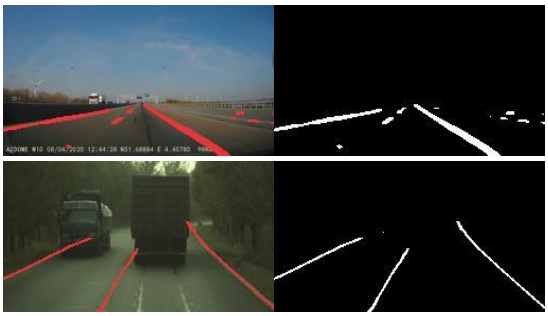
MSc. Thesis Project

Robust Lane Detection using Spatial-temporal Sequential Attention Based Transformer Model



Problem description

Reliable and accurate lane detection is undoubtedly vital for the safe performance for autonomous vehicles (AVs) and advanced driver assistance systems (ADAS). In recent years, many sophisticated visual-based lane detection methods, especially these employing Deep Learning (DL), have been proposed. However, treating it as an image instance segmentation or semantic segmentation problem, most methods focus on detecting the lane from one single image, and often lead to unsatisfactory performance in handling some extremely-challenging situations, e.g., bad lighting, marking degradation, serious vehicle occlusion. As lanes are continuous line structures, the lane that cannot be accurately detected in one single frame may potentially be inferred out by incorporating information of previous sequence of images. Therefore, end-to-end sequential DL models have been proposed as a research direction with great potential and promising prospects. Furthermore, the powerful spatio-temporal information processing ability and self-attention mechanism of Vision Transformer (ViT) could also be explored and integrated into the sequential DL model. The main aim of this research is to develop a sequential DL model incorporating Vision Transformer and sequential attention mechanism with continuous image frames as inputs to detect lanes of the last input image frame. The model need to be tested and verified through multi open sourced datasets (e.g., TuSimple, tvtLane, LLAMAS, etc.). The research group had already built the whole pipeline and collected a large scale driving scene data in the Netherlands (without labeling) for further testing the model's robustness.

Assignment

- Review and compare state-of-the-art lane detection methods using DL methods and Vision Transformer (e.g., SCNN, U-Net/SegNet-ConvLSTM, PINet, LSTR), and reproduce some of the high performing models as baseline;
- Develop the sequential DL neural architecture incorporating Vision Transformer and sequential attention mechanism, train and validate the model using various open-source datasets (tvtLane, TuSimple, LLAMAS, etc.);
- Encapsulate the developed model into a toolkit with friendly user interface for visualization and reassessment, demo codes already available.

Requirements: Experienced in Python and TensorFlow/Pytorch; Expertise in DL / Computer Vision, previous experience with Transformer is a plus.

Research group

Automated Mobility in Mixed Traffic; Transpiration AI; Transport & Planning; ISAC

Thesis supervisors: Yongqi Dong (https://yongqidong.github.io/)

External supervisor: possibility with industry partners

Information

For further information on this Master Thesis topic, please contact: yongqi.dong@rwth-aachen.de