Autonomous Vision Group
Prof. Dr.-Ing. Andreas Geiger

Eberhard Karls
Universität
Tübingen

Self-Driving Cars

Ex. 2 – Pen and Paper (Lecture 02 + 03)

2.1 Self-driving basics

Self Driving

a) In the lecture, we have discussed 3 general approaches / paradigms to self-driving. State all 3 approaches to self-driving and specify one advantage and one disadvantage for each approach.

Deep Learning

a) We pass an image of size $C = 3, W = 320, H = 320$ through the network below. Fill in values for the ?'s so that the architecture is valid. Then report the tensor size (C_{out}, H, W) after each layer when the image is passed through the network.

Layer
Conv(?, ?, 3, 1, 1)
ReLU
MaxPool(2, 2)
Conv(64, ?, 3, 1, 1)
ReLU
MaxPool(2, 2)
Conv(128, 256, 3, 1, 1)
ReLU
MaxPool(2, 2)
Conv(?, 512, 3, 1, 1)
ReLU
MaxPool(2, 2)

b) The network architecture above is the (simplified) convolutional part of a network called VGG16 by (Simonyan et al., 2014). This stack of convolutional and max-pooling layers is often followed by a few fully-connected layers. The first fully-connected layer of a VGG16 is a FC($C_{in} = 25088, C_{out} = 4096$). Calculate the number of trainable parameters for the first and the second convolutional layers and the first fully-connected layer.

2.2 Imitation Learning

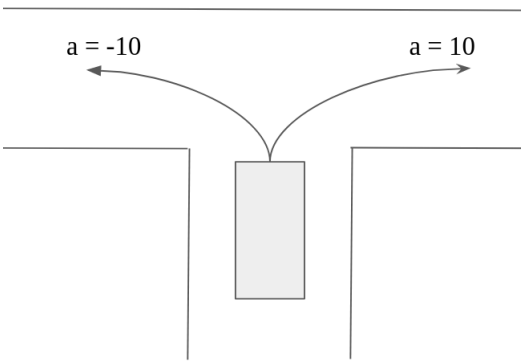
a) Describe the key idea of IL. What is a policy and how can it be represented formally?

b) Give two examples of possible environments and their states and actions - one where the state is fully known/ can be fully observed and one example where the state is only partly observable.

c) What is the main assumption Behavior Cloning (BC) is based on? What does that mean for testing the trained policy? (See also FAQ Lecture 2 and [2])

d) What is the problem when using an expert driver for collecting data? How can this problem be handeled?

e) Consider a self-driving vehicle in front of a simple intersection (see Figure). We name this state $s_i = 0$. The goal is to train a policy that maps this state to an action a_i, which is the steering angle of the vehicle.



i) Consider this policy is trained with Behavior Cloning. What is the problem with BC in this case? If you train on data from this intersection, what does the car do during inference?

ii) We collect a dataset on this intersection: $D = \{(-10, 0), (10, 0), (10, 0), (-10, 0)\}$ with action-state pairs $(a_i^\star, s_i^\star)$. We train a deterministic policy that directly maps the state to the action. For this we use an L2-Loss: $\mathcal{L}\left(a_i^*, \pi_\theta(s_i^*)\right) = \left(a_i^* - \pi_\theta(s_i^*)\right)^2$.
Derive the optimal action fo the policy to take given the above dataset and loss function.

iii) Name an option to overcome this problem.

f) What is the main difference between BC and CIL? Explain this based on the formulas for the objectives and assumptions.

2.3 Direct Perception

Affordances

a) Give a high level description of what affordances are and name four examples in the context of self-driving.

b) Which affordances are useful for i) highways ii) inner-city driving and iii) both.

Visual Abstractions

a) What are relevant properties of visual abstractions?

b) Semantic segmentation can be used as an intermediate representation. Behl et.al [3] show that the performance can improve when using i) less classes, ii) coarsely anotated images. What can be a reason for this?

c) In order to train a visual abstraction model we need two datasets: n_r (images annotated with semantic labels) and n_a (images annotated with expert actions). Why do we assume $n_r << n_a$? Do you have an intuition why it is possible to use less data to train the visual abstraction model?

Evaluation

a) Describe the difference between online and offline evaluation. Name three metrics respectively.

b) Why are offline metrics not always good proxies of online driving performance?

2.4 References

[1] https://arxiv.org/pdf/1604.07316.pdf
[2] https://arxiv.org/pdf/1011.0686.pdf
[3] http://www.cvlibs.net/publications/Behl2020IROS.pdf
[4] http://proceedings.mlr.press/v87/sauer18a/sauer18a.pdf
[5] https://arxiv.org/pdf/1809.04843.pdf