End to End Learning Self-Driving Cars

The convolutional neural network was trained to receive images and provide an specific angle to maintain centered in the road. With minimum training data from humans the system learns to drive in traffic on local roads with or without lane markings and on highways. It also operates in areas with unclear visual guidance such as in parking lots and on unpaved roads.

Authors

Andres Gamboa Baldi.
Froilan Josué Cardozo Zambrana.
Vanessa Gabriela Bustillos Nájera.
Alejandra Dayana Quelali Calcina.

Universidad Privada Boliviana

The Idea

A new effort was started at NVIDIA that sought to build on DAVE project and create a robust system for driving on public roads. The primary motivation for this work is to avoid the need to recognize specific human-designated features, such as lane markings, guard rails, or other cars.

Methodology

The first layer of the network performs image normalization.We follow the five convolutional layers with three fully connected layers leading to an output control value which is the inverse turning radius. The fully connected layers are designed to function as a controller for steering. A more detailed description of the filters is shown below.

Output: vehicule control

Fully-connected layer
Fully-connected layer
Fully-connected layer

Convolutional feature map 64@1x18

Convolutional feature map 64@3x20

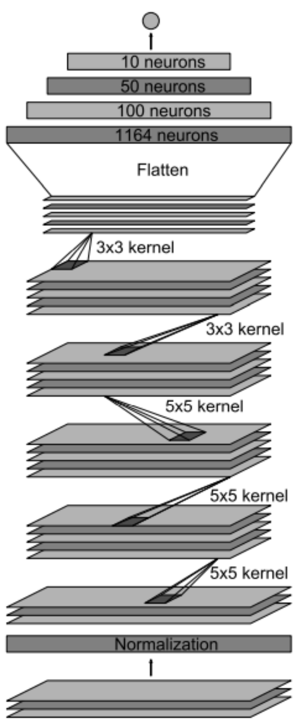
Convolutional feature map 48@5x22

Convolutional feature map 36@14x47

Convolutional feature map 24@31x98

Normalized input planes 3@66x200

Input planes 3@66x200



Architecture. From [1]

Conclusion

Empirically was demonstrated that CNNs are able to learn the entire task of lane and road following without manual decomposition into road or lane marking detection, path planning, and control. The system learns to detect the outline of a road without explicit labels during training.

Results

Estimating what percentage of the time the network could drive the car (autonomy). The metric is determined by counting simulated human interventions that occur when the simulated vehicle departs from the center line by more than one meter.

$$\text{autonomy} = \left(1 - \frac{(\text{number of interventions}) \cdot 6 \text{ seconds}}{\text{elapsed time [seconds]}}\right) \cdot 100$$

For a typical drive in Monmouth County NJ, the system is autonomous approximately 98% of the time.

Visualization of Internal CNN State

Figures 2 and 3 show the activations of the first two feature map layers for an unpaved road and a forest. In case of the unpaved road, the feature map activations clearly show the outline of the road while in case of the forest the feature maps contain mostly noise.

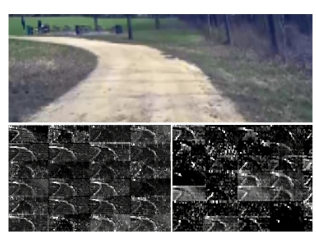


Figure 2: How the CNN “sees” an unpaved road From [1]

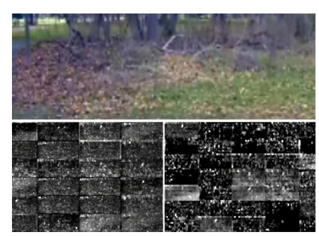


Figure 3: Example image with no road From [1]

Related Literature

[1] NVIDIA Corporation. (2016). End-to-End Learning for Self Drving Cars 1604.07316.pdf (arxiv.org).

[2] B. (2020). GitHub - berkcomba/Self-Driving-Car-Keras: Self driving car keras dave 2 CNN. GitHub. https://github.com/berkcomba/Self-Driving-Car-Keras