Human-In-The-Loop Machine Learning for Safe and Ethical Autonomous Vehicles: Principles, Challenges, and Opportunities

Yousef Emami, *Member, IEEE,* Luis Almeida, *Senior Member, IEEE,* Kai Li, *Senior Member, IEEE,* Wei Ni, *Fellow Member, IEEE,* and Zhu Han, *Fellow Member, IEEE,*

arXiv:2408.12548v2 [cs.LG] 8 Sep 2024

***Abstract*—Rapid advances in Machine Learning (ML) have triggered new trends in Autonomous Vehicles (AVs). ML algorithms play a crucial role in interpreting sensor data, predicting potential hazards, and optimizing navigation strategies. However, achieving full autonomy in cluttered and complex situations, such as intricate intersections, diverse sceneries, varied trajectories, and complex missions, is still challenging, and the cost of data labeling remains a significant bottleneck. The adaptability and robustness of humans in complex scenarios motivate the inclusion of humans in the ML process, leveraging their creativity, ethical power, and emotional intelligence to improve ML effectiveness. The scientific community knows this approach as Human-In-The-Loop Machine Learning (HITL-ML). Towards safe and ethical autonomy, we present a review of HITL-ML for AVs, focusing on Curriculum Learning (CL), Human-In-The-Loop Reinforcement Learning (HITL-RL), Active Learning (AL), and ethical principles. In CL, human experts systematically train ML models by starting with simple tasks and gradually progressing to more difficult ones. HITL-RL significantly enhances the RL process by incorporating human input through techniques like reward shaping, action injection, and interactive learning. AL streamlines the annotation process by targeting specific instances that need to be labeled with human oversight, reducing the overall time and cost associated with training. Ethical principles must be embedded in AVs to align their behavior with societal values and norms. In addition, we provide insights and specify future research directions.**

***Index Terms*—Human-In-The-Loop, Machine Learning, Curriculum Learning, Reinforcement Learning, Active Learning, Ethical Principles, Autonomous Vehicles, Unmanned Aerial Vehicles**

I. Introduction

Recent developments in Autonomous Vehicles (AVs) have been useful in improving efficiency and safety [1]. AVs have ushered in a substantial transformation in terms of automation and connectivity. The global AVs market size was valued at USD 1,500.3$ billion in 2022 and is projected to grow from USD 1,921.1$ billion in 2023 to USD 13,632.4$ billion by 2030 [2]. Commercially envisioned and available self-driving cars can perform functions such as lane-changing, highway driving, and autonomous parking by sensing the environment using built-in technologies [3]. Ground-based autonomous systems can be paired with aerial ones to drive a significant advancement in transportation and technology and offer a range of exciting possibilities, including surveillance, monitoring, traffic management, remote sensing, and autonomous control [4] [5].

Copyright (c) 2024 IEEE. Personal use of this material is permitted. However, permission to use this material for any other purposes must be obtained from the IEEE by sending a request to pubs-permissions@ieee.org.

TABLE I: List of Acronyms

Acronym	Definition
AD	Autonomous Driving
AL	Active Learning
AUV	Autonomous Underwater Vehicle
AVs	Autonomous Vehicles
CNN	Convolutional Neural Network
CL	Curriculum Learning
CPSs	Cyber-Physical Systems
CV	Connected Vehicle
DDPG	Deep Deterministic Policy Gradients
DL	Deep Learning
DQN	Deep-Q-Network
DRL	Deep Reinforcement Learning
EVs	Electric Vehicles
HITL	Human-In-The-Loop
MDP	Markov Decision Process
ML	Machine Learning
MLR	Multiple Linear Regression
PPO	Proximal Policy Optimization
RL	Reinforcement Learning
SVM	Support Vector Machine
SVR	Support Vector Regression
UAV	Unmanned Aerial Vehicle
XGBoost	eXtreme Gradient Boosting

Current trends in AVs include leveraging Machine Learning (ML) [6], tightening security [7] [8], and facilitating data and command dissemination [9] [10] to decrease accidents, fatalities, injuries, pollution, and traffic congestion while fostering a more equitable transportation infrastructure.

ML is essential for the operation of AVs because it enables AVs to learn from large datasets and create predictive models for their environment. ML algorithms facilitate the detection and classification of objects, pedestrians and other vehicles, as well as the prediction of their future trajectories. These predictive insights are critical to enabling AVs to make real-time braking, turning, acceleration and steering decisions. The use of AVs fundamentally depends on ML-driven systems to perform key functions such as perception [11], state estimation [12], path planning [13], traffic rule compliance and decision making in dynamic and uncertain environments.

Table I presents a list of the acronyms used throughout this paper for easy reference.

A. Motivation and Background

Human-In-The-Loop Machine Learning (HITL-ML) combines human creativity, ethical judgment, and emotional intelligence with the learning power of ML algorithms. Human intuition and judgment help AVs handle situations that are not

Samsara EX1060
Samsara v. Motive Tech. Inc.
IPR2026-00034

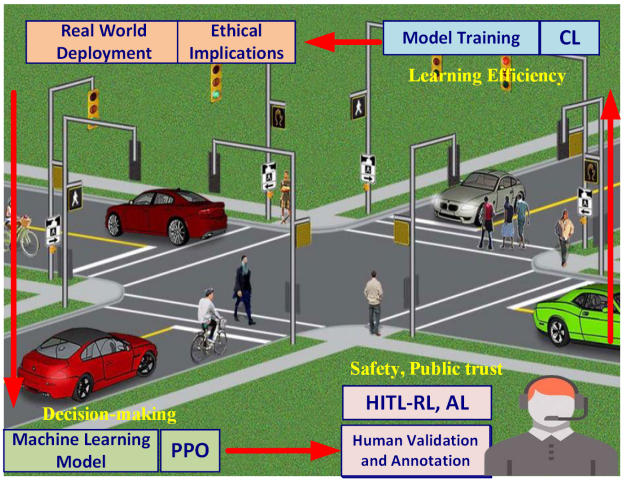

Fig. 1: The workflow of HITL-ML for AVs includes several key steps: 1) developing an ML model for decision-making and predictive modeling using PPO, 2) applying human validation and annotation through HITL-RL and AL to improve safety and foster public trust, 3) conducting model training where data is organized and refined with CL to enhance learning efficiency and performance, and 4) real-world deployment while considering ethical implications.

well-represented in training data [14]. Decision making for ethical dilemmas in unavoidable accidents requires human-like reasoning that current ML models do not possess. Human-like reasoning is crucial in these scenarios because it incorporates context, societal norms, and a nuanced understanding of moral principles, capacities that current ML models lack. Human input and guidance can enhance ML models and accelerate the process of building public trust in AVs. Humans can be incorporated into the foundations of ML through collaboration, and oversight. In collaboration, humans and ML models are near equals in performance but not in abilities. Humans excel at capturing region-specific behaviors and cultural norms, providing context and insights that ML algorithms might miss. ML models can ask for help when they are either unconfident in their performance and/or think humans can better carry out the task.

Imagine an AV driving on a local road where pedestrians frequently cross unpredictably and outside of designated crosswalks. In such a scenario, the AV's ML model may struggle to adapt, resulting in dangerous actions such as abrupt stops or failure to yield appropriately. When the AV encounters these challenges, it can request human intervention to prevent dangerous behaviors and refine its response. A human operator familiar with the local conditions can override or adjust the model's actions, supplementing the reward system to prioritize pedestrian safety, even in situations that deviate from the AV's original training data. This human input ensures that the AV can navigate the environment safely and effectively despite the limitations of its original training.

ML models are strictly better than humans at tasks such as predictive analytics and risk assessment, but we still need to monitor them to ensure safety and alignment with our goals (oversight). Tesla's Autopilot, for instance, has problems when driving on local roads in the presence of people and relies on the driver's supervision to navigate effectively [15]. Google's self-driving cars still make mistakes, such as being confused by fixed-gear cyclists at stop signs [16]. In October 2023, Cruise's permit to test AVs in California was suspended after a serious accident in which a car dragged a woman across the road. Waymo recently recalled some of its vehicles after two collided with the same tow truck within minutes [17]. This ongoing need for human involvement underscores why AVs are still several years away from being ready for widespread consumer use [18]. As demonstrated by the recent challenges faced by companies like Tesla and Waymo, the effective deployment of AVs requires a hybrid approach where human oversight continues to play a crucial role in ensuring safety and reliability.

Ethical principles must be embedded in AVs to align their behavior with societal values and norms. To this end, ethical theories must be integrated into the decision-making processes of AVs. However, ethical decision-making in AVs faces several challenges. One major challenge is the complexity of ethical dilemmas that require difficult trade-offs. For example, in urban driving scenarios that aim to minimize injuries in accidents. Another challenge is the need to adapt ethical decision-making processes in real time. Ensuring transparency and accountability in these processes is also crucial. HITL-ML can improve self-recognition and scene recognition in AVs to assist in managing complex situations such as accidents and making ethical decisions in real-time.

As shown in Fig. 1, the workflow of HITL-ML for AVs comprises several key steps [19]: 1) developing an ML model using predictive algorithms, 2) applying human validation and annotation with expert oversight, 3) conducting model training where input data is organized and refined, and 4) real-world deployment while considering ethical implications. Predictive algorithms such as Proximal Policy Optimization (PPO) [20] are initially used, with human experts monitoring their functionalities where they replace dangerous actions in critical situations and modify or supplement the reward function. Moreover, Active Learning (AL) can detect instances where AVs lack adequate confidence and refer them to human annotators [21]. The model is trained after human validation, annotation, and feedback from human experts. Meanwhile, Curriculum Learning (CL) with the knowledge of human experts provides sorted and organized input for AVs [22], facilitating and expediting the learning process. Finally, the model is deployed in the real world, considering ethical implications.

Rapid progress in user interface design [23], regulatory frameworks [24], AV security [25] [26] and explainable AI [27] have significantly improved the practicality of HITL approaches for AVs. Improved user interfaces provide intuitive controls and real-time feedback, while regulatory requirements ensure human oversight for safety and ethical compliance. Developments in AV security safeguard against cyber threats, and research in explainable AI enhances the transparency of decision-making processes, making it easier for human operators to understand and interact with AVs. Together, these advancements enable effective human oversight and intervention, crucial for the safe

and ethical operation of AVs.

B. Existing Surveys and Tutorials

In the literature, a few works consider HITL-ML. For example, Duan etal. [28] provide an up-to-date overview of the latest results on understanding and utilizing the human factor in different contexts of mobile networks. Expanding on HITL-ML strategies, Mosqueira-Rey etal. [29] present a comprehensive review of state-of-the-art techniques, highlighting different strategies for incorporating humans into the ML process. From a data perspective, Wu etal. [19] survey HITL-ML, addressing how human interaction can enhance data quality and model performance. Rajendran etal. [30] review the main concepts and methods for HITL learning, paving the way for developing frameworks that facilitate human-machine teaming through safe learning and anomaly prediction. Methodologies, challenges, and opportunities associated with real-world applications of HITL systems are reviewed by Kumar etal. [31], discussing key issues like data quality, bias, and user engagement. Wilchek etal. [32] review the literature on HITL, emphasizing the role of human interaction in enhancing computer vision models performance and accuracy.

Specifically, in the realm of CL within RL, Narvekar etal. [33] provide a framework and an overview of existing CL methods. They classify these methods based on their assumptions, capabilities, and goals, offering insights into how CL can systematically improve RL performance. Soori etal. [34] examine recent advances in AI, ML, and DL in advanced robotic systems, and discuss their applications in modifying robots. Ma etal. [35] review the usage of AI for AVs development, and elaborate on critical issues such as functional needs and challenges in implementing AI-driven solutions for AVs. From a computer vision system design perspective, Nunes etal. [36] present a review of HITL-Cyber-Physical Systems (CPSs), providing insights into how human factors can be integrated into CPSs to enhance their functionality and reliability. AD technologies have seen substantial advancements, driven by developments in ML and AI. Chib and Singh [37] review the End-to-End AD stack, categorizing automated driving tasks using neural networks in an End-to-End manner. Surveying essential components and cutting-edge technologies, Zhao etal. [38] explore the development and implementation of AD systems. This survey highlights the key technologies that enable AD and the challenges associated with their deployment. Both Wang etal. [39] and Moura etal. [40] address ethical dilemmas associated with different aspects of AVs.

C. Contribution

In this paper, we review HITL-ML to support the safe and ethical operation of AVs, facilitate the management of complex and unpredictable scenarios from busy local roads to intricate intersections, trajectories and missions, and improve their trustworthiness, enhancing public confidence on them. Our paper differs from others in the literature by emphasizing the critical role of HITL-ML in ensuring the safe and ethical operation of AVs. We focus on CL, HITL-RL and AL (see Table II for a comparative summary of recent research papers on these techniques) and then we look into ethical principles tailored to the unique challenges and opportunities in AVs.

The key contributions of this paper are listed as follows:

- We provide a comprehensive overview of the application of CL in AVs to improve generalization and convergence speed by structuring the training process from simpler to more complex tasks. We highlight the synergy between CL and RL and how their combination can reduce training time and improve AV performance on tasks such as navigation, path planning and obstacle avoidance.
- We provide an in-depth overview of the latest HITL-RL methods and their applications, highlighting how human involvement can enhance the RL process through techniques such as reward shaping, action injection and interactive learning. We also refer examples from recent studies showing various HITL-RL frameworks that utilize human feedback to refine AV learning processes and ensure real-time safety interventions.
- We review the literature on AL, focusing on how AL optimizes the annotation process, accelerates AV development, and enhances system robustness by leveraging human expertise. We also explore how integrating AL with ML techniques such as DL, SVMs, and Symmetry-based Analysis enhances the efficiency and accuracy of tasks like anomaly detection, semantic mapping, object detection, and vehicle recognition in AVs.
- We investigate the ethical considerations of HITL-ML for AVs, highlighting the integration of HITL-ML to enhance self-recognition, scene recognition, and real-time ethical decision-making.

D. Paper Structure and Organization

Section II introduces CL and its significant role in improving learning speed in AVs. Section III introduces HITL-RL and how humans can cooperate in the RL process in AVs. Section IV discusses AL and examines its role in facilitating data annotation in AVs. Section V addresses the ethical aspects of AVs and provides guidelines to achieve ethical AVs using HITL-ML. Section VI highlights open research directions and, finally, we conclude the paper in Section VII.

II. Curriculum Learning

CL is an ML technique that structures the training process by starting with simpler tasks and gradually progressing to more complex ones. This approach enhances generalization and convergence speed, making it particularly effective for training AVs [55]. Recent studies have applied CL to AVs in different ways, including curriculum-based multi-agent DRL for decentralized flocking and obstacle avoidance, collision-free flocking control in large-scale UAV swarms, and adaptive

TABLE II: Comparative Summary of Recent Research Papers Exploring HITL-ML.

Ref	Objective	ML Technique	Scenario
[41]	assist the UAV in obstacle avoidance.	DRL	Navigation / UAV
[42]	improve the agent's learning ability and performance.	DRL	Intelligence / AV
[43]	tackle sparse reward and low learning efficiency.	RL	Exploration / AUV
[44]	facilitate mapless obstacle avoidance.	DRL	Navigation / UAV swarm
[45]	improve the learning performance of RL	DRL	AD
[46]	mitigate in-flight UAV anomalies by using semi-supervised SVM	AL	Anomaly detection / UAV
[47]	develop a new path planning framework for AL	AL	Path planning / Semantic mapping
[48]	detect unknown network attacks in the UAV network.	AL	IDS / UAV
[21]	facilitate manual labeling of large-scale surveillance data.	AL	VTR / ITS
[49]	learn the decentralized flocking and obstacle avoidance policy	CL	Obstacle avoidance / UAV swarm
[50]	reduce the required training time and improve agent performance	CL	AD
[51]	enable mission planning for controlling a UAV	ML	Mission planning / UAV
[52]	address 3D object detection in AV	AL	Object detection / UAV
[53]	study collision-free flocking control problem for scalable and large-scale fixed-wing UAV	CL	Flocking problem / UAV swarm
[54]	learn an end-to-end urban driving policy for CARLA driving simulator	CL	Driving simulator

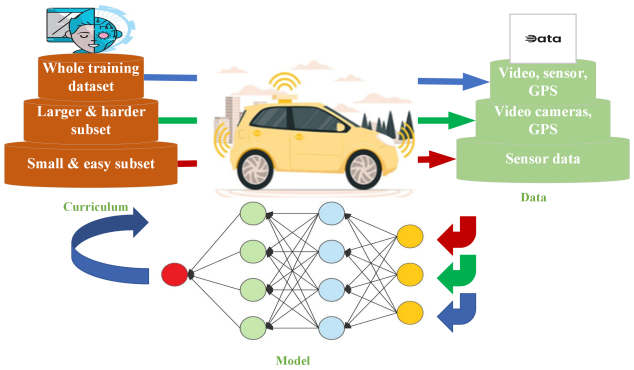

Fig. 2: With CL for AVs tasks are divided into smaller, manageable parts, progressing from easy to more difficult challenges. Data associated with each part is sequentially fed to the model to enhance learning efficiency and improve overall performance.

curriculum strategies for efficient target search in disaster scenarios using UAV swarms [56]. These applications demonstrate the effectiveness of CL in guiding autonomous agents through progressively challenging tasks to achieve optimal performance.

Fig. 2 illustrates the concept of CL for AVs. Tasks are divided into smaller, manageable parts, progressing from easy to more difficult challenges. This method allows the AVs to build foundational skills with simpler tasks before gradually tackling more complex scenarios, thereby improving learning efficiency and performance over time. The structured progression ensures that the vehicle can handle increasingly difficult situations with greater accuracy and reliability. In the realm of AVs, the combination of CL and RL can reduce the required training time, improve agent performance, and help learn better policies [50]. AVs require mastery of complex tasks such as navigation, path planning, and obstacle avoidance. CL can aid the learning process by ordering these tasks from simple to complex.

A. Literature Review

1) Navigation and Data Collection: Curriculum learning can significantly improve collision avoidance in UAV swarms [57]by gradually increasing the complexity of the tasks that the UAVs have to solve during their training process. This approach ensures that the UAVs first master simpler tasks, such as avoiding static objects or other UAVs in a sparse environment, before moving on to more challenging scenarios, such as dynamic environments with multiple moving objects or denser UAV formations. A particular challenge in UAV swarms is collision avoidance, where the UAVs must avoid other UAVs as well as fixed and moving objects. Yan etal. [49] study the flocking and obstacle avoidance problem for multiple fixed-wing UAVs in cluttered environments. A novel curriculum-based multi-agent DRL is developed to learn the decentralized flocking and obstacle avoidance policy. In this context, the collision-avoiding flocking task is decomposed into multiple subtasks and the number of subtasks to be solved is gradually increased. In the developed approach, a novel multi-agent actor-critic DRL algorithm and two knowledge transfer mechanisms are developed for online learning and offline transfer, respectively. Similarly, Yan etal. [53] study collision-free flocking control problems for scalable and large-scale fixed-wing UAV swarms by integrating CL with multi-agent DRL. The problem considers both collision avoidance and the large population of UAVs in the swarm. The key idea is to divide the training process into several learning phases and gradually expand the scale of UAV swarms. In addition, Braathen etal. [58] investigate navigating UAVs in complex 3D environments and develop a solution using curriculum-based DRL. The human expert divides the training into ten different levels with an increasing range of possible numbers of obstacles in the environment. The leveraged curriculum exposes the UAV to simpler environments initially and progressively to more complex ones.

Moreover, data collection [61] and UAV landing tasks require arranging the input. Seong etal. [59] address the collection of data from wireless sensor nodes to minimize the Age of Information during data collection, while considering the energetic sustainability of UAVs. They introduce Multi-Agent Reinforcement Learning (MARL) in combination with CL and an evolutionary strategy to improve the performance of traditional MARL setups as the number of agents increases. Their method promotes MARL scalability and integrates energy-efficient charging mechanisms that effectively improve system performance in large-scale deployments. Moreover, the

TABLE III: Key Techniques for Navigation and Data Collection

Ref	Technique	Application	Strengths and Limitations
[49]	Curriculum-based Multi-agent DRL	Flocking and obstacle avoidance for UAVs	Learns decentralized policies; requires knowledge transfer mechanisms.
[53]	Curriculum Learning with Multi-agent DRL	Collision-free flocking for large UAV swarms	Scalable to large swarms; gradually increases swarm size.
[58]	Curriculum-based DRL	UAV navigation in complex 3D environments	Improves navigation reliability; relies on human-defined curriculum.
[59]	MARL with Curriculum Learning and Evolutionary Strategy	Data collection from wireless sensor nodes with UAVs	Promotes MARL scalability; integrates energy-efficient charging mechanisms; enhances AoI minimization and energetic sustainability.
[60]	Automatic Curriculum Learning with TD3	Collaborative autonomous landing of UAV on UGV	Offers rapid localization and pose estimation; handles different UGV motions and wind interference; breaks down tasks into simpler subtasks.

collaborative autonomous landing of a quadrotor UAV on a moving Unmanned Ground Vehicle (UGV) is challenging due to the need for accurate real-time tracking of the UGV and the adjustment of the landing policy. Wang etal. of [60] propose the Landing Vision System to offer rapid localization and pose estimation of the UGV. Then, they design an Automatic CL approach in conjunction with the Twin Delayed Deterministic Policy Gradient (TD3) that breaks down UAV landing control into a series of simpler subtasks, and learn the landing tasks under different conditions of UGV motions and wind interference. Table III summarizes the above findings.

2) Autonomous Driving: Self-driving tasks such as intersection management and motion planning benefit from CL to facilitate the learning tasks. Peng etal. [22] present a reward-driven automated curriculum proximal policy optimization (RD-ACPPO) approach for self-driving tasks at unsignalized intersections. Simply, an adaptive CL technique is presented to accommodate a progressively increasing difficulty of learning tasks in AD. Their approach allows for the reasonable allocation of the importance weights for different curricula during various RL training periods. The RL policy trained by the proposed framework attained the highest task success rate. In addition, Peiss etal. [62] develop a novel CL scheme for motion planning in AD that incorporates traffic rule and map information into a GNN-RL agent whose state is extended by map and traffic rule information. They decompose the complicated learning task by applying CL for GNN-RL agent. The curriculum's sequence of sub-tasks gradually increases the level of difficulty of the motion planning problem for the agent. Each of their sub-tasks contains a set of rewards, including new rewards for temporal-logic-based traffic rules for speed, safety distance, and braking. Moreover, the authors of [50] investigate the challenge of learning AD behaviors at urban intersections with DRL. They combine CL with DRL to reduce the required training time and improve agent performance. They formed a set of tasks to consider and used an automated method to create the curriculum for the training process to avoid the burden of human handcrafting. Table IV summarizes the above findings.

TABLE IV: Key Techniques for AD and UAV Swarm Navigation

Ref	Technique	Application	Strengths and Limitations
[50]	CL with DRL	AD at urban intersections	Reduces training time; improves performance; uses automated curriculum creation; avoids human handcrafting.
[63]	Adaptive Embedded Curriculum	UAV swarm target search in disaster management	Adjusts task difficulty based on success rate; handles sparse rewards.
[54]	PPO with Curriculum Learning	Urban driving policy for CARLA simulator	Guides agent to better policy; improves learning efficiency.
[22]	RD-ACPPO	Self-driving tasks at unsignalized intersections	Adaptive curriculum learning; allocates importance weights dynamically; demonstrates effective curriculum transition timings and high task success rate.
[64]	Multi-agent scenario-based framework	CARLA simulator for AD	Practical for multi-agent learning; generates auto-curricula from scenario descriptions.
[62]	Curriculum Learning with GNN-RL	Motion planning in AD	Incorporates map and traffic rule information; decomposes learning task into sub-tasks with increasing difficulty; includes rewards for temporal-logic-based traffic rules.

CL can facilitate learning and training in CARLA [65] driving simulator and target search for UAV swarms. Anzalone etal. [54] aim to learn an end-to-end urban driving policy for the CARLA driving simulator by combining the PPO algorithm with CL. In their approach, the RL phase is divided into several stages of increasing difficulty, so that the agent is guided to learn a better driving policy. The curriculum-based agent is evaluated on different metrics, cities, weather conditions, and traffic scenarios. Furthermore, Brunnbauer etal. [64] develop a multi-agent, scenario-based AD training and evaluation framework for the CARLA simulator. The results demonstrate the proposed approach is practical for multi-agent learning and can generate auto-curricula from scenario descriptions. In a similar vein, Xiao etal. [63] use UAV swarm for target search in disaster management scenarios. An adaptive embedded curriculum can adjust the difficulty of the task based on the success rate achieved in training and is used to enable UAVs to explore and navigate cluttered 3D environments with sparse rewards efficiently.

B. Key Findings and Insights

From the analysis in the previous subsection it became clear that CL, by systematically increasing the complexity of tasks, not only lays a solid foundation for the AVs learning process but also reduces training time and improves agent performance. This structured approach is crucial for preparing the vehicle to handle complex scenarios, where HITL-RL allows further fine-tuning the vehicle decision-making abilities.

In the sequel of our literature survey we highlight the following open issues:

*1) **Merger with Transfer Learning**:* When developing a curriculum that includes multiple tasks, the intermediate tasks may have different state/action spaces, reward functions, or

transition functions compared to the final task. Consequently, transfer learning is essential to extract and pass on reusable knowledge from one task to the next. Traditionally, transfer learning focuses on how to transfer knowledge from one or more source tasks directly to a target task. However, CL extends this concept by considering training sessions in which the agent repeatedly transfers knowledge from one task to another, resulting in a series of end tasks [66]. The integration of transfer learning and CL has significant practical implications in AVs. For instance, an AV might first learn basic navigation in a simple environment and transfer the learned knowledge to complex tasks like obstacle avoidance or object manipulation.

*2) **CL vs Hard Example Mining**:* When developing a curriculum for AVs, a common approach is to sort tasks from easy to hard, which is central to CL. However, this raises the question of whether this incremental approach is always the most effective strategy. An alternative approach, Hard Example Mining (HEM) [67], shifts the focus from a structured progression to selecting the most challenging examples for training. HEM operates under the assumption that harder examples, which are often identified by the model losses or gradient magnitude, provide more informative and valuable learning opportunities.

*3) **Human Defined CL vs Human Data**:* This leads to an important consideration: should we always rely on explicitly designed curricula, or can we leverage human data to generate curricula more naturally? Using real-world driving data from humans allows AVs to be exposed to a broad spectrum of realistic scenarios and behaviors, thereby enhancing their learning process. This data-driven approach can potentially offer a more nuanced and practical training experience than a pre-defined curriculum.

*4) **Co-learning: a multi-agent approach to CL**:* Additionally, another intriguing possibility is the use of co-learning, a multi-agent approach to CL where the curriculum emerges from the interactions among multiple AV within the same environment. In this setup, the curriculum evolves organically as AVs interact, creating an implicit and adaptive learning process. This interaction-driven method fosters continuous improvement, as the agents learn from each other and adapt to increasingly complex driving tasks over time, as discussed by [66]. This dynamic approach may offer significant advantages by reflecting real-world driving challenges and enabling more robust learning outcomes.

III. Human-In-The-Loop Reinforcement Learning

HITL-RL significantly enhances the RL process by leveraging techniques like reward shaping, action injection, and interactive learning. These methods assist AVs in tasks such as lane-change, driving, obstacle avoidance, navigation, and overall performance enhancement, thereby improving safety and fostering public trust. The setup of HITL-RL for AVs is shown in Fig. 3. In this setup, human experts play a crucial role in monitoring the RL process. They can intervene by injecting actions when necessary and can enhance the system

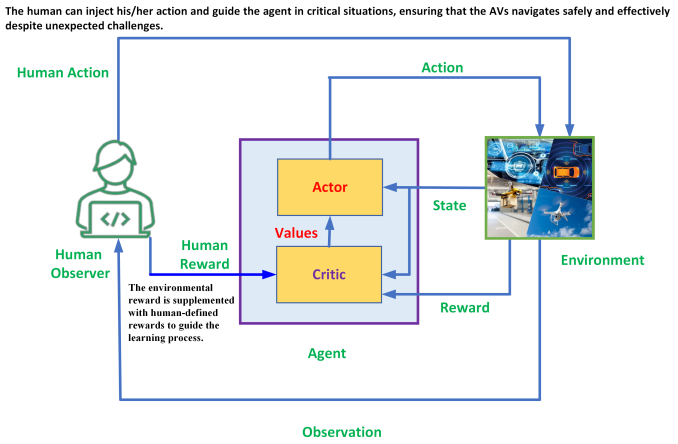

Fig. 3: The actor-critic setup of HITL-RL for AVs allows human experts to inject their actions and guide the agent in critical situations, ensuring that the AV navigates safely and effectively despite unexpected challenges. Additionally, the environmental reward is supplemented with human-defined rewards to guide the learning process. Both the environmental reward and the human reward are fed to the critic.

by providing supplementary rewards, known as human rewards. This integration ensures that the learning process remains safe, reliable, and aligned with human values and judgments.

A. Literature review

1) Reward Shaping: In reward shaping, the environmental reward is supplemented with human-defined rewards to guide learning. HITL can switch between multiple UAV reward functions in different situations and assist in constructing interactive reward functions. Humans can help to switch between different reward functions in different situations and replace the manual construction of the reward function.

Li etal. [41] develop a HITL-DRL framework for UAV navigation in large-scale 3D complex environments. This work allows the UAV to avoid obstacles automatically during flight. The role of HITL is to assist the UAV in obstacle avoidance by dynamically switching between multiple UAV reward functions in different situations. The combined approach significantly improves the success rate and navigation efficiency. In a similar vein, Karumanchi etal. [68] address the problem of autonomous UAV navigation and develop a practical solution based on Preference-based Reinforcement Learning (PbRL). In PbRL, the manual construction of the reward function is replaced with human feedback. A human is presented with a pair of trajectories followed by the RL agent and is asked to indicate a preference for one over the other. The PbRL algorithm then computes an optimal sequence of actions using only the set of preferences collected over different trajectories.

In the AD context, reward can be learned from human driving data and encoded with human prior experience. Zhuang etal. [69] address the intricate challenges in designing appropriate reward functions to manage optimizing complex goals in RL. They develop a framework that learns the reward function from human driving data to help manage the learning process. Trajectory sample, offline preference learning, and RL are

TABLE V: Key HITL-DRL Techniques with Reward Shaping

Ref	Technique	Application	Strengths and Limitations
[41]	HITL-DRL	UAV navigation	Improves obstacle avoidance and navigation efficiency; requires multiple reward functions.
[68]	PbRL	UAV navigation	Uses human feedback for reward; dependent on accurate feedback.
[69]	RL with human driving data	AV navigation	Learns reward function from human driving data; complexity in modeling human decisions.
[70]	HITL-RL with TD3	AV lane-change decisions	Encodes human experience in rewards and constraints; guided by human experts to limit unsafe exploration.

TABLE VI: Key HITL-DRL Techniques with Action Injection

Ref	Technique	Application	Strengths and Limitations
[42]	HITL-DRL with Human Guidance	Autonomous Vehicle navigation	Improves learning ability and safety; complexity in integrating human actions with DRL policy updates.
[71]	HITL-DRL Framework	UAV swarm navigation in large-scale 3D urban environments	Allows human intervention to correct dangerous actions; increased complexity in training Actor-Critic networks with human feedback.
[72]	HITL-DDPG with HCI Feedback	Autonomous UAV navigation in challenging environments	Handling of unforeseen delays due to human input processing.
[73]	Enhanced HITL-RL for Driving Policy Learning	AD in mixed-traffic environments	Improves safety and traffic flow efficiency; requires effective human supervision, intervention, and demonstration.

the comprising components, where feasible trajectories are generated by sampling end targets from a reachable state space, then a transformer network that is used to model the human driving decision-making process, and obtaining a reward function, is trained by a new offline preference learning where generated feasible trajectories are compared with human driving trajectories. Finally, the final driving decision network is obtained by incorporating the derived reward function in the RL framework. In another related effort, Huang etal. [70] aim to facilitate AVs mandatory lane-change scenarios in complex environments. In this context, they develop a HITL-RL method based on Twin Delayed Deep Deterministic policy gradient (TD3) for changing lane decisions. They encode human prior experience into reward function and safety constraints in an offline manner. Also, human experts guide the agents online to limit unsafe exploration during the training process and provide a demonstration in complex scenarios. Table V summarizes the above findings.

2) Action Injection: The human can inject his/her action and guide the agent in critical situations, ensuring that the AVs navigates safely and effectively despite unexpected challenges. Also, human guidance is used in real-time to improve the agent's learning ability and performance.

Introducing human guidance into the training loop of DRL algorithms enable AV to better handle different situations. Further, human intervention can make the UAV swarm safe. Wu etal. [42] state that the limited intelligence and capabilities of DRL-based strategies in AV applications prevent them from handling various situations and replacing human drivers. Introducing human guidance into the training loop of DRL algorithms can improve the situation. In the proposed HITL-DRL, human guidance is used in real-time to improve the agent's learning ability and performance. Specifically, the human participant observes the agent's real-time training process and assists when necessary. The provided action of the human guidance is prioritized as the output action of the agent to interact with the environment and replaces the original action of the DRL policy. The actor and critic are updated according to the human guidance and experience of the DRL agent. The actor-network can learn from both human guidance through imitation learning and the experience of the interactions through RL.

The critic network can evaluate the values of the agent's actions as well as the actions of human guidance. Similarly, Li etal. [71] develop a HITL-DRL framework for UAV swarm navigation in a large-scale 3D complex urban environment. This framework enables human intervention in the flight of the UAV swarm when necessary to correct dangerous actions. Additionally, the Actor-Critic structure of DDPG has been modified to identify and learn from positive human experiences. Human feedback can enhance UAV obstacle avoidance performance by having the human acting as a mentor to the AI agent of AD. Lee etal. [72] investigate autonomous navigation in challenging environments with the capability to avoid obstacles. The UAV navigates toward its target using the DDPG model, which can handle unexpected situations. HITL feedback is integrated with the DDPG model to enhance obstacle avoidance performance. This feedback is crucial when the pre-trained DDPG model encounters environments it is not fully adapted to, and gives the system the flexibility to handle unforeseen scenarios effectively and make real-time adjustments. For the HCI feedback, the first-person shooter concept is employed, where feedback is provided using keyboard directional keys. Through this HCI-based HITL feedback, real-time control of the UAV flight movement is achieved.

Huang etal. [73] aim to improve the safety of AD technology, ensuring traffic flow efficiency in mixed-traffic environments. They develop an enhanced HITL-RL for driving policy learning, where the human serves as a mentor to the AI agent, supervising, intervening, and demonstrating in its learning process, and integrated with DRL to guarantee the safety of the AVs, but also optimize the traffic flow efficiency. Table VI summarizes the above findings.

3) Interactive Learning: The agent and human trainer can interact and cooperate to solve tasks. For example, human can supplement environmental reward and guide learning objectives to improve learning efficiency and RL performance. Zhang etal. [43] aim to explore the oceans using autonomous underwater vehicles (AUV). RL can improve its autonomy but still has practical problems due to sparse rewards and low learning efficiency. They propose a deep interactive RL method for AUV path tracking by combining the advantages of DRL and

interactive RL. Since the human trainer cannot give human rewards to the AUV in the operational condition and the AUV needs to adapt to a changing environment when it is in the ocean, a DRL method that simultaneously learns from human rewards and environmental rewards is proposed. Wu etal. [74] improve RL performance in simulation and practice for the navigation task of UGVs using HITL-RL. The framework improves human interventions with a series of mechanisms to provide demonstrations and improve RL's capabilities. The mechanisms comprise human-guided learning objectives, prioritized human experience replay, and human intervention-based reward shaping.

Human behavior modeling mechanisms can alleviate the workload on humans, further using meta-model provides a language to describe the roles of UAVs and humans and autonomous decisions. Wu etal. [45] aim to improve the learning performance of RL by incorporating human guidance. In this regard, they develop a human guidance-based RL framework along with a novel mechanism for leveraging the experience of human guidance. A TD3-based algorithm was developed to improve algorithmic capabilities in the context of HITL-RL. The performance evaluation was conducted using AD tasks and confirmed the ability of the developed algorithm to improve the learning efficiency compared to state-of-the-art human guidance-based RL. In addition, a human behavior modeling mechanism is proposed to reduce the burden on humans. Additionally, Agrawal etal. [75] investigate semi-autonomous UAVs to support emergency response scenarios. Humans and multiple UAVs must work together as a team to successfully execute a time-critical mission. Therefore, a meta-model is proposed to describe the interaction between humans and multiple UAVs. The meta-model provides a language to describe the roles of UAVs and humans and autonomous decisions. They have used two different scenarios involving multiple semi-autonomous UAVs in emergency operations to demonstrate the use of the meta-model by formally specifying human interactions and points of intervention.

HITL-DRL can faciliate UAV operation in complex environments and enable mission planning, further HITL-DRL address high-level decision-making problems in AD. Zhang etal. [44] state that collision-free navigation for UAVs can facilitate data collection from IoT devices. However, traditional path planning methods are energy and computationally intensive, while DRL-based approaches tend to trap UAVs in complex environments. They propose a hybrid collision avoidance method for real-time UAV navigation in complex environments with unpredictable obstacles. They first develop a HITL-DRL training module for mapless obstacle avoidance and then set up a global planning module that generates some points as waypoint guidance. In addition, they have developed a novel algorithm for target updating that combines the HITL-DRL training module and the global planning module. In a similar vein, Muezzino ¨ glu etal. [51] develop an intelligent human–UAV interaction approach in real time based on ML using wearable gloves. The objective is to enable mission planning for controlling a UAV. ML-based methods process the signal from the gloves and classify appropriate commands to include in the human–UAV interaction process via the interface. The inclusion is according to the task scheduling algorithm to facilitate sequential and fast operation.

TABLE VII: Key HITL Techniques with Interactive Learning

Ref	Technique	Application	Strengths and Limitations
[43]	Deep Interactive RL	AUV path tracking	Addresses sparse rewards and low learning efficiency; practical implementation challenges.
[44]	Hybrid Collision Avoidance	UAV collision-free navigation	Effective in complex environments; requires integration of multiple modules.
[45]	Human Guidance-based RL	AD	Improves learning efficiency with human guidance; complexity in modeling human behavior.
[51]	ML-based Human-UAV Interaction	UAV mission planning	Facilitates sequential and fast operation; dependent on accurate signal processing.
[75]	Meta-Model for Human-UAV Interaction	Emergency response	Useful in time-critical missions; complexity in specifying interactions and interventions.
[76]	Value-based DRL with Human Guidance	AD	Improves training efficiency and adaptability; utilizes human guidance for high-level decision-making in multi-objective lane-change problems.
[74]	HITL-RL with Human-Guided Learning	UGV navigation task	Improves RL capabilities with human interventions; includes human-guided objectives, experience replay, and reward shaping.

Moreover, the authors of [76] aim to improve the training efficiency and adaptability of DRL methods to untrained cases. They develop a new value-based HITL-DRL to improve its performance for addressing high-level decision-making problems in AD. Specifically, the developed DRL algorithm was used to address a challenging multi-objective lane-change decision-making problem where a new learning objective for DRL is designed to value more the human policy over the undertrained DRL policy so that the DRL agent can be biased to mimic human behaviors and thereby utilizing human guidance more efficiently. They collected human guidance from a HITL driving experiment and evaluated their method in a high-fidelity simulator. Table VII summarizes the above findings.

B. Key Findings and Insights

From the discussion in the previous subsection we could see that the combination of human expertise with RL algorithms enables AVs to handle complex and diverse environments effectively. The potential AV bottleneck in handling large volumes of data, such as surveillance data and unlabeled data, can be addressed using AL to manage the situation. Below, we discuss open research questions concerning HITL-RL:

*1) **Designing AV Systems for Improved Decision-Making**:* A major open research question in the field of HITL-RL is how to design AV systems that improve decision-making over time through human interaction. To address this, a possible approach is to follow a human-centered design for DRL [77]. First,

a clear and testable hypothesis is formulated, e.g., *DRL can optimize traffic signal timings to reduce city-wide congestion by 30%*, ensuring it justifies human interaction and data collection. Second, the DRL design is aligned with human needs for safety, convenience, and sustainability, engaging stakeholders like researchers, commuters, and policymakers to assess impacts and maintain open communication. Finally, specific measurable, attainable, relevant, and timely goals are set, such as achieving a 30% reduction in traffic congestion within three years, and metrics are defined for both optimization and human-centric success.

2) ***Integration of Human Feedback in HITL-RL*** *:* The presence of actual humans in HITL-RL can significantly slow down the training process due to the time-consuming nature of integrating human feedback and the practical challenges of coordinating human interactions. Moreover, human input may introduce bias into the learning process. Simulating human involvement through expert-designed programs offers a practical solution to address this issue. These simulations can model human decision-making and behavior efficiently, allowing for rapid, consistent training without the delays of real human interaction. Simulated human involvement assists experts in reducing bias, and enhance time efficiency, thus accelerating the training process and improving the effectiveness of the HITL-RL. However, simulated human do not behave like actual humans, and results with simulated human can differ from those with human users [78].

3) ***Agent-Specific vs Agent-Agnostic Protocols****:* Agent-specific protocols, meaning protocols for human interaction that are designed for a specific HITL-RL algorithm are preferred over agent-agnostic protocols. Agent-specific protocols allow humans to provide feedback and advice that is directly relevant to the AV learning algorithm. By understanding the vehicle specific learning process and representation, the protocols enable humans to ask more precise and informative questions. This leads to higher-quality guidance that is more likely to accelerate the vehicle's learning and improve its decision-making capabilities. In contrast, agent-agnostic protocols are not specialized to the agent and therefore cannot ask human informative questions or have an observation model that accurately represents the human's process for generating advice. Similarly, humans cannot provide optimally informative advice to the agent because they are unaware of the agent's prior knowledge, exploration techniques, representation, or learning method [79].

4) ***Evaluating HITL-ML Algorithms****:* Despite of significant contributions towards HITL-ML algorithms, most studies focus on improving the performance of algorithms rather than the interaction experience, but proper evaluation is essential to create practical applications for different types of end-users. The HITL-ML necessitates incorporating subjective user evaluations and acknowledging the cyclic influence between algorithms and users, making HITL-ML performance evaluation inherently subjective and complex [80] [81].

5) ***Robustness Against Adversarial Attacks****:* HITL-ML can make the models more robust against adversarial attacks. Adversarial attacks are deliberate attempts to fool ML models by providing deceptive input. When human are engaged in the model development process, they can provide valuable insights, detect anomalies, and make adjustments that might not be apparent through automated processes alone. This human involvement can help identify and mitigate potential vulnerabilities that adversarial attacks might exploit, thus reducing the severity and impact of such attacks on the models [82] [83].

6) ***Combining with safe RL Techniques*** *:* The addition of Safe RL techniques to HITL-RL setting can further reduce the likely of dangerous actions. Safe RL techniques ensure good performance and respect safety constraints during learning and deployment [82] [84].

IV. Active Learning

AL optimizes the annotation process, accelerates development, and enhances system robustness by leveraging human expertise effectively. This can be very valuable in the scope of AVs, from UAV anomaly mitigation and semantic aerial mapping to intrusion detection and 3D object detection in AD. Annotating extensive data for training AVs is challenging due to their complex environments and unexpected situations. AL addresses these issues by reducing manual labeling through targeted instance selection, improving the quality of labeled data, and the accuracy of networks. The amount of data needed to train autonomous aerial or ground vehicles is enormous. AVs must navigate a variety of complex and diverse environments, which necessitate being represented in the training data. Moreover, AVs must deal with unexpected situations. Thus, to master these events, the model needs extensive data that captures such occurrences. Annotating such a volume of data for training autonomous aerial and ground vehicles can be an overwhelming and tedious task for humans and can pose a major challenge to AV development.

AL can be applied to this process, to cut down on the time and cost spent on training, while also increasing the accuracy of the networks. AL makes it possible to automate the selection process while choosing valuable data points. It starts by training a dedicated ML model on already-labeled data. The network then sorts through unlabeled data, uses the acquisition function to identify the instances it is most uncertain about, thereby finding data that would be challenging to the AVs algorithm and asks a human to label these specific instances which are added to training data and then the model is trained with labeled data. This approach reduces the number of instances that require manual labeling and ensures that the labeled data is as informative as possible [85].

As can be seen in Fig. 4, AVs can be confronted with unknown and unexpected scenarios and AL can be used to annotate them. The bottom right of the figure shows the AL loop including the training, object selection, annotation and appending steps. In the annotation step, the selected examples are annotated by human experts. AL can help to cope with unseen situations, reducing training time and costs and improve training efficiency.

Fig. 4: AVs face unknown and unexpected scenarios and utilize AL to annotate them. The bottom right shows the AL loop comprising the steps of training, object selection, annotating, and appending. In the annotation step, human experts annotate the selected examples. AL impacts on unseen situations and training are shown in the top right.

TABLE VIII: Key AL Techniques for UAV Security and Navigation

Ref	Technique	Application	Strengths and Limitations
[46]	Semi-Supervised SVM with AL	In-flight UAV anomaly mitigation	Addresses lack of unlabeled data; requires informative sample selection.
[47]	AL for Path Planning	Aerial semantic mapping	Adapts path to regions with informative data; depends on effective AL features.
[48]	AL for Intrusion Detection	UAV network security	Efficiently learns novel attacks with low labeling budget; dependent on least confidence strategy.

A. Literature Review

1) Security and Navigation: Pan etal. [46] aim to mitigate in-flight UAV anomalies by using a semi-supervised SVM. Since the pure semi-supervised SVM classification method may not classify well due to a lack of unlabeled data with rich content. AL is used to search the samples with rich information and significant differences for labeling to improve the quality of labeled data. Building on this concept, Ruckin etal. [47] have developed a new path planning framework for AL in aerial semantic mapping to improve UAV semantic awareness with minimal expert guidance. In an innovative approach, path planning goals are combined with AL features that allow us to adapt the UAV path to regions with informative training data. Similarly, Zhang etal. [48] design the intrusion detection system to detect unknown network attacks in the UAV network. The Open-CNN model is proposed to detect unknown attacks. Meanwhile, an AL approach is proposed for the unknown attacks based on the least confidence strategy, which enables the intrusion detection model to efficiently learn about observed novel network attacks with a low labeling budget. Table VIII summarizes the above findings.

2) Detection: Liang etal. [52] are concerned with 3D object detection in AV. DL-based object detectors require large annotated datasets, which incurs compilation costs. To address this problem, they developed diversity-based AL, which uses the multimodal information of an AV dataset to propose a novel detection function that enforces spatial and temporal diversity in the selected samples. Simply, the most informative frames and objects are automatically selected for annotation by humans. Similarly, Ding etal. [21] address vehicle type recognition (VTR) from surveillance data in intelligent transportation. To facilitate manual labeling of large-scale surveillance data, they propose a VTR system based on deep AL. A deep neural network is efficiently trained by actively, incrementally, and automatically selecting unlabeled data points for human labeling. The training procedure is repeated several times so that the neural network is re-trained with the incrementally updated training data. Unlike traditional AL where the samples with high entropy are selected as the network output, in the proposed method, samples with both high and low entropy values are selected to be included in the training data. The proposed method is evaluated using the surveillance images from the CompCars public dataset. In addition, Satzoda etal. [86] develop a two-part technique for vehicle detection using an AL system and a symmetry-based iterative analysis. A modified AL technique is proposed that automatically selects positive and negative samples from multiple parts to train adaboost classifiers. The results of the classifiers are linked using a novel iterative window search algorithm and a symmetry-based regression model to extract fully visible vehicles. The AL method combined with the symmetry-based analysis has been shown to improve detection rates and also reduce false alarms compared to existing methods. Table IX summarizes the above findings.

TABLE IX: Key AL Techniques for Vehicle Type Recognition and Detection

Ref	Technique	Application	Strengths and Limitations
[21]	Deep AL for VTR	Vehicle type recognition from surveillance data	Uses high and low entropy samples; requires multiple training iterations.
[86]	AL with Symmetry-based Iterative Analysis	Vehicle detection	Reduces false alarms; requires iterative window search and regression model.
[52]	Diversity-based AL	3D object detection in AV	Enforces spatial and temporal diversity in selected samples; reduces annotation costs by selecting the most informative frames and objects.

B. Key Findings and Insights

From the survey in the previous subsection we identify the following open issues that need further research for better leveraging the AL potential in the scope of AVs:

*1) **Batch Training vs Sequential Training**:* One critical aspect of the AL process is determining how many new instances should be labeled before retraining the model. In batch training, several examples are labeled and accumulated before the model undergoes retraining. This approach allows the system to process and learn from a collection of labeled instances in one go. Conversely, in sequential training, the model is retrained immediately after each new instance is labeled, providing immediate feedback to the AV. Although sequential training allows for continuous updates, it is often impractical due to the high cost of retraining after each labeled instance. Therefore, selecting batches of instances for labeling

is preferred, as it reduces the number of training steps and enhances overall efficiency.

*2) **Annotator Quality**:* Despite these efficiencies, AL processes are typically based on the assumption that human-labeled data is of high quality and that humans consistently act correctly. In practice, however, some instances can be challenging for both models and humans. Human annotators may introduce variability due to factors like distraction or fatigue, leading to inconsistent annotation quality. While employing multiple annotators can mitigate this issue, it still requires careful decision-making to determine the reliability of the labels.

*3) **Transfer Learning**:* Integration of transfer learning and AL can enhance the training efficiency and adaptability of AV systems. Transfer learning can significantly reduce the need for new data collection and expedite the learning process for new driving tasks, by using pre-trained models, which have already been trained on extensive and related datasets.

*4) **Uncertainty Sampling**:* Uncertainty sampling, a common method in AL, helps identify which data points require human labeling. The ML system decides whether to query for additional labels based on a predefined uncertainty threshold. For self-driving cars, this threshold must be set very low, as even minor errors in data labeling can lead to severe consequences, such as accidents.

*5) **Diverse Road Condition**:* Moreover, the diverse and variable nature of road conditions worldwide presents a significant challenge in preparing self-driving cars for all possible scenarios. AL leverages both human expertise and ML to ensure that important details are not overlooked. This integration is crucial for developing AVs that are capable of safely navigating a wide range of driving environments globally.

V. Ethical Implications of AVs

We have seen that CL, HITL-RL and AL improve the capabilities of AVs. We emphasize that these HITL-ML techniques can promote self-recognition and scene recognition in AVs to assist in managing complex situations such as accidents, but ML algorithms in AI in general bring forth critical ethical and legal considerations. At the end, AVs should act as ethically rational agents and consider the ethical and legal implications of their actions, not just the outcomes.

Therefore, ethical principles must be encoded into AVs to ensure that their behavior aligns with societal values and norms. To this end, ethical theories and principles must be integrated into the decision-making processes of AVs. For example, an autonomous UAV can be considered a rational ethical agent if it accurately perceives its environment and makes decisions based on ethical frameworks such as utilitarianism where the ethical worth of an action is evaluated based on its consequences, with a focus on maximizing utility or happiness [87], or deontology where the ethical actions should be guided by a set of ethical principles or rules, regardless of the consequences [88], and acts in a way that is consistent with ethical principles, ensuring that its actions are "right" and beneficial to all stakeholders. The following subsection reviews the state-of-the-art in this topic.

A. Literature Review

Complex situations involving human live need ethical consideration. Konert etal. [89] analyze the legal and ethical aspects of using UAVs, and state that we can achieve a higher degree of trust by formalizing the process of ensuring the safe and compatible functioning of systems based on AI and ML. However, there are some challenges in realizing ethical decision-making in AVs. One of the salient challenges is that ethical dilemmas often involve complex trade-offs that are difficult to quantify and automate. For example, how should AVs decide in accident situations in which human harm is a likely consequence? and prioritize passenger safety over pedestrian safety? For example, Vakili etal. [90] address ethical decision-making in complex urban driving scenarios in the case that human lives are at risk. They develop an ethical decision-making model based on DQN, and its main goal is to minimize injuries resulting from self-driving car crashes. The model considers factors such as the type of user involved in the accident, their age, and injury severity, the dynamic conditions of the self-driving car such as speed and acceleration, and the ethical considerations for the car. Wang etal. [91] address ethical decision-making during crashes involving humans and propose a predictive control framework for ethical decision-making in AD using rational ethics. The Lexicographic Optimization-based model predictive controller (LO-MPC) is designed to facilitate ethical rules implementation, In which obstacles and constraints are prioritized.

Different approaches can be taken for complex situations but still AVs are facing with the risk of accidents. Contissa etal. [92] delegate to the user/passenger the task of deciding which ethical approach should be taken by AVs in unavoidable accident scenarios. Indeed, passengers are enabled to ethically customize their AVs, namely, to choose between different settings corresponding to different ethical approaches or principles. Kim etal. [93] develop a computational model for building ethical AVs by learning and generalizing from human ethical judgments. They cast the problem of ethical learning for AVs as learning how to weigh the different features of the dilemma using utility calculus, to make these trade-offs reflect how people make them in a wide variety of ethical dilemmas. Martinho etal. [94] investigate the ethics of AD and conclude that AD stakeholders are aware of the ethics of AD technology, and have different approaches concerning the authority of remote operators; the scientific community prioritizes safety and cybersecurity and agrees that AVs will not eliminate the risk of accidents.

A code of ethics needs to be developed for AVs, and scientific bodies should provide recommendation. The German Ethics Commission [95] presents a code of ethics consisting of 20 ethical guidelines for Automated and Connected Driving and prioritizes safeguarding human life and rejecting discrimination based on personal traits in accidents. IEEE advocates for fair treatment and non-discrimination based on personal characteristics [96]. AVs often operate in dynamic environments where ethical considerations can change quickly. The system must be able to adapt its ethical decision-making processes in real-

TABLE X: Ethical Decision-Making Models for AVs

Ref	Technique	Application	Strengths and Limitations
[89]	Formalization of Ethical Processes	UAVs	Enhances trust by formalizing AI/ML system processes; challenges in automating complex ethical dilemmas.
[90]	DQN-Based Ethical Decision-Making Model	Urban driving scenarios for AVs	Considers factors like user type, age, injury severity; aims to minimize crash injuries; difficulty in quantifying ethical trade-offs.
[91]	LO-MPC Framework	Crash scenarios involving humans	Implements ethical rules using lexicographic optimization; prioritizes obstacles and constraints; balancing multiple ethical considerations.
[92]	User-Customizable Ethical Settings	AVs in unavoidable accident scenarios	Allows passengers to choose ethical settings; promotes customization; relies on passengers' ability to make ethical choices.
[93]	Utility Calculus for Ethical Learning	Building ethical AVs	Learns and generalizes from human ethical judgments; adapts ethical trade-offs; depends on accurate representation of human ethical decision-making.
[94]	Stakeholder Awareness	Ethical considerations in AD technology	Highlights diverse stakeholder approaches and priorities; emphasizes safety and cybersecurity; acknowledges residual accident risks.
[95]	Code of Ethics for Automated Driving	Automated and Connected Driving	Provides 20 guidelines focusing on human life and non-discrimination; sets ethical standards; may need adaptation to evolving scenarios.
[96]	Fair Treatment and Non-Discrimination	General AV ethics	Advocates for fair treatment based on characteristics; promotes non-discrimination; requires practical implementation in dynamic environments.
[97]	Real-Time Ethical Adaptation	ML algorithms in AVs	Ensures transparency and accountability; addresses technical and regulatory challenges; focuses on dynamic adaptation of ethical decision-making.

time. It must be ensured that the ML algorithms of AVs are transparent in their decision-making processes and that there is accountability for their actions. This involves not only technical challenges but also legal and regulatory considerations [97]. Table X summarizes the above findings.

B. Key Findings and Insights

To effectively address the above challenges and overcome the limitations of ML in the ethical assessment of AVs, it is important to bring humans into the loop to improve self-recognition and scene-recognition capabilities. Self-recognition refers to the AI's ability to understand its state, limitations, and role in a given context. Scene recognition is about accurately interpreting the environment, including dynamic elements such as people, obstacles, and changes in the environment [98].

However, another aspect emerges from the previous discussion, which is that of accountability and its importance in addressing the referred challenges. Accountability, in the context of AVs, is the framework of responsibilities, processes, and mechanisms that ensure these vehicles safe, ethical, and lawful operation. This includes clearly defining the roles of those involved, ensuring transparency and explainability of the AVs decisions, monitoring and evaluating performance, taking corrective action when necessary, and enforcing consequences for non-compliance or failure. This framework aims to build trust, ensure safety, and maintain ethical standards in the introduction and operation of AVs. Humans play a crucial role in ensuring accountability in AVs by providing responsive monitoring and performance evaluation.

Overall, integrating ethical considerations into the decision-making processes of AVs is critical to aligning these systems with societal norms and values. HITL-ML methods enhance the ethical decision-making capabilities of AVs by providing real-time adaptability, accountability, and improved detection capabilities. Incorporating HITL is essential for monitoring, evaluating and controlling AV operations to ensure ethical standards and safe, reliable performance in dynamic environments.

VI. Open Research Directions

In the previous sections we already pointed to several open issues emerging from the application of HITL-ML technologies to AVs. These technologies enhance learning capabilities, improve performance, and enable safer and more responsive systems in critical situations. In this section we bring up remaining opportunities to further advance HITL-ML globally within the scope of AVs.

A. *Explainable AVs*

AI approaches, currently dominated by DRL algorithms, have led to significant improvements in many essential components of AVs, including advances in navigation and control. However, these techniques fail to contribute to the safe and hazard-free navigation of AVs, and their black-box nature makes it impossible to understand the causes of actions or decisions. AVs equipped with AI will only gain social acceptance if people have confidence in their operation. Providing explanations greatly increases people's trust in AVs. Therefore, developing explainable AVs is a promising approach for trustworthy AV use. When trustworthiness is developed for AVs' intelligent decision-making, it also brings transparency and accountability to AVs technology. What methods can be used to evaluate the effectiveness of explainable AI in improving public trust and transparency in the decision-making processes of AVs? Another key question is that how can explainable AI techniques be integrated into scene understanding systems for AVs to enhance safety and hazard-free navigation in unstructured traffic environments?

B. *Timely Human Intervention*

Ensuring timely human intervention is a salient challenge in developing and adopting HITL-ML for AVs. How can advanced interfaces and real-time feedback mechanisms be developed to enhance the effectiveness of HITL-ML in AVs? Moreover, how can we design intuitive interfaces that facilitate seamless

human-AI interaction, making it easier for humans to input their knowledge into the AV to improve feedback quality and accelerate the learning cycle? One practical option is a hand gesture recognition system, which identifies the moment a hand gesture is performed and its class. Hand gesture recognition based on electromyography signals and inertial measurement unit signals has the potential to help control robots and vehicles. However, these systems should be used in conjunction with ML techniques to achieve greater accuracy and better handle complex situations.

C. *Safety*

Safety for HITL-ML in AVs includes a set of protocols, procedures and guidelines for humans and AV operations. These safety measures are designed to ensure that HITL-ML in AVs do not cause harm to users, passengers and pedestrians. At its core, the safety of HITL-ML in AVs is about using human commands, sensory data and other technologies to guide an AV through the streets while avoiding obstacles or potential traffic hazards. How can safe AV systems be designed and implemented combining human judgment and machine precision effectively? What specific roles should human operators play during critical situations? How can machine precision be used for real-time data processing, accurate decision-making, and consistent execution of safety protocols? How can these measures enhance the overall reliability and safety of AVs?

D. *Security*

Adverse conditions or potential cyber-attacks should not degrade the performance of HITL-ML for AVs, compromising their security and reliability. Security performance includes safeguarding human operators from malicious tampering and ensuring they can operate safely without causing unintended harm. Additionally, there is a risk of session hijacking, where an attacker takes over a session between a vehicle and a human operator. To prevent session hijacking, strong security measures such as encryption and two-way authentication are recommended. Authentication protocols should align with minimum overhead requirements, keep efficiency in message transmission, and cope with open unreliable media which is the case in AVs. Moreover, key agreement based on post-quantum cryptography and in particular Ring Learning with Errors (RLWE) is suggested to provide robust authentication for both humans and AVs. Overall, we need to establish rigorous testing, validation, and implementation of security protocols to prevent failures or exploits.

E. *Regulatory Frameworks*

The regulations for HITL-ML in AVs are designed to meet the public expectations around safety, legal responsibility, and privacy. The regulations ensure that HITL-ML systems in AVs are rigorously tested and meet high safety standards. How can comprehensive regulatory frameworks be developed and implemented to ensure the safe and effective integration of HITL processes into AV deployment, including defining the responsibilities for human operators, the safety standards and protocols for their intervention, and the ethical and legal boundaries of HITL systems? Moreover, how can we balance human effort and machine autonomy and regulate communication between humans and AVs to ensure effective status reporting and support in unforeseen situations? What are the most effective methods for documenting and evaluating the actions of both ML systems and human operators in HITL-ML frameworks for AVs to ensure appropriate legal accountability? How can regulations be designed to effectively protect human privacy in HITL-ML systems for AVs, ensuring transparent, consensual data collection and usage while preventing unauthorized access and breaches?

F. *Human Input*

How can we ensure that human input is reliable, ethical, and consistent in these systems? How can the integration of human feedback into AI models of AVs be managed meticulously to avoid the introduction of human biases, ensuring that the subjective nature of human judgment does not skew the AI models away from objective decision-making? What constant vigilance and sophisticated validation mechanisms are needed to address these biases and ensure that human contributions enhance rather than detract model performance?

VII. Conclusion

In this paper we reviewed the literature on HITL-ML for AVs, specifically on CL, HITL-RL and AL technologies. We noted that CL helps organize the tasks in AVs from simple to complex and facilitates the learning process. HITL-RL promotes public trust and enhances safety. AL leads to better results in tasks such as anomaly detection, semantic mapping, object detection, and vehicle detection, also contributing to the safety of AVs. We then investigated the ethical implications of HITL-ML and surveyed the related literature concerning enhancing self-recognition and scene recognition in AVs to assist in managing complex situations. We also observed the importance of accountability in AVs operation. In all these topics we point to state-of-the-art works available in the literature. Finally, we go beyond the previous studies and enumerate a few global topics in HITL-ML applied to AVs that constitute current research opportunities. Namely, we believe that developing explainable AVs, enabling timely human intervention, enhancing safety and security, defining regulatory frameworks along with proper human input are critical for effective HITL-ML adoption.

References

[1] P. Koopman and M. Wagner, "Autonomous vehicle safety: An interdisciplinary challenge," *IEEE Intelligent Transportation Systems Magazine*, vol. 9, no. 1, pp. 90–96, Jan. 2017.

[2] Fortune Business Insights, "Autonomous vehicle market size, share, growth & trends," 2024, accessed: 2024-07-03. [Online]. Available: https://www.fortunebusinessinsights.com/autonomous-vehicle-market-109045

[3] I. Yaqoob, L. U. Khan, S. M. A. Kazmi, M. Imran, N. Guizani, and C. S. Hong, "Autonomous driving cars in smart cities: Recent advances, requirements, and challenges," *IEEE Network*, vol. 34, no. 1, pp. 174–181, Aug. 2020.

[4] Y. Yang, Y. Xiao, and T. Li, "A survey of autonomous underwater vehicle formation: Performance, formation control, and communication capability," *IEEE Communications Surveys & Tutorials*, vol. 23, no. 2, pp. 815–841, Feb. 2021.

[5] J. (Jonak) Lisinska, *Autonomous vehicles: automotive, maritime and aerial: A policy landscape review*, M. Kleinman, Ed., Dec. 2021.

[6] 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, Feb. 2021.

[7] K. Li, L. Lu, W. Ni, E. Tovar, and M. Guizani, "Secret key agreement for data dissemination in vehicular platoons," *IEEE Transactions on Vehicular Technology*, vol. 68, no. 9, pp. 9060–9073, Jul. 2019.

[8] ——, "Cooperative secret key generation for platoon-based vehicular communications," in *IEEE International Conference on Communications (ICC)*, Shanghai, China, May 2019.

[9] K. Li, W. Ni, E. Tovar, and M. Guizani, "Optimal rate-adaptive data dissemination in vehicular platoons," *IEEE Transactions on Intelligent Transportation Systems*, vol. 21, no. 10, pp. 4241–4251, Oct. 2020.

[10] ——, "Lcd: Low latency command dissemination for a platoon of vehicles," in *IEEE International Conference on Communications (ICC)*, Kansas City, MO, USA, May 2018.

[11] H.-H. Jebamikyous and R. Kashef, "Autonomous vehicles perception (avp) using deep learning: Modeling, assessment, and challenges," *IEEE Access*, vol. 10, pp. 10 523–10 535, Jan. 2022.

[12] S. Tao, J. Chen, B. Zhou, and H. Zhang, "Simultaneous robust state and sensor fault estimation of autonomous vehicle via synthesized design of dynamic and learning observers," *IEEE Transactions on Vehicular Technology*, vol. 73, no. 2, pp. 1753–1764, Sep. 2024.

[13] Z. Qi, T. Wang, J. Chen, D. Narang, Y. Wang, and H. Yang, "Learning-based path planning and predictive control for autonomous vehicles with low-cost positioning," *IEEE Transactions on Intelligent Vehicles*, vol. 8, no. 2, pp. 1093–1104, Jan. 2023.

[14] T. Sigma, "Why human intuition is essential in machine learning," 2024, accessed: 2024-08-27. [Online]. Available: https://www.twosigma.com/articles/why-human-intuition-is-essential-in-machine-learning/

[15] Tesla, Inc., "Model s owner's manual," 2024, accessed: 2024-07-03. [Online]. Available: https://www.tesla.com/ownersmanual/models/en_us/GUID-E5FF5E84-6AAC-43E6-B7ED-EC1E9AEB17B7.html

[16] L. Tomaszewski, "How fixed-gear bikes can confuse google's self-driving cars," 2015, accessed: 2024-07-03. [Online]. Available: https://www.washingtonpost.com/news/innovations/wp/2015/08/26/how-fixed-gear-bikes-can-confuse-googles-self-driving-cars/

[17] CNN, "China's baidu launches robotaxi service amid public anxiety," July 2024, accessed: 2024-07-31. [Online]. Available: https://edition.cnn.com/2024/07/18/cars/china-baidu-apollo-go-robotaxi-anxiety-intl-hnk/index.html

[18] T. Litman, "Autonomous vehicle implementation predictions: Implications for transport planning," 2015. [Online]. Available: https://api.semanticscholar.org/CorpusID:264152433

[19] X. Wu, L. Xiao, Y. Sun, J. Zhang, T. Ma, and L. He, "A survey of human-in-the-loop for machine learning," *Future Generation Computer Systems*, vol. 135, pp. 364–381, Oct. 2022.

[20] P. Zhao, Z. Yuan, K. Thu, and T. Miyazaki, "Real-world autonomous driving control: An empirical study using the proximal policy optimization (ppo) algorithm," *Transdisciplinary Research and Education Center for Green Technologies …*, 2024.

[21] X. Ding, J. Wang, C. Dong, and Y. Huang, "Vehicle type recognition from surveillance data based on deep active learning," *IEEE Transactions on Vehicular Technology*, vol. 69, no. 3, pp. 2477–2486, Mar. 2020.

[22] Z. Peng, X. Zhou, L. Zheng, Y. Wang, and J. Ma, "Reward-driven automated curriculum learning for interaction-aware self-driving at unsignalized intersections," *arXiv preprint arXiv:2403.13674*, 2024.

[23] O. Benderius, C. Berger, and V. Malmsten Lundgren, "The best rated human–machine interface design for autonomous vehicles in the 2016 grand cooperative driving challenge," *IEEE Transactions on Intelligent Transportation Systems*, vol. 19, no. 4, pp. 1302–1307, Oct. 2018.

[24] J. Klasinc, "Legal framework for autonomous vehicles in eu and croatia," in *45th Jubilee International Convention on Information, Communication and Electronic Technology (MIPRO)*, Opatija, Croatia, May 2022, pp. 1253–1257.

[25] K. Li, W. Ni, Y. Emami, Y. Shen, R. Severino, D. Pereira, and E. Tovar, "Design and implementation of secret key agreement for platoon-based vehicular cyber-physical systems," vol. 4, no. 2, 2019.

[26] H. Kurunathan, K. Li, W. Ni, N. Li, E. Tovar, and M. Guizani, "Exploring lstm-assisted a2c for physical layer security in vehicular cyber-physical systems," in *2024 International Wireless Communications and Mobile Computing (IWCMC)*, Ayia Napa, Cyprus, May 2024, pp. 1177–1182.

[27] H. Mankodiya, M. S. Obaidat, R. Gupta, and S. Tanwar, "Xai-av: Explainable artificial intelligence for trust management in autonomous vehicles," in *2021 International Conference on Communications, Computing, Cybersecurity, and Informatics (CCCI)*, Oct. 2021, pp. 1–5.

[28] L. Duan, L. Huang, C. Langbort, A. Pozdnukhov, J. Walrand, and L. Zhang, "Human-in-the-loop mobile networks: A survey of recent advancements," *IEEE Journal on Selected Areas in Communications*, vol. 35, no. 4, pp. 813–831, Apr. 2017.

[29] E. Mosqueira-Rey, E. Hernández-Pereira, D. Alonso-Ríos, J. Bobes-Bascarán, and Á. Fernández-Leal, "Human-in-the-loop machine learning: A state of the art," *Artificial Intelligence Review*, vol. 56, no. 4, pp. 3005–3054, Aug. 2023.

[30] P. T. Rajendran, H. Espinoza, A. Delaborde, and C. Mraidha, "Human-in-the-loop learning methods toward safe dl-based autonomous systems: A review," in *Computer Safety, Reliability, and Security. SAFECOMP Workshops*. Cham: Springer International Publishing, Aug. 2021, pp. 251–264.

[31] S. Kumar, S. Datta, V. Singh, D. Datta, S. Kumar Singh, and R. Sharma, "Applications, challenges, and future directions of human-in-the-loop learning," *IEEE Access*, vol. 12, pp. 75 735–75 760, May 2024.

[32] M. Wilchek, W. Hanley, J. Lim, K. Luther, and F. A. Batarseh, "Human-in-the-loop for computer vision assurance: A survey," *Engineering Applications of Artificial Intelligence*, vol. 123, p. 106376, May. 2023.

[33] S. Narvekar, B. Peng, M. Leonetti, J. Sinapov, M. E. Taylor, and P. Stone, "Curriculum learning for reinforcement learning domains: A framework and survey," *J. Mach. Learn. Res.*, vol. 21, no. 1, Jan. 2020.

[34] M. Soori, B. Arezoo, and R. Dastres, "Artificial intelligence, machine learning and deep learning in advanced robotics, a review," *Cognitive Robotics*, vol. 3, pp. 54–70, Apr. 2023.

[35] Y. Ma, Z. Wang, H. Yang, and L. Yang, "Artificial intelligence applications in the development of autonomous vehicles: a survey," *IEEE/CAA Journal of Automatica Sinica*, vol. 7, no. 2, pp. 315–329, Mar. 2020.

[36] D. S. Nunes, P. Zhang, and J. Sá Silva, "A survey on human-in-the-loop applications towards an internet of all," *IEEE Communications Surveys & Tutorials*, vol. 17, no. 2, pp. 944–965, May 2015.

[37] P. S. Chib and P. Singh, "Recent advancements in end-to-end autonomous driving using deep learning: A survey," *IEEE Transactions on Intelligent Vehicles*, vol. 9, no. 1, pp. 103–118, Sep. 2024.

[38] J. Zhao, W. Zhao, B. Deng, Z. Wang, F. Zhang, W. Zheng, W. Cao, J. Nan, Y. Lian, and A. F. Burke, "Autonomous driving system: A comprehensive survey," *Expert Systems with Applications*, vol. 242, p. 122836, Dec. 2024.

[39] H. Wang, A. Khajepour, D. Cao, and T. Liu, "Ethical decision making in autonomous vehicles: Challenges and research progress," *IEEE Intelligent Transportation Systems Magazine*, vol. 14, no. 1, pp. 6–17, Mar. 2022.

[40] N. De Moura, R. Chatila, K. Evans, S. Chauvier, and E. Dogan, "Ethical decision making for autonomous vehicles," in *IEEE Intelligent Vehicles Symposium (IV)*, Las Vegas, NV, USA, Jan. 2020, pp. 2006–2013.

[41] X. Li, J. Fang, K. Du, K. Mei, and J. Xue, "Uav obstacle avoidance by human-in-the-loop reinforcement in arbitrary 3d environment," in *42nd Chinese Control Conference (CCC)*, Tianjin, China, Jul. 2023, pp. 3589–3595.

[42] J. Wu, Z. Huang, Z. Hu, and C. Lv, "Toward human-in-the-loop ai: Enhancing deep reinforcement learning via real-time human guidance for autonomous driving," *Engineering*, vol. 21, pp. 75–91, Jul. 2023.

[43] Q. Zhang, J. Lin, Q. Sha, B. He, and G. Li, "Deep interactive reinforcement learning for path following of autonomous underwater vehicle," *IEEE Access*, vol. 8, pp. 24 258–24 268, Jan. 2020.

[44] S. Zhang, Y. Li, F. Ye, X. Geng, Z. Zhou, and T. Shi, "A hybrid human-in-the-loop deep reinforcement learning method for uav motion planning for long trajectories with unpredictable obstacles," *Drones*, vol. 7, no. 5, p. 311, May 2023.

[45] J. Wu, Z. Huang, W. Huang, and C. Lv, "Prioritized experience-based reinforcement learning with human guidance for autonomous driving,"

IEEE Transactions on Neural Networks and Learning Systems, vol. 35, no. 1, pp. 1–15, Jan. 2022.

[46] D. Pan, L. Nie, W. Kang, and Z. Song, "Uav anomaly detection using active learning and improved s3vm model," in *International Conference on Sensing, Measurement & Data Analytics in the era of Artificial Intelligence (ICSMD)*, Xi'an, China, Oct. 2020, pp. 253–258.

[47] J. Rückin, F. Magistri, C. Stachniss, and M. Popović, "An informative path planning framework for active learning in uav-based semantic mapping," *IEEE Transactions on Robotics*, vol. 39, no. 6, pp. 4279–4296, Dec. 2023.

[48] Z. Zhang, Y. Zhang, J. Niu, and D. Guo, "Unknown network attack detection based on open-set recognition and active learning in drone network," *Transactions on Emerging Telecommunications Technologies*, vol. 33, no. 10, p. e4212, Jan. 2022.

[49] C. Yan, C. Wang, X. Xiang, K. H. Low, X. Wang, X. Xu, and L. Shen, "Collision-avoiding flocking with multiple fixed-wing uavs in obstacle-cluttered environments: A task-specific curriculum-based madrl approach," *IEEE Transactions on Neural Networks and Learning Systems*, vol. 35, no. 8, pp. 1–15, Feb. 2023.

[50] Z. Qiao, K. Muelling, J. M. Dolan, P. Palanisamy, and P. Mudalige, "Automatically generated curriculum based reinforcement learning for autonomous vehicles in urban environment," in *IEEE Intelligent Vehicles Symposium (IV)*, Changshu, China, Jun. 2018, pp. 1233–1238.

[51] T. Müezzinoğlu and M. Karaköse, "An intelligent human–unmanned aerial vehicle interaction approach in real time based on machine learning using wearable gloves," *Sensors*, vol. 21, no. 5, p. 1766, Mar. 2021.

[52] Z. Liang, X. Xu, S. Deng, L. Cai, T. Jiang, and K. Jia, "Exploring diversity-based active learning for 3d object detection in autonomous driving," *arXiv preprint arXiv:2205.07708*, 2022.

[53] C. Yan, X. Xiang, C. Wang, F. Li, X. Wang, X. Xu, and L. Shen, "Pascal: Population-specific curriculum-based madrl for collision-free flocking with large-scale fixed-wing uav swarms," *Aerospace Science and Technology*, vol. 133, p. 108091, Jan. 2023.

[54] L. Anzalone, P. Barra, S. Barra, A. Castiglione, and M. Nappi, "An end-to-end curriculum learning approach for autonomous driving scenarios," *IEEE Transactions on Intelligent Transportation Systems*, vol. 23, no. 10, pp. 19 817–19 826, May 2022.

[55] J. Park, S. Jang, and Y. Shin, "Indoor path planning for an unmanned aerial vehicle via curriculum learning," in *International Conference on Control, Automation and Systems (ICCAS)*, Jeju, South Korea, Oct. 2021, pp. 529–533.

[56] K. Li, W. Ni, X. Wang, R. P. Liu, S. S. Kanhere, and S. Jha, "Energy-efficient cooperative relaying for unmanned aerial vehicles," *IEEE Transactions on Mobile Computing*, vol. 15, no. 6, pp. 1377–1386, Aug. 2016.

[57] H. Kurunathan, H. Huang, K. Li, W. Ni, and E. Hossain, "Machine learning-aided operations and communications of unmanned aerial vehicles: A contemporary survey," *IEEE Communications Surveys & Tutorials*, vol. 26, no. 1, pp. 496–533, Sep. 2024.

[58] K. Braathen de Carvalho, I. R. L. de Oliveira, and A. S. Brandão, "Av navigation in 3d urban environments with curriculum-based deep reinforcement learning," in *International Conference on Unmanned Aircraft Systems (ICUAS)*, Warsaw, Poland, Jun. 2023, pp. 1249–1255.

[59] M. Seong, O. Jo, and K. Shin, "Age of information minimization in uav-assisted data harvesting networks by multi-agent deep reinforcement curriculum learning," *Expert Systems with Applications*, vol. 255, p. 124379, Jun. 2024.

[60] C. Wang, J. Wang, C. Wei, Y. Zhu, D. Yin, and J. Li, "Vision-based deep reinforcement learning of uav-ugv collaborative landing policy using automatic curriculum," *Drones*, vol. 7, no. 11, p. 676, Nov. 2023.

[61] K. Li, W. Ni, E. Tovar, and M. Guizani, "Joint flight cruise control and data collection in uav-aided internet of things: An onboard deep reinforcement learning approach," *IEEE Internet of Things Journal*, vol. 8, no. 12, pp. 9787–9799, Aug. 2021.

[62] L. F. Peiss, E. Wohlgemuth, F. Xue, E. Meyer, L. Gressenbuch, and M. Althoff, "Graph-based autonomous driving with traffic-rule-enhanced curriculum learning," in *IEEE 26th International Conference on Intelligent Transportation Systems (ITSC)*, Bilbao, Spain, Feb. 2023, pp. 4239–4246.

[63] J. Xiao, P. Pisutsin, and M. Feroskhan, "Collaborative target search with a visual drone swarm: An adaptive curriculum embedded multistage reinforcement learning approach," *IEEE Transactions on Neural Networks and Learning Systems*, pp. 1–15, Nov. 2023.

[64] A. Brunnbauer, L. Berducci, P. Priller, D. Nickovic, and R. Grosu, "Scenario-based curriculum generation for multi-agent autonomous driving," *arXiv preprint arXiv:2403.17805*, 2024.

[65] 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.

[66] S. Narvekar, B. Peng, M. Leonetti, J. Sinapov, M. E. Taylor, and P. Stone, "Curriculum learning for reinforcement learning domains: A framework and survey," *Journal of Machine Learning Research*, vol. 21, no. 181, pp. 1–50, Jan. 2020.

[67] A. Shrivastava, A. Gupta, and R. Girshick, "Training region-based object detectors with online hard example mining," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, Las Vegas, NV, USA, Jun 2016, pp. 761–769.

[68] S. H. Karumanchi, R. B. Diddigi, K. Prabuchandran, and S. Bhatnagar, "Autonomous uav navigation in complex environments using human feedback," in *32nd IEEE International Conference on Robot and Human Interactive Communication (RO-MAN)*, Busan, Korea, Republic of, Aug. 2023, pp. 499–506.

[69] H. Zhuang, H. Chu, Y. Wang, B. Gao, and H. Chen, "Hgrl: Human-driving-data guided reinforcement learning for autonomous driving," *IEEE Transactions on Intelligent Vehicles*, pp. 1–15, May. 2024.

[70] Y. Huang, Y. Gu, K. Yuan, S. Yang, T. Liu, and H. Chen, "Human knowledge enhanced reinforcement learning for mandatory lane-change of autonomous vehicles in congested traffic," *IEEE Transactions on Intelligent Vehicles*, vol. 9, no. 2, pp. 3509–3519, Nov. 2024.

[71] X. Li, J. Fang, K. Du, and J. Xue, "Group uav navigation by qualifying human-machine decisions in hybrid reinforcement learning," in *China Automation Congress (CAC)*, Chongqing, China, Nov. 2023, pp. 4238–4243.

[72] H. Lee and S. Park, "Sensing-aware deep reinforcement learning with hci-based human-in-the-loop feedback for autonomous nonlinear drone mobility control," *IEEE Access*, vol. 12, pp. 1727–1736, Jan. 2024.

[73] Z. Huang, Z. Sheng, C. Ma, and S. Chen, "Human as ai mentor: Enhanced human-in-the-loop reinforcement learning for safe and efficient autonomous driving," *Communications in Transportation Research*, vol. 4, p. 100127, 2024.

[74] J. Wu, Y. Zhou, H. Yang, Z. Huang, and C. Lv, "Human-guided reinforcement learning with sim-to-real transfer for autonomous navigation," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 45, no. 12, pp. 14 745–14 759, Sep. 2023.

[75] A. Agrawal, J. Cleland-Huang, and J.-P. Steghöfer, "Model-driven requirements for humans-on-the-loop multi-uav missions," in *IEEE Tenth International Model-Driven Requirements Engineering (MoDRE)*, Zurich, Switzerland, Nov. 2020.

[76] J. Wu, H. Yang, L. Yang, Y. Huang, X. He, and C. Lv, "Human-guided deep reinforcement learning for optimal decision making of autonomous vehicles," *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, pp. 1–15, Apr. 2024.

[77] K. W. Mathewson and P. M. Pilarski, "A brief guide to designing and evaluating human-centered interactive machine learning," *arXiv preprint arXiv:2204.09622*, 2022.

[78] S. Krening and K. M. Feigh, "Interaction algorithm effect on human experience with reinforcement learning," *ACM Transactions on Human-Robot Interaction (THRI)*, vol. 7, no. 2, pp. 1–22, Oct. 2018.

[79] D. Abel, J. Salvatier, A. Stuhlmüller, and O. Evans, "Agent-agnostic human-in-the-loop reinforcement learning," *arXiv preprint arXiv:1701.04079*, 2017.

[80] N. A. Wondimu, C. Buche, and U. Visser, "Interactive machine learning: A state of the art review," *arXiv preprint arXiv:2207.06196*, 2022.

[81] S. Krening and K. M. Feigh, "Effect of interaction design on the human experience with interactive reinforcement learning," in *Proceedings of ACM Designing Interactive Systems Conference*, California, USA, Jun. 2019, pp. 1089–1100.

[82] C. Arzate Cruz and T. Igarashi, "A survey on interactive reinforcement learning: Design principles and open challenges," in *Proceedings of the ACM Designing Interactive Systems Conference*, Eindhoven, Netherlands, Jul. 2020, pp. 1195–1209.

[83] H. Kurunathan, K. Li, E. Tovar, A. M. Jorge, W. Ni, and A. Jamalipour, "Drl-keyagree: An intelligent combinatorial deep reinforcement learning-based vehicular platooning secret key generation," *IEEE Transactions on Intelligent Transportation Systems*, pp. 1–16, Jun. 2024.

[84] K. Li, Y. Cui, W. Li, T. Lv, X. Yuan, S. Li, W. Ni, M. Simsek, and F. Dressler, "When internet of things meets metaverse: Convergence of

physical and cyber worlds," *IEEE Internet of Things Journal*, vol. 10, no. 5, pp. 4148–4173, Dec. 2022.
[85] D. SHAPIRO, "What is active learning?" 2023, accessed: 2023-06-012. [Online]. Available: https://blogs.nvidia.com/blog/2020/01/16/what-is-active-learning/
[86] R. K. Satzoda and M. M. Trivedi, "Multipart vehicle detection using symmetry-derived analysis and active learning," *IEEE Transactions on Intelligent Transportation Systems*, vol. 17, no. 4, pp. 926–937, Apr. 2016.
[87] N. JafariNaimi, "Our bodies in the trolley's path, or why self-driving cars must not be programmed to kill," *Science, Technology, & Human Values*, vol. 43, no. 2, pp. 302–323, Jul. 2018.
[88] J. K. Uleman, *An introduction to Kant's moral philosophy*. Cambridge University Press, 2010.
[89] F. A. A. Konert and B. T. Balcerzak, "Legal and ethical aspects of rules for the operation of autonomous unmanned aircraft with artificial intelligence," in *International Conference on Unmanned Aircraft Systems (ICUAS)*, Athens, Greece, Jul. 2021, pp. 602–609.
[90] E. Vakili, A. Amirkhani, and B. Mashadi, "Dqn-based ethical decision-making for self-driving cars in unavoidable crashes: An applied ethical knob," *Expert Systems with Applications*, vol. 255, p. 124569, Jun. 2024.
[91] H. Wang, Y. Huang, A. Khajepour, D. Cao, and C. Lv, "Ethical decision-making platform in autonomous vehicles with lexicographic optimization based model predictive controller," *IEEE Transactions on Vehicular Technology*, vol. 69, no. 8, pp. 8164–8175, May. 2020.
[92] G. Contissa, F. Lagioia, and G. Sartor, "The ethical knob: ethically-customisable automated vehicles and the law," *Artificial Intelligence and Law*, vol. 25, pp. 365–378, Sep. 2017.
[93] R. Kim, M. Kleiman-Weiner, A. Abeliuk, E. Awad, S. Dsouza, J. B. Tenenbaum, and I. Rahwan, "A computational model of commonsense moral decision making," in *Proceedings of the AAAI/ACM Conference on AI, Ethics, and Society*, ser. AIES '18, New York, NY, Dec. 2018, p. 197–203.
[94] A. Martinho, N. Herber, M. Kroesen, and C. Chorus, "Ethical issues in focus by the autonomous vehicles industry," *Transport reviews*, vol. 41, no. 5, pp. 556–577, Jan. 2021.
[95] C. Luetge, "The german ethics code for automated and connected driving," *Philosophy & Technology*, vol. 30, pp. 547–558, Sep. 2017.
[96] P. Lin, "Why ethics matters for autonomous cars," *Autonomous driving: Technical, Legal and Social Aspects*, pp. 69–85, 2016.
[97] A. S. Halibas, T. A. R. M. Al Bulushi, R. C. Soriano, and A. S. M. Al Shaqsi, "Ethical design perspectives of intelligent uavs," in *1st International Conference on Unmanned Vehicle Systems-Oman (UVS)*, Muscat, Oman, Feb. 2019.
[98] X. Chen, "Ethical governance of ai: An integrated approach via human-in-the-loop machine learning," *Computer Sciences & Mathematics Forum*, vol. 8, no. 1, Aug. 2023.