Using LiDAR as Camera for End-to-End driving

Jan Aare van Gent
Computer Science Curriculum
Supervisor: Tambet Matiisen
University of Tartu, Institute of Computer Science, Estonia

Introduction

- One of the approaches in autonomous driving research is known as end-to-end (E2E) driving. The idea is to train a neural network that predicts control signals directly from input sensors. Usually, the main sensor used for E2E driving is a regular front-facing camera.
- The Light Detection And Ranging (LiDAR) instrument is another sensor used for self-driving. In comparison to cameras, LiDARs give accurate distance estimations and can be more robust to weather and lighting conditions.

Objective

This project **aims to test the feasibility of LiDAR as a camera for E2E driving and data collection.** Specifically, the sensor examined is the Ouster OS1-128 LiDAR instrument, which can output measurements as a 360-degree raster image with range, intensity and ambience channels.

Methods

1. **Data Collection using Ouster LiDAR.** In total, 43826 images were collected from different road types, predominantly dirt roads and some low-traffic highways as well, over a period spanning from February to May 2021. Around 37% of the data is recovery data from dangerous situations. Example of images from dataset is in Figure 1.
2. **Train a model that predicts steering angle.** The architecture of the network is based on the DAVE-2 PilotNet network (Figure 2). In addition, batch normalization was added and cropping is done done on the input layers.
3. **System integration.** A ROS node was created to read the LiDAR images, run the network with the TensorRT [1] runtime and output network predictions as steering commands.
4. **Evaluation** was done with **closed-loop metrics** that require online testing with a real car. They include **number of interventions, lateral error** (reduced to RMSE and MAE)**, failure rate and whiteness.** The real-world evaluation was done on a 5 km dirt road track driving both ways (total 10 km) at 30 km/h with the UT ADL test vehicle shown in Figure 3.

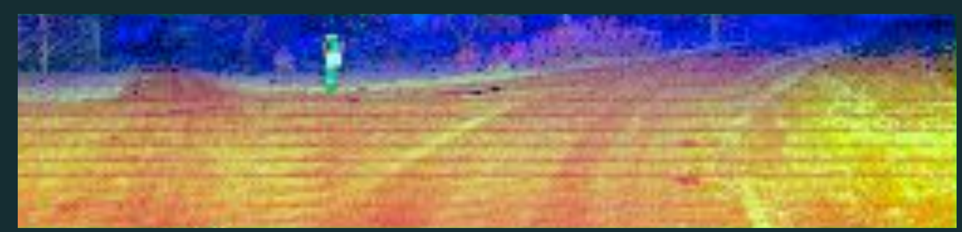

Figure 1. Example LiDAR image from the dataset.

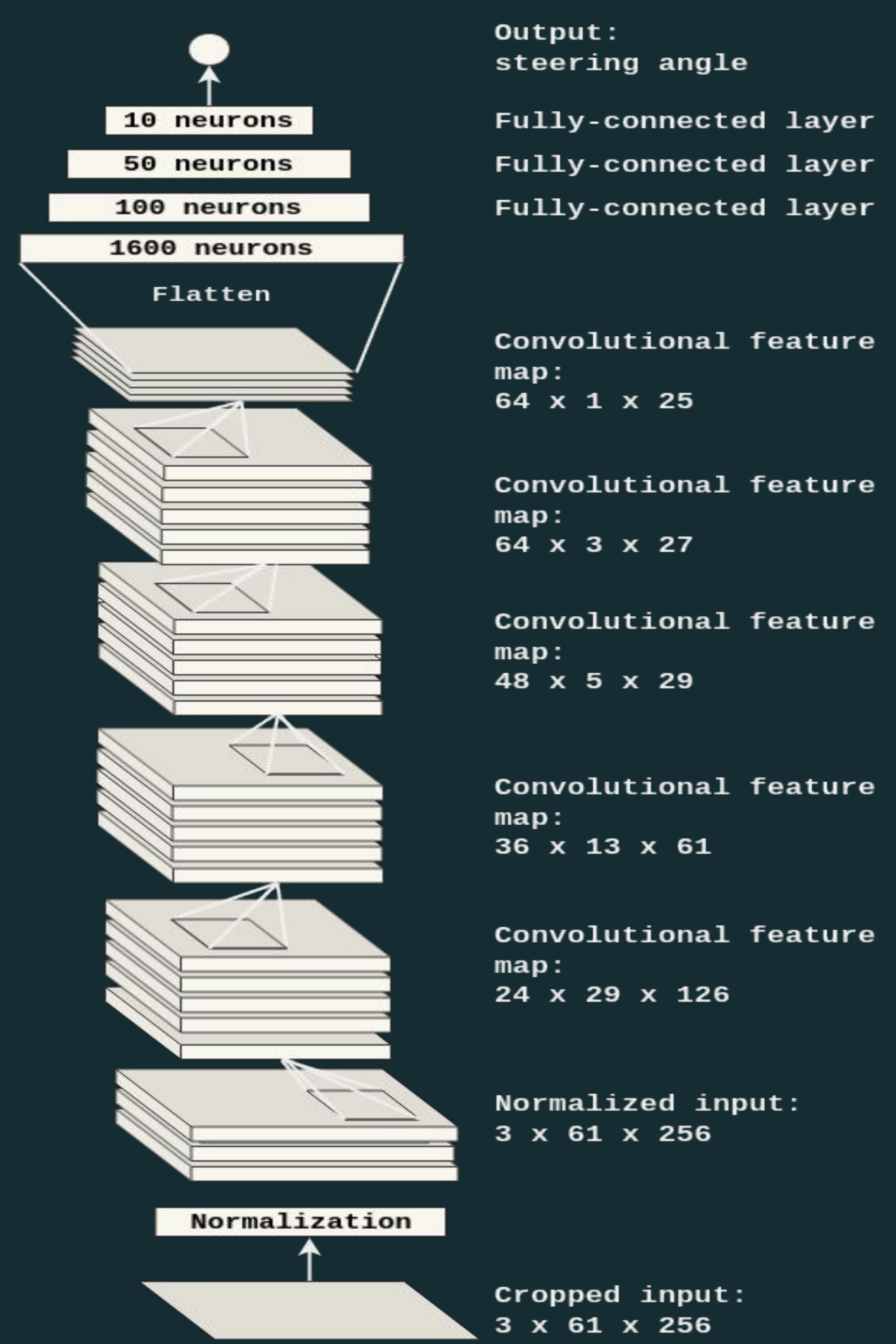

Figure 2. Network architecture inspired by Nvidia's PilotNet [2].

Figure 3. The Lexus 450h RX test vehicle used by UT Autonomous Driving Lab [3].

Experimental Results

The following models were trained and evaluated.

a. Model **BASE** is the main outcome of this project and is used as a reference model to compare experimental results.
b. Model **FUT** predicts the current steering angle and also two additional steering angles, 1 second and 2 seconds into the future respectively. This could force model to learn more robust road curvature features.
c. Models **AMB, INT, RNG** were trained for each individual channel (ambience, intensity, range).
d. 2-channel models **RNG+AMB, RNG+INT, INT+AMB** were trained to show which combinations of 2 image channels are the most useful.
e. Model **NO-INT** is trained without recovery data. This experiment attempts to quantify the importance of recovery data.
f. Model **SINGLE-TRACK** is trained only on the track the models were evaluated on later. Since around 60% of the training data was collected from this track, the model is trained on less varied data. The experiment measures the importance of varied data.

Model	Interventions	MAE (meters)	RMSE (meters)	Whiteness (degrees/second)	Failure rate (%)
RNG+AMB	3	0.24	0.32	1.91	0.67
INT+AMB	1	0.28	0.36	0.80	0.78
AMB	4	0.28	0.38	3.13	1.04
INT	1	0.29	0.37	1.07	1.21
RNG+INT	**0**	0.31	0.42	0.87	3.44
FUT	**0**	0.31	0.41	0.71	2.72
RNG	**0**	0.33	0.42	2.29	2.41
SINGLE-TRACK	2	**0.23**	**0.30**	**0.54**	0.42
NO-INT	6	0.48	0.58	1.08	8.03
BASE	1	0.25	0.33	0.77	**0.28**
Human driver	N/A	N/A	N/A	0.67	N/A

To keep the number of interventions low, it seems more data and channels other than ambience are needed.
The best performance for MAE and RMSE of the lateral errors is achieved by training only on the track where the evaluation was done. Other than that, models that made use of ambience seem to outperform models that do not. This suggests that ambience is a good source of information but not robust enough to different lighting and weather conditions.
Whiteness was used to measure the smoothness with lower whiteness corresponding to a smoother drive. By limiting the training set to the track where evaluation is done, the model was able to outperform a human driver in terms of driving smoothness. For other models, low lateral errors did not correlate to low whiteness.
Finally, failure rate measures the percentage of time the model trajectory had a lateral error larger than 1 meter compared to the expert trajectory. As can be expected, it is highly correlated to the RMSE and MAE of the lateral errors and models that had low lateral errors also had low failure rates.

References

[1] NVIDIA. Nvidia TensorRT. https://developer.nvidia.com/tensorrt
[2] Bojarski, Mariusz, et al. "End to end learning for self-driving cars." arXiv preprint arXiv:1604.07316 (2016).
[3] Autonomous Driving Lab https://www.cs.ut.ee/en/autonomous-driving-lab

GitHub

Bolt

UNIVERSITY OF TARTU
Institute of Computer Science