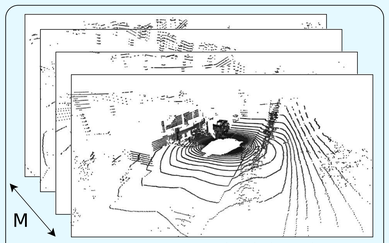
Inference datatset (Target)



3D point clouds (scans)

+

Sensor poses $\{(R_i, T_i)\}_{1 \le i \le M}$

Align

3D Scene

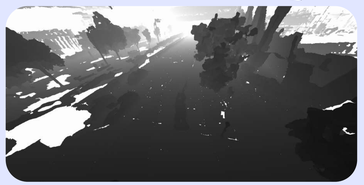
Modality Processing

Pick Your Modalities :

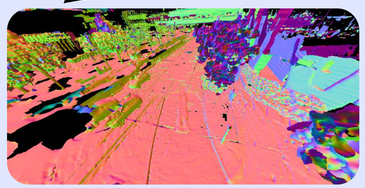
Pointcloud

Mesh

Change colors ?



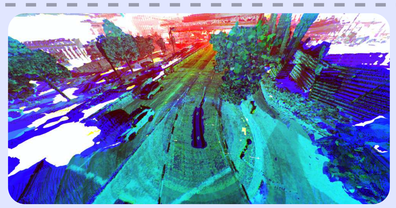
Depths



Normals

Exemple Features

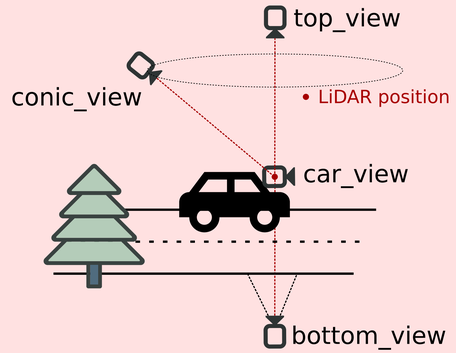
- R: Pcd Intensities
- G: Mesh Normals_X
- B: Mesh Depths



Sample features

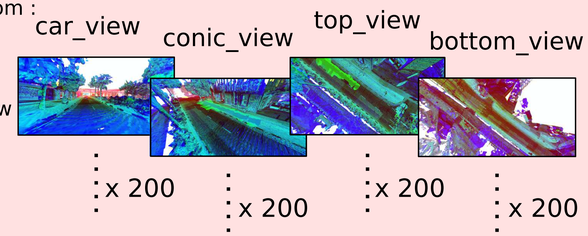
Multi-View Image Generation

Pick your visualization views :



Exemple Features

car_conic_top_bottom :
- 200x car_view
- 200x conic_view
- 200x top_view
- 200x bottom_view



2D-3D projection and votes

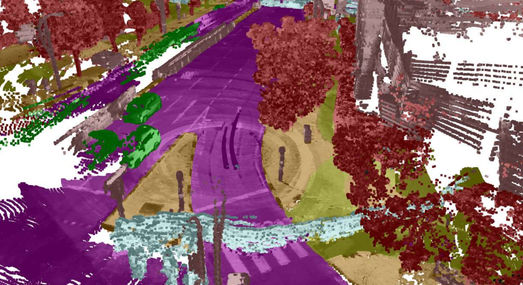
Pick Your projection and voting system :

Accumulate

Segmentation Masks

Logits

- Use occlusion ?



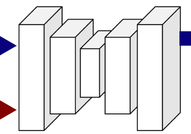
Inference Pipeline

Training Pipeline

MODO

VGO

Train



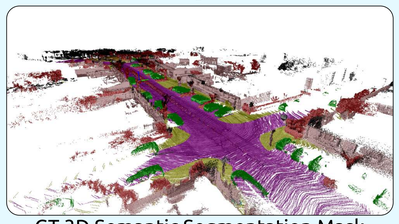
2D semantic segmentation

2D-3D Proj

Training dataset (Source)

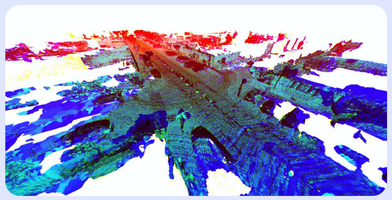
3D Scene

+



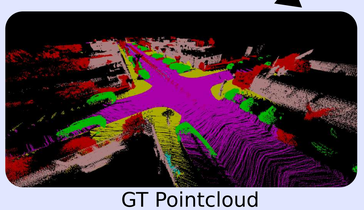
GT 3D Semantic Segmentation Mask

Re-use same modalities for training :

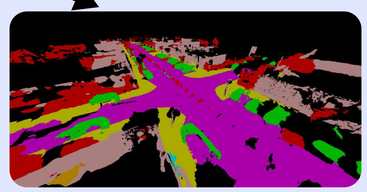


Sample features

Pick Ground Truth modality :

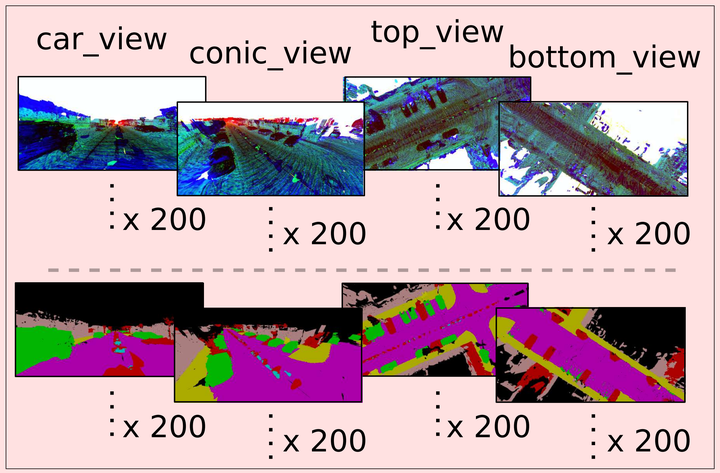


GT Pointcloud



GT Mesh

Re-use the same views for training :



Predicted 3D Semantic Segmentation Mask