TD1 by Team 4: Imitation Learning

Varshitha Janavi*, Pranay Palem†, and Puneet Sai Naru‡
Emails: vpurra@asu.edu*, ppalem1@asu.edu†, pnaru@asu.edu‡

***Abstract*—We are team 4 and address the "What is Imitation Learning?" prompt. Imitation Learning(IL) is an alternative to traditional reinforcement learning (RL). Our discussion focuses on applying IL in autonomous driving, where vehicles learn to navigate complex environments by mimicking human drivers. By leveraging techniques such as Behavior Cloning, Dataset Aggregation (DAgger), and Inverse Reinforcement Learning (IRL), IL allows for developing policies that generalize across diverse driving scenarios, including lane changes, intersections, and highway merging. IL in this context demonstrates its potential to enhance the safety and efficiency of autonomous systems by enabling them to learn from expert behavior while adapting to real-world complexities.**

***Index Terms*—Imitation Learning(IL), Reinforcement Learning(RL), Inverse Reinforcement Learning(IRL), Behavioral Cloning(BC), Autonomous Vehicles, Dataset Aggregation (DAgger), Expert Demonstrations**

I. Introduction

A. Overview

IL is widely used in autonomous vehicle development, robotics, gaming, and healthcare for training models based on expert performance. This report focuses on autonomous driving, where IL trains vehicles to navigate safely and efficiently in dynamic environments. Traditional RL methods often face impractical safety concerns due to the high number of required interactions. In contrast, IL learns from demonstrations, replicating human driving behavior to handle tasks like lane-keeping, highway merging, intersection navigation, and reacting to other vehicles and pedestrians. IL techniques such as Behavior Cloning (BC), Dataset Aggregation (DAgger), and Inverse Reinforcement Learning (IRL) address specific challenges in autonomous driving.

B. Additional Support Materials

We provide additional support materials in the following accessible link):

- Imitation Learning ICML 2018 Lecture - https://www.youtube.com/watch?v=WjFdD7PDGw0&t=7s
- Imitation Learning ICML 2018 Slides - https://www.slideshare.net/slideshow/imitation-learning-tutorial/105192070
- ALVINN Driver Less Car CMU 1998 Video - https://www.youtube.com/watch?v=2KMAAmkz9go

II. Context and Challenges in Autonomous Driving

The primary challenge in autonomous driving is ensuring safety while making quick decisions. Vehicles must process large amounts of sensor data in real-time and respond to changing environments. Imitation Learning (IL) helps address these challenges by using expert demonstrations to mimic human driving. Here are common IL approaches:

- **Behavioral Cloning**: This method maps sensor inputs directly to control actions by copying expert decisions. For example, Pomerleau's early work with ALVINN (Pomerleau 1988) utilized a neural network to learn steering commands from visual input.
- **Inverse Reinforcement Learning (IRL)**: Instead of copying actions, IRL learns the reward function behind expert behaviors and then optimizes actions based on this learned reward.
- **Reinforcement Learning with Imitation**: Combining reinforcement learning (RL) with imitation allows the system to learn from both expert demonstrations and its own experiences. For instance, recent approaches like SDR (Tippannavar, S D, and K M 2023) use behavioral cloning initially but then rely on RL to fine-tune the system, providing better performance in unfamiliar scenarios

There are many challenges associated with IL, some of which are:

- **Generalization**: If the training data does not cover a diverse set of situations, the model may not be able to generalize from the expert demonstrations to unseen scenarios.
- **Resource Constraints**: Collecting and labeling large datasets of high-quality expert driving behavior is resource intensive.
- **Errors**: In IL techniques like behaviour cloning(BC), small errors accumulate over time

In the next section, we will delve into the How aspect of imitation learning techniques, focusing on recent advancements and specific methodologies.

III. How Imitation Learning is Implemented

Imitation learning is essential for training autonomous systems, such as self-driving cars, by enabling them to replicate expert behavior. Key techniques include Behavioral Cloning (BC), DAgger (Dataset Aggregation), and Inverse Reinforcement Learning (IRL) which are explained in depth in the paper "A Survey of Imitation Learning: Algorithms, Recent Developments, and Challenges" (Zare et al. 2024). BC learns driving policies by mapping observations directly to actions based on expert data. DAgger refines the model through iterative corrections by incorporating feedback from its actions. IRL infers the underlying reward functions that drive expert behavior, enhancing decision-making. In the paper "A Survey on Imitation Learning Techniques for End-to-End Autonomous Vehicles"(Le Mero et al. 2022) these techniques are referred with more detail where they discuss their algorithms, advancements, and challenges in the context of autonomous driving.

A. *Behavior Cloning (BC)*

BC is a supervised learning approach where the agent learns a policy by directly mimicking the expert's actions through minimizing a loss function between the agent's and expert's actions.

$$\mathcal{L}(\theta) = \mathbb{E}_{(s,a^*)\sim\mathcal{D}}[\ell(\pi_\theta(s), a^*)] \tag{1}$$

Where $\ell(\pi_\theta(s), a^*)$ is the loss function (e.g., cross-entropy, mean squared error), and $\mathcal{D}$ is the dataset of state-action pairs (s, a^*) provided by the expert. The goal is to minimize the loss function to ensure that $\pi_\theta(s)$ closely approximates the expert's behavior.

B. *DAgger (Dataset Aggregation)*

DAgger iteratively updates the agent's policy by collecting new states encountered during policy execution and querying the expert for corrective actions to mitigate covariate shift.

Algorithm 1 **DAgger Algorithm**

Initialize dataset $\mathcal{D} \leftarrow \{(s_i, a_i^*)\}$
for each iteration $i = 1, 2, \ldots, N$ **do**
  Train policy π_i on dataset $\mathcal{D}_i$
  Collect new states $s_i \sim \pi_i$
  Query expert for actions a_i^* on s_i
  Update dataset $\mathcal{D}_{i+1} = \mathcal{D}_i \cup \{(s_i, a_i^*)\}$
end for

At each iteration, the agent improves by learning from its own state distribution and receiving corrections from the expert.

C. *Inverse Reinforcement Learning (IRL)*

IRL focuses on learning a reward function from expert demonstrations, which the agent uses to optimize its policy through reinforcement learning to replicate expert behavior.

$$r^* = \arg\max_r \sum_{(s,a^*)\in\mathcal{D}} r(s, a^*) \tag{2}$$

After learning the reward function, the agent solves a reinforcement learning problem to find the optimal policy π^*:

$$\pi^* = \arg\max_\pi \mathbb{E}\left[\sum_{t=0}^{T} r(s_t, a_t)\right] \tag{3}$$

IV. Why Imitation Learning

Mimicry of expert policy with the given states s and actions a is at the core of imitation learning. Imitation learning enables us to capture expert data and infer expert policy π^* to train a system to imitate the expert. As mentioned above, there are multiple ways of training a system to imitate the expert policy. Imitation learning significantly reduced the time and effort required to map the state action space for the autonomous vehicle by comparing and inferring the optimal policy of the expert as expanded in (Pomerleau 1988). IL also allows the system to learn without having to decide a policy manually and generalize states that differ from the available training data (Le Mero et al. 2022).

Taking an example of Behavioural Cloning (BC) in Autonomous system, we can see that the neural network receiving the desired steering command from the expert and is comparing it to the output from a neural network as shown in Figure 1. Then the system adjusts weights of the CNN to minimize the loss (similar to equation 1) between the desired steering control and the predicted control by the CNN (Tippannavar, S D, and K M 2023). The ease of capturing the desired expert data from a human driver to train a system makes Imitation Learning a preferred method to train a network to control autonomous cars.

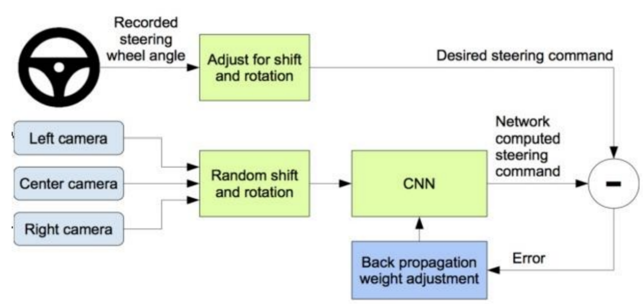

Fig. 1. Behavioural Cloning System in Autonomous Vehicle with three camera inputs to the CNN

V. Future Work

Having a conceptual clarity on what imitation learning is is essential for implementing the idea on any application. As the next step, we could read more literature and to understand the pros and cons of various methods mentioned in this paper. We could further explore why minimization of loss function of maximization of reward function leads to a better performing robot.

VI. Links

Here is a link to the presentation: Presentation Link

REFERENCES

Reference	Description
Luc Le Mero et al. (2022). "A Survey on Imitation Learning Techniques for End-to-End Autonomous Vehicles". In: *IEEE Transactions on Intelligent Transportation Systems* 23.9, pp. 14128–14147. DOI: 10.1109/TITS.2022.3144867	Our discussion relied heavily on this paper. (Sections: Introduction, How and Why)
Maryam Zare et al. (2024). "A Survey of Imitation Learning: Algorithms, Recent Developments, and Challenges". In: *IEEE Transactions on Cybernetics*, pp. 1–14. DOI: 10.1109/TCYB.2024.3395626	We used this paper to explore the concepts of Imitation learning better (Sections: How and Why).
Sanjay S Tippannavar, Yashwanth S D, and Puneeth K M (2023). "SDR – Self Driving Car Implemented using Reinforcement Learning & Behavioural Cloning". In: *Proceedings of the 2023 International Conference on Recent Trends in Electronics and Communication (ICRTEC)*, pp. 1–7. DOI: 10.1109/ICRTEC56977.2023.10111893	We referred to this paper to gain an understanding on Behavioural Learning in Autonomous Vehicles.
Dean A Pomerleau (1988). "Alvinn: An autonomous land vehicle in a neural network". In: *Advances in neural information processing systems* 1	This paper is one of the foundational papers of Imitation learning and most of the autonomous driving literature builds on the work presented in this paper.