



High-level camera-LiDAR fusion for 3D object detection with Machine Learning

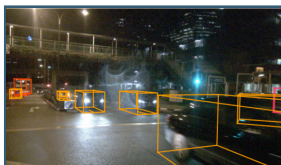
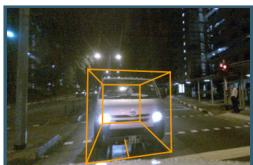
Gustavo A. Salazar-Gomez, Miguel A. Saavedra-Ruiz & Víctor Romero-Cano

Universidad Autónoma de Occidente, Colombia



3D object detection

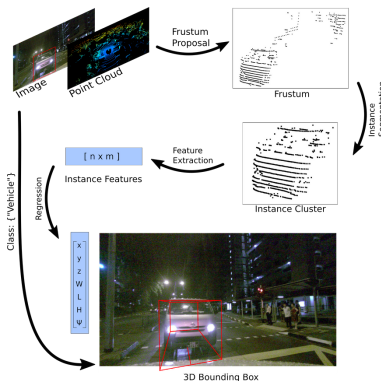
Allows autonomous agents to estimate the relative pose of multiple objects neighbouring an ego-vehicle.



The astonishing results presented by Deep Learning models to address this issue, are usually occluded by the vast need of computational resources required to deploy them.

Our proposed solution

- A frustum region proposal is assembled, taking advantage of 2D object detectors.
- The point cloud instance inside the frustum is segmented using DBSCAN.
- A global feature representation encoding the given segmented instance.
- A Support Vector Regressor (SVR) estimates the 3D bounding box parameters.

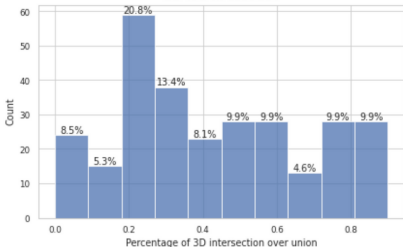


Why is our work relevant

- Model deployment under conditions of low computational resources.
- Acceptable accuracy for low complexity tasks.
- Promising results in terms of accuracy based on processing times for training and inference working on CPU only.

Results

3D IoU in test set



Set	x	y	z	ψ	w	l	h	Avg.	Avg. 3D	Avg. BEV
Training	98.8	98.2	99.9	78.0	96.8	94.2	99.8	95.1	62.0	68.4
Test	96.6	97.8	95.4	55.7	80.7	88.4	95.0	87.1	42.7	47.8

Table 1: Evaluation metrics of the proposed method.

Module	Instance Segmentation	Feature Extraction	Regression	Total
Training	11.1s \pm 2.3s	19.1s \pm 0.5s	2.6s \pm 8.3ms	32.8s \pm 2.35s
Inference	4.7ms \pm 4.8ms	13ms \pm 1.4ms	0.7ms \pm 1.2ms	18.4ms \pm 5.1ms

Table 2: Processing times through training and inference stages per module.

- Accurate estimation of centroid coordinates x, y, z and the boundingbox dimensions w, l, h.
- Training of the whole system with 1136 images and LiDAR sweeps in roughly 32.8s.
- Process a new data sample in approximately 18.4ms or 55FPS in a CPU-only setup.

Conclusions and future work

- Our framework is capable of predicting 3D bounding boxes for vehicles and shows promising results estimating its parameters using classic ML techniques.
- With improvements as MultiBin for heading estimation, the evaluation metrics could be considerable boosted toward better results.

