Semantic 3D Object Detection

Alp Emek and Marcelo Chulek

Overview

- 3D Object detection is necessary for Autonomous Driving
- Goal is detection of Cars, Pedestrians and Bicyclists
- Predicts 2D and 3D bounding boxes around objects
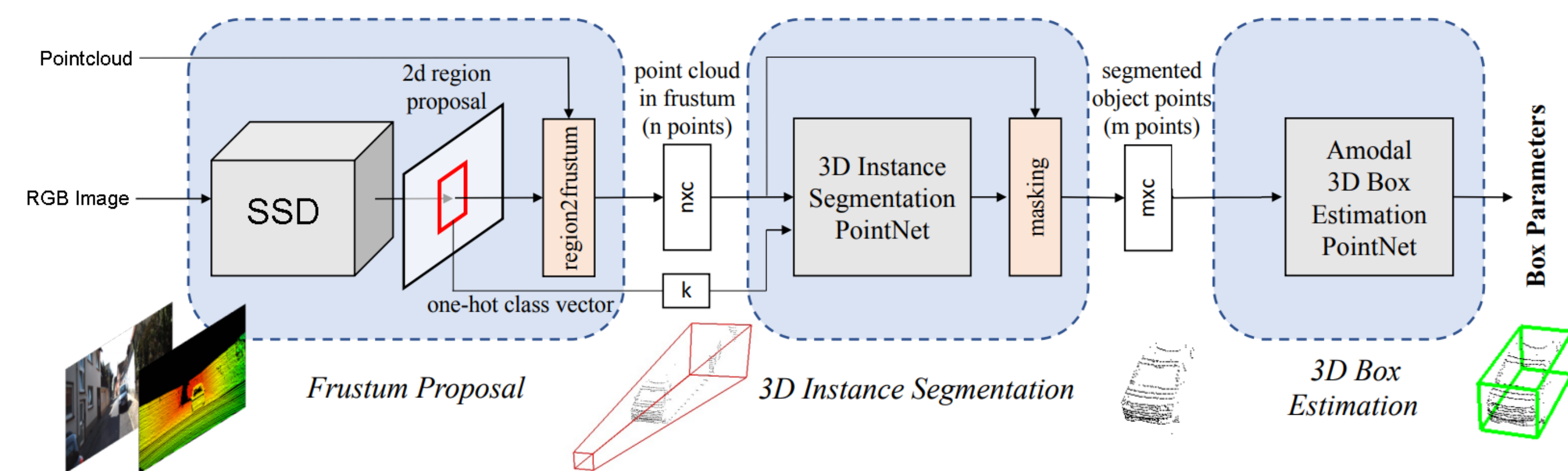- Training and Validation performed on KITTI dataset [1]

Figure 1: Architecture used for 3D object detection [2]

Prediction of 2D Bounding Boxes
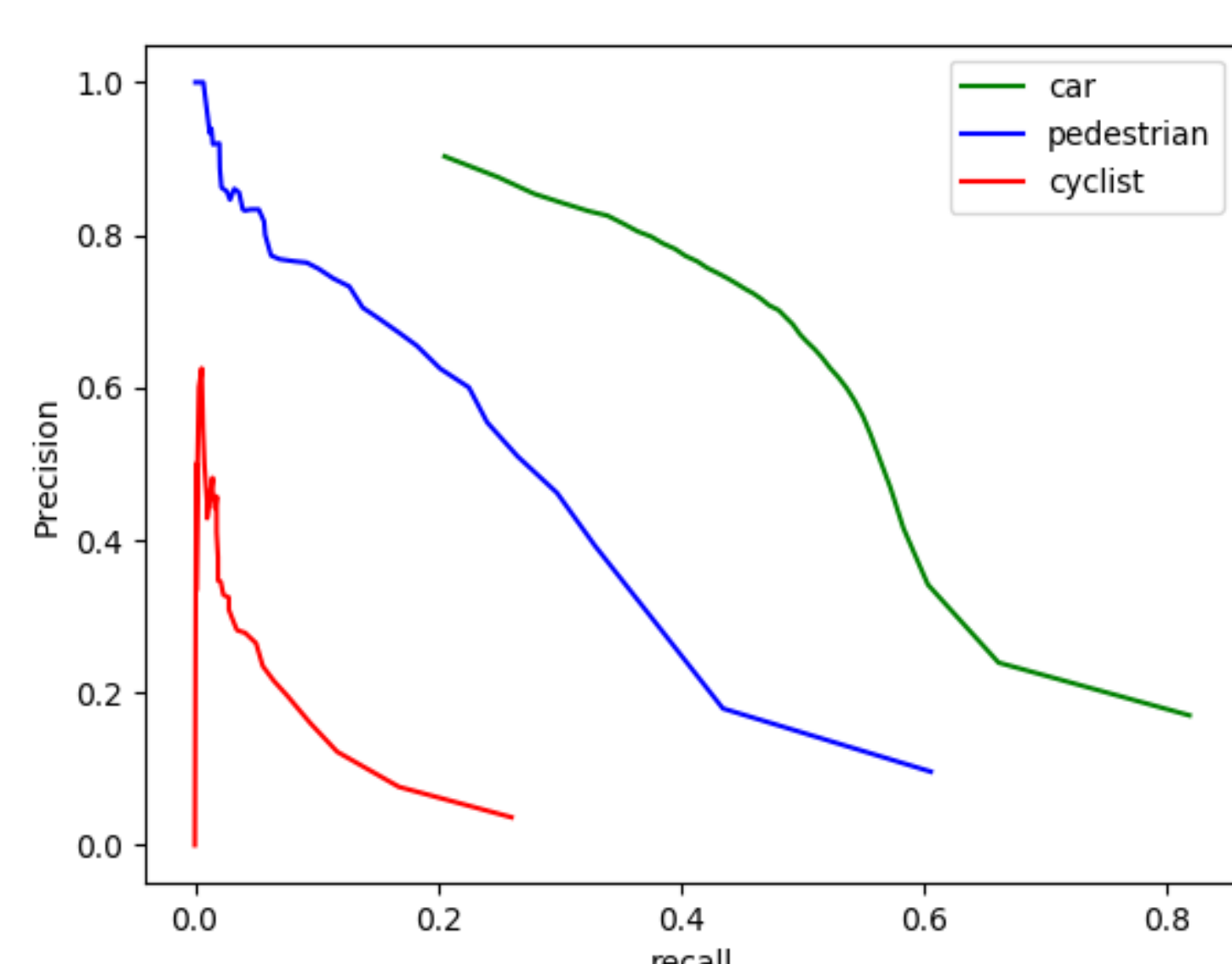
- Single Shot Detection
- Works in real-time

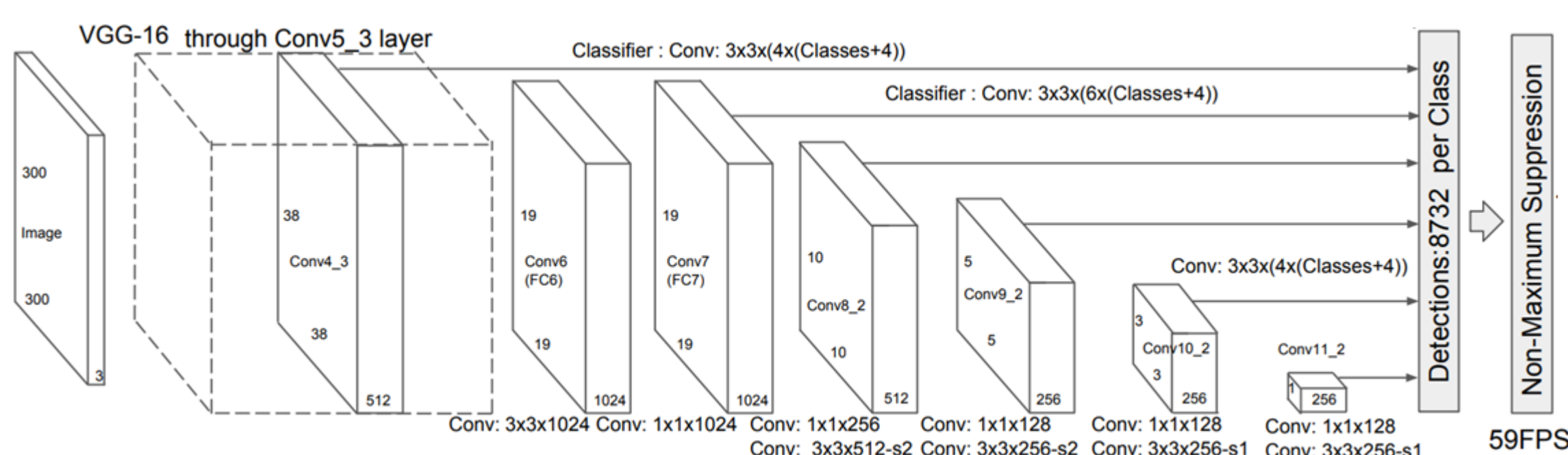**Figure 2: Precision - Recall Curve**

Figure 3: SSD Architecture [3]

Frustum Proposal

- Input is Raw LIDAR scan
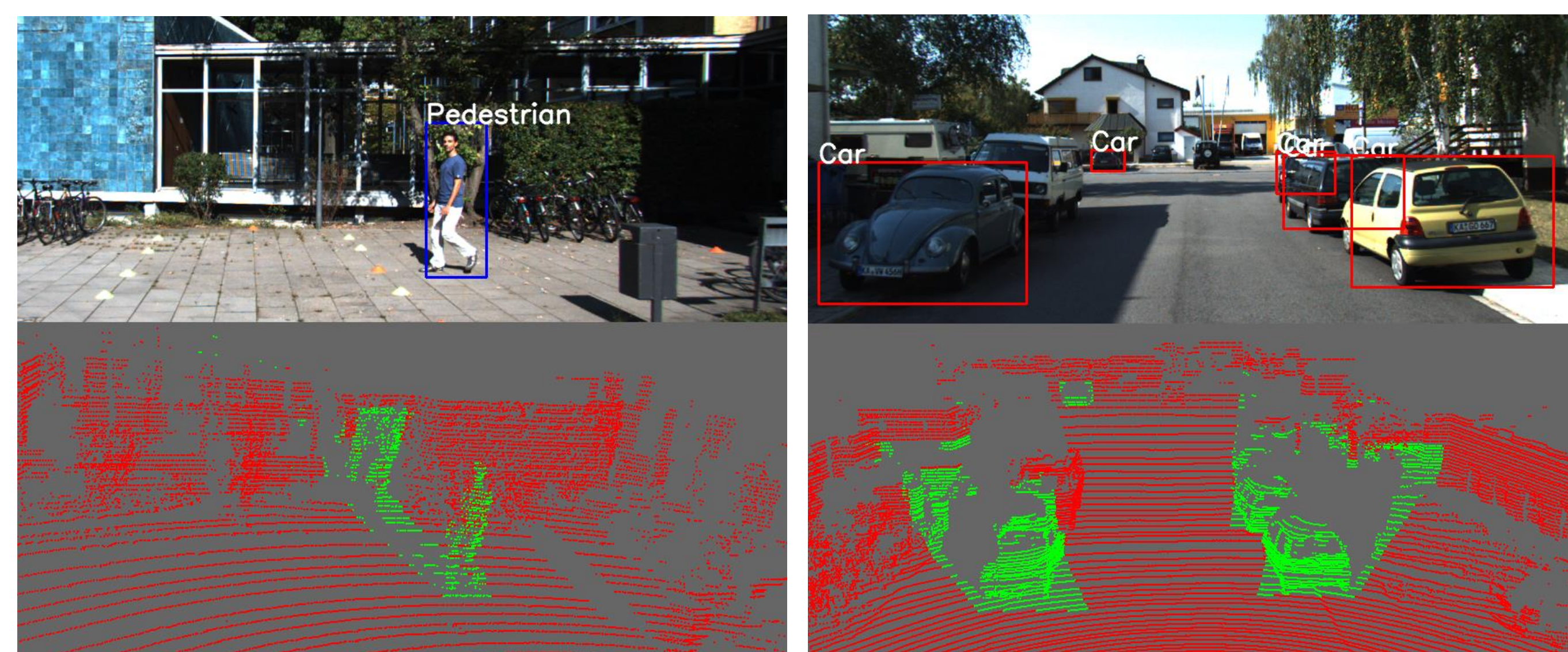- Creates a Frustum based on the 2D Prediction

Figure 4: Frustum Proposals

3D Segmentation Network

- We used PointNet Architecture
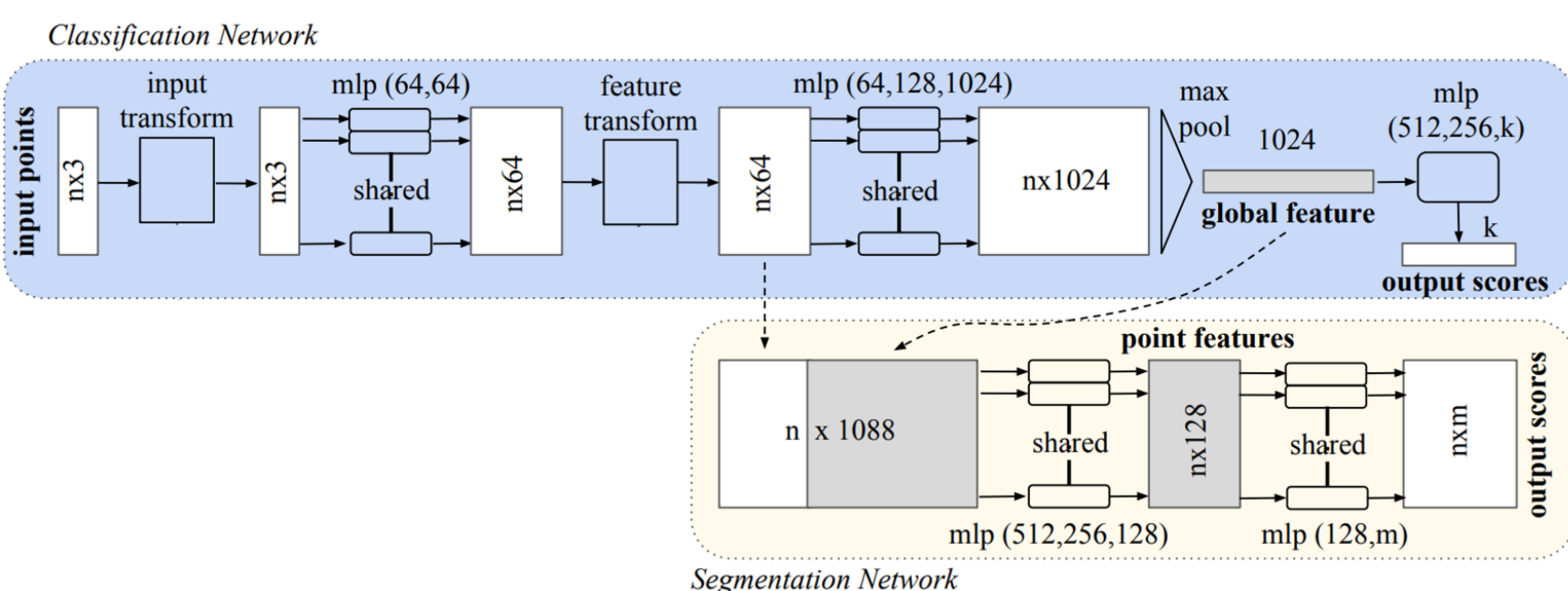- Trained with labeled points inside the Frustums

Figure 5: PointNet Architecture [4]

Figure 6: 3D Segmentation

3D Bounding Box Regression Network

- Trained with the segmented Point Cloud (x, y, z)
- Predicts the bounding box parameters (height, width, length, rotation, x_{center}, y_{center}, z_{center})
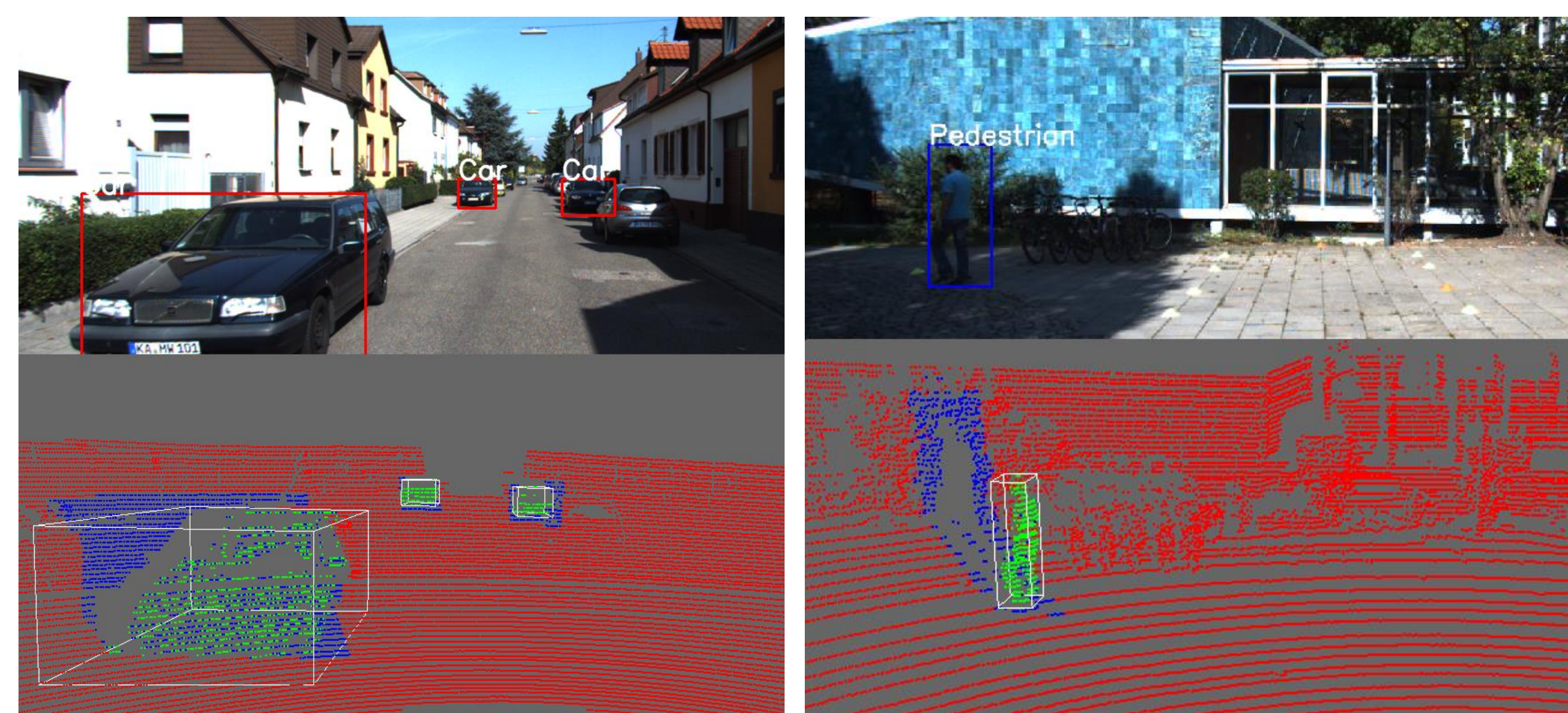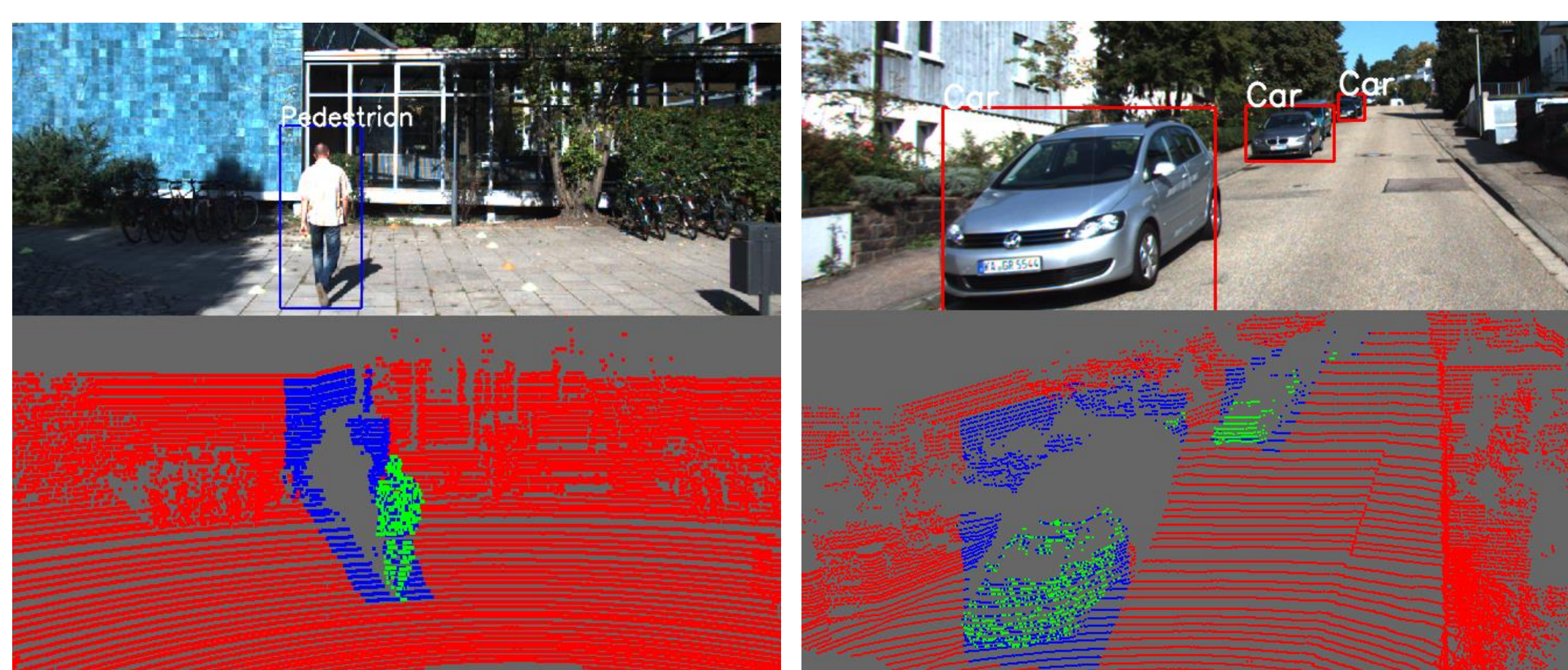- Tests using Min-Max of segmented Point Cloud

Figure 7: 3D Bounding Boxes

References

1 – The KITTI Vision Benchmark Suite - http://www.cvlibs.net/datasets/kitti/

2 - QI C. et al. "Frustum PointNets for 3D Object Detection from RGB-D Data"

3 - LIU W. et al. "SSD: Single Shot MultiBox Detector"

4 - QI C. et al. "PointNet: Deep Learning on Point Sets for 3D Classification and Segmentation"