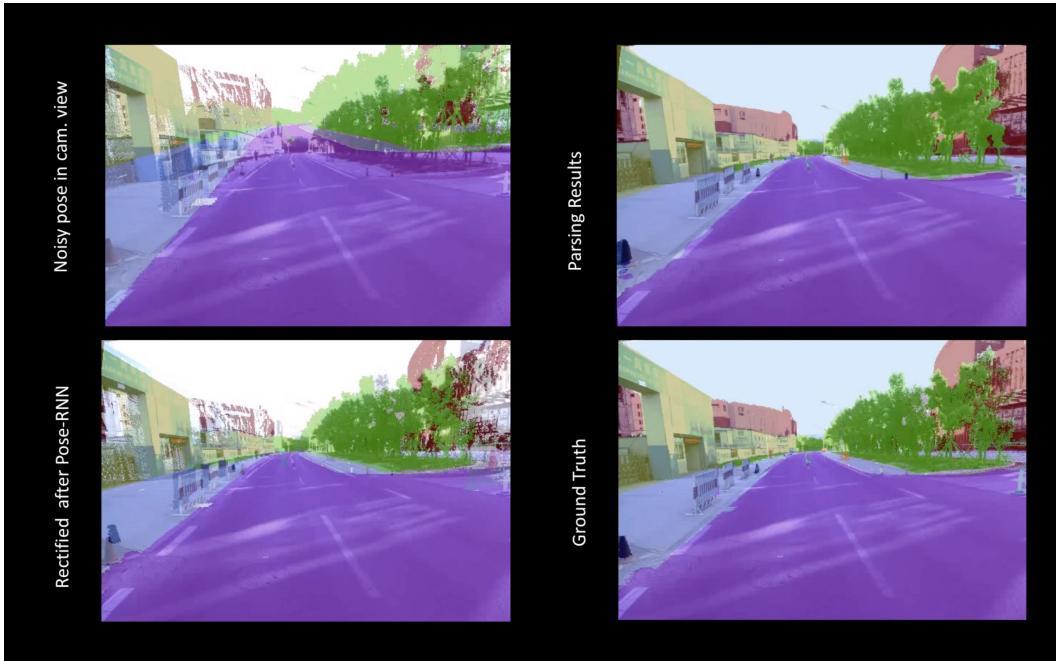


Figure 2: The results of a driving video in Zpark dataset.

**Results for Drake dataset.** Youtube link <https://youtu.be/ApyqPnvmJAs>

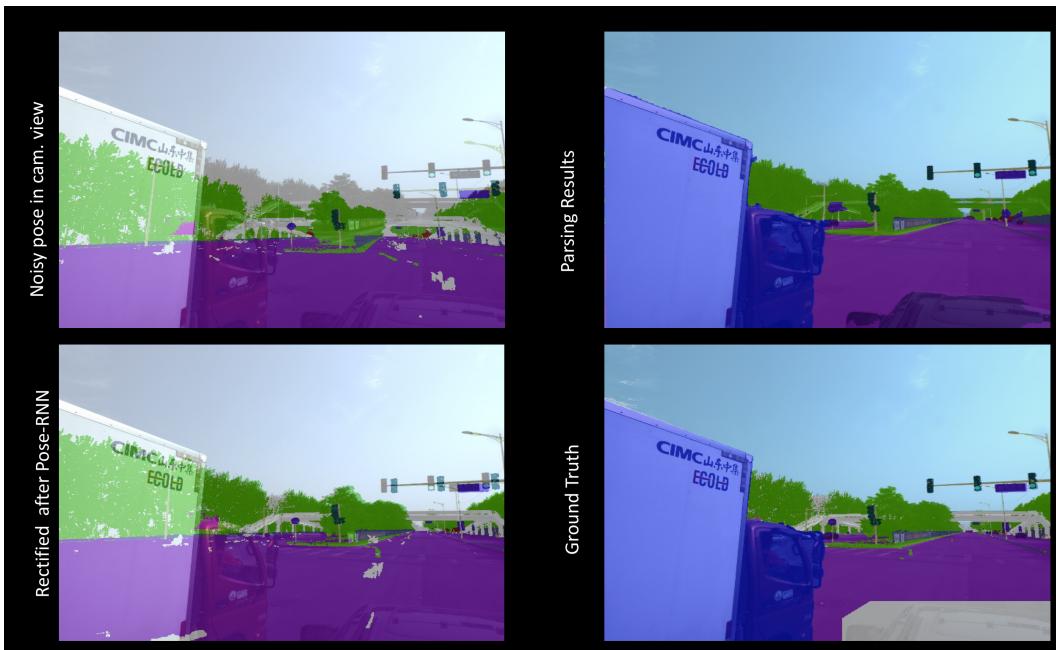


Figure 3: The results of a driving video in Drake dataset.