

Learning Autonomous Control Policy for Intersection Navigation with Pedestrian Interaction

Zeyu Zhu, *Member, IEEE*, Huijing Zhao *Member, IEEE*,

Abstract—In recent years, great efforts have been devoted to deep imitation learning for autonomous driving control, where raw sensory inputs are directly mapped to control actions. However, navigating through densely populated intersections remains a challenging task due to uncertainty caused by uncertain traffic participants. We focus on autonomous navigation at crowded intersections that require interaction with pedestrians. A multi-task conditional imitation learning framework is proposed to adapt both lateral and longitudinal control tasks for safe and efficient interaction. A new benchmark called IntersectNav is developed and human demonstrations are provided. Empirical results show that the proposed method can achieve a success rate gain of up to 30% compared to the state-of-the-art.

Index Terms—deep imitation learning, multi-task learning, autonomous driving control, interaction with pedestrians

I. INTRODUCTION

Interest in autonomous driving is growing at a rapid pace [1] and navigating through crowded intersections is one of the most challenging tasks in autonomous driving [2], [3]. At such scenes, an autonomous vehicle is by no means an isolated agent driving in a static environment, but in a complex dynamic transportation system. It needs to adjust its controls to interact with other road users and navigate the scene safely and efficiently. The situation is more challenging when interacting with pedestrians, who have more free and unpredictable movements that introduce great uncertainties [4], [5]. The autonomous vehicle must follow socially compliant rules in order to be understood and accepted by pedestrians [6].

Current autonomous driving systems are mainly modular [7], [8], which are composed of individual modules for perception, localization, decision-making, planning and control tasks. Given pre-planned driving routes and mission points, the agent generates driving policies at multiple levels of abstraction. To interact with other road users, modules such as object detection and tracking [9], trajectory prediction [10], collision and risk analysis [11] are usually involved. The results are fed into a driving behavior generation module, which seeks the optimal policy by evaluating safety, efficiency, etc., using hand-crafted rules and fine-tuned parameters. In such cases, today's autonomous driving systems choose to tune its parameters to guarantee safety first, leading many to complain about conservative behavior, inefficiency and inhuman driving.

This work was supported in part by the National Natural Science Foundation of China under Grant 61973004 and U22A2061. Correspondence: H. Zhao, zhaohj@pku.edu.cn.

Z. Zhu and H. Zhao are with the Key Lab of Machine Perception (MOE) and also with the School of Intelligence Science and Technology, Peking University, Beijing, China. (e-mail: {zhuzeyu; zhaohj}@pku.edu.cn).

As an emerging trend, end-to-end autonomous driving has attracted great attention in the latest research [12]–[14], thanks to the rapid development and impressive success of deep learning techniques. Compared with traditional microscopic vehicle behavioral models, deep learning methods can better model the highly non-linear procedure of driving at complex scenes. Such systems learn a deep model that maps sensory input to control actions in an end-to-end fashion, where deep reinforcement learning (DRL) [15]–[17] and deep imitation learning (DIL) [18]–[22] are the most representative learning methods. DRL learns from online trial and error (i.e., interaction with the environment), which can be dangerous in the real world. Therefore, most current DRL methods [15]–[17], [23], [24] rely heavily on simulators. On the other hand, DIL learns from expert demonstrations and can be executed offline, which is important for safety-critical applications such as autonomous driving [19], [20]. Furthermore, DIL can learn from human drivers' data, where large amount of demonstration data can be easily collected using low-cost on-board sensors during road driving by human experts.

This work investigates DIL-based end-to-end autonomous control policy learning for intersection navigation with pedestrian interaction. In such tasks, the autonomous agent needs to perform both lateral and longitudinal controls to navigate through the intersection safely and efficiently, and interact friendly with pedestrians on crosswalks by imitating human experts. When encountering a pedestrian, the autonomous agent can choose to yield or pass through longitudinal control, and meanwhile steer through lateral control. Such behavior is modeled implicitly through imitative learning of human driver data. However, lateral and longitudinal control are two tasks of very different properties. Various scene features have different importance in accomplishing each task. Besides, lateral and longitudinal control are manipulated by the human driver's feet and hands respectively, showing different control vibration tolerances in the expert's demonstration data. Furthermore, the need for friendly interaction requires that unfriendly events should be rare, such as forcing pedestrians to abruptly stop due to inappropriate driving strategies of automated agents. There are also literature works on DIL-based autonomous control policy learning that addressed intersection scenarios [19], [20], [25]. However, none of these works focus on the interaction with pedestrians, and the different properties of lateral and longitudinal control are ignored.

To this end, a multi-task conditional imitation learning (MTCIL) approach is proposed by extending the popular conditional imitation learning (CIL) [19] framework. Although lateral and longitudinal controls can be related, many au-

TABLE I
REPRESENTATIVE DEEP IMITATION LEARNING METHODS FOR AUTONOMOUS DRIVING CONTROL POLICY LEARNING

Category	Model	Scenario			Control	
		Main Scenario	Intersection Navigation	Interaction with Pedestrian	Type	Separate Modeling
Direct Perception	[26]	Highway Driving	✗	None	Lat. & Lon.	✓
	[25]	Urban Navigation	✓	None	Lat. & Lon.	✓
End-to-end	[18]	Road Following	✗	None	Lat.	✗
	[19]	Urban Navigation	✓	Few	Lat. & Lon.	✗
	[27]	Urban Navigation	✓	Few	Lat. & Lon.	✗
	[17]	Urban Navigation	✓	Few	Lat. & Lon.	✗

* Lat. and Lon. are abbreviations for “Lateral” and “Longitudinal” respectively.

Autonomous driving systems manage them separately, but share the same perception results [12], [28]. Following such structures, this work designs a multi-task learning framework, in which both controls share the same scene encoding module, while they map scene descriptors to actions separately. Homoscedastic uncertainties [29] that are inherent to both tasks are estimated and used for weighting the loss in training. Evaluating control models on static datasets is not enough due to compounding error caused by covariate shift [30], closed-loop evaluation is essential. This research uses the high-fidelity CARLA simulator [31] in experiments, where a new benchmark called IntersectNav was developed, in which about 2900 human driving trajectories on 41 routes were collected at six intersections under different conditions. In addition, new evaluation protocols and metrics are defined to enrich the criteria of previous benchmarks. The performance of the proposed method is extensively studied. Although we focus on interaction with pedestrians, the generalizability to vehicles is demonstrated in experiments. Experimental results show that our model achieves up to 30% success rate gain compared to the state-of-the-art. The benchmark, collected dataset and videos are available at <https://github.com/zhackzey/IntersectNav>. Our main contributions are:

- A multi-task conditional imitation learning framework is proposed for autonomous navigation at crowded intersections that require interaction with pedestrians, which leverages homoscedastic uncertainties of the lateral and longitudinal control tasks.
- A new benchmark IntersectNav is developed on three aspects: Ped-Only for interaction with pedestrians, Ped-Veh for interaction with both pedestrians and environmental vehicles and Mul-Dri for multiple driving styles. New evaluation protocols and metrics are proposed.
- Extensive experiments are conducted to evaluate the proposed method on the autonomous agent’s reliability in task completion, control quality and generalization. Superior performance of the proposed method is demonstrated compared to the state-of-the-art.

Our paper is organized as follows. Section II related work. Section III the proposed method. Section IV the proposed benchmark. Section V experimental results and Section VI our conclusion.

II. RELATED WORK

A. DIL for Autonomous Driving Control Policy Learning

Methods can be divided into two categories: direct perception methods and end-to-end methods. Comparisons between representative methods are shown in Tab. I.

Direct perception methods [25], [26] utilize neural networks to extract compact intermediate representations which are then passed to subsequent decision and control modules. Intermediate representations can be chosen as predicted affordances presented in CAL [25], such as distance to the preceding vehicle and distance to centerline. Typically, direct perception methods rely on carefully designed functions or controllers to map the predicted mediated perception results to the final control. Therefore, the different characteristics of lateral and longitudinal control can be taken in consideration through different control functions. However, well design of both mediated representations and control functions requires system expertise, which is often case-by-case and sub-optimal.

End-to-end methods [18], [19], [32] learn to map raw sensor input (e.g., images) directly to control signals (e.g., acceleration, steer angle). Bojarski et al. [18] successfully learned a steering policy. However, their model is only suitable for lane keeping and has difficulty in addressing complex scenarios. Codevilla et al. proposed Conditional Imitation Learning (CIL) [19], where the output is conditioned on high-level commands. Furthermore, CILRS [20] is an improved version of CIL. In general, offline imitation learning has difficulty in generalizing to dense traffic due to the instinctive covariate shift problem [30]. Another problem is causal confusion [33], where the model cannot distinguish spurious correlations from true causes in observed training demonstration patterns. A large body of CIL-based work has been proposed to address these issues. Privileged supervisions such as road maps (LBC [21]) or BEV representations (Roach [17]) are used as input. Object-level detections such as vehicles and pedestrians can be integrated into the input, reducing the perceptual burden on DNNs compared to front-view images. Although privileged information can be easily and efficiently accessed in the simulator, retrieving it from real-world observations is non-trivial. To overcome the covariate shift problem, some works [21], [27] employ DAgger [30] to transfer offline imitation learning to online refinement. Alternatively, online/on-policy reinforcement learning is utilized for more exploration, where an offline trained IL agent serves as the initialization of the RL agent (CIRL [16], LSD [34]), or an IL agent imitates a well-trained RL agent (Roach [17]). However, both DAgger and

online RL can only perform effectively in simulation because accessing online demonstrations in the real world is non-trivial. They also suffer from expensive training costs. Besides, a well-designed reward function is crucial for the learned policy [13], which may not reflect real human driving behavior. To overcome the difficulty of defining optimal reward functions, inverse reinforcement learning (IRL) approaches have also been investigated to imitate human driving behaviors [35]–[39]. However, model-based IRL [35], [37], [39] requires perfect knowledge of the system dynamics or transition function, model-free one [38] requires online interaction with environment, and the computational burden of IRL can be heavy as they often iteratively solves forward RL problems with each new reward function derived.

Our work differs from the above works in several ways. First, as is shown in Tab. I, few methods focus on interaction with pedestrians at intersections. Although methods such as [17], [19], [27] deal with urban scenarios containing pedestrians, most pedestrians are generated across the large-scale town in simulation. The probability of autonomous agents interacting with pedestrians at crosswalks is small. However, our work focuses on intersection scenarios where interaction with pedestrians is unavoidable. Second, we take into account the intrinsic difference between lateral and longitudinal control through separate modeling, which is neglected by current end-to-end DIL methods. Third, we learn from the human expert's rather than an autopilot agent's demonstrations in previous work. Furthermore, our strict offline learning scheme is more suitable for real-world deployment.

B. Intersection Navigation with Pedestrian Interaction

Pedestrians are one of the most vulnerable traffic participants at intersections, whose microscopic movement behavior can be modeled by force-based concepts [40]–[43]. It is hard to accurately predict the pedestrians' intentions due to their complex internal and external stimuli (e.g., they may be distracted [44] or suddenly change their mind [45]). Therefore, the interaction between autonomous driving and pedestrians is an important research topic receiving increasing attention in motion planning [46], traffic simulation [43], [47] and intelligent vehicles [5], [48]. Traditional methods analyze the vehicle-pedestrian conflicts using integrated microscopic simulation models [49]. We focus on deep learning methods, which are briefly reviewed below.

Some methods build on Partially Observable Markov Decision Process (POMDP) to solve the navigation problem amongst pedestrians [46], [50]. Bai et al. [50] modeled pedestrian intentions as their target locations. The behavior of the autonomous agent is conditioned on the intention hidden variable. Luo et al. [46] proposed a POMDP approach that models both pedestrian intentions and interactions through a pedestrian motion model predictor. However, online POMDP planners are complained of high computational burden and time inefficiency, limiting their application in real-world autonomous driving platforms. Different from POMDP that explicitly models the interaction, some methods utilize special neural network architectures to capture the interplay between

agents. Yao et al. [51] proposed to use Graph Attention Network (GAT) to model the vehicle-vehicle interaction and vehicle-motorcycle interaction at mixed-flow intersections.

Reinforcement learning is an alternative approach [15], [45], [52]. Bouton et al. [15] presented the combination of a model-checker and deep Q-learning to derive efficient policies with probabilistic safety guarantees. A scene decomposition method is proposed to generalize to more pedestrians. Papini et al. [45] leveraged reinforcement learning to learn a function that specifies the safe speed limit of an autonomous vehicle when interacting with pedestrians. The safe speed function acts as a high-level behavioral directive for the vehicle. However, these methods may be limited by their few agents settings [45] and simple 2D simulation [15], [52]. Although some deep imitation learning works [17], [19], [27] also consider intersection navigation, they don't focus on interaction with pedestrians.

Our work differs from above works in several ways. First, we leverage DIL through human demonstrations to learn interaction without explicitly modeling and inferring the pedestrian intentions (which can be difficult and imprecise). Second, our end-to-end framework avoids the time-consuming computation in online POMDP planning. Third, with the help of high-fidelity 3D simulation, our work can handle more complex and diverse vehicle-pedestrian interaction scenarios.

C. Multi-task Learning

Multi-task learning [53] aims to improve learning efficiency by learning multiple complimentary tasks from shared representations. Many multi-task methods have been proposed for computer vision. For semantic tasks, classification and semantic segmentation are learned in [54]. For geometry and regression tasks, depth, surface normals and semantic segmentation are learned in [55].

Some works build on multi-task learning and learn autonomous driving policies [56]–[58]. [57] used prediction of future actions and states as side tasks and learned together with primary control task. [56], [58] trained the policy together with semantic segmentation task to obtain a meaningful and generic feature space. Our method differs from these methods in several ways. Instead of introducing side tasks that increases training cost, we split the primary task into lateral and longitudinal tasks, which are learned together in multi-task setting. Since two tasks have different properties and units of measure, we estimate their inherent task uncertainty (i.e., homoscedastic uncertainty). Building upon the modeled uncertainty, weights in multi-task loss are learned adaptively. Empirical results demonstrate our effectiveness.

III. METHODOLOGY

A. Scenario

This work studies the scenario of an autonomous driving agent navigating through a densely populated intersection, where it needs to adjust its controls and interact safely with other traffic participants. In this research, we focus on interaction with pedestrians, the generalizability to vehicles is demonstrated in experiments. As shown in Fig. 1, the

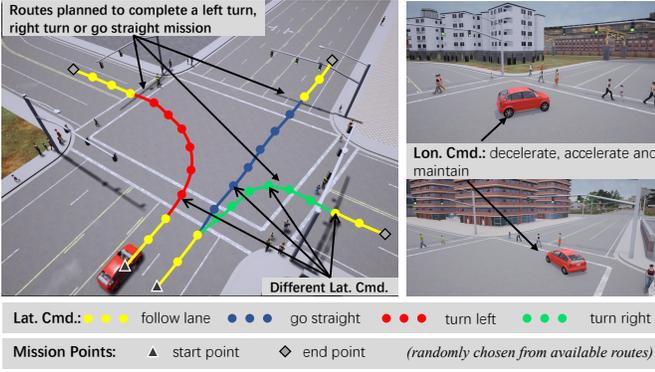


Fig. 1. Illustration of intersection scenarios. Given a planned route and high-level commands, the agent needs to complete three kinds of missions.

autonomous vehicle completes the missions of left turn, right turn and go straight at the intersection, guided by the route from a start point to an end point and commands issued by a higher-level module. To complete a mission, the agent needs to perform a sequence of driving behaviors, hereinafter referred to as commands, each of which is completed by a sequence of control actions. Specifically, lateral commands include follow lane, go straight, turn left and turn right. Longitudinal commands are decelerate, maintain and accelerate.

B. Conditional Imitation Learning (CIL)

This research follows the Conditional Imitation Learning [19] framework to formulate the problem as follows: Human driving demonstration dataset $\mathcal{D} = \{\xi_i\}_{i=1}^N$ consist of N trajectories. Each trajectory ξ_i is composed of a sequence of observation-action pairs $\{(o_t^i, a_t^i, c_t^i)\}_{t=1}^T$, where o_t^i , a_t^i and c_t^i denote the observation, action, and high-level command, respectively. The observations are tuples which include an onboard front-view RGB image I_i^t and scalar value ego speed v_i^t . The actions contain steer angle $a_i^{t, str} \in [-1, 1]$ and acceleration $a_i^{t, acc} \in [-1, 1]$. The goal is to learn a deep neural network policy π parameterized by θ that imitates human driving behavior. The optimal parameters θ^* are obtained by minimizing the imitation cost \mathcal{L} :

$$\theta^* = \arg \min_{\theta} \sum_j \mathcal{L}(\pi(o_j, c_j; \theta), a_j) \quad (1)$$

C. Multi-task Learning

Lateral and longitudinal control are two tasks of very different properties. On the one hand, scene features have different importance in accomplishing each task. Lane markings and road structures are more important for lateral control task while obstacles ahead and ego speed have significant influence on the longitudinal control task. On the other hand, lateral and longitudinal control have different tolerances for vibration in the control actions. Faced with the same scenario, the confidence levels of the lateral and longitudinal controls differ, reflecting the various uncertainties inherent in these tasks.

In multi-task learning, separate deep models are learned for each task and different learning objectives are combined in one loss function [53], [59]. Linear combination is typically

applied by weighting the losses for each individual task using the hand-tuned hyper-parameters [29]. However, the search and tuning of hyper-parameters is not trivial. Since model performance is often sensitive to hyper-parameter, its versatility may be limited in various scenarios.

Following [29], this work formulates simultaneous lateral and longitudinal control learning in a multi-task learning framework, where task-dependent uncertainties are used to weight tasks. These uncertainties are also learned from data and optimized simultaneously with model parameters.

D. Task-dependent Uncertainty Loss

We derive from a single regression task such as learning only lateral or longitudinal control. Let $\pi_{\theta}(s)$ be a DNN policy model with parameter θ , which takes input data s and outputs control action a . The likelihood is modeled as a Gaussian with the mean given by the model output, and the noise scalar σ^2 represents task-dependent uncertainty:

$$p(a|\pi_{\theta}(s)) = \mathcal{N}(\pi_{\theta}(s), \sigma^2) \quad (2)$$

$$-\log p(a|\pi_{\theta}(s)) \propto \frac{1}{2\sigma^2} \|a - \pi_{\theta}(s)\|^2 + \log \sigma \quad (3)$$

Now consider a multi-task problem that yields two outputs a_1 and a_2 . Approximately assuming the independence of two tasks, we have:

$$p(a_1, a_2|\pi_{\theta}(s)) = p(a_1|\pi_{\theta}(s)) \cdot p(a_2|\pi_{\theta}(s))$$

$$= \mathcal{N}(a_1; \pi_{\theta}(s), \sigma_1^2) \cdot \mathcal{N}(a_2; \pi_{\theta}(s), \sigma_2^2) \quad (4)$$

$$-\log p(a_1, a_2|\pi_{\theta}(s)) \propto \frac{1}{2\sigma_1^2} \|a_1 - \pi_{\theta}(s)\|^2$$

$$+ \frac{1}{2\sigma_2^2} \|a_2 - \pi_{\theta}(s)\|^2 + \log \sigma_1 \sigma_2 \quad (5)$$

Consequently, we have the task-dependent uncertainty loss for the multi-task learning of lateral and longitudinal controls:

$$\mathcal{L}(\theta, \sigma_{lat}, \sigma_{lon}) = \frac{1}{2\sigma_{lat}^2} \|a^{str} - \pi_{\theta}^{str}(s)\|^2$$

$$+ \frac{1}{2\sigma_{lon}^2} \|a^{acc} - \pi_{\theta}^{acc}(s)\|^2 + \log \sigma_{lat} \sigma_{lon}$$

where $\pi_{\theta}(s)$ denotes $\pi(o, c; \theta)$, which is composed of three sub DNN models, i.e., a feature encoder π_{θ}^{feat} shared by the lateral and longitudinal conditional modules π_{θ}^{str} and π_{θ}^{acc} . σ_{lat} and σ_{lon} denote the task-dependent uncertainties of lateral and longitudinal controls, respectively. We can interpret the first and second terms in the loss function as the objectives of each individual task, which are weighted by σ_{lat} and σ_{lon} , respectively. Minimizing the loss function with respect to σ_{lat} and σ_{lon} can learn their relative weights from data. For example, large σ_{lat} implies that the lateral control task is inherently more uncertain, then we have a smaller weight of the task, and vice versa. Different from literature where the weights of steer angle and acceleration losses are manually tuned hyper-parameters, our method adaptively learns to balance between them. The last term $\log \sigma_{lat} \sigma_{lon}$ serves as a regularization for preventing σ_{lat} and σ_{lon} from increasing too much.

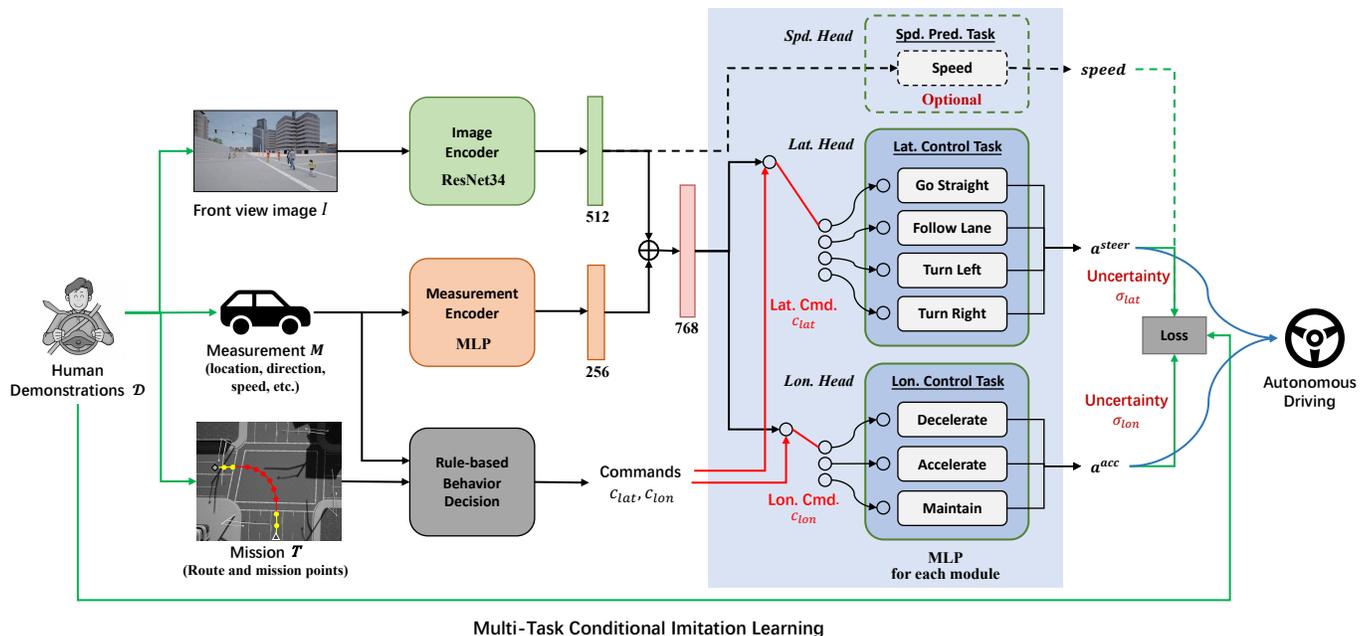


Fig. 2. Our proposed multi-task conditional imitation learning (MTCIL) framework. Two separate branches predict lateral and longitudinal control actions, respectively. Both branches share the same perception representation. For each task, corresponding high-level commands are given by rule-based decision module to select the target submodules. Task-dependent uncertainties are learned to adaptively adjust task weights.

E. Multi-Task Conditional Imitation Learning

The proposed Multi-Task Conditional Imitation Learning (MTCIL) architecture is shown in Fig. 2. We take the single-frame front view image I and the ego velocity value v as the input to the image encoder and measurement encoder, respectively. For image encoders, we evaluate the performance of CarlaNet [19] and ResNet34 [60] in the experiments. The measurement encoder is a multi-layer perceptron (MLP) consisting of three fully connected layers. Following [17], [19], [20], features from two encoders that are at different sizes are concatenated and passed to the control modules. The lateral and longitudinal control tasks are completed by a conditional module, which contains multiple MLPs corresponding to each lateral or longitudinal command. Given current commands c_{lat} and c_{lon} determined by a rule-based model, the corresponding modules are switched on and responsible for predicting control actions a^{str} and a^{acc} . Similar to [20], an optional branch predicting the current speed can be added in our framework (see Fig. 2), which encourages the perception module to extract visual cues that reflect the scene dynamics.

Compared with literature work that uses a single deep model to output both lateral and longitudinal control actions, separate modeling can greatly improve the performance of longitudinal control, which is crucial for dense intersections with pedestrian interactions, as shown in experiments. Furthermore, combining both controls into a multi-task framework can improve efficiency by sharing encoders. Meanwhile, the performance is balanced by weighting tasks according to task-dependent uncertainties, which can be learned automatically from data. Note that this framework can be easily extended to allow more side-tasks such as semantic segmentation.

IV. A NEW BENCHMARK: INTERSECTNAV

We propose a new benchmark named IntersectNav in this section. As shown in Tab. II, unlike previous benchmarks [20], [31], [61], we focus on intersections that extensively challenge the capability of the autonomous agent to interact with pedestrians. Although some benchmarks include a certain number of pedestrians, they are randomly scattered throughout the city (e.g., on pavements far away from ego car), resulting in much less interaction with pedestrians at intersections than ours. Specifically, we use CARLA [31] driving simulator 0.9.7 for realistic 3D simulation. Compared to the 0.8.X version used in previous benchmarks [20], [31], the graphics and simulation behavior in 0.9.7 has been further improved, making it more complex and realistic.

A. Scenarios

1) *Scene and Mission*: Demonstrated in Fig. 3, 6 different US-style intersections from two towns are selected for evaluation. Four scenes are used for train and validation while the other two are reserved for test. We configure the available start and goal points, based on which 41 reference paths in total are generated through standard A* algorithm. Reference paths in test scenes and train scenes are not necessarily the same due to their different road structures. The benchmark adopts an episodic setup. For each episode at an intersection, given a route randomly chosen from available configurations, the ego vehicle is generated at the start point of the route. The weather is randomly selected from train weathers {ClearNoon, CloudyNoon, WetNoon, HardRainNoon} or new weathers {ClearSunset, CloudySunset, WetSunset, HardRainSunset}. Three missions are considered, i.e., performing left turn/go straight/right turn and navigate through the intersection

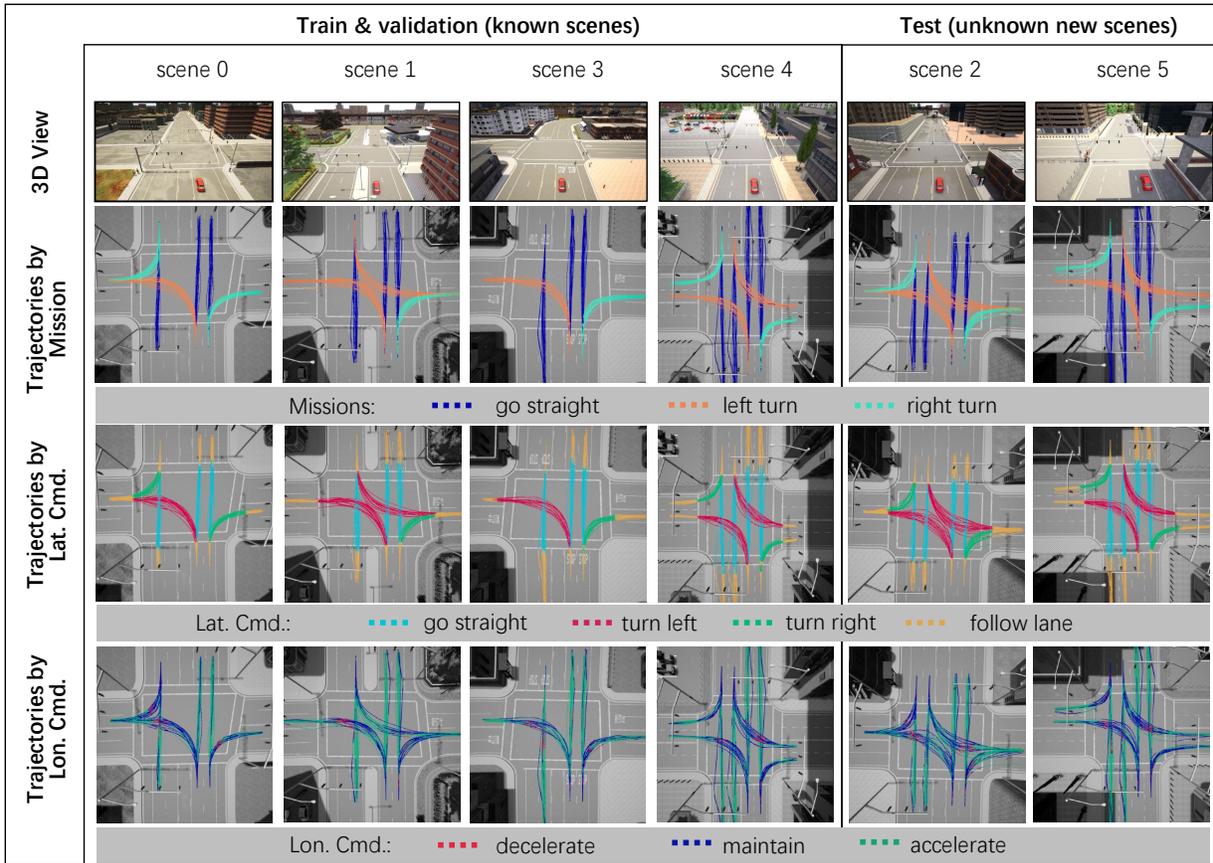


Fig. 3. Benchmark scenes and human demonstration trajectories of dataset **Ped-Only**.

(c.f. Fig. 3 row 2). The benchmark supports simulation of pedestrians and environmental vehicles, with which the autonomous agent needs to interact.

2) *Simulation of Pedestrians*: CARLA provides APIs for users to specify the pedestrians' appearances, destinations and walking speeds. We refer to the former work [19] and set our parameters as follows. In each episode, a random number of 20-30 pedestrians are generated to walk through the crosswalks to reach randomly assigned goals. The maximum walking speed of each pedestrian is randomly chosen in the range [1.3, 1.8] m/s. For now, the pedestrians' navigation and their interactions with vehicles are handled by CARLA built-in AI controllers. We assign each pedestrian a random target point to lead a path across the street. If the path is blocked, the pedestrian will wait a few seconds. If the path is still blocked, a new target point is assigned randomly by CARLA, leading to a new path across the road or back.

3) *Simulation of Environmental Vehicles*: Apart from pedestrians, the simulation environment can be configured to introduce environmental vehicles. In each episode, a random number of 6-15 environmental vehicles are generated at random spawn points around the intersection area, which are controlled by CARLA built-in autopilot agents. During the navigation through intersections, they would avoid collision, follow traffic lights and drive at target speed of 25Km/h, leaving at least 5m between themselves and leading vehicles. In this setting, the autonomous agent needs not only interact

with pedestrians at crosswalks, but also other vehicles nearby.

B. Evaluation Metrics

During the closed-loop simulation for evaluation, the autonomous agent and pedestrians are initialized according to protocols described above. At each simulation step, current observations and commands are fed into the control model. Since CARLA simulator only accepts steer angle and acceleration values in the range of $[-1.0, 1.0]$, the network's control outputs are clipped by this range and passed to the actuators in CARLA. The backend engine simulates the world dynamics and moves on to the next time step. This process iterates until an episode is terminated. Since human drivers may operate differently in the same or similar situations, i.e., there is stochasticity. There is no single correct answer for human driver control. Therefore, this work uses the following metrics to assess the reliability of the learned model.

1) *Metrics of Task Completeness*: The performance in completing the missions is one important evaluation metric. We consider four possible events that the episode ends with: collision, lane invasion, timeout and success. Detailed information can be found in Tab. II. The success rate is calculated as the ratio of the number of successful episodes to the total number of simulated episodes. The other rates are calculated similarly. The collision rate, timeout rate and lane invasion rate reflect an agent's performance in terms of safety, efficiency and traffic rules, respectively.

TABLE II
CONSIDERED EVENTS IN OUR BENCHMARK AND COMPARISON TO OTHER BENCHMARKS

Benchmark	Pedestrians Number	Failure Events	Definition of Success	Metrics
Original CARLA Benchmark [31]	rare pedestrians at each intersection	1. collision with static object/car/pedestrian 2. opposite lane 3. sidewalk	the agent reaches the goal regardless of what happened during the episode	1. success rate 2. avg. distance travelled between infractions
NoCrash Benchmark [20]	few pedestrians at each intersection	1. collision with static object/car/pedestrian 2. timeout 3. traffic light violations	the agent reaches the goal under a time limit without colliding with any object	1. success rate 2. collision rate 3. timeout rate
CARLA Leaderboard [61]	not mentioned	1. collision with static object/car/pedestrian 2. running a red light/stop sign 3. timeout	not applicable	1. driving score 2. route completion rate 3. infraction penalty
IntersectNav Benchmark (Ours)	20-30 pedestrians at each intersection	1. collision with static object/pedestrian 2. lane invasion (invading other lanes more than five times or drive out of the road) 3. timeout (failure to arrive at the goal within 1000 steps)	the agent reaches the goal under a time limit without any failure events happened	1. success rate 2. collision rate 3. timeout rate 4. lane invasion rate 5. other metrics reflecting control quality (see Tab.III-E)

TABLE III
METRICS OF CONTROL QUALITY

Metric(Unit)	Description	Formula
Intense Actions(#)	Average times of the autonomous agent's intense actions that are too large	$\frac{1}{N} \sum_{i=1}^N \sum_{t=1}^{T_i} \mathbb{1}[a_t^{str} \notin (\epsilon_{str}^{low}, \epsilon_{str}^{up}) \text{ or } a_t^{acc} \notin (\epsilon_{acc}^{low}, \epsilon_{acc}^{up})]$
Disruption to Pedestrians(#)	Average times of pedestrians $p_j, j = 1 \dots M_i$, disrupted by the autonomous agent (e.g., emergent stop in close range)	$\frac{1}{N} \sum_{i=1}^N \sum_{t=1}^{T_i} \sum_{j=1}^{M_i} \mathbb{1}[p_j.get_disrupted() = True]$
Deviation from Waypoint(m)	Mean location \vec{loc}_t 's deviation from centerline represented by the current nearest waypoint \vec{wp}_t^c and next waypoint \vec{wp}_t^n	$\frac{1}{N} \sum_{i=1}^N \sum_{t=1}^{T_i} \frac{(\vec{wp}_t^n - \vec{wp}_t^c) \times (\vec{loc}_t - \vec{wp}_t^c)}{ \vec{wp}_t^n - \vec{wp}_t^c }$
Deviation from Destination(m)	Mean final location \vec{loc}_{T_i} 's deviation from the goal location \vec{g}_i	$\frac{1}{N} \sum_{i=1}^N \vec{loc}_{T_i} - \vec{g}_i $
Heading Angle Deviation(°)	Mean final heading θ_{T_i} 's deviation from lane direction δ_i at the episode ending	$\frac{1}{N} \sum_{i=1}^N \theta_{T_i} - \delta_i $
Total Step(#)	Average total steps for each episode	$\frac{1}{N} \sum_{i=1}^N T_i$

2) *Metrics of Control Quality*: Aside from above metrics that consider task completion, we also define metrics to evaluate the model's control quality. The details are provided in Tab. III-E. By counting the number of intense actions and pedestrians getting disrupted, we can further analyze the driving comfort of the autonomous agent and its influence on pedestrians. The deviations consider the control precision while total steps measures the efficiency of the learned model.

C. Human Demonstration Dataset

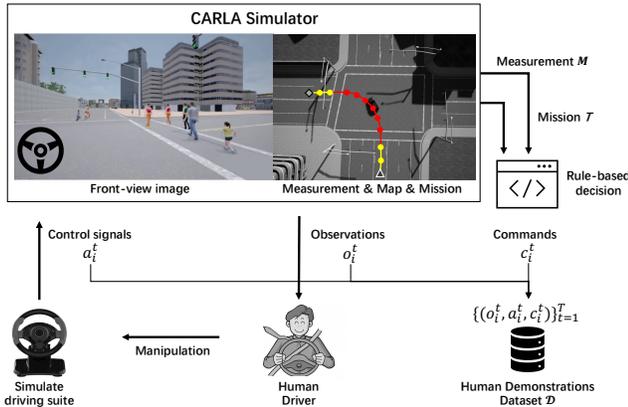


Fig. 4. Data collection procedure. The human operator manipulates the driving suite to demonstrate the mission in CARLA simulator.

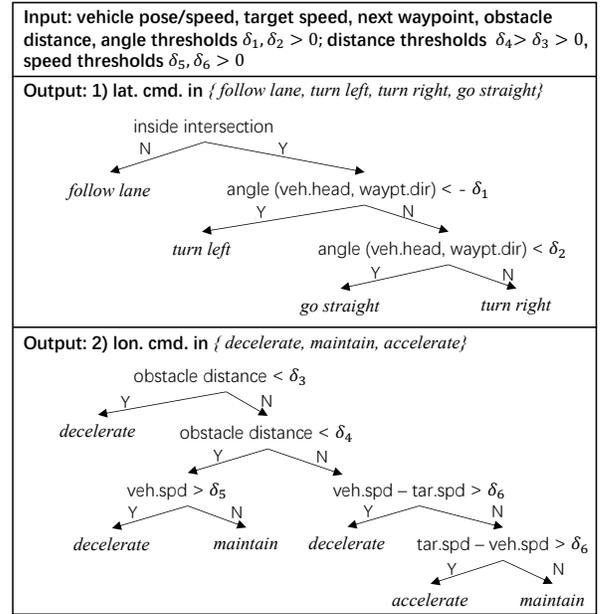


Fig. 5. Rule-based decision module.

1) *Data Collection Procedure*: As is shown in Fig. 4, we collect human driving demonstrations in CARLA through the driving suite that includes a dual-motor force feedback wheel

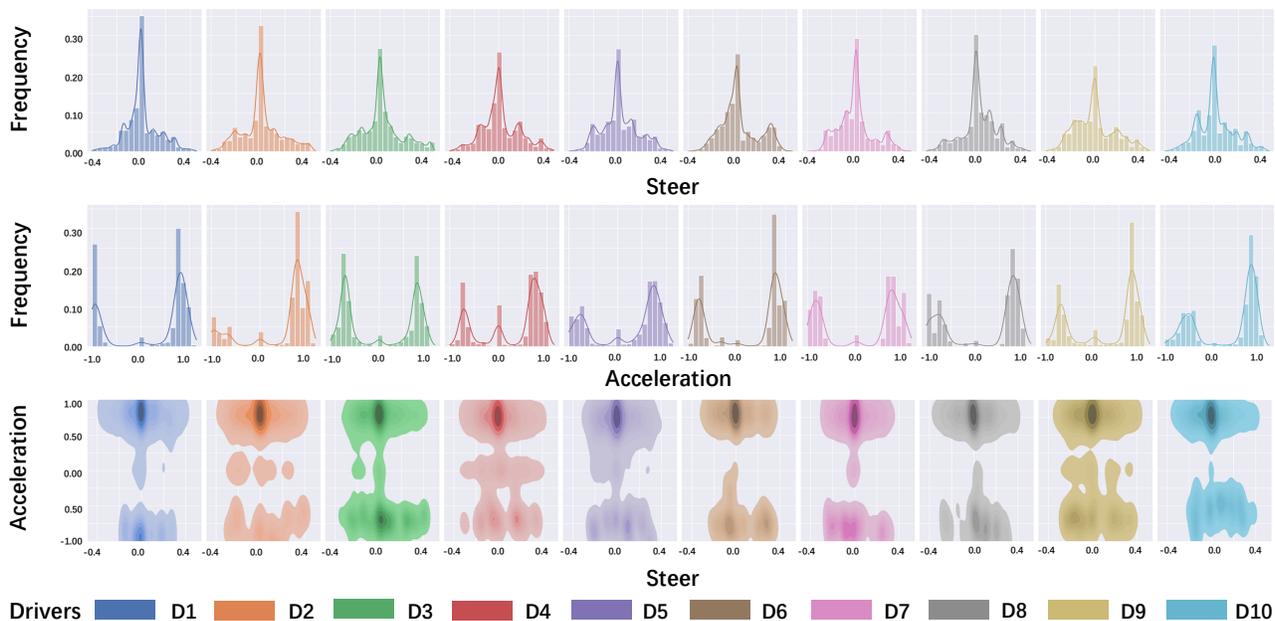


Fig. 6. The driver profile of dataset Mul-Dri. The top and middle are histograms of steer and acceleration, respectively. The bottom is the bivariate distributions of steer and acceleration. The darker the color, the higher the kernel density.

TABLE IV
STATISTICS OF DATASETS

Category	Datasets		
	Ped-Only	Ped-Veh	Mul-Dri
Frames (Trajectories) by Scene			
Scene0	12113 (150)	13980 (167)	16263 (234)
Scene1	10589 (147)	11572 (135)	14949 (236)
Scene2	16737 (199)	-	23400 (313)
Scene3	8312 (119)	-	12278 (200)
Scene4	13350 (186)	-	23202 (305)
Scene5	14682 (187)	-	19568 (301)
Frames (Trajectories) by Mission			
Left turn	25952 (229)	8609 (70)	35314 (383)
Go straight	25253 (515)	9750 (152)	42885 (821)
Right turn	24578 (244)	7193 (80)	31461 (385)
Frames by Lat. Cmd.			
Follow lane	34812	11549	46427
Turn left	18015	6081	25101
Turn right	13719	4587	18591
Go straight	9237	3335	19541
Frames by Lon. Cmd.			
Decelerate	16258	5723	27784
Maintain	25432	4529	5793
Accelerate	34093	15300	76083

* Datasets **Ped-Only** and **Ped-Veh** were collected from one driver. Dataset **Mul-Dri** was collected from ten drivers.

and a floor pedal. The human driver is provided with real-time front-view RGB images and bird-view images. Reference routes are projected onto the bird-view map to provide the mission information. Real-time high-level driving commands from a rule-based decision module (cf. Fig. 5) are provided for reference. In each episode, the operator is asked to keep a preferred 20 km/h speed and drive through the intersection following the high-level commands.

At random time steps, a triangular perturbation is added to the human's steer angles with probability 0.1. This technique aims to collect experts' demonstrations that recover from perturbations. Once an episode is over, the operator can review

this episode's metrics in Tab. III-E. Data from successful episodes with good control metrics is stored. We record raw sensor data (e.g., RGB/depth images, ego's speed and poses, etc.) along with the expert's demonstrations (e.g., control steer angle/throttle/brake, corresponding high-level commands). The observation o^t , expert action a^t and high-level commands $c^t = (c_{lat}^t, c_{lon}^t)$ are bounded together as one tuple (o^t, a^t, c^t) , which serves as a training sample. Meta task information such as town/scene/pose index and weather are also recorded.

2) *Dataset Statistics:* Focusing on interactions with pedestrians, a dataset "**Ped-Only**" was developed in scenarios containing only pedestrians. Over 30 hours of driving data was collected by an experienced human driver at six intersections, containing more than 950 trajectories. The collected trajectories are shown in Fig. 3 and the detailed statistics are given in Tab. IV.

To evaluate the method at more general scenes, a supplementary dataset "**Ped-Veh**" was developed containing both pedestrians and environmental vehicles. Over 9 hours of driving data was collected by the human driver at two intersections, i.e., scene 0 & 1, containing more than 300 trajectories. The statistics are given in Tab. IV.

To evaluate the effect of various driving styles, a dataset "**Mul-Dri**" was developed. It contains data from 10 human drivers, who need to interact with pedestrians and environmental vehicles while driving through intersections. The 10 drivers range in age from 20s to 50s, with one female and nine male drivers. By statistically analyzing the steer and acceleration parameters of each driver in Fig. 6, we can see that 10 drivers demonstrate different driving profiles. Near 1600 valid trajectories were collected, where each driver drove about 5 hours and provided 160 trajectories under six intersections. The statistics are given in Tab. IV.

All datasets were developed under four weather conditions,

i.e., *ClearNoon* : *CloudyNoon* : *WetNoon* : *HardRainNoon* in a ratio of 0.45 : 0.17 : 0.18 : 0.19. The total size of our datasets is much larger than earlier ones [19], [20], which contain 2 and 10 hours of driving data, respectively. However, deep imitation learning may still face changes in the distribution between training and testing, which reduces its reliability in new scenes that are very different from the training ones. We leave this issue for future work.

D. Dataset Bias Cleaning

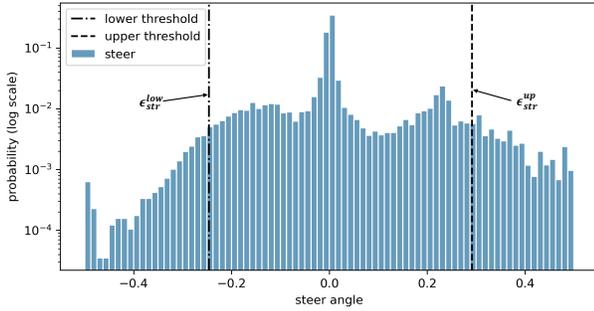


Fig. 7. Histogram of steer data. The value is normalized to $[-1, 1]$.

During data collection, biased data happens due to careless driving, such as large deviation from centerlines, too close to pedestrians, etc. Therefore, after data collection, each episode is checked for bias avoidance. On the one hand, driving parameters such as steer angle a_{str} , acceleration a_{acc} , centerline-deviation d_{ctr} of each frame are examined. Bias-frame-numbers are counted if one of them is out of range (i.e., $a_{str} \notin (\epsilon_{str}^{low}, \epsilon_{str}^{up})$, $a_{acc} \notin (\epsilon_{acc}^{low}, \epsilon_{acc}^{up})$, $d_{ctr} > \epsilon_{ctr}$). If the bias-frame-number exceeds a predefined limitation ($> \epsilon_{bad_frames}$), the episode is invalidated. On the other hand, if the episode length is too long ($> \epsilon_{len}$) representing inefficient driving, or at any a frame, the ego vehicle has a too close distance ($< \epsilon_{dis2ped}$) with a pedestrian representing dangerous driving, the episode is invalidated.

The thresholds used above are predefined by analyzing the statistics of the collected driving data. For a certain parameter k , e.g., steer angle in Fig. 7, a histogram ρ_k is generated to profile the data distribution. The upper and lower bounds of the confidence interval (e.g., ϵ_{str}^{up} and ϵ_{str}^{low}) corresponding to the confidence level 0.95 are chosen as the thresholds. The episode length threshold ϵ_{len} differs across missions such as left- or right-turn and intersection sizes, whereas it is correlated with the length of reference route l_{ref} . Therefore the threshold $\epsilon_{len}(i, j)$ of intersection i and mission j is estimated as $2 * l_{ref}(i, j)$. As for $\epsilon_{dis2ped}$, its histogram is generated on data frames which have pedestrians in the front and within a certain distance $d_{max} = 10m$.

E. Safety Criticality Analysis of the Scenes

Referring to the works [62], [63], we evaluate safety criticality of the generated scenes by focusing on pedestrian interaction during each drive through an intersection scene. The distance (dis) between the ego vehicle and the nearest pedestrian is used to assess the safety criticality of pedestrian

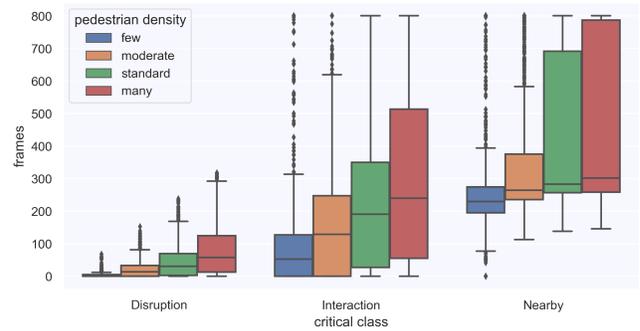


Fig. 8. Safety criticality analysis of the generated scenes.

interactions, where criticality is empirically divided into three levels: pedestrian is nearby (dis $< 10m$), pedestrian interaction is needed (dis $< 5m$) and pedestrian is disrupted (CARLA alarm). For each driving trajectory, the number of the frames at each criticality level is counted. The average numbers are used to evaluate the criticality of each generated driving scene. To evaluate the performance of intersection navigation with pedestrian interaction, scenes of different pedestrian numbers M_{ped} are designed. They are:

- **Few:** $M_{ped} \in [0, 10)$.
- **Moderate:** $M_{ped} \in [10, 20)$.
- **Standard:** $M_{ped} \in [20, 30)$, same setting as we used for data collection and closed-loop evaluation in Section V-C.
- **Many:** $M_{ped} \in [35, 50)$, which requires the agent to strike a good balance between safety and efficiency.

Safety criticality is evaluated at Fig. 8 of the generated scenes with above pedestrian densities. Comparing to the existing CARLA benchmarks [20], [31], whose pedestrian densities are similar to few, the scenes of the proposed IntersectNav benchmark are more critical.

TABLE V
DATASET AND SCENE CONFIGURATION IN EXPERIMENTS

Experiment	Dataset	Train scenes	Evaluation scenes
Exp. 1,2,5	Ped-Only	scenes 0,1,3,4	train scenes 0,1,3,4 new scenes 2,5
Exp. 3	Ped-Veh	scenes 0,1	train scenes 0,1 new scene 2,5
Exp. 4	Mul-Dri	scenes 0,1,3,4	train scenes 0,1,3,4 new scenes 2,5

V. EXPERIMENT

A. Experiment Design

Experiments are designed to evaluate the autonomous agent's reliability on the following aspects: 1) performance in completing driving tasks 2) control quality 3) generalization to new conditions, which are introduced below. Tab. V gives detailed scene configuration in experiments. Train scene denotes the scene where the training data is collected. New scene denotes the scene that is not experienced in training.

1) *Exp.1 Interaction with Pedestrians:* This experiment focuses on interaction with pedestrians, where models are trained on **dataset Ped-Only** and evaluated on scenes with only pedestrians. The performance, namely, the ability to complete tasks and control quality are evaluated. Generalization to new scenes and weathers is also analyzed.

2) *Exp.2 Generalization to Different Numbers of Pedestrians*: It aims to evaluate the models' generalization ability to different numbers of pedestrians. Models are trained on **dataset Ped-Only** and evaluated on scenes having various pedestrians' density, e.g., simulating scenes from different countries.

3) *Exp.3 Dealing with both Pedestrians and Vehicles*: It is designed to examine the generalizability of the proposed method to handle scenarios with both pedestrians and vehicles. Models are trained on **dataset Ped-Veh** and evaluated on scenes where pedestrians and environmental vehicles are both present.

4) *Exp.4 Learning from Multiple Drivers*: It is designed to examine the generalizability of the proposed method to a variety of driving styles. Unlike other experiments that learn models from a single driver's data, this experiment uses **dataset Mul-Dri** and learns from 10 drivers with various behavioral styles. As in Exp. 3, the evaluation is performed in scenes with both pedestrians and environmental vehicles.

5) *Exp.5 Ablation Studies*: It is designed to investigate the importance of different components of the proposed method. Models are trained on **dataset Ped-Only** and evaluated on scenes with only pedestrians.

B. Training Details

All models are trained using Adam optimizer [64] with an initial learning rate $2e-4$, which will be divided by 10 if validation loss stops decreasing for more than 5 epochs. Dropout is used after fully connected layers with a probability of 0.5. Each mini-batch contains 120 samples, which are randomly sampled from the shuffled train set. We follow Codevilla et al. and employ a 200×88 image resolution for CarlaNet [19] perception backbone. For ResNet34 backbone, we resize the image to resolution 224×224 . If specified, online image augmentation is performed during training, which includes Gaussian blur and noise, dropout, adjust of brightness, and contrast, etc. We follow the authors of CIL [19] and use the same parameters regarding with image augmentation.

TABLE VI
EXP.1 - TASK COMPLETION RESULTS

Models	Succ. Rt.	Time. Rt.	Lane. Rt.	Colli. Rt.
<i>TS & TW</i>	(%) \uparrow	(%) \downarrow	(%) \downarrow	(%) \downarrow
CIL	57.3 ± 2.5	21.1 ± 2.6	5.3 ± 1.7	16.3 ± 2.6
CILRS	67.5 ± 2.7	9.1 ± 3.5	6.4 ± 2.2	17.0 ± 2.8
Ours	91.2 ± 2.0	1.6 ± 2.6	4.5 ± 1.8	2.7 ± 1.9
<i>TS & NW</i>				
CIL	52.5 ± 3.1	21.9 ± 3.4	4.8 ± 1.4	20.8 ± 2.5
CILRS	46.7 ± 5.3	33.9 ± 5.6	2.1 ± 1.4	17.3 ± 2.2
Ours	88.6 ± 2.0	1.9 ± 2.6	6.6 ± 1.8	2.9 ± 1.9
<i>NS & TW</i>				
CIL	50.4 ± 2.4	23.3 ± 4.2	4.6 ± 3.1	21.7 ± 3.1
CILRS	53.3 ± 2.8	28.7 ± 3.3	2.5 ± 2.4	15.5 ± 2.8
Ours	88.8 ± 3.7	3.1 ± 2.6	4.6 ± 1.6	3.5 ± 1.6
<i>NS & NW</i>				
CIL	40.8 ± 4.9	27.1 ± 8.5	1.3 ± 1.7	30.8 ± 5.7
CILRS	32.1 ± 5.5	46.3 ± 6.1	0.4 ± 0.8	21.2 ± 3.3
Ours	86.8 ± 3.6	3.8 ± 2.5	4.7 ± 1.6	4.7 ± 1.6

¹ Mean and standard deviation over 5 evaluation seeds.

² **Succ. Rt.**: Success Rate, **Time. Rt.**: Timeout Rate, **Lane. Rt.**: Lane Invasion Rate, **Colli. Rt.**: Collision Rate.

³ TS/NS: Train/New Scene, TW/NW: Train/New Weather.

TABLE VII
EXP. 1 - SUCCESS RATES OF THREE NAVIGATION TASKS

Model	Left Turn	Go Straight	Right Turn
CIL	29.3 ± 4.5	93.8 ± 3.5	35.6 ± 13.9
CILRS	39.6 ± 6.5	96.4 ± 2.0	54.4 ± 9.6
Ours	74.2 ± 6.5	100.0 ± 0.0	93.3 ± 6.7

¹ Mean and standard deviation over 5 evaluation seeds.

C. Exp.1 Interaction with Pedestrians

Since offline and online methods cannot be directly compared, this work focuses on offline methods. We choose CIL [19] and CILRS [20] as our baselines where no additional supervisions (e.g., reconstructions, BEV representations) apart from expert demonstrations are used. Our reported multi-task model uses ResNet34 backbone and uncertainty loss. We conduct five evaluation runs using five specified random seeds. During each run, multiple episodes for each route in our benchmark are simulated to calculate the average metrics.

1) *Results of Task Completion*: The task completion evaluation results are shown in Tab. VI, where the values represent mean and standard deviation over five evaluation seeds. Under train scenes and weathers, CIL and CILRS have the highest success rates (57.3% and 67.5%) compared to other conditions. However, both models suffer from high collision rates ($\sim 16\%$). The timeout rate of CIL is even twice higher than that of CILRS. We regard this as the inertia problem [20], where the model creates a spurious correlation between low speed and no acceleration, inducing excessive stopping and difficult restarting. CILRS alleviates this problem by introducing the speed prediction branch. These failures show that baselines have difficulty in learning longitudinal control under interactive scenarios. Compared with baselines, our method has much lower timeout rate and collision rate and achieves 90% success rate, demonstrating the effectiveness of separate modeling of lateral and longitudinal control.

We further report success rates of three navigation tasks under train scene and train weather in Tab.VII. Intuitively, left turn is the most difficult due to its longest decision and control procedure. The right turn is next while go straight is the easiest. The statistics show that the success rate gap between turn tasks and go straight tasks is much bigger than that between left turn and right turn. Three tasks all achieve high performance in go straight under train conditions. However, baselines perform poorly in turn tasks, resulting in lower overall success rates than ours.

Shown in Tab.VI, all models experience performance degradation to varying degrees when generalizing to new scenarios (new scenes or new weathers). This can be interpreted as the inherent covariate shift problem of imitation learning. Interestingly, the success rate reduction of CILRS is more sensitive to weather changes (20.8%) than scene changes (14.2%). Besides, CILRS's performance degradation problem is the most severe among the compared methods. When transferring to new scene and new weathers, the success rate reduction of CILRS (35.4%) is much higher than that of CIL (16.5%). Our method are more robust to new scenes and weathers with a drop in success rate of only 4.4% compared to the baselines, which have significantly degraded performance.

TABLE VIII
EXP. 1 - CONTROL QUALITY RESULTS

Models	Intense Actions	Disruption to Pedestrians	Deviation from Waypoint	Deviation from Destination	Heading Angle Deviation	Total Steps
	#, ↓	#, ↓	m, ↓	m, ↓	°, ↓	#, ↓
<i>TS & TW</i>						
CIL	0.440 ± 0.108	88.848 ± 35.771	1.988 ± 0.233	7.365 ± 0.272	17.187 ± 3.157	385.013 ± 18.176
CILRS	1.300 ± 0.503	75.448 ± 29.667	1.107 ± 0.085	4.598 ± 0.375	14.271 ± 1.615	310.642 ± 33.795
Ours	0.000 ± 0.000	17.219 ± 17.671	0.520 ± 0.029	1.253 ± 0.396	4.368 ± 0.466	308.376 ± 14.951
<i>TS & NW</i>						
CIL	0.064 ± 0.035	34.435 ± 7.286	1.527 ± 0.147	9.114 ± 0.814	17.467 ± 1.143	375.848 ± 24.455
CILRS	0.003 ± 0.005	124.845 ± 54.668	1.117 ± 0.126	9.123 ± 0.949	20.901 ± 1.420	482.261 ± 39.933
Ours	0.000 ± 0.000	17.224 ± 17.680	0.520 ± 0.029	1.253 ± 0.396	4.369 ± 0.464	308.392 ± 14.944
<i>NS & TW</i>						
CIL	0.167 ± 0.137	110.867 ± 39.536	1.853 ± 0.376	9.502 ± 0.600	21.407 ± 1.808	376.825 ± 16.704
CILRS	0.267 ± 0.077	192.037 ± 36.298	1.361 ± 0.166	7.922 ± 1.098	19.161 ± 3.063	514.867 ± 29.989
Ours	0.000 ± 0.000	36.458 ± 31.388	0.581 ± 0.012	1.390 ± 0.472	5.437 ± 0.679	339.292 ± 16.309
<i>NS & NW</i>						
CIL	0.104 ± 0.077	35.142 ± 8.137	1.159 ± 0.148	10.471 ± 0.942	24.189 ± 2.445	425.917 ± 64.926
CILRS	0.004 ± 0.008	144.762 ± 46.157	0.868 ± 0.165	14.279 ± 1.106	26.983 ± 0.170	578.217 ± 40.494
Ours	0.000 ± 0.000	37.492 ± 32.246	0.582 ± 0.012	1.426 ± 0.381	5.988 ± 0.111	346.625 ± 15.805

¹ Mean and standard deviation over 5 evaluation seeds.

³ TS/NS: Train/New Scene, TW/NW: Train/New Weather.

2) *Results of Control Quality*: Evaluation results of control quality are provided in Tab. VIII, which demonstrate that our method has better control quality than baselines.

On the one hand, driving comfort is important to passengers. Our model performs best in intense actions under either train or new scenarios, which means rare violent driving behaviors. While CILRS also has small intense actions values under new weathers, it doesn't perform as well in train weathers. Autonomous agents should follow social norms to interact friendly with pedestrians. The disruption to pedestrians is an indicator that reflects the degree of influence of autonomous agents on pedestrians. Tab. VIII shows that our proposed method achieves the best performance under most conditions.

On the other hand, in all settings, our method deviates from waypoints by nearly 0.55 m, which is about half of the baselines. Meanwhile, our deviation from destination remains low when generalizing to new scenes and weathers, while other models exhibit a large increase. As for heading angle deviation, our model outperforms others, achieving a small error under 6 degrees. These results demonstrate that our model can learn more precise control policies than the baselines. Since other models fail to generalize and result in large timeout rates, their total steps under test settings are much more than ours, indicating the higher efficiency of our method.

3) *Case Study*: Fig. 9 illustrates a case study of closed-loop simulated trajectories under train and new scenarios. In each subfigure, eight-episode trajectories are plotted in different colors, representing success and failure events. The trajectory points are plotted every four time steps in simulation. The positions of pedestrians disrupted by the autonomous agent are also marked with black squares. The trajectories of our model perform the best in terms of task completion and control precision. At train scenes, baselines tend to collide into pedestrians or drive off the road. Besides, they cause more disturbance to pedestrians than our method. When dealing with new scenes, CIL and CILRS fail to generalize due to the high probability of timeout.

Fig. 10 further shows two episodes of CILRS and ours from Fig. 9 for comparison. The tasks are right turn at train

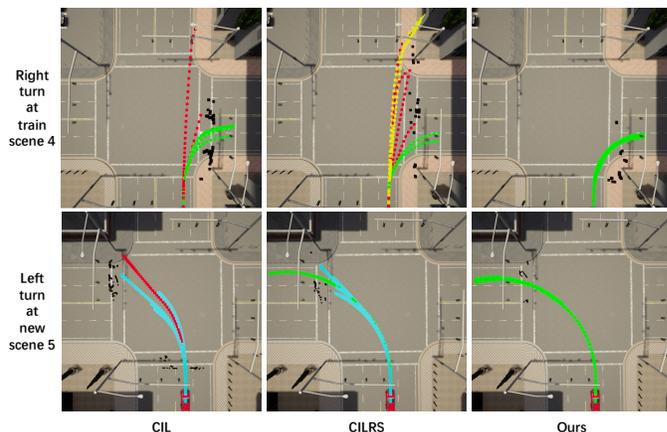
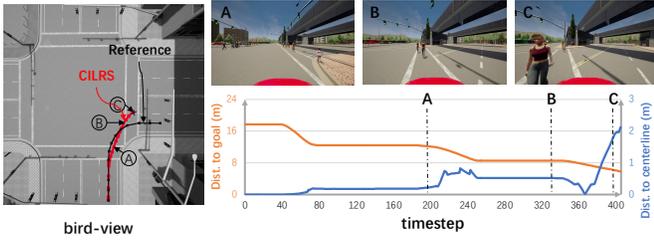


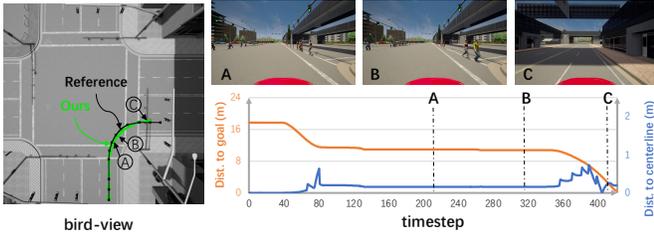
Fig. 9. Case study of the driving trajectories in Exp. 1 at train and new scenes. Eight trajectories are plotted (may be stacked) in each subfigure for illustration. Colors of trajectories: success, timeout, lane invasion, collision. Black squares mark the location of pedestrians that are disrupted by the autonomous vehicle.

scene 4. Each subfigure plots the bird-view trajectories and shows the front-view camera images in chronological order. The plot of deviations (i.e, distance to goal and distance to centerline) is also provided. In Fig. 10(a), CILRS failed to react to the front pedestrians who were very close, resulting in collision. Besides, its deviation from the reference centerline increases to 2m. In Fig. 10(b), our model waited for the pedestrians to avoid collision and then completed the mission. For comparison, our distance to centerline kept at a relatively low-level ($< 0.5m$).

However, it is difficult to analytically explain the failure cases. On the one hand, deep learning models suffer from poor interpretability due to their black-box nature. On the other hand, closed-loop evaluation is essential for the control-level driving policy study, whereas the stochasticity of closed-loop evaluation makes it hard to repeat failure cases. Furthermore, failure is not caused by a single frame's prediction but by the compounding error in the driving sequence.



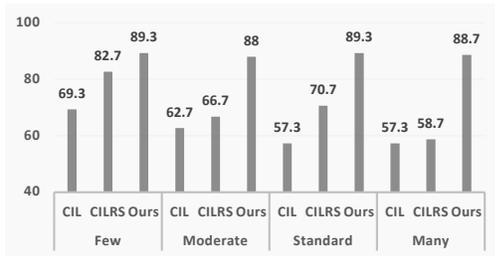
(a) collision episode of CILRS, right turn



(b) success episode of Ours, right turn

Fig. 10. Case study of the collision and success episodes for right-turn at train scene 4 in Fig. 9. In each subfigure, bird-view trajectories are plotted in left. Front-view camera images in chronological order are shown in right-top. The plot of distance to goal and centerline is shown in right-bottom.

D. Exp.2 Generalization to Different Numbers of Pedestrians



(a) Success rate



(b) Failure rates

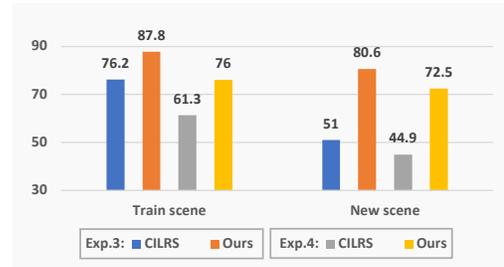
Fig. 11. Exp.2 - The results of interaction with various pedestrian density.

In this experiment, models trained on dataset Ped are tested on scenes having various pedestrians' densities (c.f. IV-E) to examine their generalization ability.

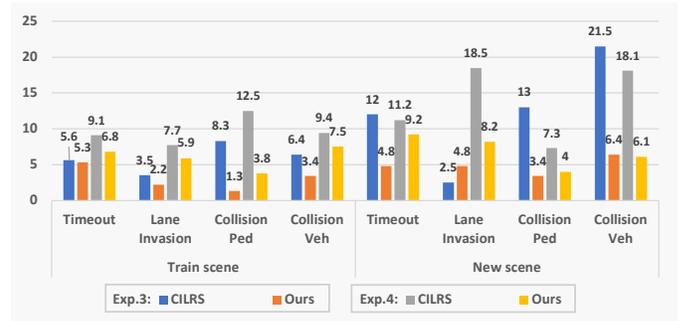
Fig. 11(a) compares between models their success rates under different pedestrians number settings. Failure events rates are given in Fig. 11(b). The reported values are mean over 3 evaluation seeds from train scenes and train weathers. Our model achieved the highest success rates, lowest timeout and collision rates across the different settings. This reflects

that our model is safer and more efficient than the other two. One interesting phenomenon is that models have different sensitivity to different pedestrian numbers. As the pedestrians' number increases from few to many, the performance of baselines shows a nearly monotonic decreasing trend. Our model's performance does not decrease too much when generalizing to either fewer or more pedestrians.

E. Exp.3 Dealing with both Pedestrians and Vehicles



(a) Success rate



(b) Failure rates

Fig. 12. The results of learning from a single driver (Exp.3) and multiple drivers (Exp.4).

This experiment is conducted to examine the generalizability of our method to handle scenarios with pedestrians and vehicles' both presence. Models were trained from scratch on the dataset Ped-Veh and evaluated through closed-loop simulation at train scenes 0 & 1 and new scene 2. Fig. 12 reports the evaluation results on scenarios with both pedestrians and vehicles. The reported values are mean over 5 evaluation seeds. Despite the slight decrease in performance when generalizing to vehicles, our method still outperforms CILRS in success rate. This demonstrates our method's ability to generalize to more realistic complex environments. However, both methods may collide with environmental vehicles. The problem is distinct when testing CILRS on new scenes, whose collision into vehicles rate increases a lot.

F. Exp.4 Learning from Multiple Drivers

Driver heterogeneity has been confirmed to exist in human driving behaviors by many studies [65]. Different drivers may follow sub-optimal policies, exhibiting different driving styles. In the same situation, their actions may conflict with each other since different drivers may lead to different decisions. This experiment further explores the performance of the proposed

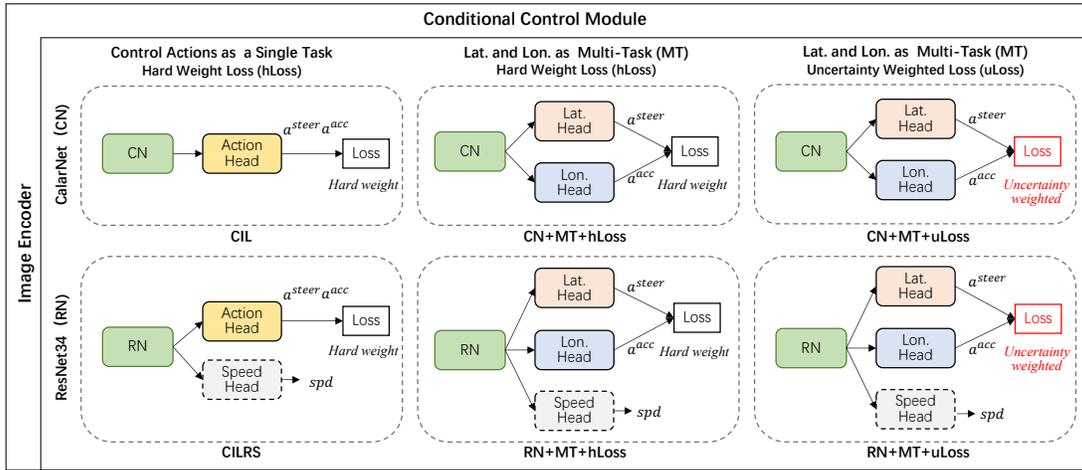


Fig. 13. Models for comparison in Exp.5 ablation studies.

TABLE IX
EXP. 5 - ABLATION STUDY TASK COMPLETION RESULTS

Models	Succ. Rt. (%) \uparrow	Time. Rt. (%) \downarrow	Lane. Rt. (%) \downarrow	Colli. Rt. (%) \downarrow
Baselines				
CIL \ Aug	20.1 \pm 1.0	61.1 \pm 2.6	0.0 \pm 0.0	18.8 \pm 2.9
CILRS \ Aug	40.3 \pm 2.6	45.8 \pm 4.5	0.7 \pm 1.0	13.2 \pm 3.5
CIL	40.8 \pm 4.9	27.1 \pm 8.5	1.3 \pm 1.7	30.8 \pm 5.7
CILRS	32.1 \pm 5.5	46.3 \pm 6.1	0.4 \pm 0.8	21.3 \pm 3.3
MT+hLoss				
CN+MT+hLoss \ Aug	64.6 \pm 1.7	10.4 \pm 0.0	8.3 \pm 1.7	16.7 \pm 1.7
RN+MT+hLoss \ Aug	78.5 \pm 2.6	4.2 \pm 1.7	2.1 \pm 1.7	15.3 \pm 1.0
CN+MT+hLoss	72.2 \pm 1.0	1.4 \pm 1.0	2.1 \pm 1.7	24.3 \pm 1.0
RN+MT+hLoss	79.5 \pm 1.7	1.4 \pm 1.0	1.4 \pm 1.0	17.7 \pm 2.0
MT+uLoss				
CN+MT+uLoss	84.0 \pm 1.9	4.9 \pm 2.3	8.9 \pm 0.6	2.2 \pm 0.6
RN+MT+uLoss(Ours)	86.8 \pm 3.6	3.8 \pm 2.5	4.7 \pm 1.6	4.7 \pm 1.6

¹ Mean and standard deviation over 5 evaluation seeds.

² **Succ. Rt.:** Success Rate, **Time. Rt.:** Timeout Rate, **Lane. Rt.:** Lane Invasion Rate, **Colli. Rt.:** Collision Rate.

³ \ means **without**.

method on learning from a variety of driving styles. The models are trained on dataset Mul-Dri that are collected from 10 drivers with various behavioral styles. The evaluation is performed as in Exp.3 where the scenes have both pedestrians and environmental vehicles. We present the mean evaluation results over 5 seeds in Fig. 12. Compared to the results of Exp.3 that were learnt from a single driver, we see some performance loss in Exp.4, which reflects the challenge of learning from a variety of driving styles [66]. However, our method still performs better than baseline, demonstrating its effectiveness. Addressing driver heterogeneity in imitation learning is beyond the scope of this work, we leave this topic to future work.

G. Exp.5 Ablation Studies

Ablation experiments in Fig. 13 are conducted to further investigate the importance of three components: backbone image encoder (CN for CarlaNet [19] or RN for ResNet34), multi-task learning (MT) and loss (hLoss for hard weight loss and uLoss for uncertainty weighted loss). The influence of data augmentation (Aug) is also evaluated. All models are trained on dataset Ped and tested on scenes with only pedestrians. Detailed results of task completion on new scenes and new weathers are provided in Tab.IX.

Experiments in the first group compares between different backbones and demonstrate that data augmentation is of vital importance to CarlaNet, whose structure is smaller than ResNet. Without data augmentation, CIL has a poor performance due to high timeout rate. Through multi-task modeling of lateral and longitudinal control, performance of models in the second group greatly exceeds that of single-task baselines with respect to success rate and timeout rate.

The last group, which uses uncertainty weighted loss instead of hard weight loss, achieves the best testing performance. Our model adaptively learns to balance between lateral and longitudinal control tasks and further reduces the relatively high collision rates in the second group.

VI. DISCUSSION, CONCLUSION AND FUTURE WORKS

The traditional microscopic vehicle behavior models have been studied since the middle of last century [67]. With over sixty years of research, great results have been achieved. These models are mainly developed on expert knowledge, have simple structures and a few explainable parameters, have been incorporated into ADAS and autonomous driving systems today. However, these models are too simplified to describe the highly non-linear procedure of human drivers at complex driving scenes like intersections.

On the other hand, end-to-end autonomous driving has been studied since NVIDIA's pioneering work [18], and has attracted great attentions in the intelligent vehicle societies [68]. One of the main advantages is that a deep model can represent the highly non-linear procedure of a human driver's decision-making at complex scenes, and can be learned in an end-to-end way. Especially for DIL, the cost of dataset generation is very low as data can be collected during human drivers' daily driving. The learned model has the potential to mimic human drivers' behavior to achieve human-like driving.

This work studies DIL-based autonomous control for intersection navigation with pedestrians interaction. In order to navigate through the intersection safely and efficiently, and interact friendly with the pedestrians on crosswalks, this research propose a multi-task conditional imitation learning method to adapt both lateral and longitudinal control tasks

simultaneously. Homoscedastic uncertainties that are inherent to both tasks are learned to weight the loss in training. A new benchmark called IntersectNav is developed and human demonstrations are collected. Experimental results show that our method can achieve a high performance.

Although end-to-end autonomous driving has demonstrated promising results during the last decade, it still faces many challenges such as poor interpretability, generalization to new scenes (e.g., domain gap [69]) and the open world problem [68]. It will take more time and effort for this new but promising technique to mature. In addition, high-fidelity simulators are important for closed-loop driving policy studies. For the task of autonomous navigation at crowded intersections, a sophisticated pedestrian model that realistically reproduces pedestrian-vehicle interactions is essential to reduce the gap between simulation and reality, and to improve the effectiveness of driving policy learning. For this work, learning from real-world dataset and closing the sim-to-real loop [70] are among the most important topics in future studies. On the other hand, learning driving policies at both decision- and control-levels, achieving human-like driving by using such as inverse reinforcement learning (IRL) [71], [72] will also be addressed.

REFERENCES

- [1] L. Chen, Y. Li, C. Huang, *et al.*, "Milestones in autonomous driving and intelligent vehicles: Survey of surveys," *IEEE Transactions on Intelligent Vehicles*, 2022.
- [2] S. Li, K. Shu, C. Chen, and D. Cao, "Planning and decision-making for connected autonomous vehicles at road intersections: A review," *Chinese Journal of Mechanical Engineering*, vol. 34, no. 1, pp. 1–18, 2021.
- [3] L. Wei, Z. Li, J. Gong, C. Gong, and J. Li, "Autonomous driving strategies at intersections: Scenarios, state-of-the-art, and future outlooks," in *2021 IEEE International Intelligent Transportation Systems Conference (ITSC)*, IEEE, 2021, pp. 44–51.
- [4] A. Rasouli and J. K. Tsotsos, "Autonomous vehicles that interact with pedestrians: A survey of theory and practice," *IEEE Transactions on Intelligent Transportation Systems*, vol. 21, no. 3, pp. 900–918, 2019.
- [5] R. Ezzati Amini, C. Katrakazas, A. Riener, and C. Antoniou, "Interaction of automated driving systems with pedestrians: Challenges, current solutions, and recommendations for ehmis," *Transport Reviews*, vol. 41, no. 6, pp. 788–813, 2021.
- [6] M. Predhumeau, A. Spalanzani, and J. Dugdale, "Pedestrian behavior in shared spaces with autonomous vehicles: An integrated framework and review," *IEEE Transactions on Intelligent Vehicles*, pp. 1–1, 2021.
- [7] S. Grigorescu, B. Trasnea, T. Cocias, and G. Macesanu, "A survey of deep learning techniques for autonomous driving," *Journal of Field Robotics*, vol. 37, no. 3, pp. 362–386, 2020.
- [8] E. Yurtsever, J. Lambert, A. Carballo, and K. Takeda, "A survey of autonomous driving: Common practices and emerging technologies," *IEEE access*, vol. 8, pp. 58 443–58 469, 2020.
- [9] J. E. Hoffmann, H. G. Tosso, M. M. D. Santos, *et al.*, "Real-time adaptive object detection and tracking for autonomous vehicles," *IEEE Transactions on Intelligent Vehicles*, vol. 6, no. 3, pp. 450–459, 2021.
- [10] K. Messaoud, I. Yahiaoui, A. Verroust-Blondet, and F. Nashashibi, "Attention based vehicle trajectory prediction," *IEEE Transactions on Intelligent Vehicles*, vol. 6, no. 1, pp. 175–185, 2021.
- [11] M. Roth, J. Stapel, R. Happee, and D. M. Gavrila, "Driver and pedestrian mutual awareness for path prediction and collision risk estimation," *IEEE Transactions on Intelligent Vehicles*, pp. 1–1, 2021.
- [12] S. Kuutti, R. Bowden, Y. Jin, P. Barber, and S. Fallah, "A survey of deep learning applications to autonomous vehicle control," *IEEE Transactions on Intelligent Transportation Systems*, vol. 22, no. 2, pp. 712–733, 2020.
- [13] Z. Zhu and H. Zhao, "A survey of deep rl and il for autonomous driving policy learning," *TITS*, 2021.
- [14] A. O. Ly and M. Akhloufi, "Learning to drive by imitation: An overview of deep behavior cloning methods," *IEEE Transactions on Intelligent Vehicles*, vol. 6, no. 2, pp. 195–209, 2021.
- [15] M. Bouton, A. Nakhaei, K. Fujimura, and M. J. Kochenderfer, "Safe reinforcement learning with scene decomposition for navigating complex urban environments," in *2019 IEEE Intelligent Vehicles Symposium (IV)*, IEEE, 2019, pp. 1469–1476.
- [16] X. Liang, T. Wang, L. Yang, and E. Xing, "Cirl: Controllable imitative reinforcement learning for vision-based self-driving," in *Proceedings of the European Conference on Computer Vision*, 2018, pp. 584–599.
- [17] Z. Zhang, A. Liniger, D. Dai, F. Yu, and L. Van Gool, "End-to-end urban driving by imitating a reinforcement learning coach," in *Proceedings of the IEEE International Conference on Computer Vision*, 2021, pp. 15 222–15 232.
- [18] M. Bojarski, D. Del Testa, D. Dworakowski, *et al.*, "End to end learning for self-driving cars," *arXiv preprint arXiv:1604.07316*, 2016.
- [19] F. Codevilla, M. Müller, A. López, V. Koltun, and A. Dosovitskiy, "End-to-end driving via conditional imitation learning," in *IEEE International Conference on Robotics and Automation*, IEEE, 2018, pp. 4693–4700.
- [20] F. Codevilla, E. Santana, A. M. López, and A. Gaidon, "Exploring the limitations of behavior cloning for autonomous driving," in *Proceedings of the IEEE International Conference on Computer Vision*, 2019, pp. 9329–9338.
- [21] D. Chen, B. Zhou, V. Koltun, and P. Krähenbühl, "Learning by cheating," in *Conference on Robot Learning*, PMLR, 2020, pp. 66–75.
- [22] S. Teng, L. Chen, Y. Ai, *et al.*, "Hierarchical interpretable imitation learning for end-to-end autonomous driving," *IEEE Transactions on Intelligent Vehicles*, pp. 1–11, 2022.
- [23] A. E. Sallab, M. Abdou, E. Perot, and S. Yogamani, "End-to-end deep reinforcement learning for lane keeping assist," *arXiv preprint arXiv:1612.04340*, 2016.
- [24] H. Porav and P. Newman, "Imminent collision mitigation with reinforcement learning and vision," in *2018 21st International Conference on Intelligent Transportation Systems (ITSC)*, IEEE, 2018, pp. 958–964.
- [25] A. Sauer, N. Savinov, and A. Geiger, "Conditional affordance learning for driving in urban environments," in *Conference on Robot Learning*, PMLR, 2018, pp. 237–252.
- [26] C. Chen, A. Seff, A. Kornhauser, and J. Xiao, "Deepdriving: Learning affordance for direct perception in autonomous driving," in *Proceedings of the IEEE International Conference on Computer Vision*, 2015, pp. 2722–2730.
- [27] A. Prakash, A. Behl, E. Ohn-Bar, K. Chitta, and A. Geiger, "Exploring data aggregation in policy learning for vision-based urban autonomous driving," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2020, pp. 11 763–11 773.
- [28] A. Eskandarian, "Handbook of intelligent vehicles," in Springer, 2012, vol. 2, ch. Vehicle Longitudinal and Lateral Control, pp. 165–232.
- [29] A. Kendall, Y. Gal, and R. Cipolla, "Multi-task learning using uncertainty to weigh losses for scene geometry and semantics," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2018, pp. 7482–7491.
- [30] S. Ross and D. Bagnell, "Efficient reductions for imitation learning," in *Proceedings of the International Conference on Artificial Intelligence and Statistics*, JMLR Workshop and Conference Proceedings, 2010, pp. 661–668.
- [31] A. Dosovitskiy, G. Ros, F. Codevilla, A. Lopez, and V. Koltun, "Carla: An open urban driving simulator," in *Conference on robot learning*, PMLR, 2017, pp. 1–16.
- [32] D. Pomerleau, "ALVINN: an autonomous land vehicle in a neural network," in *Advances in Neural Information Processing Systems*, 1988, pp. 305–313.
- [33] P. de Haan, D. Jayaraman, and S. Levine, "Causal confusion in imitation learning," *Advances in Neural Information Processing Systems*, vol. 32, pp. 11 698–11 709, 2019.
- [34] E. Ohn-Bar, A. Prakash, A. Behl, K. Chitta, and A. Geiger, "Learning situational driving," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2020, pp. 11 296–11 305.
- [35] M. Kuderer, S. Gulati, and W. Burgard, "Learning driving styles for autonomous vehicles from demonstration," in *2015 IEEE International Conference on Robotics and Automation (ICRA)*, 2015, pp. 2641–2646.
- [36] S. Sharifzadeh, I. Chiotellis, R. Triebel, and D. Cremers, "Learning to drive using inverse reinforcement learning and deep q-networks," *arXiv preprint arXiv:1612.03653*, 2016.
- [37] M. Wulfmeier, D. Rao, D. Z. Wang, P. Ondruska, and I. Posner, "Large-scale cost function learning for path planning using deep inverse reinforcement learning," *Int. J. Robotics Res.*, vol. 36, no. 10, pp. 1073–1087, 2017.

- [38] P. Wang, D. Liu, J. Chen, H. Li, and C.-Y. Chan, "Decision making for autonomous driving via augmented adversarial inverse reinforcement learning," in *2021 IEEE International Conference on Robotics and Automation (ICRA)*, 2021, pp. 1036–1042.
- [39] Z. Zhao, Z. Wang, K. Han, et al., "Personalized car following for autonomous driving with inverse reinforcement learning," in *2022 International Conference on Robotics and Automation (ICRA)*, 2022, pp. 2891–2897.
- [40] W. Zeng, P. Chen, H. Nakamura, and M. Iryo-Asano, "Application of social force model to pedestrian behavior analysis at signalized crosswalk," *Transportation research part C: emerging technologies*, vol. 40, pp. 143–159, 2014.
- [41] W. Zeng, P. Chen, G. Yu, and Y. Wang, "Specification and calibration of a microscopic model for pedestrian dynamic simulation at signalized intersections: A hybrid approach," *Transportation Research Part C: Emerging Technologies*, vol. 80, pp. 37–70, 2017.
- [42] Q. Chao, P. Liu, Y. Han, et al., "A calibrated force-based model for mixed traffic simulation," *IEEE Transactions on Visualization and Computer Graphics*, pp. 1–1, 2021.
- [43] Y. Han, Q. Chao, and X. Jin, "A simplified force model for mixed traffic simulation," *Computer Animation and Virtual Worlds*, vol. 32, no. 1, e1974, 2021.
- [44] H. Zhu, T. Han, W. K. Alhajyaseen, M. Iryo-Asano, and H. Nakamura, "Can automated driving prevent crashes with distracted pedestrians? an exploration of motion planning at unsignalized mid-block crosswalks," *Accident Analysis & Prevention*, vol. 173, p. 106711, 2022.
- [45] G. P. R. Papini, A. Plebe, M. D. Lio, and R. Dona, "A reinforcement learning approach for enacting cautious behaviours in autonomous driving system: Safe speed choice in the interaction with distracted pedestrians," *IEEE Transactions on Intelligent Transportation Systems*, pp. 1–18, 2021.
- [46] Y. Luo, P. Cai, A. Bera, et al., "Porca: Modeling and planning for autonomous driving among many pedestrians," *IEEE Robotics and Automation Letters*, vol. 3, no. 4, pp. 3418–3425, 2018.
- [47] Q. Chao, X. Jin, H.-W. Huang, et al., "Force-based heterogeneous traffic simulation for autonomous vehicle testing," in *2019 International Conference on Robotics and Automation (ICRA)*, 2019, pp. 8298–8304.
- [48] A. Rasouli and J. K. Tsotsos, "Autonomous vehicles that interact with pedestrians: A survey of theory and practice," *IEEE Transactions on Intelligent Transportation Systems*, vol. 21, no. 3, pp. 900–918, 2020.
- [49] P. Chen, W. Zeng, and G. Yu, "Assessing right-turning vehicle-pedestrian conflicts at intersections using an integrated microscopic simulation model," *Accident Analysis & Prevention*, vol. 129, pp. 211–224, 2019.
- [50] H. Bai, S. Cai, N. Ye, D. Hsu, and W. S. Lee, "Intention-aware online pomdp planning for autonomous driving in a crowd," in *2015 IEEE International Conference on Robotics and Automation (ICRA)*, IEEE, 2015, pp. 454–460.
- [51] R. Yao, W. Zeng, Y. Chen, and Z. He, "A deep learning framework for modelling left-turning vehicle behaviour considering diagonal-crossing motorcycle conflicts at mixed-flow intersections," *Transportation research part C: emerging technologies*, vol. 132, p. 103415, 2021.
- [52] N. Deshpande and A. Spalanzani, "Deep reinforcement learning based vehicle navigation amongst pedestrians using a grid-based state representation," in *2019 IEEE Intelligent Transportation Systems Conference (ITSC)*, IEEE, 2019, pp. 2081–2086.
- [53] Y. Zhang and Q. Yang, "A survey on multi-task learning," *IEEE Transactions on Knowledge and Data Engineering*, 2021.
- [54] Y. Liao, S. Kodagoda, Y. Wang, L. Shi, and Y. Liu, "Understand scene categories by objects: A semantic regularized scene classifier using convolutional neural networks," in *IEEE international Conference on Robotics and Automation*, IEEE, 2016, pp. 2318–2325.
- [55] D. Eigen and R. Fergus, "Predicting depth, surface normals and semantic labels with a common multi-scale convolutional architecture," in *Proceedings of the IEEE International Conference on Computer Vision*, 2015, pp. 2650–2658.
- [56] H. Xu, Y. Gao, F. Yu, and T. Darrell, "End-to-end learning of driving models from large-scale video datasets," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2017, pp. 2174–2182.
- [57] I. Kim, H. Lee, J. Lee, E. Lee, and D. Kim, "Multi-task learning with future states for vision-based autonomous driving," in *Proceedings of the Asian Conference on Computer Vision*, 2020.
- [58] K. Ishihara, A. Kanervisto, J. Miura, and V. Hautamaki, "Multi-task learning with attention for end-to-end autonomous driving," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2021, pp. 2902–2911.
- [59] S. Ruder, "An overview of multi-task learning in deep neural networks," *arXiv preprint arXiv:1706.05098*, 2017.
- [60] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2016, pp. 770–778.
- [61] C. team, *Carla autonomous driving leaderboard*, <https://leaderboard.carla.org>, CARLA team.
- [62] M. Klischat, E. I. Liu, F. Holtke, and M. Althoff, "Scenario factory: Creating safety-critical traffic scenarios for automated vehicles," in *2020 IEEE 23rd International Conference on Intelligent Transportation Systems (ITSC)*, IEEE, 2020, pp. 1–7.
- [63] S. Feng, Y. Feng, C. Yu, Y. Zhang, and H. X. Liu, "Testing scenario library generation for connected and automated vehicles, part i: Methodology," *IEEE Transactions on Intelligent Transportation Systems*, vol. 22, no. 3, pp. 1573–1582, 2020.
- [64] D. P. Kingma and J. Ba, "Adam: A method for stochastic optimization," *arXiv preprint arXiv:1412.6980*, 2014.
- [65] Z. Ding, D. Xu, C. Tu, et al., "Driver identification through heterogeneity modeling in car-following sequences," *IEEE Transactions on Intelligent Transportation Systems*, 2022.
- [66] C. Rui and D. Michie, "Behavioral cloning a correction," *Ai Magazine*, vol. 16, p. 92, 1995.
- [67] N. AbuAli and H. Abou-zeid, "Driver behavior modeling: Developments and future directions," *International Journal of Vehicular Technology*, pp. 1–12, 2016.
- [68] L. Le Mero, D. Yi, M. Dianati, and A. Mouzakitis, "A survey on imitation learning techniques for end-to-end autonomous vehicles," *IEEE Transactions on Intelligent Transportation Systems*, vol. 23, no. 9, pp. 14128–14147, 2022.
- [69] M. Schutera, M. Hussein, J. Abhau, R. Mikut, and M. Reischl, "Night-to-day: Online image-to-image translation for object detection within autonomous driving by night," *IEEE Transactions on Intelligent Vehicles*, vol. 6, no. 3, pp. 480–489, 2021.
- [70] Y. Chebotar, A. Handa, V. Makovychuk, et al., "Closing the sim-to-real loop: Adapting simulation randomization with real world experience," in *2019 International Conference on Robotics and Automation (ICRA)*, 2019, pp. 8973–8979.
- [71] M. Menner, K. Berntorp, M. N. Zeilinger, and S. D. Cairano, "Inverse learning for data-driven calibration of model-based statistical path planning," *IEEE Transactions on Intelligent Vehicles*, vol. 6, no. 1, pp. 131–145, 2021.
- [72] L. Guo and Y. Jia, "Inverse model predictive control (impc) based modeling and prediction of human-driven vehicles in mixed traffic," *IEEE Transactions on Intelligent Vehicles*, vol. 6, no. 3, pp. 501–512, 2021.



Zeyu Zhu received B.S. degree in computer science from Peking University, Beijing, China, in 2019, where he is currently pursuing the Ph.D. degree with the Key Laboratory of Machine Perception (MOE), Peking University. His research interests include intelligent vehicles, imitation learning and reinforcement learning.



Huijing Zhao received B.S. degree in computer science from Peking University in 1991. She obtained M.E. degree in 1996 and Ph.D. degree in 1999 in civil engineering from the University of Tokyo, Japan. From 1999 to 2007, she was a postdoctoral researcher and visiting associate professor at the Center for Space Information Science, University of Tokyo. In 2007, she joined Peking University as a tenure-track professor at the School of Electronics Engineering and Computer Science and became an associate professor with tenure on 2013. She is

now a full professor with tenure at the School of Intelligence Science and Technology, Peking University. She has research interest in several areas in connection with intelligent vehicle and mobile robot, such as machine perception, behavior learning and motion planning, and she has special interests on the studies through real world data collection.