Project: Explainable AI for Self-Driving Cars

Introduction

As the technology advances towards vehicles of higher automation level, the utilization of AI (and artificial neural networks in particular) becomes more and more prevalent. In such cases, it is important to understand how these networks produce results. One method to understand this is by visualizing the attention areas in input images. This project aims to build an end-to-end driving system that highlights the attention area in an image.

For this work, the aim is to train a DNN with attention visualization that takes input image and outputs the steering angle. Existing work[1] has already established its validity and the implementation[2] is available. This project aims to expand it into a real-time demonstrator using CARLA.

Objective

Using visual attention network to build a demonstrator of XAI on CARLA.

Tasks

- Survey of the existing work in XAI (Especially from UC Berkley).
- Set up google Colab to use the free available resources.
- Train a convolutional network end-to-end (from images to angles).
- Test it on existing data. Record the causal explanations.
- Setup CARLA to take inputs from the trained networks and show the explanation.
- Use the already trained network in CARLA to drive car autonomously and compare the results.
- Build a real-time demonstrator that highlights the attention leading to result.
- (Optional) Record dataset to train on CARLA.
- (Optional) Unfreeze last layers of the network and retrain the network on CARLA dataset.

Required skills

- Scientific literature survey
- Knowledge of Deep learning techniques, CNN, LSTM
- Python
- High performance computing for deep learning

Time Schedule

- Project start: as soon as possible
- Project implementation end: Four months after starting
- Project end: Six months after starting

Supervisor: Naveed Akram
Embedded Systems Quality Assurance (ESQ)
Fraunhofer IESE
Mail: naveed.akram@iese.fraunhofer.de

[1] "Interpretable Learning for Self-Driving Cars by Visualizing Causal Attention", Jinkyu Kim and John Canny
[2] https://github.com/JinkyuKimUCB/explainable-deep-driving