providentia ++

**Master's Thesis Initial Topic Presentation**

# Monocular 3D Traffic Perception Using HD Maps as an Auxiliary Feature

**Joseph Birkner**

Chair of Robotics, Artificial Intelligence and Real-Time Systems

Technical University of Munich (TUM)

I6 Robotics Lab

Schleißheimer Straße 90A

85748 Garching-Hochbrück

June 23, 2022

# About me

- B.Sc. Natural Language Processing

- Studying M.Sc. Informatics (50%) since 2018

  Also did lots of NLP for my IDP and practical courses.

- Senior Developer for Navigation Data Standard

  Work on 3D HD map rendering and modeling.

- I would like to bring HD Maps and Differentiable

  Rendering into my Master's thesis ☺



Figure 1: Joseph in his natural habitat.

# Monocular 3D Detection in Providentia++



Figure 2: Live views of RGB cameras from the B471:S110 intersection
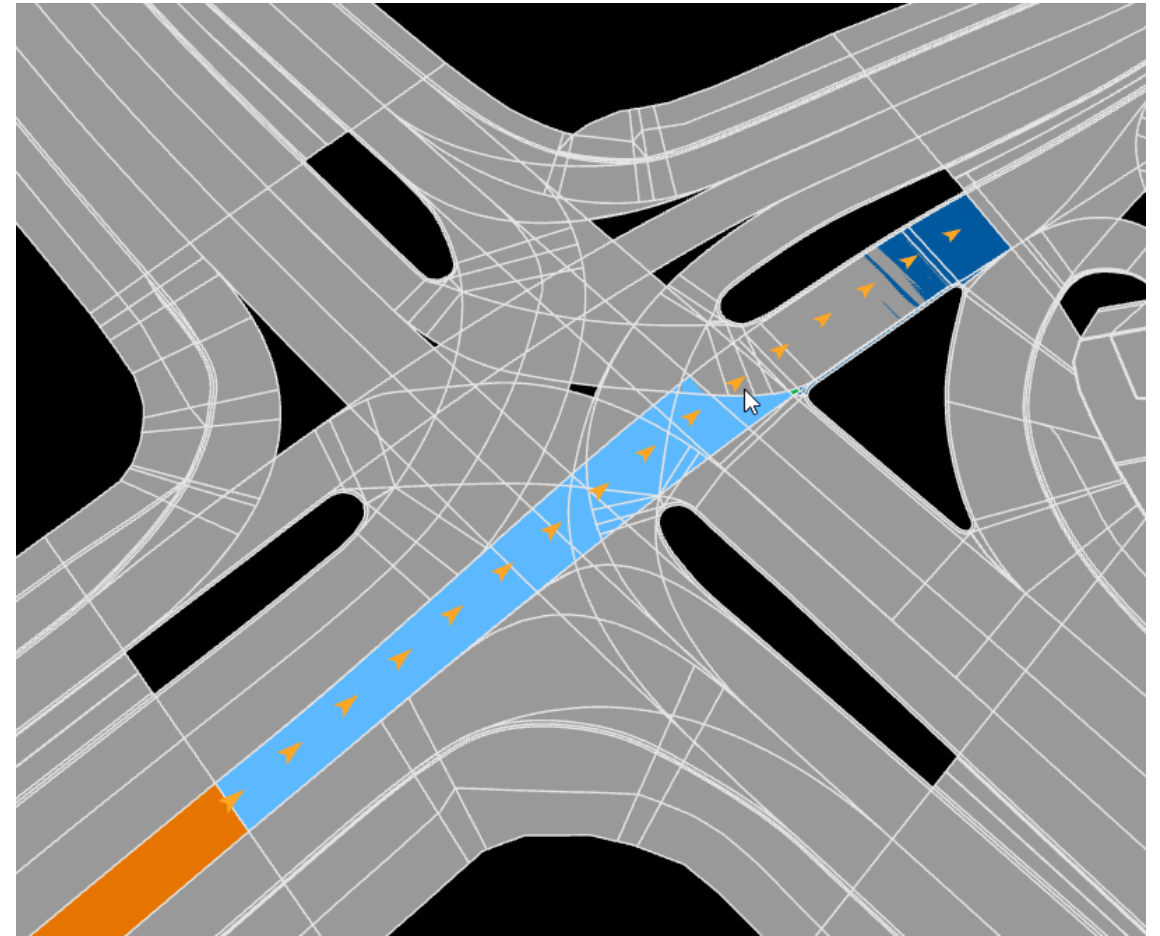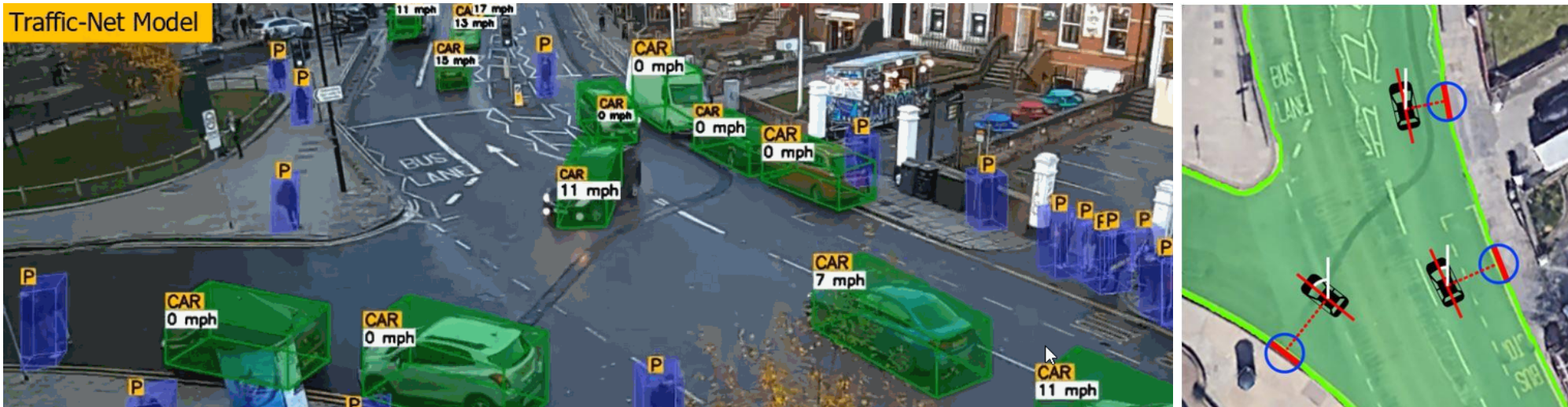


Figure 3: OpenDRIVE map of the same intersection

# Maps as an Auxiliary Feature (1)

- **Traffic-Net**: 3D Traffic Monitoring Using a Single Camera (Sep 21)



Parts of Figure 1 and 3 in the Traffic-Net Paper

- Strong prior knowledge on object width/height used to estimate 3D BBs

- Effective use of Kalman filters for trajectory/heading estimation

# Maps as an Auxiliary Feature (2)

- **UrbanNet**: Leveraging Urban Maps for Long Range 3D Object Detection (Oct 21)
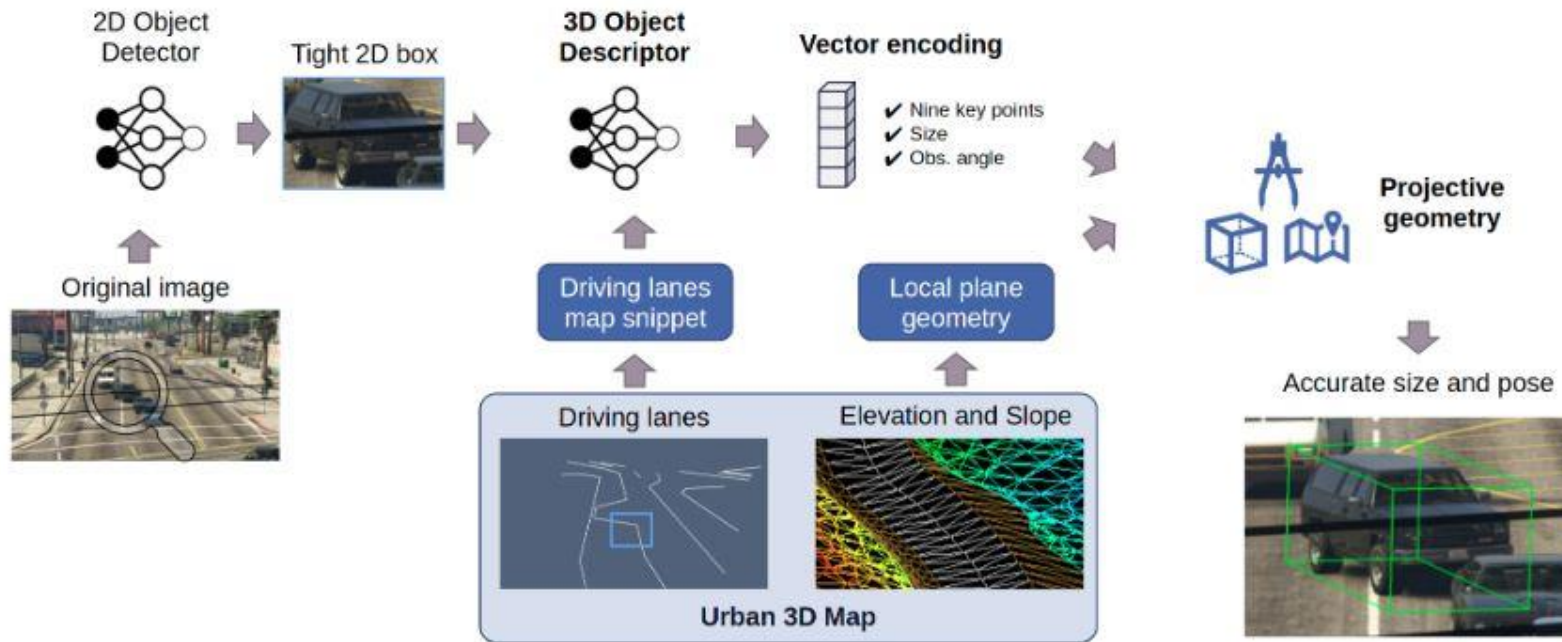


Figure 2 from UrbanNet

- Bounding Box estimation network supervised using data from GTA V ☺

# Neural Rendering

Self-supervision by differentiable rendering of the NNs proposed object labels and comparing with (some channel of) the input image.
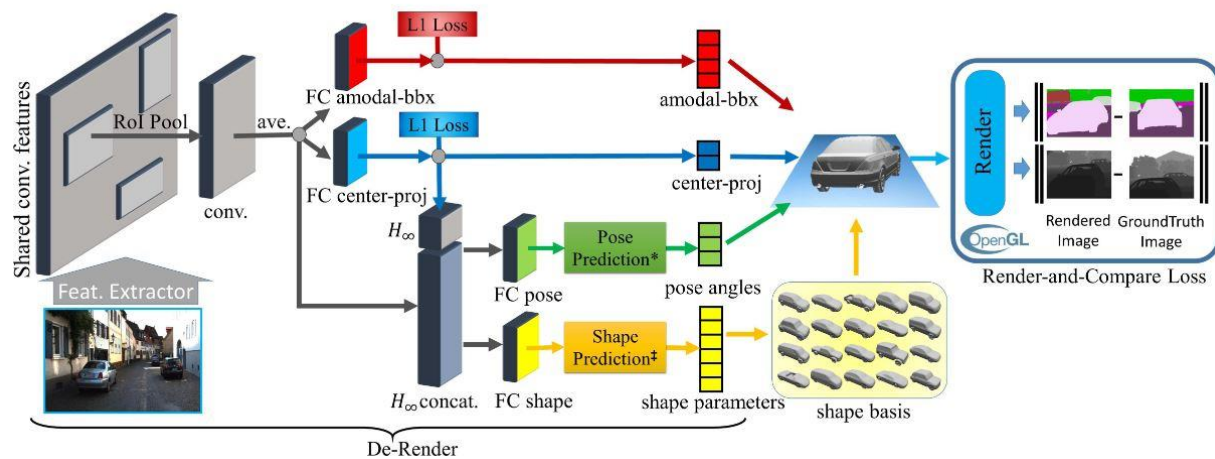


Figure 1 from 3D-RCNN

- **3D-RCNN**: Instance-level 3D Object Reconstruction via Render-and-Compare (CVPR 2018)

- Could also predict heading and velocity change on top of x/y/theta/w/h/d/category

- **Could pixel-wise lane orientation annotations help?**

# References (so far)

Rezaei, Mahdi, Mohsen Azarmi, and Farzam Mohammad Pour Mir. "Traffic-Net: 3D Traffic Monitoring Using a Single Camera." *arXiv preprint arXiv:2109.09165* (2021).

Mouawad, Issa, et al. "Time-to-Label: Temporal Consistency for Self-Supervised Monocular 3D Object Detection." *arXiv preprint arXiv:2203.02193* (2022).

Marinello, Nicola, Marc Proesmans, and Luc Van Gool. "TripletTrack: 3D Object Tracking Using Triplet Embeddings and LSTM." *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*. 2022.

Heylen, Jonas, et al. "MonoCInIS: Camera Independent Monocular 3D Object Detection using Instance Segmentation." *Proceedings of the IEEE/CVF International Conference on Computer Vision*. 2021.

# References (so far)

Carrillo, Juan, and Steven Waslander. "Urbannet: Leveraging urban maps for long range 3d object detection." *2021 IEEE International Intelligent Transportation Systems Conference (ITSC)*. IEEE, 2021.

Ye, Xiaoqing, et al. "Rope3D: The Roadside Perception Dataset for Autonomous Driving and Monocular 3D Object Detection Task." *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*. 2022.

Li, Peixuan, and Jieyu Jin. "Time3D: End-to-End Joint Monocular 3D Object Detection and Tracking for Autonomous Driving." *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*. 2022.