

Learning by cheating

Vision-based autonomous driving consists of simultaneous perception and action in the environment. This presents a nontrivial challenge through the lens of learning in high-dimensional spaces. The work decomposes these two stages by means of an agent which has access to privileged information. In the first stage, a teacher agent learns from an expert by cheating. The agent has access to ground-truth layout and positions of participants. In the second stage, the agent acts as a teacher for a sensorimotor agent which does not have access to any privileged information. This two-stage Learning By Cheating (LBC) procedure allows the sensorimotor agent to learn by treating the teacher as a query-expert. LBC, when evaluated for autonomous driving on the CARLA and NoCrash benchmark, presents improvements in success rates and reduced number of infractions.

The LBC framework consists of two stages which abstract perceptual and sensorimotor learning by means of privileged information. The privileged agent cheats by having access to ground-truth layouts and positions of participants in the scene. This privileged information is provided to the agent via imitation wherein the agent learns to mimic a simulation expert. In the second stage, the privileged agent is used to train a sensorimotor agent which does have access to privileged information. This teaching setup can be thought of as a querying mechanism wherein the agent, in every state, asks the question to the teacher *"What would you do here?"*. Both agents consist of a feature extraction module followed by a low-level controller. In the case of privileged agent, the feature extraction module selects waypoints from the bird's eye view as input. On the other hand, the sensorimotor agent receives images from the forward-facing camera which yield egocentric waypoints. Waypoints are provided to the low-level controller which is used to select actions.

The LBC framework demonstrates improved performance in the form of success rates on the CARLA autonomous driving benchmark. Additionally, LBC presents lower number of infractions such as red light violations and collisions on the NoCrash benchmark in CARLA. While LBC presents a suitable scheme for learning perception and action in conjunction, its experimental setup presents a few caveats. Firstly, the privileged and sensorimotor agents comprise of different architecture backbones. The privileged agent uses a randomly initialized ResNet-18 while the sensorimotor agent uses ResNet-34 agent pretrained on the ImageNet dataset. One could argue that the sensorimotor agent has a much richer representation module which is qualitatively equivalent to the privileged information provided to teacher agent. Secondly, the teacher agent requires 4-6 hours of driving data for training which is in contrast to modern learning methods which aim to reduce the data-dependency of algorithms. Moreover, this presents a nontrivial challenge in practical driving scenarios wherein data collection is significantly expensive.

The work presents a suitable scheme for abstraction of learning perceptual and sensorimotor control signals by means of privileged information. This presents two new directions for future work. Firstly, the framework may be extended towards Reinforcement Learning settings wherein the agent has to learn sensorimotor control by itself. Secondly, extensions of LBC can aim to improve its high data requirements from a practical standpoint.

Vision-based autonomous driving presents learning perception and action simultaneously. The work aims to abstract these aspects in light of privileged information provided to a teacher agent. The agent learns from an expert by gaining access to ground-truth layouts. In the second stage, the agent teaches a sensorimotor agent which does not cheat. Abstraction of learning in a bi-level framework demonstrates improved performance and lower infractions on the CARLA benchmark.