

# Computer Vision and Deep Learning for Autonomous Driving

Introductory meeting

Dr. Habil. Federico Tombari, Prof. Nassir Navab,  
Fabian Manhardt, Johanna Wald

# Seminar contents

- The seminar includes a selection of the most recent and relevant papers in the field of computer vision and deep learning aimed at autonomous driving
- Papers are selected to cover different aspects of the topic:
  - Semantic Segmentation
  - Scene Understanding
  - Multi-modal sensor fusion
  - SLAM and 3D reconstruction
  - Synthetic dataset creation
  - Car Simulators
  - Semi-supervised large-scale dataset annotations
  - Trajectory planning and lane change prediction

# Goals

- You are going to learn:
  - about relevant works in the field of Computer Vision and Deep Learning
  - what is Autonomous Driving and why it is a relevant technology for future applications
  - how to read and understand a scientific article
  - how to write a scientific report
  - how to give a talk to an audience, and deal with related questions afterwards

# Seminar Schedule

- 6 sessions, 1 every Thursday, 11am-12.30pm
- Two presentations per session
- Seminarraum 03.13.010
  
- Paper assignments:
  - selected students can express up to 3 preferences
  - We will then match them to a paper and tutor trying to maximize global happiness

# Presentation

- Each presentation is 20 minutes + 10 minutes for Q&A
- Slides templates (Powerpoints, Latex, ..) provided on website
- The presentation should cover all relevant aspects of the paper
  - Introduction and state of the art
  - Main contribution(s)
  - Experimental results
  - Discussion, summary and future work
- The presentation should be self-contained
- All students are expected to attend all presentations and interact during Q&A (this will influence your final mark)

# Report

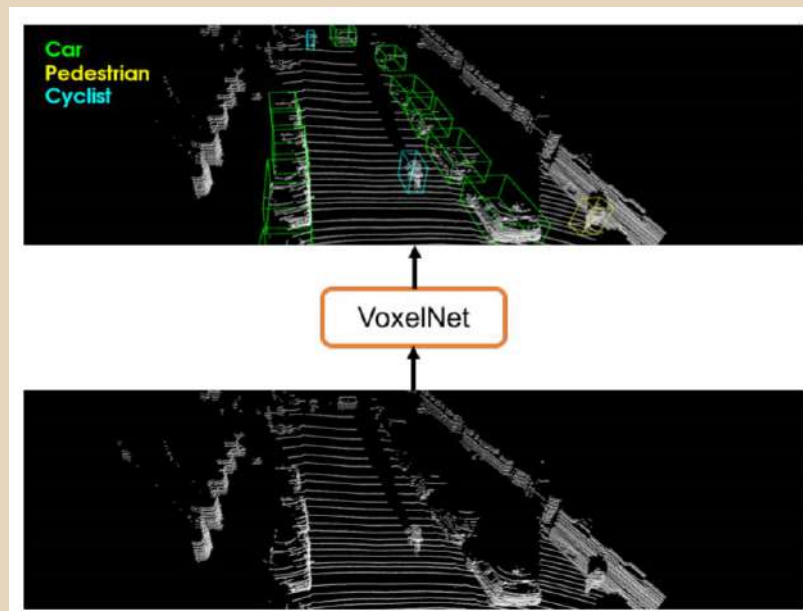
- The report should summarize the paper in the way it has been presented during the talk, and provide the student's opinion concerning the main contributions and impact
- Language: English
- Max 8 pages
- Template on course website
- Once ready, send the report to supervisor, within **two weeks** from the day of the presentation

# Evaluation criteria

- Quality of presentation (both regarding slides and speech)
- Quality of the report
- Comprehension of the scientific contents of the presented work
- Interaction and participation during the other talks

# Y. Zhou et al, “VoxelNet: End-to-End Learning for Point Cloud Based 3D Object Detection”, 2018

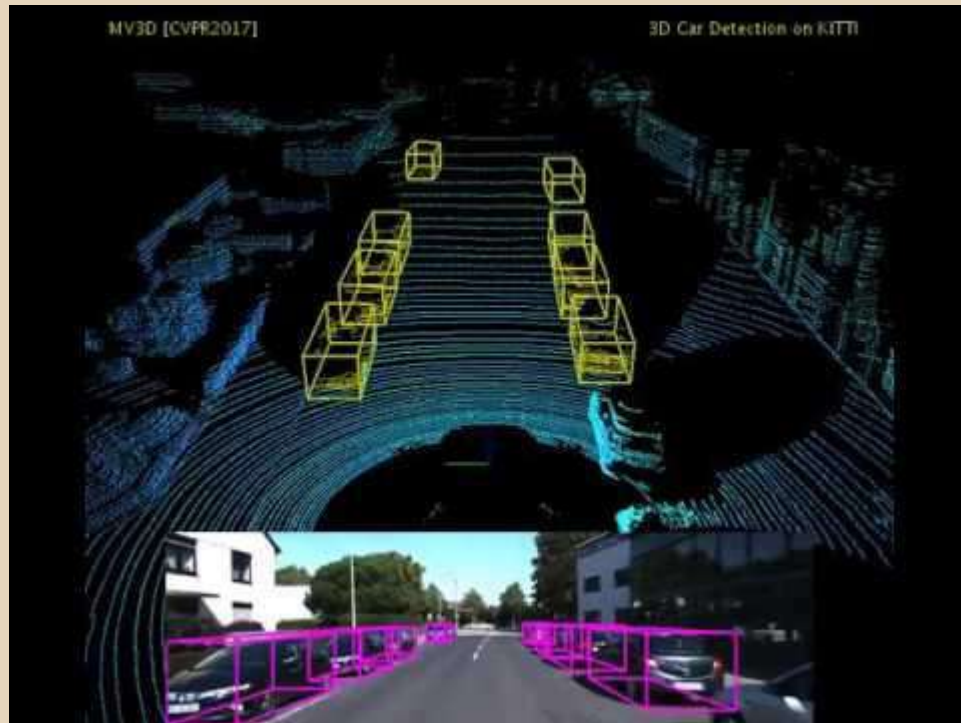
- Object Detection in point clouds from outdoor data
- Convolutional Neural Network used as classifier for 3 classes
- Designed to work with sparse LIDAR data





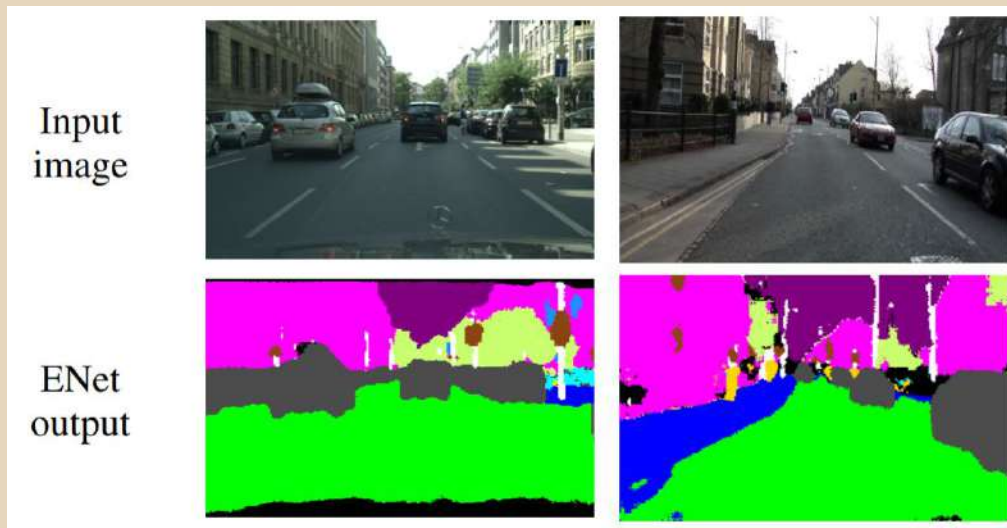
# X. Chen et al, “Multi-View 3D Object Detection Network for autonomous Driving”, CVPR 2017

- Sensor-fusion framework that takes both LIDAR point cloud and RGB to predict 3D bounding boxes
- Generates 3D Proposals from LIDAR data
- Projects proposal on RGB image and conducts data fusion
- Classification of proposal together with refinement for tighter fit



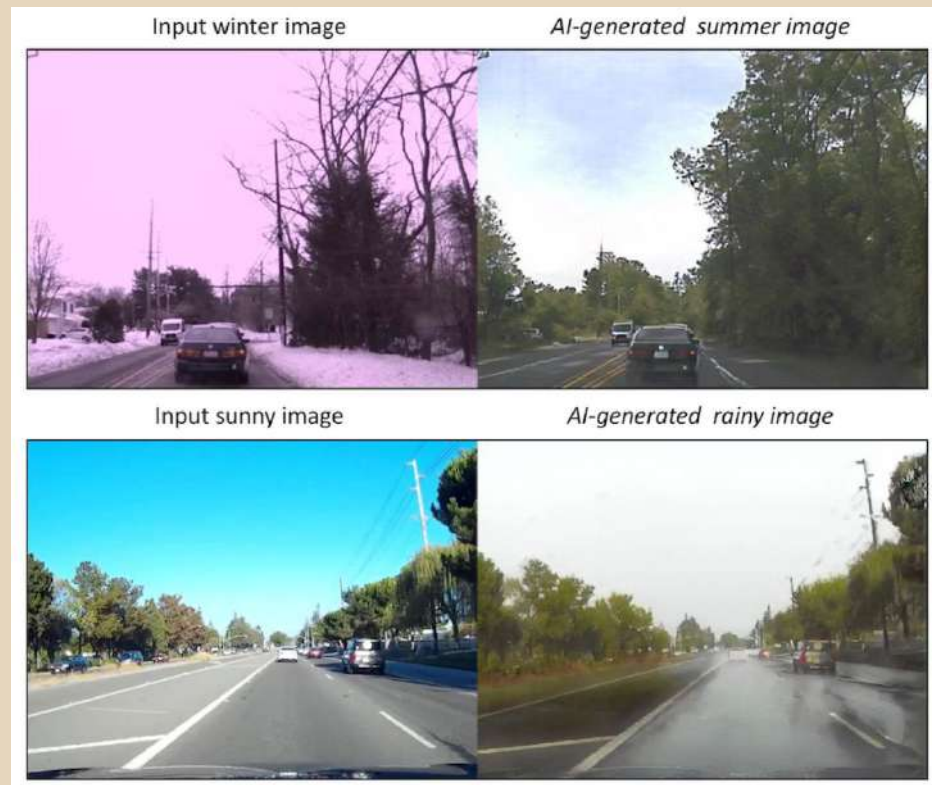
## A. Paszke et al., “ENet: A Deep Neural Network Architecture for Real-Time Semantic Segmentation”, 2016

- “Lightweight” Deep Learning for embedded/mobile applications
- Goal: reduce memory footprint and increase feed-forward time while keeping a relatively good segmentation accuracy



## M. Liu et al., “Unsupervised Image-to-Image Translation Networks”, 2017

- Use of Generative Adversarial Networks (GANs) to generate novel illumination/weather conditions and image modalities
- Works in an unsupervised manner
- Applied to street scenes for autonomous driving applications



## A. Dosovitskiy et al, “CARLA: An Open Urban Driving Simulator”, 2017

- Simulator for autonomous driving research
- Flexible: different sensing modalities, different weather and traffic conditions
- Completely open source

