Gaze Supervision for Mitigating Causal Confusion in Driving Agents

Extended Abstract

Abhijat Biswas
Carnegie Mellon University
Pittsburgh, USA
abhijat@cmu.edu

Badal Arun Pardhi
Carnegie Mellon University
Pittsburgh, USA
bpardhi@alumni.cmu.edu

Caleb Chuck
University of Texas at Austin
Austin, USA
calebc@cs.utexas.edu

Jarrett Holtz
Robert Bosch LLC &
University of Texas at Austin
Austin, USA

Scott Niekum
University of Massachusets Amherst
Amherst, USA
sniekum@cs.umass.edu

Henny Admoni
Carnegie Mellon University
Pittsburgh, USA
henny@cmu.edu

Alessandro Allievi
Robert Bosch LLC &
University of Texas at Austin
Austin, USA
alessandro.allievi@us.bosch.com

ABSTRACT

Imitation Learning (IL) algorithms show promise in learning human-level driving behavior, but they often suffer from "causal confusion," a phenomenon where the lack of explicit inference of the underlying causal structure can result in misattribution of the relative importance of scene elements, especially pronounced in complex scenarios like urban driving with abundant information per time step. Our key idea is that while driving, human drivers naturally exhibit an easily obtained, continuous signal that is highly correlated with causal elements of the state space: eye gaze. We collect human driver demonstrations in a CARLA-based VR driving simulator, allowing us to capture eye gaze in the same simulation environment commonly used in prior work. Further, we propose a method to use gaze-based supervision to mitigate causal confusion in driving IL agents — exploiting the relative importance of gazed-at and not-gazed-at scene elements for driving decision-making. We present quantitative results demonstrating the promise of gaze-based supervision improving the driving performance of IL agents.

KEYWORDS

Causal confusion; Eye gaze; Urban driving; Imitation learning

ACM Reference Format:
Abhijat Biswas, Badal Arun Pardhi, Caleb Chuck, Jarrett Holtz, Scott Niekum, Henny Admoni, and Alessandro Allievi. 2024. Gaze Supervision for Mitigating Causal Confusion in Driving Agents: Extended Abstract. In *Proc. of the 23rd International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2024), Auckland, New Zealand, May 6 – 10, 2024*, IFAAMAS, 3 pages.

Proc. of the 23rd International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2024), N. Alechina, V. Dignum, M. Dastani, J.S. Sichman (eds.), May 6 – 10, 2024, Auckland, New Zealand. © 2024 International Foundation for Autonomous Agents and Multiagent Systems (www.ifaamas.org).

1 INTRODUCTION

Imitation learning (IL) is a popular method for learning urban driving policies due to its ease of implementation and de-coupling of the data collection/action step and the training step by allowing offline learning of control, among other factors. However, it does not explicitly model underlying causal structures of tasks, instead inferring causality from strongly correlated elements of the state space that occur before specific actions are performed. This results in a policy that does the right things for the wrong reasons in the training distribution and thus doesn't generalize well at test time. The most straightforward resolution to the causal confusion problem would be to simply learn the correct underlying causal structure. De Haan *et al* [4] propose methods using targeted interventions to prune a set of 2^N causal hypotheses where N is the dimensionality of the state space. This is a huge search space for visuomotor tasks like driving. Moreover, expert queries or environment interaction in the training loop can often be unfeasible.

Taking a complementary approach, we seek to use a signal that human drivers naturally exhibit while operating vehicles which is highly correlated with causal parts of the state space — eye gaze. Our idea is to use driver eye gaze as a supervisory signal, alongside driving control, to highlight the lower dimensional parts of the (very high-dimensional) visual state space that the driver fixated on before making their driving decision. Specifically, we use a contrastive learning formulation to encourage visuomotor IL driving policies to change driving decisions based on visual information in the fixated-at regions. This gaze supervision seeks to mitigate causal confusion by directing the causal function of the policy towards the variables of the observation (clusters of pixels), which correspond to an underlying state variable that the human believes is causal to the optimal behavior.

This work was funded in part by the National Science Foundation (IIS-1943072, IIS-1749204, IIS-1925082), the Tang Family AI Innovation Fund, the Link Foundation, Bosch, AFOSR (FA9550-20-1-0077), and ARO (78372-CS, W911NF-19-2-0333).

The benefit of using eye gaze from human driving demonstrators is that it is essentially "free", *i.e.* it is a signal that is naturally exhibited by humans as they drive. Importantly, it does not require additional labeling or intervention from human experts and is non-intrusive, with gaze data being able to be collected with a pair of wearable glasses or even in-cabin sensors. In fact, some data-collection vehicles are already instrumented with cabin-facing visual or infrared sensors, that can be used to obtain traffic-scene registered eye gaze directly.

We propose a gaze-based contrastive supervision method to incorporate driver gaze into policy training and show that fine-tuning a pre-trained IL driving policy using our method results in better driving performance than the pre-trained model. Our formulation encourages the trained policy network's driving actions to be affected by gazed-at regions. Further, the fine-tuned method's saliency better matches drivers' attention as indicated by their gaze. In summary, we investigate the utility of natural driver eye gaze-based supervision as a tool for mitigating causal confusion in imitation learning-based driving agents.

2 METHOD

Causal confusion in IL driving agents: As an example algorithm for exploring causal confusion in IL-based driving agents, we consider the Learning by Cheating (LBC) [2] model for autonomous urban driving in CARLA. As one may expect, the LBC model also shows symptoms of suffering from causal confusion [1]. These problems seem to occur primarily in the absence of surrounding vehicles which may be wrongly used as causal cues. We especially notice traffic light infractions where the LBC agent either does not stop for or fails to restart after stopping at, a red light. We also observe cases where the agent stops at a red light but restarts when opposing traffic moves, even though the red light has not changed.

To investigate the relative importance of regions of the input state space in making decisions, we used a saliency method to investigate the decision-making process of the LBC model. Specifically, we used the blur-based saliency method by Greydanus *et al.* [5]. The method is network architecture agnostic and works by blurring different regions of the given visual input and measuring the difference in output with the original input. It reasons that input image regions which, when blurred, cause the greatest difference in the agent's policy network output, are the most salient.

Thus, we can generate saliency maps for the LBC method. Applied to a vehicle stopped at a red light the frame before it turns green, it shows that most of the salience lies erroneously on non-causal parts of the input image, such as the base of the traffic light.

Gaze based supervision We collect driving demonstrations and driver eye gaze in a VR based driving simulation and incorporate gaze supervision as a contrastive loss to existing driving IL policies.

Human demonstrations were collected in the DReyeVR simulator [6], which allows human driving in CARLA in VR. Eye gaze movements were pre-processed to filter out high-frequency noise, filtering out saccades, and aggregated into attention maps using a Gaussian distribution over a 15-second window. 7 drivers completed five routes each, with some routes and participants excluded due to motion sickness or data recording issues, resulting in a dataset of 17 routes ≈ 70 minutes (henceforth: DRVR dataset).

Table 1: Driving performance and Model saliency IoU on the Longest6 [3] benchmark. Base model for all rows is LBC [2]

Training approach	Training data	Loss used	DS (↑)	IoU (↑)
Pre-trained [2]	RBE	LBC	7.01	0.13
Mixed (control only)	Mix	LBC	7.81	0.12
Mixed (control & gaze)	Mix	LBC+Triplet	9.61	0.18

Our idea to provide gaze supervision comes from correcting misplaced salience using a triplet loss and gaze-based attention. In our formulation, the original set of input images (*left, center, right*, waypoint) constitutes the triplet's anchor data point. The negative input is constructed by applying Gaussian blur (same parameters as [5]) to important gazed-at scene locations in the same set of images (above x_+, our formulation: x_{gaze}). The corresponding positive point has the same blur applied to the unimportant scene regions (above x_-, our formulation: complement of attention maps $x_{!gaze}$). The reasoning for this formulation is as follows: the most important regions for decision making for actions lie in the gazed-at regions (as indicated by attention maps) and the non-gazed-at regions do not contain information that would change the driving decision. Hence, we can write the loss as follows $L_t(x_a, x_+, x_-) = \max(\,||\phi(x_a)-\phi(x_+)||_2-||\phi(x_a)-\phi(x_-)||_2+\alpha,\ 0)$ This loss enforces that visual inputs blurred in locations unimportant to driving should lead to a smaller change in network output than the same blur applied in important regions.

We use two primary datasets: data from the rule-based expert (RBE) and data from humans in the DReyeVR simulator (DRVR).

3 EXPERIMENTAL RESULTS

We evaluate the effectiveness of our gaze-based supervision method in two parts: First, we show that after applying gaze supervision, the saliency of the model matches the attention maps from human drivers' gaze more closely. Second, we show that the imitation agent's driving performance improves after applying gaze supervision.

To investigate the effect of gaze-based supervision on the driving agent's saliency, we use a modified version of Greydanus *et al.* [5] to compute saliency maps. We use DRVR routes since those have associated ground truth gaze (and hence, attention) available for comparison. In our results (Table 1), we see that fine-tuning with gaze supervision does indeed improve the Intersection over Union of model saliency maps (*i.e.*, they better match the true attention maps from human demonstrators). Fine-tuning with Mixed data and both gaze and control supervision achieves a good balance of both IOU and driving performance.

To evaluate the driving performance of our fine-tuned models, we used the Longest6 benchmark [3] and their $DrivingScore(DS)$ metric. From experiments investigating the agent's driving performance, using just control supervision on the Mixed (RBE + DRVR) dataset does improve performance over the vanilla pre-trained LBC model since it sees more training data than just the RBE dataset. However, finetuning using Mixed data using the both gaze and control losses in conjuction leads to the better driving performance than using control loss alone.

Societal Impact Statement

In our assessment, this work helps autonomous driving agents' decisions to be based on parts of the environment human drivers pay attention to. However, not all human drivers are equal and it is important that we only distill those human behaviors into driving policies that are deemed exemplarily safe. Here, we excluded drivers who had not held a valid driver's license for at least a year. For on-road deployments a much higher bar should be set with multiple independent raters verifying demonstrations for safety.

REFERENCES

[1] [n.d.]. Causal confusion in Learning-by-Cheating github issue. Accessed Sept 20, 2022. https://github.com/bradyz/2020_CARLA_challenge/issues/16.

[2] Dian Chen, Brady Zhou, Vladlen Koltun, and Philipp Krähenbühl. 2020. Learning by cheating. In *Conference on Robot Learning*. PMLR, 66–75.

[3] Kashyap Chitta, Aditya Prakash, Bernhard Jaeger, Zehao Yu, Katrin Renz, and Andreas Geiger. 2022. TransFuser: Imitation with Transformer-Based Sensor Fusion for Autonomous Driving. *Pattern Analysis and Machine Intelligence (PAMI)* (2022).

[4] Pim De Haan, Dinesh Jayaraman, and Sergey Levine. 2019. Causal confusion in imitation learning. *Advances in Neural Information Processing Systems* 32 (2019).

[5] Samuel Greydanus, Anurag Koul, Jonathan Dodge, and Alan Fern. 2018. Visualizing and understanding atari agents. In *International conference on machine learning*. PMLR, 1792–1801.

[6] Gustavo Silvera, Abhijat Biswas, and Henny Admoni. 2022. DReyeVR: Democratizing Virtual Reality Driving Simulation for Behavioural & Interaction Research. In *Proceedings of the 2022 ACM/IEEE International Conference on Human-Robot Interaction* (Sapporo, Hokkaido, Japan) *(HRI '22)*. IEEE Press, 639–643.