IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received 16 October 2023, accepted 23 November 2023, date of publication 30 November 2023,
date of current version 11 December 2023.

Digital Object Identifier 10.1109/ACCESS.2023.3338156

SURVEY

A Survey on Cyber-Physical Security of Autonomous Vehicles Using a Context Awareness Method

AYDIN ZABOLI[1], (Graduate Student Member, IEEE), JUNHO HONG[1], (Senior Member, IEEE), JAEROCK KWON[1], (Senior Member, IEEE), AND JOHN MOORE[2], (Member, IEEE)

[1]Department of Electrical and Computer Engineering, University of Michigan–Dearborn, Dearborn, MI 48128, USA
[2]Ford Motor Company, Dearborn, MI 48128, USA

Corresponding author: Junho Hong (jhwr@umich.edu)

This work was supported by the Ford/University of Michigan Research Alliance Program.

ABSTRACT Autonomous vehicles face challenges in ensuring cyber-physical security due to their reliance on image data from cameras processed by machine learning. These algorithms, however, are vulnerable to anomalies in the imagery, leading to decreased recognition accuracy and presenting security concerns. Current machine learning models struggle to predict unexpected vehicular situations, particularly with unpredictable objects and unexpected anomalies. To combat this, scholars are focusing on active inference, a method that can adapt models based on human cognition. This paper aims to incorporate active inference into autonomous vehicle systems. Multiple studies have delved into this approach, showing its potential to address security gaps in this field. Specifically, these frameworks have proven effective in handling unforeseen vehicular anomalies.

INDEX TERMS Autonomous vehicles, cyber-physical security, active inference, context awareness, abnormal scenarios.

I. INTRODUCTION

Conventional manual driving involves steering and employing the pedals in the vehicle, which can lead to hazardous actions in complicated traffic conditions and cause many different types of driving accidents. These human failings include a lack of expertise, distractibility, and reaction time. Annually, more than 1.3 million people die in road accidents around the world, and more than 90 percent of these fatalities are caused by the behavior of humans. Moreover, many drivers have poor driving habits, which in turn contributes to the erratic driving behaviors responsible for the congestion on the roads and the decreased effectiveness of the system as a whole. On the other hand, it is anticipated that autonomous vehicles (AVs) could either entirely minimize accidents caused by human behaviors/actions or improve the effectiveness of roads. Concurrently, the concept of the ''internet of vehicles'' has increasingly received endorsement for enhancing the integration of AVs. The internet of vehicles enables automobiles to communicate a variety of datasets, including data on sensory inputs, locations, and the vehicles' perceptions of the surrounding environment [1], [2], [3], [4], [5], [6].

Making the roadways secure has always been a top priority. One of the most actively studied areas in the adaptability of both human-driven vehicles and AVs is finding the best way to reduce vehicle-related traffic accidents in our modern world. AVs are required to formulate decisions in scenarios where various determinants (e.g., weather conditions) are in a continuous state of transformation [7]. A vehicle must be able to reliably anticipate the behavior of other vehicles or objects in order to reduce the risk of collisions. In the conventional traffic scenario, all vehicles are operated by humans, who (due to their training and experience)

The associate editor coordinating the review of this manuscript and approving it for publication was Kashif Saleem.

can make accurate predictions about the next actions of the other vehicles/objects around them. All drivers make instantaneous adjustments to their behavior based on this assessment, improving traffic flow and safety. In a mixed-traffic scenario, however, AVs must predict the behavior of human drivers based on their observations of the road. Using vehicle-to-vehicle (V2V) and vehicle-to-infrastructure (V2I) communications, a vehicle can learn about the current driving conditions of other objects on the road. Also, vehicle-to-everything (V2X) communication plays a crucial role in improving the decision-making processes of AVs by allowing them to exchange information with surrounding entities, leading to safer and more efficient driving [8]. The importance of V2X communication, specifically V2V communication, in enhancing the decision-making processes of AVs highlighted in [9]. The authors proposed a machine learning (ML)-based steering control approach that leverages V2X communication to improve autonomous driving in urban environments. Using V2X data, the proposed model can anticipate potential hazards and adjust the vehicle's trajectory accordingly, ensuring safer navigation through complex urban scenarios. Substantial investigations have gone into the communication methods used in vehicular networks that allow for secure connections between autonomous and human-operated vehicles. By communicating with other vehicles and gathering information about their current driving situations, an AV may anticipate the behavior of other nearby objects [10], [11], [12], [13].

The evolution of AVs promises a transformative impact on transportation. It also presents multifaceted challenges, particularly in the domain of cyber-physical security. Central to the operation of these vehicles is their ability to accurately recognize traffic signs (TSs) and classify objects in their environment'a function primarily achieved through advanced ML algorithms. However, the integrity of these algorithms is paramount. Misinterpretations or malicious interventions can lead to severe traffic mishaps. While traffic sign recognition (TSR) is crucial for ensuring that the vehicle adheres to traffic rules, object classification (OC) plays a vital role in distinguishing between a pedestrian, another vehicle, or any other potential obstacle. However, as with any sophisticated technology, vulnerabilities exist. Threat actors can exploit these vulnerabilities to deceive autonomous driving systems, possibly leading to catastrophic consequences. Hence, a comprehensive understanding of cyber-physical security, especially focused on TSR and OC algorithms, is essential to foster the safe integration of AVs into our transportation matrix [14], [15], [16], [17], [18].

In light of these vulnerabilities, the field of autonomous driving research has heavily emphasized refining these algorithms. Deceptive adversarial attacks, where small perturbations can mislead deep learning models, are of particular concern in the context of TSR and OC [19]. For instance, an attacker might subtly modify a TS's appearance, causing a TSR system to misinterpret it. Similarly, alterations to the surrounding environment might result in the OC algorithm misidentifying an object. Such misleading classifications underscore the importance of constantly improving the robustness of these systems. In addition to the algorithmic challenges, these issues clearly showed the need for collecting diverse and comprehensive datasets. These datasets should encapsulate various real-world scenarios that an AV might encounter, ensuring the training process accounts for a broad spectrum of challenges, from adverse weather conditions to vandalized TSs [20], [21]. Given the dynamic and unpredictable nature of road scenarios, the integrity and expansiveness of training data become paramount in shaping the robustness of these algorithms, which can be explained in the following. Feature-extraction-based ML techniques (particularly supervised learning) depend substantially on training data. It is highly improbable that all feasible scenarios will be present in the training data. Therefore, there must be mechanisms in place in an AV's control system based on an algorithm to mitigate object misclassification and block out noise (e.g., bright and direct sunlight to vision sensors). Any misrecognition of objects by the ML algorithm used for image recognition in AVs is a potential safety risk. A failure in recognition or classification has occurred if, say, the STOP sign is misinterpreted or a pedestrian is mistaken for a tiny animal. Weather conditions, as well as direct sunlight on the camera sensor or obstructing a portion of the object, can all lead to these issues [22]. Understanding TSs located on the roadway, as well as all other traffic indicators that provide navigation instructions (e.g., traffic lights, lane markings, or markers) is known as TSR. This entails detecting numerous signs from camera sensors depending on their geometry, texts, logos, and colors. A front camera can take TSs, and TSR enables AVs to fully comprehend road rules and laws for the security of either vehicles or pedestrians. TSR simply needs ML technology and a camera, whereas the Advanced Driving Assistant System (ADAS) contains an additional sensor (e.g., ultrasonic or radar device). For acceptable performance and real-time modeling of this recognition system, it is indispensable to install advanced hardware in the vehicle system [23]. Moreover, Unsupervised learning algorithms for TSR can have some challenges including data labeling, a lack of ground truth, interpretability, and scalability. These algorithms require large amounts of unlabeled data. They do not have access to explicit ground truth labels, making it challenging to evaluate their accuracy and reliability, and require significant computational resources and time for training, posing challenges in real-time applications. AVs are designed to perceive their surroundings analogous to the human visual system's processing capabilities, thus facilitating cognitive architectures capable of learning, recalling, and executing actions. This will enable object detection, cognition, and scene recognition. AVs must be able to recognize their surroundings, analyze 3D world representations, and distinguish the motion of objects, people, and other vehicles. Furthermore, they need to cope with human emotional responses in order to be similar to human driving. These vehicles attempt to process data/information

captured by cameras and apply deep learning algorithms for route planning, collision detection, sophisticated decision-making, and problem-solving.

The concept of an individual's driving proficiency remains multifaceted and hinges upon the formulation of indicators representing driving behavior and their subsequent analytical methodologies. Such proficiency is shaped by factors including, but not limited to, the driver's style, skill set, and a combination of their physical and cognitive attributes. Due to the inherently stochastic and varied nature of a driver's attributes, coupled with the intricate dynamics of traffic scenarios, the characterization of driving competence takes on a random and fluid dimension. As a consequence, arriving at a holistic assessment of driving capability through a single metric presents significant challenges [24], [25], [26]. Although these methods have demonstrated a large amount of potential for modeling human activities, they have several significant drawbacks when it comes to representing cognitive functions and applying them in situations where there are many possibilities. These limitations can be applied particularly to AVs because drivers' responses to failures are significantly affected by their own internal states. The most popular method for applying models of human perceptions and behaviors to the domain of autonomous driving is active inference (ActInf) from the disciplines of cognitive science and neuroscience, which has been used by scholars to address these gaps. Considering generative models of the environment, ActInf can be highlighted as self-evidencing, in which case both perception and action are interpreted as maximization of evidence for a Bayesian model. More information related to this concept will be provided in the upcoming section based on the principle, comparison with other algorithms, and related works [27], [28], [29], [30]. Because of the vulnerability of these algorithms, this paper proposes a statistical method called the ActInf model based on partial observation, in which human behaviors and different conditions in the vehicle's environment can be considered. This method is based on brain neurons and considers partial observations in comparison to full observations in the environment. The following sections present the TSR and OC algorithms, their challenges and applicability in different conditions, the applicability of the ActInf model in autonomous driving, and proposed driving scenarios with ML algorithms' weaknesses. The main contributions of this review paper can be summarized as follows:

- The literature critically examines current security concerns within AVs, specifically addressing the incomplete environmental observations and diverse behavioral patterns of road objects.
- A pivotal analysis of TSR and OC algorithms is offered, elucidating their efficacy and precision within entirely observable frameworks.
- An in-depth assessment of ActInf models, anchored in a partially observable Markov decision process (POMDP), is delineated. This examination contrasts various environmental observation methods and Context Awareness (CAW) within AV-centric applications, further emphasizing the superior attributes of this model relative to conventional ML algorithms.
- Some abnormal scenarios proposed in the area of autonomous navigation where prevailing ML algorithms falter, yielding predictions with diminished accuracy regarding road-bound entities.

The rest of this paper is organized as follows: Section II provides information regarding the ActInf background and architecture of AVs considering the CAW. Section III outlines POMDP principles and applicability and makes a comparison with other decision-making processes. Furthermore, it discusses collision avoidance in autonomous driving based on the ActInf model. Section IV represents the applicability of the ActInf model with a connection between the ML algorithms and this model to provide a clear understanding. Reviews of the literature on AV applications of the ActInf model are presented in this section. Abnormal driving scenarios with ML algorithms' disabilities are proposed and discussed in Section V. A discussion on different parts of the scenario considering the risk assessment analysis and the authors' perspectives is mentioned in Section VI. Finally, this paper is concluded in Section VII.

II. ACTIVE INFERENCE BACKGROUND

The ActInf model presents a unique mathematical model for representing cognition in all its perceptual, cognitive, and inferential (i.e., decision-making) forms. The framework views these psychological/cognitive procedures and their interconnections as complementary modes of inference [31], [32], [33], [34]. Within this framework, agents are posited to possess a generative model of their environment, and perception and learning occur through variational inference processes applied to this model. These processes aim to minimize variational free energy (VFE), an information-theoretic measure. Furthermore, the ActInf method suggests that action selection can be conceptualized as an inference process, facilitated by the same mechanisms involved in perceptual inference and learning. The implementation of this framework demonstrates a degree of biological plausibility and finds support in substantial empirical evidence. Decisions are based on ''active'' inferences, where the agent predicts which behaviors will result in the most desirable sensory information. Agents are able to infer the actions to take that will most effectively minimize ambiguities or uncertainties and promote knowledge acquisition. Because of this, it is preferable to choose actions that achieve a balance between maximizing reward and maximizing information acquisition (i.e., increasing situational awareness). ActInf operates under the assumption that the generative model inherently possesses a built-in bias towards states or observations that hold intrinsic value [28], [35]. In recent years, the fields of

psychology, neurophysiology, ML, and even transportation science have all benefited from this concept.

Recent research has showcased the applicability of ActInf in complex tasks and environments with high dimensionality. On the other hand, reinforcement learning (RL), which focuses on adaptive action selection, traditionally seeks to maximize the expected cumulative rewards. However, the framework of control as inference has recently reconceptualized RL within the framework of variational inference. This reformulation generalizes and contextualizes prior work on stochastic optimal control, highlighting the inherent relationship between control and inference. Instead of maximizing rewards, agents are now required to infer actions that lead to optimal trajectories. This paradigm shift enables the utilization of powerful inference algorithms in RL while naturally promoting exploratory behavior [36], [37], [38].

AVs rely on their situational awareness, or CAW, to have good performance in their internal algorithms. The terms ''active'' and ''inference'' serve as the basis for the ActInf concept. The former part of this concept expresses the idea that organisms actively interact with their surroundings to learn about them, find ''preferred'' observations, and prevent ''non-preferred'' perceptions. The latter term refers to Bayesian inference, a statistical method that explains how best to adjust one's prior beliefs in light of newly acquired evidence (i.e., sensory perception).

As mentioned, an increase in the usage of the ActInf concept as a viable method for representing neurocognitive processes has been observed in recent years, especially in the context of its latest developments as a POMDP model. The adaptability of this approach makes it ideal for modeling a wide variety of cognitive tasks, and it can also be used to imitate the neuronal responses to those tasks in light of the theory underlying those responses. Due to the assumption that states can be determined from partial observations only probabilistically, POMDP features are a state-based extension of MDP frameworks for uncertain environments. In this method, states are used to indicate patterns or clusters of occurrences, and system behaviors are represented as mappings (e.g., policies) from possible beliefs to accessible actions. These frameworks are generally constructed for specific tasks within a particular context, and the policies that arise are maximized with consideration of the expected outcome. In the real world, though, the agent will certainly be confronted with a variety of tasks in environments with remarkably similar but never identical characteristics [39].

An ActInf approach has other internal processing parameters in addition to the POMDP method that should be considered in the model. These concepts can be understood based on perception, learning, action selection, planning, and decision-making (VFE is one of the parameters). Its goal is to change the intractable summation needed for model inversion into a mathematically tractable optimization model that can be considered a posterior distribution. Another parameter is the ''expected free energy'' (EFE) which determines the difference between the probability of desired states and the probability of predicted states [40]. In the ActInf framework, VFE serves as a core quantity. It represents a bound on the surprise or improbability of sensory data, given a particular model or belief about the world, and EFE is central to ActInf as well, guiding decision-making by considering future outcomes and the uncertainty that goes along with them. Nevertheless, POMDPs do not inherently employ VFE or EFE but instead utilize belief states and reward structures to guide decisions [41].

It is preferable to use accurate inferences when calculating posterior distributions for hidden states (variables). Although effective for a discrete model, accurate inference procedures are intractable for everything except straightforward models. It is necessary to use methods of inference that are approximate in the majority of practical scenarios. One such approximate inference technique is called ''variational message passing''. With the aid of this method, the calculus of variations can be put to use. Many problems can be seen as optimization problems that can be solved by trying out different input functions until one is found that maximizes or minimizes the goal. Approximate solutions can be found using variational techniques, despite the fact that these methods are not inherently approximate.

To recapitulate, the goal of perception and learning is to minimize VFE by determining appropriate posterior beliefs in response to each new observation. Obtaining the best estimate on each attempt/trial is not the only way to minimize VFE. There is noise in the sensory information as an input, so identifying the best posterior for each trial can be achieved with overfitting. At the same time, this issue will not happen in the VFE minimization procedure. The purposes of the VFE minimization process can be divided into two categories, namely ''complexity'' and ''accuracy.'' The term ''complexity'' describes the degree to which a model's beliefs must vary in response to new sensory data while retaining a high degree of precision. The term ''accuracy'' reflects how effectively those beliefs can predict that sensory information [40]. Choosing actions that will lead to future observations that minimize VFE is the goal of planning and action selection. The issue is, however, that it is hard to identify what will happen because the future is unknown. As a result, decisions need to be made to get the lowest possible EFE. Importantly, EFE assigns a score based on the expected cost (a lower score denotes a higher reward) minus the expected acquisition of new information from a certain action. Thus, decisions attempt to remove uncertainty and maximize rewards for EFE minimization. When beliefs regarding states are ambiguous or uncertain, actions are more probable to be taken toward finding information. When confidence in beliefs about states is strong, on the other hand, determined actions seek the reward. This means that the agent knows exactly what to do to achieve the goals and is free from any remaining ambiguity/uncertainty [40].

As aforementioned, there are uncertainties and unknown situations on the roads that need to be managed by AVs. A POMDP can provide a better representation of the

uncertainties than other methods, considering the possible observations. The upcoming sections present a mathematical framework for a POMDP to model partial observations based on uncertainties in the surrounding environment and the advantages of this approach, along with an ActInf-based POMDP model for collision avoidance in AVs. Hence, it is necessary to mention the AV system architecture, including motion planning, decision-making, and other components, in the following section to get familiar with different AVs' concepts regarding their environment. However, the most important part of the ActInf model is CAW which distinguishes it from other concepts, which can be described in the following section.

A. SYSTEM ARCHITECTURE OF AUTONOMOUS VEHICLES

System architectures of AVs can vary based on a wide range of circumstances, such as the presence of other vehicles on the road and the number of pedestrians. Thus, a generic system architecture of AVs can be depicted in Fig. 1 to help with comprehension of the framework within a fully autonomous system. It mainly consists of route planning, perception,

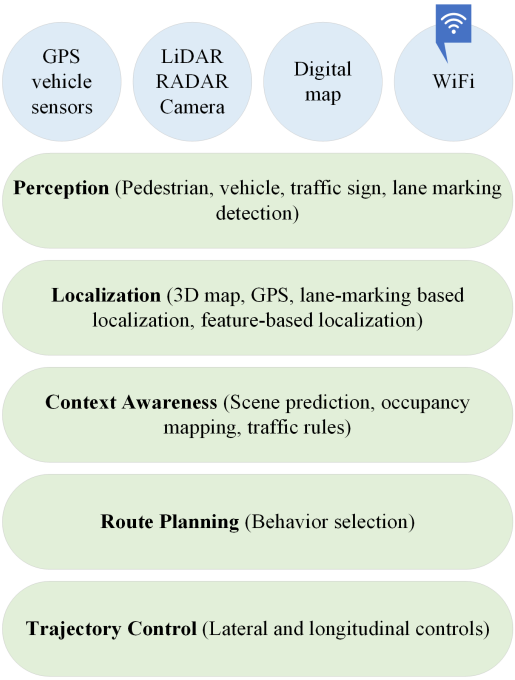


FIGURE 1. A system architecture of AVs.

localization, CAW, and trajectory control. As mentioned, because of the importance of CAW, this section explains this term. Other parts of the architecture are summarized in the figure and can be found in previous research [42], [43].

1) CONTEXT AWARENESS

Real-world road traffic presents a wide range of scenarios, necessitating robust situational awareness and analytic methods. The overall context, from the road conditions to the driving situations inside or outside the vehicle, is part of this situational awareness [44], [45]. In order to enhance the situational awareness of ADAS, the acquired context information is transmitted back to the vehicle. An AV's context is the physical location in which it is used to perform its specified tasks. There are many different factors that contribute to context. These items are referred to as elements of context (EOC). In its operational context, an AV operates as an active agent. Based on observations, it can be referred to as the primary EOC, while the others are classified as the secondary EOC. An EOC can be further classified as active, moderately active, or inactive according to its level of interaction with the surrounding environment. Each EOC gives what is called contextual information (CXI), which describes information in the context of its current operation. These data represent a wide range of prospective insights about the characteristics, states, and attitudes of EOCs [5], [45], [46], [47], [48]. Therefore, instead of assuming the worst scenario, as was the case with previous safety measures, CAW-assisted safety measures consider the alternative possible significant cases of the context to guarantee the system's secure behavior. This avoids the need for AVs to undertake excessively cautious maneuvers in situations where only moderate actions are essential, eliminating unnecessary constraints on the AV's capabilities and boosting its performance and efficiency. The context can be logically categorized into relevant and irrelevant classes. This removes the part of the context that does not affect the AV's functionality. There is still a vast context space with many possible EOCs that are significant. Various characteristics, including environment, traffic actors (other vehicles and people), climate, and time of day, are used to classify EOCs in this domain in order to minimize the complexity that goes along with them. A conceptual classification of context focused on safety-relevant components is shown in Fig. 2. The CXI without consideration of safety is ignored since not all information concerning EOCs is crucial to the AV's security [46], [49], [50].

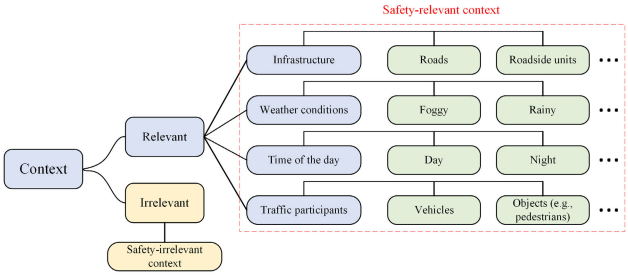


FIGURE 2. Classification of context considering the safety-relevant components.

III. PARTIALLY OBSERVABLE MARKOV DECISION PROCESS

According to the Markov assumption, which asserts that the future state of the system based on its current condition is independent of any past occurrences, Markov models (MM)

present the development of randomly variable architectures. MMs can be divided into two classes, namely, autonomous systems (Markov chains and hidden Markov models, [HMMs]) and controlled systems (MDPs and POMDPs). The states of a system modeled by a Markov chain are a series of random variables in which each state is independent of the others. Both observable and hidden states exist in an HMM. However, neither the hidden nor the observed states map onto the observable symbols in a one-to-one manner. As a result, the generated observation symbol can no longer be used as a reliable indicator of the model's hidden state. The MDP is a dynamic system that can move between any two states. At each step, the decision-maker must choose between specific options. This decision results in a gain immediately and another gain subsequently since it affects the next move's transition probabilities. The challenge for the decision-maker is to find a sequence of actions that will result in the greatest possible gain. Ultimately, a POMDP is an MDP in which the agent has no direct access to the hidden state variables. Since the agent cannot observe through the underlying states, the concept of a "belief state" is applicable. The belief state presents an option for dealing with the model's intrinsic uncertainty. To recap, Markov chains primarily emphasize the transitions between states. On the other hand, MDPs incorporate elements of actions and their corresponding rewards. POMDPs introduce an additional dimension of complexity by accounting for uncertainties in observations. Meanwhile, HMMs are centered around extracting concealed states from data that can be observed. Table 1 summarizes the various models' distinctions with further information regarding different models [51].

In autonomous driving, the ability to make consistent and robust decisions is vital despite the presence of uncertainties. To be more specific, an AV has to determine how to behave tactically in order to complete its tasks, even though it can be affected by a variety of failures and perturbations that have an effect on its sensors and perception system as well as by the fact that it does not have enough information or perception regarding itself or its surroundings. The errors and the insufficient information lead to the effects of taking actions to be non-deterministic from the viewpoint of the AV. Additionally, the errors and the lack of information allow the AV's state to be only partially observable, which shows that the AV never knows its correct state [52], [53].

Numerous societal and biological challenges can be viewed as decisions to be made in sequence while facing some levels of uncertainty. Because the observable action outcomes reveal only a portion of the environment's state, they can be viewed as extensions of the MDP that refer to the POMDP framework. This mathematical approach can be used to express an extensive variety of sequential decision-making issues. For example, agents take actions in an MDP that change the state of the system in order to maximize the rewards it receives by going through various states. This is considerably more challenging in a POMDP since the agent can gain information about the state from noisy observations.

A tuple $(S, A, T, r, O, Z, \gamma)$ can be specified as a POMDP, where S and A represent all possible states and actions, respectively. A probability of moving to state s' can be recorded after an action a is performed in state s with the help of the transition model $T(s'|s, a)$. Whenever an action a is taken, the reward for transferring from state s to state s' is defined by the reward function $r(s, a, s')$. $Z(o|s)$ is the probability density of getting observation o in state s, where O and Z are the observation space and the model, respectively. Furthermore, the rate (at which future rewards are discounted) is controlled by $\gamma \in [0, 1]$.

A. MATHEMATICAL MODELING OF MDP AND POMDP

The following tuple provides the definition of an MDP [54]:

$$(S, A, T, r, \gamma),$$

where

- S: Finite sets of states
- A: Finite sets of actions
- T: Probability function for state transitions with a state $s \in S$ and action $a \in A$ in which $T(s, a, s') = \Pr(s'|s, a)$
- r: Reward function, $r(s, a)$, for taking an action a while in state s
- γ: Discount factor, [0, 1]

The state transition probability function determines the agent's future state (s') after taking action in the initial state (s). The agent also gets a reward in accordance with the state and the action taken in that state. The Markov property can be followed by the environment in order to employ this decision-making method. This model can be written as follows:

$$\Pr(s_{t+1}|s_t, a_t, s_{t-1}, a_{t-1}, \ldots) = \Pr(s_{t+1}|s_t, a_t) \quad (1)$$

According to this approach, future states are determined solely by the current state. A policy (π) provides the mapping between states and actions. It is responsible for the actual generation of actions. This mapping can be specified by a stochastic or deterministic function, in which the stochastic policy makes multiple actions with a certain probability distribution for the agent to choose from, and the latter makes a single action for a given state. An MDP aims to find a strategy that yields the greatest predicted return over time. A distinction is considered between problems with a finite horizon, in which the agent will stop after time steps of K. Also, $r_K(s_K)$ depicts the final reward at K. In problems with an infinite horizon, the agent continues taking action and getting rewarded endlessly. For finite and infinite time horizons, Eqs. (2) and (3), respectively, can be written for the expected discounted summation of rewards:

$$R_{\pi,finite} = E_\pi \cdot \left(\sum_{t=0}^{K-1} \gamma^t r_t(s_t, a_t) + r_K(s_K) \right) \quad (2)$$

$$R_{\pi,infinite} = E_\pi \cdot \left(\sum_{t=0}^{\infty} \gamma^t r(s_t, a_t) \right) \quad (3)$$

TABLE 1. A comparison of different Markov models based on their features.

Feature	Markov Chain	HMM	MDP	POMDP
Definition	A stochastic model describing a sequence of possible events where the probability of each event depends only on the state attained in the previous event.	A statistical Markov model in which the system being modeled is assumed to be a Markov process with unobserved (hidden) states.	An extension of Markov Chains that includes decisions in the model.	An extension of MDPs where the agent can't always observe the true state but receives observations correlated with the state.
States	Observable	Hidden (not directly observable)	Observable	Not always observable (partially observable)
Decisions/Actions	No	No	Yes	Yes
Observations	Not Applicable	Yes (related to states)	Not Applicable	Yes (related to states)
Rewards	No	No	Yes	Yes
Applicability	Probability theory, Queueing theory	Speech recognition, Finance	Operations research, AI	Robotics, AI where there is uncertainty in observation

The optimal policy (π^*) can be expressed as Eq. (4), showing the greatest possible discounted reward over time. A value function at state $V(s)$, as defined by the policy, is depicted in Eq. (5). By computing the expected cumulative reward, this function illustrates how favorable the given state is while the policy is being followed. The optimal policy can be attained by considering the optimal value function. The first step in policy analysis is determining the policy's value function. The next step is to apply Eq. (6) to all possible states and to the generated value function to enhance the policy further. These two processes are repeated frequently through policy iterations unless no further improvements can be achieved.

$$\pi^* = \arg\max_{\pi} R_{\pi} \tag{4}$$

$$V^{\pi}(s) = r(s,a) + \gamma \sum_{s' \in S} T(s,a,s') V^{\pi}(s') \tag{5}$$

$$\pi_{k+1}(s) = \arg\max_{a} \left(r(s,a) + \gamma \sum_{s' \in S} T(s,a,s') V^{\pi}(s') \right) \tag{6}$$

However, the environment may have too many chaotic noisy characteristics (e.g., GPS signal) to detect an agent's location. Hence, the agent may never be confident of the accurate state of the environment at time t, (s_t). Uncertainties can be modeled by observations, observation probabilities, and beliefs, and a POMDP can be used to address these parameters. An extensive tuple as, $(S, A, T, r, O, Z, \gamma)$, can be determined for a POMDP. The collection of all possible observations that an agent can take in is denoted by the O as an observation space. The probability of a given observation Z determines the occurrence of the uncertainty of observing a specific observation o when taking action. In addition to all information above regarding the MDP, the agent receives an observation simultaneously in which $o \in O$. Actions a and observations o at time t can be considered as a sequence of events represented by a set of action-observation combinations as follows [54]:

$$\begin{aligned} h_t &= \text{set of (action, observation)} \\ &= \{(a_0, o_1), \ldots, (a_{t-3}, o_{t-2}), (a_{t-2}, o_{t-1}), (a_{t-1}, o_t)\} \end{aligned}$$

The agent must keep track of a probability distribution over all possible states at any given time because it has no idea what state it is currently in. The belief state (b) denotes this particular probability distribution in which if the agent is in the state, the probability is $b(s)$. Furthermore, at time step $t = 0$, the initial belief is represented by b_0. The belief space B is the space represented by all possible belief states. One can depend on either historical information or beliefs while making a decision in light of uncertainties. Different strategies for belief updating are applicable for either discrete or continuous state spaces. Here, an example of a discrete state filter can be stated using Bayes' rule as Eq. (7):

$$\begin{aligned} b_{t+1}(s') &= b_{t+1}(s'|b,a,o) = \Pr(s'|b,a,o) \\ &= \frac{\Pr(o|s',a,b) \cdot \Pr(s'|b,a)}{\Pr(o|b,a)} \\ &= \frac{O(s',a,o) \sum_{s \in S} T(s,a,s') b_t(s)}{\Pr(o|b,a)} \end{aligned} \tag{7}$$

The optimal value function of a POMDP for a given belief state b can be written as Eq. (8):

$$\begin{aligned} V^*(b) = \max_{a \in A} \Bigg(& \sum_{s \in S} r(s,a) b(s) \\ & + \gamma \sum_{o \in O} \Pr(o|b,a) V^*(b_{t+1}(s'|b,a,o)) \Bigg) \end{aligned} \tag{8}$$

The optimum policy, denoted by π^*, is obtained by solving the POMDP so that it takes the best possible actions

considering the current belief. For POMDPs, solutions can be classified into two groups, namely, accurate and approximate solutions. In this section, the value iteration method as an accurate solution will be discussed. This approach aims to find an optimal policy in a POMDP based on a belief state so that the value function maximizes the value for all belief states. The optimal policy can be calculated by maximizing Eq. (9) for all belief states.

$$\pi^*(b) = \arg\max_a \left(\sum_{s \in S} r(s,a) b(s) + \gamma \sum_{o \in O} \Pr(o|b,a) V^*(b_{t+1}(s'|b,a,o)) \right) \quad (9)$$

Once the largest error is under a user-defined threshold (ψ) based on Eq. (8), the iterative process of calculating the value function for time step $t+1$ stops (i.e., $|V_t^*(b) - V_{t+1}^*(b)| < \psi$) [54].

B. WHEN ARE POMDP MODELS APPLICABLE?

Whenever a decision maker lacks all relevant information about the current status of the system, POMDPs provide an influential framework for addressing optimization problems based on sequential decision-making processes. Recognizing the circumstances that require POMDPs is a complicated step. Some criteria for what makes a problem favorable for modeling as a POMDP can be regarded by considering the following aspects [51]:

- Decision makers have limited awareness (observation) of the system's real conditions. In contrast to the perfectly observable scenario, making the optimal planning decision is more challenging when only partial information about the system state is available. An MDP becomes a POMDP when the state is ambiguous and there are uncertainties.
- Irregularity is introduced into the dynamics of the system by means of inherent variability and/or control unpredictability. It is assumed that the transition probability to a state at time $t+1$ is independent of anything other than the current status of the system (state) and the action taken at time t. The success of the results of management operations is expected to be uncertain because of the probabilistic nature of the population structures. The mechanism of the transition between states conforms to the Markov property, provided that a suitable time step is used and the relevant information is included within the description of each state.
- It is possible to classify the system into different states, and no single decision or action is appropriate for every possible state. For endangered species, a natural description of system states might be provided by the expected upper and lower population density estimations under various environmental situations and hazards.

C. ACTINF-BASED POMDP MODEL FOR COLLISION AVOIDANCE IN AVS

A decision-making process in partially observable scenarios can be defined using an ActInf-based POMDP model. Given that the brain must interpret the actual state of the world from noisy or chaotic sensory information, incomplete observability is to be expected. The key concepts of the ActInf model, including states, actions, preferences, free energy concepts, variational inference, and a POMDP procedure, can be presented as a representation of human decision-making. The POMDP model $(S, A, T, r, O, Z, \gamma)$ can be described in terms of the model for collision avoidance as follows [54]:

- **State** (S): A tuple including the position, velocity, and angle of the vehicle can be represented by (x_v, y_v), V_v, and θ_v, respectively. Based on the research, there are different types of objects that need to be considered. Hence, "o" can denote different objects (e.g., pedestrians, vehicles, animals, traffic obstacles). In this case, (x_o, y_o), V_o, and g_o can show the position, velocity, and current goal of objects, respectively. Furthermore, updating of beliefs for every time step can be denoted by t. A specific location of an object within the environment is known as a goal. Because the purpose of the objects is to achieve the goals, the agent in the POMDP model needs to be able to anticipate the objects' movements. On the other hand, the goal is partially observable. Hence, it is part of the belief (there is no goal for stationary objects, e.g., traffic barriers and signs). In this way, a state space can be mapped onto a set as follows:

$$S = \{s|(x_v, y_v, V_v, \theta_v, x_o, y_o, V_o, g_o, t)\}$$
$$S = \{s|(\text{vehicle position, vehicle velocity, vehicle angle/orientation, object position, object velocity, object goal, time step})\}$$

- **Action,** (A): The action will affect the vehicle's velocity by a fixed value. This value will serve as an acceleration that can be negative, zero, or positive. Thus, three actions in the action space can be expressed as the following set:

$$A = \{\text{accelerate (positive value), stop (zero), decelerate (negative value)}\}$$

- **Transition probability function,** (T): Based on their predetermined application, transitions of states can be represented by simple and complex motion models where $T_{ijk} = P(s_k|s_i, a_j)$ represents the probability of transitioning from state s_i to state s_k after taking the action a_j. Eqs. (10) and (11) determine the next position (x') and updated velocity (v'), respectively, based on a simple model, where Δt is the sample time for this transition and acc represents the acceleration of the vehicle.

$$x' = x - v\Delta t - \frac{(acc)\Delta t^2}{2} \quad (10)$$
$$v' = v + (acc)\Delta t \quad (11)$$

- **Reward function**, $r(s, a)$: The current environmental circumstances will determine whether the reward model returns a positive or negative reward. Generally, the following rewards can be mentioned in autonomous driving:
  - If objects (e.g., pedestrians, vehicles, or animals) get near the target vehicle, it will receive a negative reward. In this case, a distance based on the radius of a circle with the target vehicle in the center can be considered.
  - Accelerating and decelerating too quickly can lead to a negative reward that prohibits erratic driving.
  - Based on the current velocity and maximum velocity, a negative reward of $c_p = (v_{current} - v_{max})/v_{max}$ is given to the vehicle as a penalty to encourage smooth/cautious driving.
  - Finally, when the vehicle gets closer to its goal, it will be positively rewarded. If the vehicle comes within a specific range of the goal, the reward will be provided.
- **Observation,** (O): Contains the position, orientation, and velocity of the target/ego vehicle as well as the positions of objects as $O = \{o|(x_v, y_v, V_v, \theta_v, x_o, y_o)\}$.
- **Observation probability function,** $Z(s', a, o)$: After an action is taken, this function will return the possibility to observe once, resulting in the next state (s'). It will assess noisy sensor information and the uncertain motions of objects.
- **Discount factor,** (γ): A low γ indicates that the agent prioritizes immediate rewards over those that will be received in the future. If there is a large γ, it will be favored for the future reward.
- **Variational Inference:** It is a method used to approximate complex probability distributions with more tractable ones. In the context of autonomous driving, it can help in approximating the true state of the environment (e.g., the positions and intentions of other vehicles) based on sensor observations. The encoder in a variational autoencoder uses this method to make inferences from the input data and generates an approximate posterior. This can be particularly useful in predicting the possible actions of nearby vehicles or pedestrians, aiding in decision-making.
- **Preferences:** They can be formally introduced as a part of the generative model in active inference. If s represents states and o represents observations, then preferences can be encoded in the likelihood $P(o|s)$. For collision avoidance, the likelihood of observing a collision given a state that leads to a collision would be low, indicating a non-preferred state. Conversely, states that lead to safe trajectories would be associated with high likelihoods, reflecting preferred outcomes. They are not static and can dynamically evolve based on the vehicle's experiences, changing environmental conditions, or updated objectives. For instance, in congested traffic scenarios, the preference might shift from reaching the destination quickly to ensuring safe maneuvering among closely packed vehicles [30], [55].
- **Free Energy Principle:** The principle posits that any adaptive system (similar to an AV) minimizes a certain quantity known as VFE to maintain stability and adapt to its environment. The vehicle can reduce the discrepancy between its predictions (based on the approximated posterior) and the real-world observations by continually minimizing the free energy. This allows for better decision-making and, safer driving practices. [56].
- **Free Energy Minimization,** (F): To minimize EFE, Eq. (12) can be written:

$$F = \sum_{s} P(s|o) log \frac{P(s|o)}{Q(s)} \tag{12}$$

  where $P(s|o)$ is the true posterior and $Q(s)$ is the approximated posterior distribution over states given observations, which can be obtained through variational inference. The goal is to achieve better and safer predictions, especially in scenarios such as collision avoidance.
- **Policy,** (π): The optimal policy, π^*, is derived by minimizing the EFE over a series of actions. This can be represented as Eq. (13):

$$\pi^* = argmin_{\pi} F \tag{13}$$

Incorporating the principles of ActInf into the POMDP framework can enhance decision-making in complex environments. This integration uses a generative model to infer hidden environmental states based on observations. Preferences, representing desired outcomes, are embedded within the POMDP's reward function. Variational inference is applied to approximate posterior beliefs about hidden states, especially when direct inference is computationally challenging. Actions are selected to minimize EFE, which represents a balance between maximizing expected rewards (preferences) and reducing uncertainty about the environment. This enriched approach to decision-making encourages agents to act adaptively and robustly in uncertain scenarios, which is vital for autonomous driving applications.

IV. ACTIVE INFERENCE, CONTEXT AWARENESS AND OTHER FULL OBSERVABLE MODELS IN AUTONOMOUS VEHICLES

This section delineates the transition from prior sections, underscoring the relevance of the ActInf model in scenarios where ML techniques (e.g., TSR and OC) falter in their precision in predicting TSs or objects in the AV's area. Fig. 3 offers an expository blueprint of the proffered TSR and OC algorithmic structure, particularly when they struggle with accurate predictions. The ActInf model augments prediction precision by leveraging context-awareness of anomalous situations. Contemplating the myriad TSs and roadway objects susceptible to deliberate or accidental disruptions (e.g., digital intrusions, signage defacement, extreme weather conditions, lighting variations), these ML models, although aspiring for optimal prediction, often yield erroneous or

imprecise outputs. As elucidated in **Task 1**, enhancing prediction accuracy is feasible by training diverse TS and object classifications, factoring in image reconfiguration and segmentation techniques. Nonetheless, when these methods confront challenges such as cyberattacks or total object obstructions, their efficacy diminishes. Consequently, the ActInf model, drawing from contextual data and auxiliary inputs (e.g., V2X metrics such as proximate vehicle speed, geolocation, and navigational insights), as delineated in **Task 2**, bolsters prediction precision. Our model incorporates these inputs and refines predictions by iteratively updating beliefs through the ActInf's internal computational mechanism.

In the area of autonomous driving applications, ideal vehicular routing strategies can be formulated through standard optimization methodologies, especially when uncertainties within the transportation system are addressed. However, the potential shortcomings of these methodologies, particularly in efficiently navigating system dynamics for instantaneous decision-making, may limit their practical utility. Given the extended duration these strategies require to derive solutions, their adaptability to swift alterations in system dynamics is compromised. Consequently, the challenges of instantaneous computational requirements and the navigation of environmental ambiguities emerge as significant impediments to decision-making and routing optimization. Moreover, while a majority of ML techniques offer complete observability, AVs might only partially observe future scenarios due to unpredictable environmental variables [57], [58], [59]. This section also references various scholarly studies, highlighting existing lacunae related to ML strategies and the advantages conferred by ActInf models.

In terms of the enhancement of the prediction accuracy by the ActInf method, consider TSR process using conventional ML models. If there is an abnormality (e.g., putting sticker, painting, occluded by snow in winter) on a TS, the ML model cannot recognize this sign or recognize it with a low prediction accuracy. Therefore, the Av can make a decision related to the recognized TS with a low probability which is a wrong action. However, the ActInf can enrich the prediction accuracy in TSR process, specifically in abnormal scenarios to make a correct prediction with an accurate action. This can happen by gaining additional information (e.g., GPS/navigation data, speed of nearby vehicles) from environment which can be considered as partial observations. Liu et al. [60] developed a novel RL approach to managing the route of vehicles driving in emergency scenarios in light of significant challenges to the timely dispatch of vehicle repositioning. Tang et al. [61] proposed an RL method for resolving the online routing challenges facing fully-automated electric taxi fleets. A new RL technique was devised by Koh et al. [62] to address the issue of real-time vehicle navigation. When faced with the challenge of improving an existing taxi dispatching service, Mao et al. [63] developed a novel RL approach. The RL framework is one example; it is involved with finding the best policy in situations where the MDP method is insufficient. Therefore, the RL agent will adjust the control strategy for the transport network through direct communication with the system itself, rather than relying on any prior knowledge of the model. Al-Abbasi et al. introduced the distributed model-free pool (DeepPool) [64], which used the Q-learning method to learn optimal allocation approaches through observation of the environment. Due to the complex metropolitan context and unpredictable social occurrences (e.g., taxi drivers), Guo et al. [65] focused on building a real-time routing system to optimize the applicability of automobile service providers. They introduced an RL model that integrates a dynamic attention mechanism and a deep neural network (DNN) to address the problem of uncertainty in sensor data and the sophistication of interactions between vehicles and other road users. To the best of our knowledge, academic research on the utilization of the RL approach to deal with the dynamics and uncertainty concerns in AVs is quite rare. Methods used in RL typically rely on having a full observation, while future states necessitate some partial events. In order to formulate the partially observable ambiguity and dynamics, a POMDP model can be integrated into an ActInf model to cover the partial observations. The remaining parts show the previous research on ActInf models in AVs and their potential challenges.

Wei et al. [27] suggested an ActInf model (based on a POMDP model) integrated by visual looming as a new model of driving behavior. They considered the braking reaction scenarios in AVs to predict the AV's behavior. When compared with other models, this one is preferable since it evaluates cognitive patterns explicitly and can be easily scaled to represent some limited driving scenarios (e.g., vehicle following), while other driving scenarios for accelerating/decelerating are ignored. Their limited scenarios cannot provide a large driving dataset. For instance, in one of the abnormal scenarios related to TSs, it can be assumed that there is a STOP sign back (SSB) across the street. Hence, different maneuver scenarios are necessary to have a good prediction for an action. To tackle the necessity of large amounts of training data in addition to the inherent uncertainties in driving habits, Nozari et al. presented a combination structure of imitation learning and the ActInf model. By using a Bayesian network to take into account normal and abnormal scenarios, they were able to optimize the learning policy through a trade-off between return and abnormal values. Another novel aspect of their technique that demonstrated its superiority over RL methods was the ability to optimize action planning [30], [66]. Sensory data from two AVs were used as inputs to verify their model. This research did not consider the abnormal scenarios regarding TSs and different types of objects. Their validation was based on two AVs to show their implementation; however, different objects (e.g., pedestrians, TSs) have multiple states, and lexical concepts in TSs could be examples in this case [67].

A CAW navigation protocol was developed by Mugabarigira et al. [44] to improve the security of urban roadways for vehicles. The protocol's collision avoidance functionality

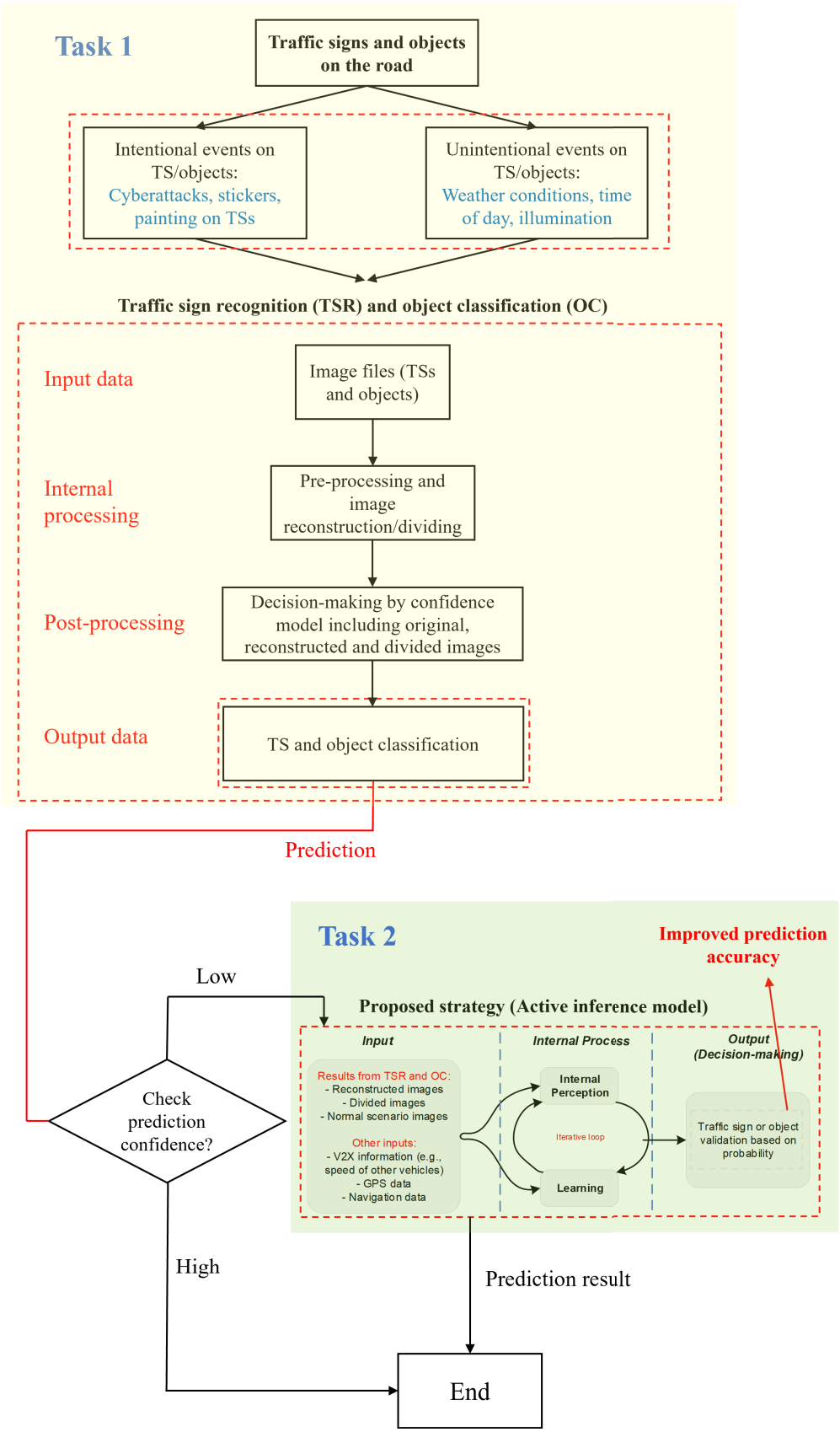


FIGURE 3. A proposed framework considering the ActInf model in abnormal scenarios to enhance the prediction accuracy.

enables AVs to navigate around barriers and roadside accidents in a safe mode. In this paper, the authors merely focused on different driving scenarios for collision avoidance, and even though they considered CAW in terms of states, actions, and driving behavior patterns, they did not bring up partial observation as an important aspect in decision-making procedures to enhance the prediction accuracy using partial observations. In [22], it was demonstrated that the multi-layer perceptron (MLP) algorithm failed to capture information from the obstructed TS images, and it was resolved by a hybrid MLP-POMDP classification model. The model provided a significantly more accurate classification of the TS images filtered with arbitrary fog masks for weather conditions. Based on their research, it appears that a POMDP model can be trained to interpret TSs correctly by giving it a dataset of previously observed and labeled images. The proposed algorithm considerably reduced the training time and achieved a precision of 93% for TS images. The challenges in this model include the following: they trained only seven TSs, which based on their accuracy values is not very applicable to a huge number of TSs and classes. Furthermore, some abnormal scenarios in TSs based on environmental behavior (i.e., perceptions from other objects) are not considered (e.g., fallen TSs and a sign back across the street). In addition, only a particular weather pattern was taken into account, which was replicated using images of fog.

For complex system planning processes, an extendable POMDP was employed by Pouya and Madni [68] in flexible and partially observable situations. The key benefit of this strategy was that it provided a basic framework (i.e., probabilistic state representation) in contrast to other data-driven approaches that rely on large datasets. This technique also initiated an action when belief in a state was very strong, resulting in moderate behavior. However, this model does not account for the variety of abnormal driving situations. In order to perform the POMDP solver algorithm in a matter of seconds while maintaining accuracy, the search time for locating the optimal policy was considerably decreased. An MDP model considering the velocity and position of the vehicle is considered in [1] for designing a behavior decision strategy based on the RL-based method. However, they did not consider the different patterns of other objects and TSs as a comprehensive model. Furthermore, the decision-making process in an ActInf model tries to minimize EFE and maximize the information gained from the environment (e.g., vehicles, TSs, and other objects) simultaneously. In an RL-based algorithm, it is time-consuming to maximize the rewards, while selective actions will help the model get the maximum probability of observations in the ActInf model.

V. PROPOSED ABNORMAL SCENARIOS WITH MACHINE LEARNING METHODS' DISABILITY IN AUTONOMOUS VEHICLES

This section highlights the abnormal occurrences, scenarios, and challenges related to each drivability aspect, including obvious and incidental factors and their difficulties. In contrast, for optimal or controlled driving situations typically considered in investigations of AVs, abnormal scenarios are those that reduce drivability and are thus more challenging to deal with. In these cases, ML methods are not able to handle these scenarios or can control with only low probability and less accuracy. All of these scenarios are based on the partial observation that ML approaches cannot show good accuracy in their predictions of results. Hence, a POMDP model based on an ActInf model can control and improve the accuracy of prediction for TSR and OC processes. To recapitulate, full observation models cannot consider some hazardous scenarios in autonomous driving, but a POMDP model can control these scenarios by considering the environmental behavior and CAW [69].

The research challenges associated with dealing with each factor are distinct, and current methods have not yet fulfilled their promise of providing comprehensive solutions. In order to keep track of the wide range of potential research topics, two types of factors can be mentioned, namely, obvious and incidental factors. Both factors can lead to unexpected and abnormal scenarios featuring high complexity, controllability, and uncertainty in autonomous driving. A summary of these factors can be outlined as follows:

- **Obvious Factors**
  - Illumination: Perception is substantially affected by changes in lighting conditions brought on by the time of day (night, dawn, dusk), the landscape (shades), and the individual light sources.
  - Weather Conditions: Many vision-based tasks, including road recognition and object tracking, can be challenging in extreme weather conditions, including fog and snow. How different climates impact the models' robustness in tackling perceptual tasks apart from road detection is, however, not very well elaborated [70].
  - Road Conditions: Road damage, irregular and rough surfaces, and construction are all examples of potentially unsafe road conditions. Due to the scarcity of documentation on potentially hazardous road situations, a shortage of labeled data, therefore, presents a significant challenge to recognizing them.
  - Road Geometry: Driving through intersections and roundabouts is more challenging than driving on a straight roadway. Time restrictions on yielding and merging movements make roundabouts much more problematic than intersections.
  - Traffic Conditions: Situations with heavy traffic, a high-speed limit, and the possibility of an accident require more attention. Furthermore, it is not apparent whether or how AVs' performance is affected by traffic circumstances or when an accident is most probable to occur because there are not enough datasets to make such assessments.
  - Static & Dynamic Objects: Existing approaches still show significant inaccuracy when confronted with

TABLE 2. A summary of ActInf and CAW research in AVs and their challenges.

Authors	Description
Wei *et al.* [27]	- Considered braking reaction scenarios by an ActInf-based POMDP. - Visual looming as a novel model of driving behavior. - Neglected some potential scenarios (e.g., TS back across the street).
Nozari *et al.* [30], [66]	- Considered combined ActInf model and imitation learning without a POMDP. - Superiority over RL techniques by optimizing action planning. - No consideration of abnormal scenarios related to TS and objects (e.g., vehicles, pedestrians, animals) blocking. - Validated based on two AVs, while TSs have lexical concepts.
Mugabarigira *et al.* [44]	- Developed a CAW navigation service to enhance roadway security. - No consideration of an ActInf model based on partial observations in scenarios.
Shahryari *et al.* [22]	- Captured environmental information by a combined model based on POMDP for obstructed TS images. - Enhanced accuracy prediction for a specific weather condition. - Limited number of TS images/classes. - Neglected TS abnormal scenarios (e.g., fallen TS, TS back) that ML methods can not recognize.
Pouya *et al.* [68]	- Developed an extendable POMDP that is not dependent on a large dataset. - Limited number of abnormal driving situations.
Zheng *et al.* [1]	- Considered an MDP model with only position and velocity of vehicles based on RL methods. - Ignored different objects and TSs in roadways. - Did not mention partial observation, including the information gained from the environment. - RL-based algorithms are time-consuming to maximize the reward while an ActInf model uses only selective actions.

unknown or difficult-to-detect objects, especially those of small size and severe occlusion. More research into various objects (e.g., animals, sudden obstructions) is required. Many of these obstructions (e.g., fallen trees or TSs) can also cause abrupt and sudden changes to the map necessary for vehicle positioning. These variations may reduce localization precision, which could complicate AV planning and decision-making procedures.

- Lane Markings: Despite the importance of lane markings as visual information for lane or road detection, a number of roads in urban and rural regions are unorganized, with damaged or missing lane markings or unusual lane shapes, making road detection challenging due to its dependence on lane marking identification [71].

- **Incidental Factors**

  - Object Behaviors (e.g., vehicle, pedestrian, animal, driver, bicyclist/motorcyclist): The behavior and actions of different objects in their interactions with AVs are examples of incidental factors. Every object can have a specific action, or interactions can happen for some static objects in abnormal cases (e.g., fallen TSs, unusual shapes of traffic barriers). Some of these unusual behaviors can include overtaking and failure to obey the traffic laws for vehicles, failure to comply with the laws for pedestrians, and so on.

As mentioned, some abnormal scenarios can be considered that cannot be predicted with ML approaches or can be predicted only with low accuracy and probability [72], [73]. The main reason is that partial observations in terms of objects' behavior and optimization of the number of actions are neglected. Hence, an ActInf model based on a POMDP model can enhance the accuracy of prediction and consider the posterior beliefs about the road's occurrences. These abnormal scenarios are outlined as follows:

A. SCENARIO I: FALLEN TRAFFIC SIGNS

One of the problematic scenarios for TSR through ML algorithms is a fallen TS, particularly the STOP sign. As shown in Fig. 4, for an AV (i.e., ego vehicle), particularly one depicted in green, forecasting potential scenarios on the road remains a challenging endeavor. When relying on comprehensive current observations, the TSR system struggles to anticipate instances such as a toppled TS. Consequently, its ability to discern or forecast the intended TS is compromised in terms of precision. Additionally, representations of the ego vehicle's vision, radar, and LiDAR capabilities are depicted in Fig. 5.

Conventional models might fail to recognize the missing sign if the sign has not encountered such a scenario during training. As a practical solution, the ActInf model, however, can incorporate behaviors of other road users, e.g., sudden stops or cautious driving near the intersection, as partial observations, allowing it to infer the potential presence of a STOP sign even if it is not physically visible. This method not only reacts to the current environment but also predicts potential future actions. For example, if it detects cars slowing down at the intersection and pedestrians looking both ways more frequently, it can infer the need to halt

even without a visible STOP sign. That means the ActInf-equipped vehicle, noticing unusual behavior from pedestrians and other vehicles, would infer the need to stop and proceed with caution, ensuring safety. Therefore, it does not require extensive retraining for new scenarios, and it adjusts to real-time changes efficiently.

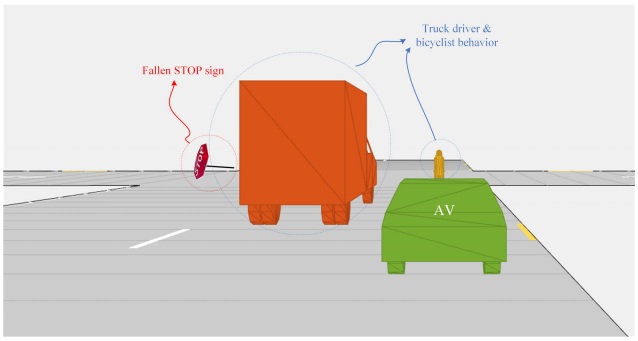


FIGURE 4. A fallen TS regarding other objects' behaviors on the road (Scenario I).

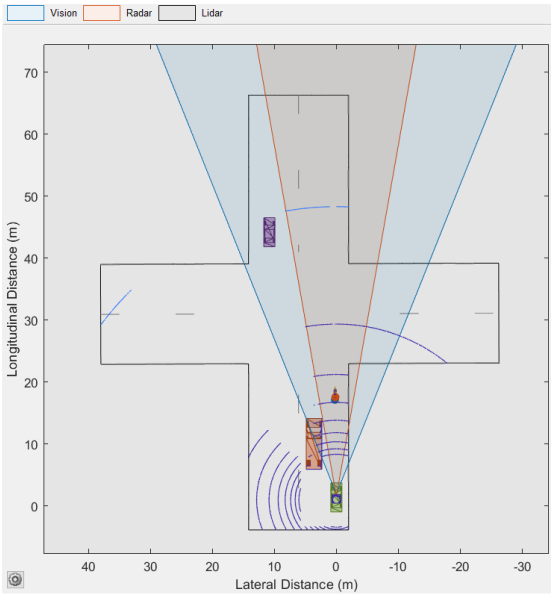


FIGURE 5. A sample coverage of camera, radar, and LiDAR sensors for an AV in an intersection with a fallen TS (Scenario I).

B. SCENARIO II: ROAD WORKERS' ALERTS BY TRAFFIC SIGNS

This scenario shows an emergency alert from road workers to obey their instructions. As can be seen in Fig. 6, a *65 mph Speed Limit* TS is on the roadside to continue forward on the route. Because a barrier exists on the road ahead, however, a road worker holds a different TS in his/her hand to show new instructions. It is a priority for all drivers to obey the workers in an emergency situation; however, an AV cannot follow this type of change. There is no difference between the presence/absence of barriers on the road, and it is the responsibility of drivers just to follow the new conditions. This scenario can be confusing for ML algorithms handling this issue, and a TSR model cannot make a highly accurate prediction of the appropriate TS. Even though the ego vehicle has full sensor coverage, it cannot recognize the correct route. In the context of roadwork scenarios where workers use various TSs to direct traffic, an AV can employ the ActInf framework to handle this scenario. The vehicle maintains a generative model of the world, which includes the probability of encountering roadwork and the associated signs. When it detects a road worker holding a sign, this observation updates the belief state. Based on the updated beliefs, the vehicle predicts the most probable TS the road worker might display next, considering the current context (e.g., ongoing construction, lane closures). The vehicle then selects an action (e.g., slowing down or changing lanes) that minimizes the discrepancy between its predictions and subsequent observations. This approach's superiority over traditional ML models lies in its ability to continuously refine its predictions based on feedback, allowing for more adaptive and context-aware decision-making.

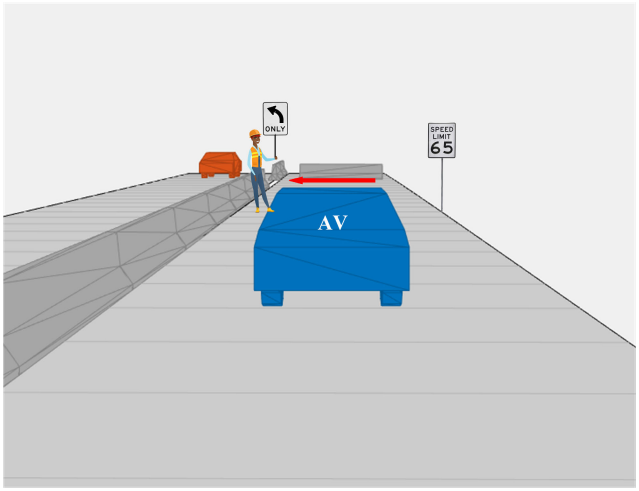


FIGURE 6. A misleading scenario for TSR regarding a new TS held by a road worker (Scenario II).

C. SCENARIO III: TRAFFIC SIGN BACK/ROTATED ACROSS THE ROAD

This scenario explains a TS back or a rotated TS on another side of an intersection (Fig. 7). This could be similar to "fallen TS," with the same analysis. The ego vehicle needs to gain information from the environment to perform a good action. Nevertheless, ML methods cannot predict the TSs because there is no lexical or CXI gained from the signs. Using ActInf model, the vehicle predicts that it could be a STOP or YIELD sign based on the intersection context. Thus, it slows down, preparing for both scenarios. As it gets closer, it identifies the sign as STOP and halts. In certain scenarios, the TSR model can accurately predict based on distinct TS shapes, such as the STOP sign. However, its efficacy

diminishes in the majority of cases due to variances in shapes and resemblances among certain numerals, among other factors. Moreover, the existing dataset for driving situations fails to encompass these atypical cases, which can potentially lead to mishaps when deploying standard ML models.

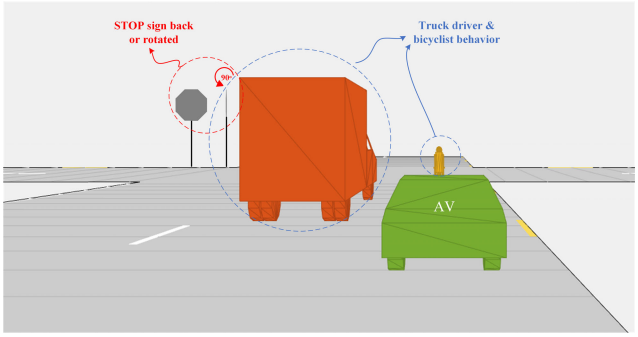


FIGURE 7. A scenario of STOP sign back or rotated sign on another side of an intersection (Scenario III).

D. SCENARIO IV: SPIRAL MOVEMENTS OF VEHICLES AHEAD IN ANOTHER LANE

The spiral movements of the vehicles ahead in another lane can make it impossible for the ego vehicle to make a good decision. Because it is an emergency situation, ML algorithms cannot make a correct decision, and the AV keeps going with the specified TS (i.e., 65 mph Speed Limit). By driving on a straight route, the AV can have an accident with the hazardous vehicle ahead, which makes it difficult for intelligent algorithms to make a correct decision. An up-view portion of this scenario associated with the route illustration for both vehicles is shown in Fig. 8. Consider another vehicle in a different lane that starts spiraling due to a tire burst. Traditional ML models might recognize the vehicle but not predict its erratic trajectory. An ActInf model, noticing the unexpected spiral movement, would increase the prediction error associated with that vehicle. By simulating possible future positions of the spiraling vehicle, the AV decides to slow down, allowing the spiraling vehicle to pass safely, and then continues its path, ensuring the safety of all road users.

E. SCENARIO V: AMBIGUOUS LANE MARKINGS

Low lane marking visibility or no marking is another scenario that can be regarded in considering ML algorithms' disability in AVs, as can be shown in Fig. 9. Visibility can be reduced by severe weather conditions (e.g., fog, snow, heavy rain) and can cause AVs to acquire misleading information [71]. Because of these crucial algorithms (e.g., TSR and OC) in AVs, any complexity or ambiguity in the lane marking can reduce the prediction accuracy of the AV's decision-making, and the AV needs to gain other behavioral/environmental information to predict the objects or TSs more accurately. As a potential solution, ActInf models emphasize prediction and adaptability, which can be crucial when lane markings are faded or non-existent. It integrates high-definition maps

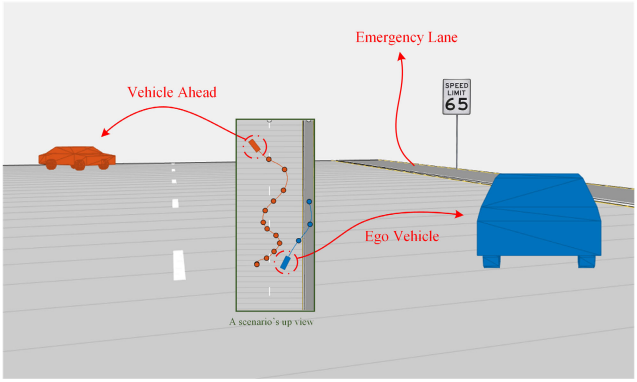


FIGURE 8. An emergency situation caused by spiral driving of a vehicle ahead (Scenario IV).

and real-time observations. If lane markings are missing, the model can rely more heavily on the map data to navigate safely. Also, if a road segment previously had clear lane markings, the model can use this historical data to predict where the lanes should be, even if current observations are unclear.

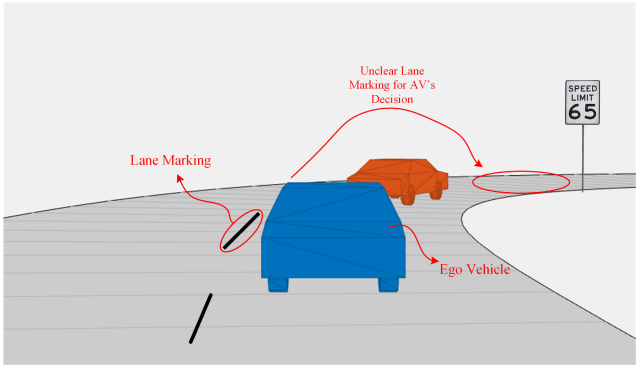


FIGURE 9. An impact of unclear lane markings on an AV's decision making and consideration of other objects' behaviors (Scenario V).

F. SCENARIO VI: CYBERATTACK ON STOP SIGN AT A 4-WAY STOP INTERSECTION

In this scenario, a 4-way stop intersection involving the AV (ego vehicle) is shown in Fig. 10 in which an attacker tries to change the STOP sign recognized by the ego vehicle. By intruding into the STOP sign, an attacker can change this sign, and the ego vehicle is not able to recognize the correct sign. The attacker can make a wide range of misleading TSs, leading the AV to approach other vehicles on the road and cause an accident. In the context of vehicular intersections, there exists a potential risk for collisions involving all vehicles present. The TSR algorithm's limitations are evident, as it fails to accurately predict signs due to its lack of integration with the environmental dynamics of surrounding entities. However, utilizing an ActInf model, which incorporates partial observational data, offers a more robust countermeasure against cyber threats. This enhanced

prediction capability stems from its ability to assimilate information from adjacent vehicles, their respective motion patterns within the same traffic lane, and the context of a 4-way intersection. The vehicle's sensors detect a sign that looks similar to a Turn Left Only sign, but other sensory data (e.g., road layout and intersection structure) suggest a 4-way stop. So, the vehicle anticipates the presence of a STOP sign based on its generative model, even though the visual data is inconsistent with this prediction. Instead of relying solely on the altered visual cue, the vehicle cross-references its prediction with other sensor data and decides stop, thus avoiding a potential accident. Then, the vehicle updates its internal model to account for the possibility of sign tampering, making it more resilient to similar attacks in the future.

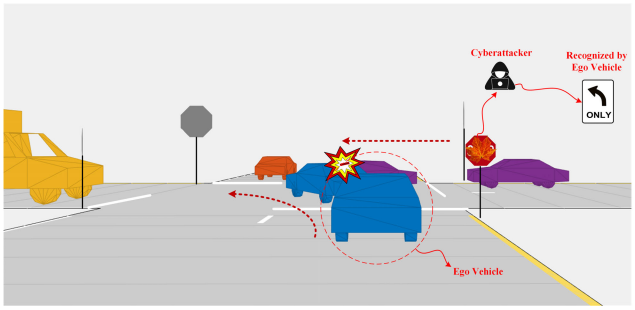


FIGURE 10. A misleading representation of STOP sign by a cyberattack on the sign (Scenario VI).

G. SCENARIO VII: CYBERATTACK ON DIGITAL HIGHWAY SIGNS

This scenario demonstrates an abnormal case caused by an attacker. As it can be observed in Fig. 11, the ego vehicle is moving forward in a lane and can recognize the "SPEED LIMIT 85" in a good way. After crossing this TS, it confronts an unexpected event (i.e., an accident ahead) on the road in which the digital highway sign shows more information and instructions for all vehicles. This digital sign should illustrate the "ACCIDENT AHEAD, REDUCE SPEED" in normal conditions; however, a cyber attacker intrudes and changes the instructions to a new announcement, as "WEAR A MASK. SAVE LIVES." which can be considered a cyberattack. Generally, drivers can observe the accident ahead, figure out that this digital sign shows misleading information, and make the correct decision. Nevertheless, the TSR algorithm cannot recognize the accident ahead and just receives normal instruction, which is not necessary to take any action. Therefore, the ML algorithm cannot handle this attack, and the AV approaches the accident location and can cause another catastrophic event. Instead of relying solely on historical data, the ActInf method utilizes the current environment and ongoing observations to generate a predictive model of the highway system and follows these procedures to have a good prediction and action simultaneously:

- The system continuously observes traffic patterns, sign behaviors, and driver reactions in real-time.
- The model updates its predictions based on new observations, factoring in the motion of vehicles, behavior of nearby signs, and any feedback from drivers or other infrastructure.
- If a digital sign displays unexpected information not aligning with the model's predictions, ActInf recognizes this anomaly more efficiently than traditional ML methods.
- The system can then alert authorities or even initiate corrective measures such as broadcasting alerts to nearby vehicles or resetting the affected sign.

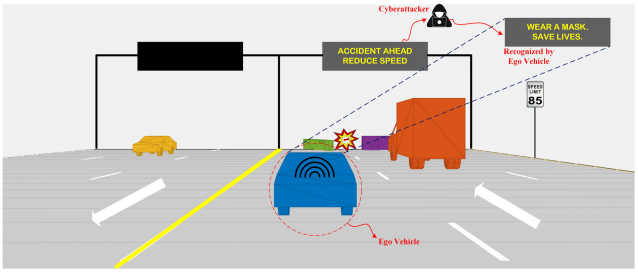


FIGURE 11. A cyberattacks on digital highway sign (Scenario VII).

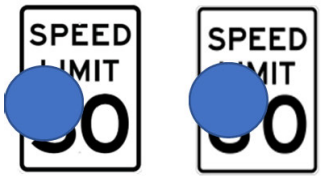


FIGURE 12. Blocked TSs with blue circles (left: SPEED LIMIT 30, right: SPEED LIMIT 60).

VI. DISCUSSION

The ActInf model, leveraging partial observation, emerges as a potential solution to address TSR in autonomous driving. Nonetheless, its implementation is not without challenges. A detailed risk assessment reveals several areas of concern as follows:

- **Abnormal Scenarios:** While the ActInf model may exhibit improved prediction accuracy in abnormal situations, the true extent of this improvement needs validation. Autonomous driving operates in dynamic environments where unexpected events are frequent. Ensuring the model's robustness in these situations is paramount.
- **Dataset Limitations:** A significant risk lies in the available datasets. Comprehensive naturalistic TS and object datasets are essential for training, but many datasets might not capture the full spectrum of abnormal or rare scenarios. This lack of representation can affect the model's efficiency in real-world applications.

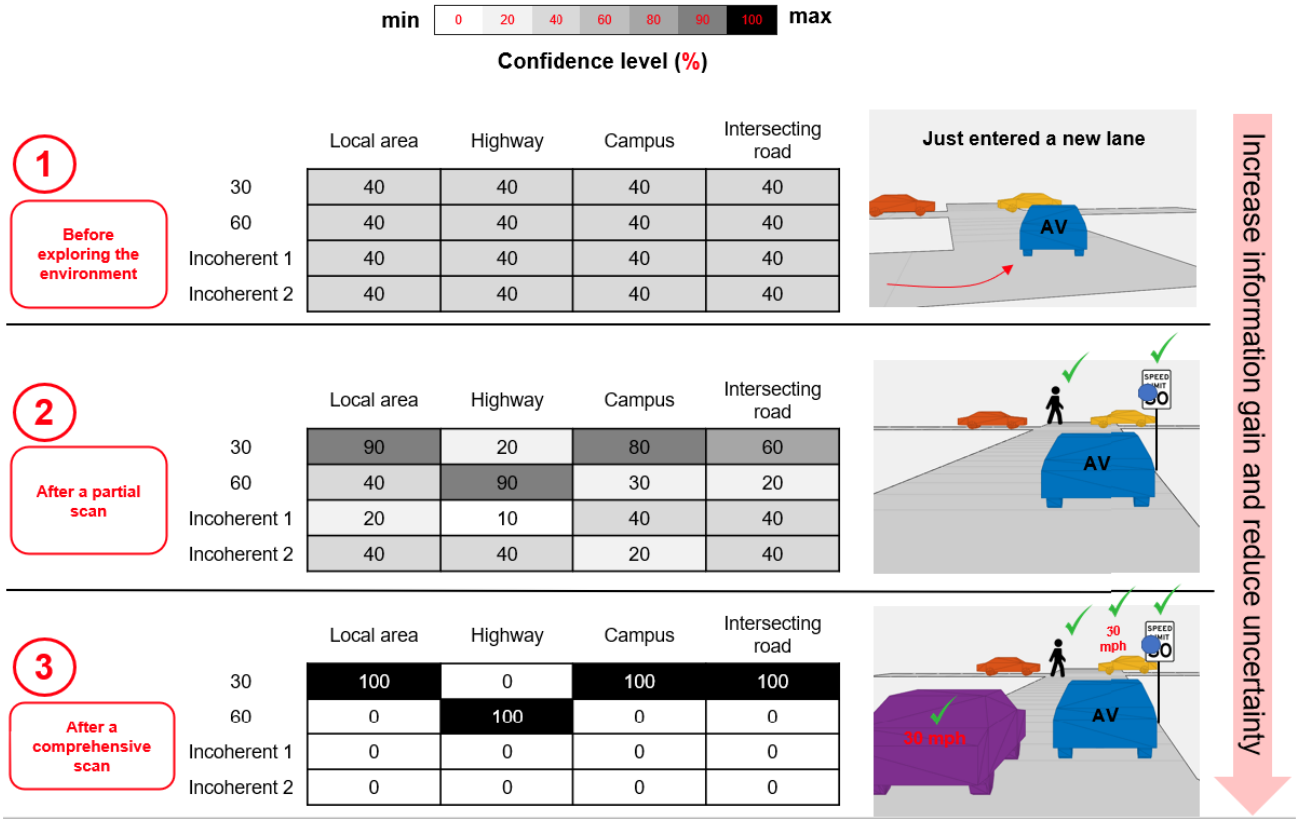


FIGURE 13. A sample learning process of mapping matrix between hidden states and observations in an ActInf model.

- **Object Variability:** The autonomous driving landscape involves a myriad of objects, both static and dynamic. Distinguishing between them and predicting their behaviors, especially in crowded or cluttered environments, can complicate the model. Incorporating 3D object detection methods might assist in better environmental understanding.
- **Acceptance Rate Analysis:** Beyond its technical capabilities, the societal and practical acceptance of the model is crucial. It is necessary to assess how frequently the model's predictions align with real-world outcomes and user expectations. A high acceptance rate would signify not only technical proficiency but also user trust.

A. AUTHORS' PERSPECTIVES

The authors believed that the ActInf model can efficiently improve the TSR and OC algorithms' accuracy using the CAW and situational awareness considering the proposed abnormal scenario. For more clarification, a sample example with a part of internal processing is demonstrated to show the better performance of this suggested model. Consider blocks on SPEED LIMIT 30 and SPEED LIMIT 60 signs by blue circles that an AV has been faced with, as illustrated in Fig. 12.

In this case, the TSR algorithm, considering the image reconstruction and image dividing methods, cannot make a good prediction for the recognition of the TS on the road. The major difference between these signs (i.e., numerical parts) are blocked with the same size, and the reminder parts are analogous. Hence, the ActInf model employs other data from the AV environment, including the speed of adjacent vehicles, GPS data, and so on to make an estimation in each step, and updates the beliefs for each time for hidden states and observations. Fig. 13 demonstrates a sample learning process in the internal processing of the ActInf model which can happen.

According to Fig. 13, the confidence level bar shows the minimum and maximum probability of a prediction based on an accuracy value. Three tables within this figure show mapping matrices between hidden states and observations in which the rows and columns depict observations and hidden states, respectively. It is noteworthy to mention that we can define different hidden states and observations and these tables merely illustrate sample matrices to assess the probability values in different steps. Step 1 presents a scenario in which an AV just entered a new street from another one, and this vehicle has not any perception of this new environment. Hence, all probability values can be adjusted by 40% as an example. If we want to mention an array in the matrix, we can consider 1×1 array which expresses the AV has a speed of 30 at a local area with a probability of 40%. Other arrays can be similarly described. The terms "Incoherent 1" and

"Incoherent 2" represent that prediction has some features of SPEED LIMIT 30 or SPEED LIMIT 60 signs, so it is unknown based on a visual illustration to distinguish between these two signs. In Step 2, a partial scan of the AV surrounding has happened and the AV can gain some information from the environment to improve the knowledge. Based on this step, this AV can get information on a presence of a pedestrian and a blocked speed limit sign. This new information helps the AV to enhance the accuracy of confidence as shown in the second matrix (e.g., 1×1 array). As can be observed, the confidence level increased from 40% to 90% from Step 1 to Step 2 after a partial scan of the environment occurred. The same learning process can occur for Step 3 by gaining more comprehensive information (i.e., speed of adjacent vehicles, presence of pedestrians, speed limit) which can lead to the highest confidence in the prediction of speed limit signs.

B. TIME EFFICIENCY OF AN ACTINF MODEL IN AV APPLICATIONS

The primary strength of ActInf lies in its ability to provide quicker decision-making processes, which is crucial in the rapidly changing environments that AVs navigate. These models differentiate themselves from POMDPs by reducing the need for extensive calculations that consider various observations and actions in uncertain situations. The ActInf model employs immediate sensory data and predictive modeling to simplify the decision-making process. This approach enables quicker adaptation to immediate circumstances, enhancing the efficiency and practicality of ActInf models in situations demanding prompt action. It directly utilizes real-time data without the exhaustive calculation of every potential scenario.

Conversely, POMDPs entail significant computational demands and extended processing times. This complexity stems from the POMDPs requirement to meticulously analyze each potential action sequence and its consequent effects to optimize future outcomes. Such extensive computational requirements can lead to inefficiencies, particularly in complex and real-world driving environments where swift decision-making is preferable. Therefore, while POMDPs present a detailed method for decision-making in uncertain conditions, their high computational load renders them less viable for real-time operations in AVs. In these instances, ActInf models emerge as a more time-efficient solution.

VII. CONCLUSION

A comprehensive study evaluating TSR and OC techniques is presented to illustrate their effectiveness and precision in fully observable models, which are different from partial observable models, without regarding uncertainties and objects' behaviors. In this article, a comprehensive overview of ActInf models based on a POMDP model is presented for assessment with various kinds of environmental occurrences and CAW in different applications. This article also discusses the capabilities and superiority of this model in comparison with ML algorithms. A thorough analysis of the existing safety concerns associated with AVs is described, focusing on partial observations of the surrounding environment and the behavior of various objects. In the context of autonomous driving, some anomalous scenarios have been described in which ML algorithms are unable to address these circumstances or in which they obtain a prediction based on the objects on the road with less accuracy. Potential enhancements for this methodology warrant exploration in future work. Undertaking a comprehensive risk evaluation of the ActInf model, particularly when rooted in partial observations, is pivotal. There is an imperative to probe the augmented predictive accuracy this innovative approach brings, especially in atypical circumstances. It would be judicious to assess expansive datasets encompassing naturalistic TSs and other vehicular objects within the autonomous driving domain to ascertain the efficacy of this strategy. Real-world driving conditions inevitably present a plethora of unforeseen and exigent situations. The myriad of objects, be they static or in motion, amplifies the intricacy of the model in deciphering environmental dynamics. Consequently, a rigorous scrutiny of the model's acceptance rate is indispensable.

REFERENCES

[1] K. Zheng, H. Yang, S. Liu, K. Zhang, and L. Lei, "A behavior decision method based on reinforcement learning for autonomous driving," *IEEE Internet Things J.*, vol. 9, no. 24, pp. 25386–25394, Dec. 2022.

[2] C. Xia, M. Xing, and S. He, "Interactive planning for autonomous driving in intersection scenarios without traffic signs," *IEEE Trans. Intell. Transp. Syst.*, vol. 23, no. 12, pp. 24818–24828, Dec. 2022.

[3] Y. Ma, C. Sun, J. Chen, D. Cao, and L. Xiong, "Verification and validation methods for decision-making and planning of automated vehicles: A review," *IEEE Trans. Intell. Vehicles*, vol. 7, no. 3, pp. 480–498, Sep. 2022.

[4] J. Müller, J. Strohbeck, M. Herrmann, and M. Buchholz, "Motion planning for connected automated vehicles at occluded intersections with infrastructure sensors," *IEEE Trans. Intell. Transp. Syst.*, vol. 23, no. 10, pp. 17479–17490, Oct. 2022.

[5] M. Buchholz, J. Müller, M. Herrmann, J. Strohbeck, B. Völz, M. Maier, J. Paczia, O. Stein, H. Rehborn, and R.-W. Henn, "Handling occlusions in automated driving using a multiaccess edge computing server-based environment model from infrastructure sensors," *IEEE Intell. Transp. Syst. Mag.*, vol. 14, no. 3, pp. 106–120, May 2022.

[6] W. Wang, W. Han, X. Na, J. Gong, and J. Xi, "A probabilistic approach to measuring driving behavior similarity with driving primitives," *IEEE Trans. Intell. Vehicles*, vol. 5, no. 1, pp. 127–138, Mar. 2020.

[7] H. A. Ignatious, H. El-Sayed, M. A. Khan, and B. M. Mokhtar, "Analyzing factors influencing situation awareness in autonomous vehicles—A survey," *Sensors*, vol. 23, no. 8, p. 4075, Apr. 2023.

[8] S. S. Avedisov, C. R. He, D. Takács, and G. Orosz, "Machine learning-based steering control for automated vehicles utilizing V2X communication," in *Proc. IEEE Conf. Control Technol. Appl. (CCTA)*, Aug. 2021, pp. 253–258.

[9] W. Zhang and W. Wang, "Learning V2V interactive driving patterns at signalized intersections," *Transp. Res. C, Emerg. Technol.*, vol. 108, pp. 151–166, Nov. 2019.

[10] S. Liu, K. Zheng, L. Zhao, and P. Fan, "A driving intention prediction method based on hidden Markov model for autonomous driving," *Comput. Commun.*, vol. 157, pp. 143–149, May 2020.

[11] E. Ohn-Bar and M. M. Trivedi, "Looking at humans in the age of self-driving and highly automated vehicles," *IEEE Trans. Intell. Vehicles*, vol. 1, no. 1, pp. 90–104, Mar. 2016.

[12] F.-C. Chou, T.-H. Lin, H. Cui, V. Radosavljevic, T. Nguyen, T.-K. Huang, M. Niedoba, J. Schneider, and N. Djuric, "Predicting motion of vulnerable road users using high-definition maps and efficient ConvNets," in *Proc. IEEE Intell. Vehicles Symp. (IV)*, Oct. 2020, pp. 1655–1662.

[13] M. Hasan, S. Mohan, T. Shimizu, and H. Lu, "Securing vehicle-to-everything (V2X) communication platforms," *IEEE Trans. Intell. Vehicles*, vol. 5, no. 4, pp. 693–713, Dec. 2020.

[14] M. Girdhar, J. Hong, and J. Moore, "Cybersecurity of autonomous vehicles: A systematic literature review of adversarial attacks and defense models," *IEEE Open J. Veh. Technol.*, vol. 4, pp. 417–437, 2023.

[15] A. Khadka, P. Karypidis, A. Lytos, and G. Efstathopoulos, "A benchmarking framework for cyber-attacks on autonomous vehicles," *Transp. Res. Proc.*, vol. 52, pp. 323–330, Jan. 2021.

[16] A. Giannaros, A. Karras, L. Theodorakopoulos, C. Karras, P. Kranias, N. Schizas, G. Kalogeratos, and D. Tsolis, "Autonomous vehicles: Sophisticated attacks, safety issues, challenges, open topics, blockchain, and future directions," *J. Cybersecur. Privacy*, vol. 3, no. 3, pp. 493–543, Aug. 2023.

[17] M. N. Teli and S. Oh, "Resilience of autonomous vehicle object category detection to universal adversarial perturbations," in *Proc. IEEE Int. IoT, Electron. Mechatronics Conf. (IEMTRONICS)*, Apr. 2021, pp. 1–6.

[18] G. Bathla, K. Bhadane, R. K. Singh, R. Kumar, R. Aluvalu, R. Krishnamurthi, A. Kumar, R. N. Thakur, and S. Basheer, "Autonomous vehicles and intelligent automation: Applications, challenges, and opportunities," *Mobile Inf. Syst.*, vol. 2022, Jun. 2022, Art. no. 7632892.

[19] C.-I. Kim, J. Park, Y. Park, W. Jung, and Y.-S. Lim, "Deep learning-based real-time traffic sign recognition system for urban environments," *Infrastructures*, vol. 8, no. 2, p. 20, Jan. 2023.

[20] A. F. Magnussen, N. Le, L. Hu, and W. E. Wong, "A survey of the inadequacies in traffic sign recognition systems for autonomous vehicles," *Int. J. Performability Eng.*, vol. 16, no. 10, p. 1588, 2020.

[21] S. Ghosh, A. Zaboli, J. Hong, and J. Kwon, "An integrated approach of threat analysis for autonomous vehicles perception system," *IEEE Access*, vol. 11, pp. 14752–14777, 2023.

[22] S. Shahryari and C. Hamilton, "Neural network-POMDP-based traffic sign classification under weather conditions," in *Proc. Can. Conf. Artif. Intell.* Cham, Switzerland: Springer, 2016, pp. 122–127.

[23] K. Lim, Y. Hong, Y. Choi, and H. Byun, "Real-time traffic sign recognition based on a general purpose GPU and deep-learning," *PLoS One*, vol. 12, no. 3, Mar. 2017, Art. no. e0173317.

[24] T. Wang, Y. Chen, X. Yan, W. Li, and D. Shi, "Assessment of drivers' comprehensive driving capability under man-computer cooperative driving conditions," *IEEE Access*, vol. 8, pp. 152909–152923, 2020.

[25] F. Wang, J. Zhang, S. Wang, S. Li, and W. Hou, "Analysis of driving behavior based on dynamic changes of personality states," *Int. J. Environ. Res. Public Health*, vol. 17, no. 2, p. 430, Jan. 2020.

[26] T. Wang, Y. Chen, X. Yan, J. Chen, and W. Li, "The relationship between bus drivers' improper driving behaviors and abnormal vehicle states based on advanced driver assistance systems in naturalistic driving," *Math. Problems Eng.*, vol. 2020, Aug. 2020, Art. no. 9743504.

[27] R. Wei, A. D. McDonald, A. Garcia, and H. Alambeigi, "Modeling driver responses to automation failures with active inference," *IEEE Trans. Intell. Transp. Syst.*, vol. 23, no. 10, pp. 18064–18075, Oct. 2022.

[28] R. Wei, A. Garcia, A. McDonald, G. Markkula, J. Engström, I. Supeene, and M. O'Kelly, "World model learning from demonstrations with active inference: Application to driving behavior," in *Proc. 3rd Int. Workshop Act. Inference*, Grenoble, France, 2022, pp. 130–142.

[29] D. A. Friedman, A. Tschantz, M. J. D. Ramstead, K. Friston, and A. Constant, "Active inferants: An active inference framework for ant colony behavior," *Frontiers Behav. Neurosci.*, vol. 15, p. 126, Jun. 2021.

[30] S. Nozari, A. Krayani, P. Marin-Plaza, L. Marcenaro, D. M. Gómez, and C. Regazzoni, "Active inference integrated with imitation learning for autonomous driving," *IEEE Access*, vol. 10, pp. 49738–49756, 2022.

[31] B. Millidge, A. Tschantz, A. K. Seth, and C. L. Buckley, "On the relationship between active inference and control as inference," in *Proc. 1st Int. Workshop Act. Inference*, Ghent, Belgium. Cham, Switzerland: Springer, 2020, pp. 3–11.

[32] K. J. Friston, J. Daunizeau, and S. J. Kiebel, "Reinforcement learning or active inference?" *PLoS One*, vol. 4, no. 7, p. e6421, Jul. 2009.

[33] K. J. Friston, T. Parr, and B. de Vries, "The graphical brain: Belief propagation and active inference," *Netw. Neurosci.*, vol. 1, no. 4, pp. 381–414, Dec. 2017.

[34] M. J. Wainwright and M. I. Jordan, "Graphical models, exponential families, and variational inference," *Found. Trends Mach. Learn.*, vol. 1, nos. 1–2, pp. 1–305, 2007.

[35] A. Imohiosen, J. Watson, and J. Peters, "Active inference or control as inference? A unifying view," in *Proc. 1st Int. Workshop Act. Inference*, Ghent, Belgium. Cham, Switzerland: Springer, 2020, pp. 12–19.

[36] B. Millidge and C. L. Buckley, "Active inference successor representations," in *Proc. 3rd Int. Workshop Act. Inference*, Grenoble, France. Cham, Switzerland: Springer, 2023, pp. 151–161.

[37] S. T. Wauthier, B. Vanhecke, T. Verbelen, and B. Dhoedt, "Learning generative models for active inference using tensor networks," in *Proc. 3rd Int. Workshop Act. Inference*, Grenoble, France. Cham, Switzerland: Springer, 2023, pp. 285–297.

[38] E. Sennesh, J. Theriault, J.-W. van de Meent, L. F. Barrett, and K. Quigley, "Deriving time-averaged active inference from control principles," in *Proc. 3rd Int. Workshop Act. Inference*, Grenoble, France. Cham, Switzerland: Springer, 2023, pp. 355–370.

[39] P. Pouya and A. M. Madni, "Policy transfer in POMDP models for safety-critical autonomous vehicles applications," in *Proc. IEEE Int. Conf. Syst., Man, Cybern. (SMC)*, Oct. 2022, pp. 1350–1355.

[40] R. Smith, K. J. Friston, and C. J. Whyte, "A step-by-step tutorial on active inference and its application to empirical data," *J. Math. Psychol.*, vol. 107, Apr. 2022, Art. no. 102632.

[41] B. Millidge, A. Tschantz, and C. L. Buckley, "Whence the expected free energy?" *Neural Comput.*, vol. 33, no. 2, pp. 447–482, Feb. 2021.

[42] B. Okumura, M. R. James, Y. Kanzawa, M. Derry, K. Sakai, T. Nishi, and D. Prokhorov, "Challenges in perception and decision making for intelligent automotive vehicles: A case study," *IEEE Trans. Intell. Vehicles*, vol. 1, no. 1, pp. 20–32, Mar. 2016.

[43] P. Hang, C. Huang, Z. Hu, Y. Xing, and C. Lv, "Decision making of connected automated vehicles at an unsignalized roundabout considering personalized driving behaviours," *IEEE Trans. Veh. Technol.*, vol. 70, no. 5, pp. 4051–4064, May 2021.

[44] B. A. Mugabarigira, Y. Shen, J. Jeong, T. Oh, and H.-Y. Jeong, "Context-aware navigation protocol for safe driving in vehicular cyber-physical systems," *IEEE Trans. Intell. Transp. Syst.*, vol. 24, no. 1, pp. 128–138, Jan. 2023.

[45] C. Skoglund, "Risk-aware autonomous driving using POMDPs and responsibility-sensitive safety," M.S. thesis, KTH Roy. Inst. Technol., Stockholm, Sweden, 2021.

[46] N. B. Haupt and P. Liggesmeyer, "Towards context-awareness for enhanced safety of autonomous vehicles," in *Proc. SAI Intell. Syst. Conf.* Cham, Switzerland: Springer, 2021, pp. 549–563.

[47] W. Schwarting, J. Alonso-Mora, and D. Rus, "Planning and decision-making for autonomous vehicles," *Annu. Rev. Control, Robot., Auto. Syst.*, vol. 1, no. 1, pp. 187–210, May 2018.

[48] B. Paden, M. Cáp, S. Z. Yong, D. Yershov, and E. Frazzoli, "A survey of motion planning and control techniques for self-driving urban vehicles," *IEEE Trans. Intell. Vehicles*, vol. 1, no. 1, pp. 33–55, Mar. 2016.

[49] K. Sonoda and T. Wada, "Displaying system situation awareness increases driver trust in automated driving," *IEEE Trans. Intell. Vehicles*, vol. 2, no. 3, pp. 185–193, Sep. 2017.

[50] N. Wang, S. Lv, J. E. Meng, and W.-H. Chen, "Fast and accurate trajectory tracking control of an autonomous surface vehicle with unmodeled dynamics and disturbances," *IEEE Trans. Intell. Vehicles*, vol. 1, no. 3, pp. 230–243, Sep. 2016.

[51] I. Chadès, L. V. Pascal, S. Nicol, C. S. Fletcher, and J. Ferrer-Mestres, "A primer on partially observable Markov decision processes (POMDPs)," *Methods Ecol. Evol.*, vol. 12, no. 11, pp. 2058–2072, Nov. 2021.

[52] H. Kurniawati, "Partially observable Markov decision processes and robotics," *Annu. Rev. Control, Robot., Auto. Syst.*, vol. 5, no. 1, pp. 253–277, May 2022.

[53] C.-J. Hoel, K. Driggs-Campbell, K. Wolff, L. Laine, and M. J. Kochenderfer, "Combining planning and deep reinforcement learning in tactical decision making for autonomous driving," *IEEE Trans. Intell. Vehicles*, vol. 5, no. 2, pp. 294–305, Jun. 2020.

[54] A. Schotschneider. (2018). *Collision Avoidance in Uncertain Environments for Autonomous Vehicles Using POMDPs*. [Online]. Available: http://www.ausy.tu-darmstadt.de/uploads/Theses/Abschlussarbeiten/albert_schotschneider_bsc.pdf

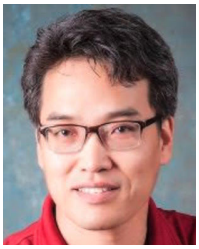
[55] A. J. M. Muzahid, S. F. Kamarulzaman, M. A. Rahman, S. A. Murad, M. A. S. Kamal, and A. H. Alenezi, "Multiple vehicle cooperation and collision avoidance in automated vehicles: Survey and an AI-enabled conceptual framework," *Sci. Rep.*, vol. 13, no. 1, p. 603, Jan. 2023.

[56] M. Cullen, "Active inference in simulated cortical circuits," Ph.D. dissertation, Dept. Eng. Math., Univ. Bristol, Bristol, England, 2021.

[57] J. Ochelska-Mierzejewska, A. Poniszewska-Marańda, and W. Marańda, "Selected genetic algorithms for vehicle routing problem solving," *Electronics*, vol. 10, no. 24, p. 3147, Dec. 2021.
[58] J. Zhang, K. Luo, A. M. Florio, and T. Van Woensel, "Solving large-scale dynamic vehicle routing problems with stochastic requests," *Eur. J. Oper. Res.*, vol. 306, no. 2, pp. 596–614, Apr. 2023.
[59] T. Mustakhov, Y. Akhmetbek, and A. Bogyrbayeva, "Deep reinforcement learning for stochastic dynamic vehicle routing problem," in *Proc. 17th Int. Conf. Electron. Comput. Comput. (ICECCO)*, Jun. 2023, pp. 1–5.
[60] X. Liu, D. Zhang, T. Zhang, Y. Cui, L. Chen, and S. Liu, "Novel best path selection approach based on hybrid improved A* algorithm and reinforcement learning," *Int. J. Speech Technol.*, vol. 51, no. 12, pp. 9015–9029, Dec. 2021.
[61] X. Tang, M. Li, X. Lin, and F. He, "Online operations of automated electric taxi fleets: An advisor-student reinforcement learning framework," *Transp. Res. C, Emerg. Technol.*, vol. 121, Dec. 2020, Art. no. 102844.
[62] S. Koh, B. Zhou, H. Fang, P. Yang, Z. Yang, Q. Yang, L. Guan, and Z. Ji, "Real-time deep reinforcement learning based vehicle navigation," *Appl. Soft Comput.*, vol. 96, Nov. 2020, Art. no. 106694.
[63] C. Mao, Y. Liu, and Z.-J. Shen, "Dispatch of autonomous vehicles for taxi services: A deep reinforcement learning approach," *Transp. Res. C, Emerg. Technol.*, vol. 115, Jun. 2020, Art. no. 102626.
[64] A. O. Al-Abbasi, A. Ghosh, and V. Aggarwal, "DeepPool: Distributed model-free algorithm for ride-sharing using deep reinforcement learning," *IEEE Trans. Intell. Transp. Syst.*, vol. 20, no. 12, pp. 4714–4727, Dec. 2019.
[65] P. Guo, K. Xiao, Z. Ye, and W. Zhu, "Route optimization via environment-aware deep network and reinforcement learning," *ACM Trans. Intell. Syst. Technol.*, vol. 12, no. 6, pp. 1–21, Dec. 2021.
[66] S. Nozari, A. Krayani, P. Marin, L. Marcenaro, D. Martin, and C. Regazzoni, "Autonomous driving based on imitation and active inference," in *Advances in System-Integrated Intelligence*, M. Valle, D. Lehmhus, C. Gianoglio, E. Ragusa, L. Seminara, S. Bosse, A. Ibrahim, and K.-D. Thoben, Eds. Cham, Switzerland: Springer, 2023, pp. 13–22.
[67] A. Rasouli, I. Kotseruba, and J. K. Tsotsos, "Understanding pedestrian behavior in complex traffic scenes," *IEEE Trans. Intell. Vehicles*, vol. 3, no. 1, pp. 61–70, Mar. 2018.
[68] P. Pouya and A. M. Madni, "Expandable-partially observable Markov decision-process framework for modeling and analysis of autonomous vehicle behavior," *IEEE Syst. J.*, vol. 15, no. 3, pp. 3714–3725, Sep. 2021.
[69] J. Guo, U. Kurup, and M. Shah, "Is it safe to drive? An overview of factors, metrics, and datasets for driveability assessment in autonomous driving," *IEEE Trans. Intell. Transp. Syst.*, vol. 21, no. 8, pp. 3135–3151, Aug. 2020.
[70] D. Bevly, X. Cao, M. Gordon, G. Ozbilgin, D. Kari, B. Nelson, J. Woodruff, M. Barth, C. Murray, A. Kurt, K. Redmill, and U. Ozguner, "Lane change and merge maneuvers for connected and automated vehicles: A survey," *IEEE Trans. Intell. Vehicles*, vol. 1, no. 1, pp. 105–120, Mar. 2016.
[71] Z. He, L. Zheng, L. Lu, and W. Guan, "Erasing lane changes from roads: A design of future road intersections," *IEEE Trans. Intell. Vehicles*, vol. 3, no. 2, pp. 173–184, Jun. 2018.
[72] L. Li, W.-L. Huang, Y. Liu, N.-N. Zheng, and F.-Y. Wang, "Intelligence testing for autonomous vehicles: A new approach," *IEEE Trans. Intell. Vehicles*, vol. 1, no. 2, pp. 158–166, Jun. 2016.
[73] Y. Li, A. Møgelmose, and M. M. Trivedi, "Pushing the 'speed limit': High-accuracy us traffic sign recognition with convolutional neural networks," *IEEE Trans. Intell. Vehicles*, vol. 1, no. 2, pp. 167–176, Jun. 2016.

AYDIN ZABOLI (Graduate Student Member, IEEE) received the B.S. degree in electrical engineering from the Babol Noshirvani University of Technology, Babol, Iran, in 2012, and the M.S. degree in electrical engineering from the Amirkabir University of Technology (Tehran Polytechnic), Tehran, Iran, in 2015. He is currently pursuing the Ph.D. degree in electrical, electronics, and computer engineering with the University of Michigan–Dearborn, Dearborn, MI, USA. From 2015 to 2021, he was a Research Assistant with the Amirkabir University of Technology. His research interests include the security of smart grids and autonomous vehicles, transportation electrification, power system planning, and load forecasting.

JUNHO HONG (Senior Member, IEEE) received the Ph.D. degree in cyber-security of substation automation systems in electrical engineering from Washington State University, Pullman, WA, USA, in 2014. From 2014 to 2019, he was with ABB, where he provided technical project leadership and supported strategic corporate technology development/productization in the areas related to cyber-physical security for substations, power grid control and protection, renewable integration, and utility communications. He has been the Principal Investigator (PI) in cyber-security of energy delivery systems of the Department of Energy (DOE) and the Co-PI in the areas of the substation, microgrid, HVDC, FACTS, and high-power EV charger. He is currently an Assistant Professor with the Department of Electrical and Computer Engineering, University of Michigan–Dearborn. He serves in Cigre WG D2.50 "Electric Power Utilities' Cyber-Security for Contingency Operations."



JAEROCK KWON (Senior Member, IEEE) received the B.S. and M.S. degrees from Hanyang University, Seoul, South Korea, in 1992 and 1994, respectively, and the Ph.D. degree in computer engineering from Texas A&M University, College Station, TX, USA, in 2009. From 1994 to 2004, he was with LG Electronics, SK Teletech, and Qualcomm Internet Service. From 2009 to 2010, he was a Professor with the Department of Electrical and Computer Engineering, Kettering University, Flint, MI, USA. Since 2010, he has been a Professor with the Department of Electrical and Computer Engineering, University of Michigan–Dearborn, Dearborn, MI, USA. His research interests include mobile robotics, autonomous vehicles, and artificial intelligence. His awards and honors, including the Outstanding Researcher Award, the Faculty Research Fellowship from Kettering University, and the SK Excellent Employee by SK Teletech. He served as the President for The Korean Computer Scientists and Engineers Association in America (KOCSEA) in 2020 and 2021.

JOHN MOORE (Member, IEEE) received the Bachelor of General Studies (B.G.S.) degree from the University of Michigan–Ann Arbor and the Master of Computer and Information Science (C.I.S.) degree from the University of Michigan–Dearborn. He works for Ford Motor Company as a Next Generation Architecture Vehicle Cybersecurity Technical Specialist for advanced engineering and research within product development. He holds a CISSP. His current work focuses on advanced computation, sensing, platform trust, cyber resiliency, and AI security architectures within the vehicle ecosystem.