A Review of Cockpit-driving Integration for Human-centric Autonomous Driving

Zhongpan Zhu 1, Shuaijie Zhao 2, Cheng Wang 2, and Aimin Du 2

1Tongji University
2Affiliation not available

November 28, 2023

Abstract

This review focuses on the integration of intelligent driving and intelligent cockpit systems. With the advancement of intelligent levels, autonomous vehicles are capable of driving in complex driving context. However, the lack of systems that consider both driver and driving scenarios information reduces the success probability of decision-making of autonomous vehicles in driving context where driver and driving scenarios information are tightly coupled during the human-vehicle cooperation stage. To solve this problem, we present the Cockpit-Driving Integration system (CDI) and review perception and decision-making algorithms for CDI systems. Additionally, to achieve human-centric autonomous vehicles, we propose that the CDI system should consider the personalized characteristics of drivers. Finally, we present a framework for CDI systems.
A Review of Cockpit-driving Integration for Human-centric Autonomous Driving

Zhongpan Zhu, Member, IEEE, Shuaijie Zhao, Cheng Wang, Member, IEEE, and Aimin Du

Abstract—Intelligent driving aims to handle dynamic driving tasks (DDTs) in complex environments, while driver behavior on-board is less focused. In contrast, an intelligent cockpit primarily concentrates on interacting with a driver, with limited connection to the driving scenarios. Since the driver on-board could affect the driving strategy significantly and thus have non-negligible safety implications on an autonomous vehicle, a cockpit-driving integration (CDI) is generally essential to take the driver’s behavior and intention into account when shaping the driving strategy. However, no comprehensive review of current existing CDI technologies is conducted despite the significant role of CDI in safe driving. Therefore, we are motivated to summarize the state-of-the-art of CDI methods and investigate the development trends of CDI. To this end, we identify thoroughly current applications of CDI for the perception and decision-making of autonomous vehicles, and highlight critical issues that urgently need to be addressed. Additionally, we propose a lifelong learning framework based on evolvable neural networks as solutions for future CDI. Finally, challenges and future work are discussed. The work provides useful insights for developers regarding designing safe and human-centric autonomous vehicles.

Index Terms—Cockpit-driving integration, autonomous vehicles, multiple functional perception, decision-making, human-vehicle interaction.

I. INTRODUCTION

AUTONOMOUS vehicles (AVs) are a hot research topic crossing multiple disciplines. They are considered key technologies to address significant issues in the transportation domain, such as traffic safety, congestion, and energy consumption [1]. Due to the recent intensive study driven by the automotive industry and software giants, AV technologies have rapidly advanced. As a result, the conventional driver-centric control is gradually transitioning to AV systems [2].

However, due to the complexity of traffic environments [3], the safety of expected AV functionalities [4], the maturity of legal regulations [5], ethical dilemmas [6], and other challenges, AVs are expected to remain at the L2 and L3 [7] for a long predictable time before advancing to higher levels of intelligence, such as L4 and L5, where a driver is completely out of the dynamic driving tasks (DDT). In particular, the long-tail challenge is a common sense problem that AVs are now facing [8]. To tackle this problem, a properly designed operational design domain (ODD) is necessary, which could facilitate the safety validation of AVs.

To navigate safely within a predefined ODD, intelligent driving plays a vital role. It focuses on achieving an accurate and comprehensive perception of surrounding vehicles, lane markings, traffic signs, pedestrians, and other elements within the driving scenarios [9] and autonomously accomplishing DDT by effectively utilizing the information obtained through perception [10] [11]. In addition to intelligent driving, intelligent cockpits also attract much attention. It enables diverse interactions with drivers through visual [12], auditory [13], and gesture [14] interactive technologies. Meanwhile, it emphasizes detection technologies for driver fatigue [15], distraction [16], emotions [17], and behaviors [18] to enhance drivers’ attention and response capabilities, ensuring driving safety. The different focuses of the two systems are summarized in Figure 1.

The distinct focuses of intelligent cockpits and intelligent driving result in their limited consideration for utilizing each other’s information. For instance, intelligent driving aims to enable drivers to engage in non-driving related tasks (NDRT), which could result in more significant driver attention diversion [19]. As a result, when faced with takeover scenarios, intelligent driving without considering the driver’s state could exacerbate driving safety [20]. Therefore, driver-state monitoring is essential for a safer transition process once an AV with L3 or below exits the ODD and a takeover request is issued. Such takeover request is fundamental for vehicles with L3 and below, where human-vehicle cooperation is usually the case. Human-vehicle cooperation means that a driver and an AV system have shared authority [21], enabling the vehicle to engage in driving tasks under specific conditions parallel to a driver’s control.

Even in the ODD, a driver may want to take back the control while an AV system is running. In this case, understanding the driver’s intention is beneficial for the system to adapt its driving strategy correspondingly [22]. Consequently, except for driving scenarios, a driver’s behavior and intention also shape the driving strategy. In fact, there is a wide range of driving contexts where the driver and driving scenario information are closely coupled [23]. Thus, driver behavior and intention become increasingly crucial for driving safety. In other words, the disconnection in perception between the
intelligent cockpit and intelligent driving leads to an increased probability of failure in human-vehicle cooperative decision-making [24]. As illustrated in Figure 1, the lack of simultaneous utilization of information obtained from an intelligent cockpit and intelligent driving poses a significant challenge for safe driving in contexts where driver and driving scenario information are tightly coupled.

Therefore, some studies integrate driver and driving scenario information to ensure driving safety [25] [26]. Leveraging driver behaviors, states, and driving scenario data, AVs can boost a driver’s situation awareness (SA). This keeps the driver engaged, minimizes lane deviation or collision warnings, and calibrates trust in AVs, fostering seamless control transitions between driver and vehicle. The human-centric integration of driver and driving scenario information for perception and decision-making in AVs effectively address various challenges in the human-vehicle cooperation stage, highlighting the necessity of the development of cockpit-driving integration (CDI).

Although driver behavior perception and driving scenario perception and decision-making are hot topics in the fields of the intelligent cockpit and intelligent driving, and relevant research is relatively mature [27] [28] [29] [30], the numerous bottlenecks faced by CDI have led to a current lack of comprehensive review articles guiding the development of CDI systems. Li et al. [31] have surveyed the research of AVs from the perspectives of driver perception and AVs, proposing that future AVs should consider both the driver and driving scenarios to make appropriate decisions. However, the paper mentions relevant concepts without systematically proposing how to develop the CDI system.

Therefore, we are motivated to provide a detailed and comprehensive review of the currently available perception and decision-making techniques in terms of their application for CDI by considering the following three research questions:

- **RQ1**: What are the considerations to build multi-modal fusion perception algorithms for CDI?
- **RQ2**: How to address the consistency of driver behavior and intention in different spatial and temporal driving contexts?
- **RQ3**: How to deal with driving process uncertainty in CDI?

To the best of the authors’ knowledge, there is a lack of a comprehensive review focuses on simultaneously utilizing driving scenario and driver information for perception and decision-making. After addressing the three questions above, we have ultimately proposed a framework for perception and decision-making in CDI based on dynamically expandable neural networks. We also discuss the current gaps and future research directions. Therefore, the contributions of this paper are as follows:

- three key issues that need to be addressed in CDI have been summarized through literature research and analysis to provide guidance for the development of human-centric AV systems.
- current perception algorithms for CDI were thoroughly investigated, and it was analyzed that the existing perception algorithms lack the ability to perceive the personalized characteristics of drivers.
- an investigation was carried out on the current research of decision-making algorithms for CDI, and it was proposed that CDI decision systems should be capable of making personalized decisions and handling uncertainty in the driving process.
- we proposed a CDI framework based on lifelong learning, incorporating dynamically expandable neural networks and a memory pool.

Section II focuses on the current research and development trends of CDI, analyzing the critical issues faced by CDI. Based on these key issues, Section III discusses state-of-the-art (SOTA) algorithms for achieving CDI perception systems. Section IV systematically discusses the technical approaches for CDI decision-making systems. Subsequently, Section V proposes a CDI framework based on lifelong learning with dynamically expandable neural networks. Drawing on the aforementioned analysis, Section VI conducts a thorough discussion and highlights the limitations of existing research. Finally, the entire paper is summarized, and future work directions are identified in Section VII.

**II. BACKGROUND AND SCOPE**

In this section, we first introduce the current status of CDI perception and decision-making studies and then summarize the key issues for CDI. Based on this, we define the scope of our survey for a comprehensive investigation of CDI algorithms.

**A. Background**

To complete DDTs within an ODD, previous studies primarily focused on sensing external factors and environmental conditions in a driving scenario to understand situations and make corresponding decisions [32] [33] [34]. For instance, Ko et al. [35] introduced a lane marking detection approach leveraging key points within driving scenarios, aiming to facilitate localization of the drivable road area. Similarly, Park et al. [36] employ historical trajectory data from surrounding vehicles to generate the future trajectory sequence of surrounding vehicles, which enables AVs to effectively plan driving routes.

These studies are valuable and contribute significantly to the autonomy of vehicles. However, due to the increasing interactions among road users and uncertain external environments, human-vehicle cooperation is considered a measure to leverage the advantages of autonomy while ensuring an acceptable level of risk [37]. This necessitates a human-centric CDI system. To this end, various factors, including driver behaviors and states, driving scenarios, vehicle states, and their close interconnections, are needed to be considered in the decision-making process of AVs.

Motivated by this, researchers have explored CDI systems from various perspectives, primarily including CDI perception and CDI decision-making. Regarding CDI perception, Rong et al. [38] predicted the driver’s lane-changing intentions using the correlation between driving scenarios and the driver’s head
and eye information. Gao et al. \[39\] discovered a correlation between the driver’s ECG signals and the driving scenarios, incorporating these features to predict the driver’s lane-changing intentions. Compared to using information solely from driving scenarios, lane-changing intention prediction based on CDI perception can provide more accurate and earlier predictions \[40\]. In addition to driver behavior prediction, Mou et al. \[41\] considered the various types of driver distractions for CDI perception by proposing a driver state detection method. Obviously, driver behavior prediction and driver state monitoring are important considerations for CDI perception.

Based on the acquired driver information, CDI decision-making aims to adjust its driving policy for safe and human-centric navigation. For instance, Dahl et al. \[42\] proposed lane-keeping assistance, which assesses whether to initiate lane-keeping based on the driver’s visual gaze and the vehicle’s distance to the lane markings. Another example is that Kim et al. \[43\] integrated driver-specific parameters, such as reaction time and heart rate, into the decision-making process to determine an appropriate takeover request time. The findings suggest that accounting for driver-specific factors can significantly mitigate the driver’s reaction time during a takeover. These studies indicate that for effective CDI decision-making, integrating perceived driver information is crucial. This integration enhances an AV’s capability to manage critical situations in the driver’s absence and better address the driver’s needs when they intervene in driving.

As a result, driver information plays a crucial role in CDI systems. Moreover, different drivers exhibit individual physiological characteristics that influence the detecting accuracy of driver states and behaviors \[44\]. Furthermore, psychological characteristics of drivers, such as driving styles, impact driving safety \[45\]. Therefore, a CDI perception system needs to be capable of personalizing driver information for CDI decision-making. This facilitates a CDI decision-making system generating adaptive decisions based on a driver’s personalized characteristics to ensure driving safety and improve collaboration between humans and AVs, since personalized decision-making for AVs can enhance drivers’ trust in AVs by reducing conflicts between driver preferences and AVs \[46\]. Additionally, during the driving process, driver behaviors can exhibit large randomness. Wagner et al. \[47\] indicated that in car-following behavior, the fluctuations in time headway caused by internal driver stochasticity could be as large as the mean time headway itself. To this end, by capturing and modeling the stochastic nature of driver behaviors, the CDI decision-making system can anticipate and respond to unexpected variations, ensuring robust decision-making in dynamic driving scenarios.

In summary, CDI perception and decision-making are two fundamental elements in building human-centric AVs, where driver information plays a significant role. Due to the fact that drivers exhibit different driving characteristics, driver personality detection is a prerequisite to generating driver-dependent decisions using a CDI decision-making system, with the goal of increasing safety and drivers’ trust in AVs. Despite the significance of CDI perception and decision-making systems, few surveys have discussed the essential considerations for developing these systems. In particular, little research has summarized and analyzed comprehensively the impact of drivers’ personalized characteristics and behavioral uncertainties on CDI systems. Moreover, a crucial gap in the existing literature is the lack of a discussion of suitable frameworks for CDI systems. Our paper bridges this gap by elaborating on the important aspects of developing CDI perception and decision-making systems, and proposing a promising framework.

**B. Scope**

Based on the above analysis, we define the review scope of the paper, as illustrated in Figure 2. First, we investigated perception algorithms complying with the CDI concept. These algorithms include driver behavior prediction, state detection, and takeover identification. Meanwhile, personalized driver state detection and style recognition methods are included to address driver personalization issues.

In addition, based on previous research, we examined algorithms relevant to CDI decision-making systems, which
can be classified as lane-keeping, collision avoidance, and SA enhancement. Following that, we introduce personalized decision-making to accommodate differences existing among various driver behaviors. Finally, we introduce approaches to address internal stochasticity within the same driver during driving.

III. CDI PERCEPTION SYSTEM

This section provides a detailed discussion of CDI perception systems, including the algorithms used in CDI perception and the impact of a driver’s personalized characteristics on a CDI perception system. Based on this investigation, we aim to determine the considerations for building CDI perception systems to answer the RQ1.

A. Perception algorithms for CDI

The multi-modal perception algorithm is a critical prerequisite for constructing a CDI perception system. It simultaneously perceives driver and driving scenario information, integrates various types of information for prediction, and provides robust support for decision-making systems. Driver behavior prediction, driver state detection, and takeover identification are three tasks where CDI perception algorithms are usually applied. They are introduced in detail in the following.

1) Driver behavior prediction: Drivers typically behave differently based on driving scenarios, vehicle states, and personal factors. Based on the CDI concept, behavior prediction integrates these three factors to anticipate driver behaviors such as lane-changing and speeding in advance. Predicting driving behaviors such as lane-changing and turning is often based on driver information or achieved by modeling driving scenarios. Xu et al. [75] have integrated vehicle states with driving scenarios to predict driver behaviors. However, the vehicle states only implicitly represent the driver’s operational actions without directly encompassing the driver’s states or behaviors. Unlike conventional approaches, lane-changing and turning behavior prediction based on CDI emphasizes the simultaneous inclusion of driver and driving scenario information. Trivedi et al. [76] proposed the Environment-Vehicle-Driver framework, which predicts a driver’s lane-changing and other behavioral intentions using Sparse Bayesian Learning (SBL) by incorporating information from driver head-motion, lane position, and vehicle states.

Time-series information is widely utilized to improve prediction accuracy by considering contextual information. Doshi et al. [77] considered a driver’s head movements, visual attention, road curvature, lateral distance to lane markings, and vehicle state information within a 1s time window. They concatenated the acquired information into a vector and utilized the Relevance Vector Machine (RVM) to infer a driver’s lane-changing intentions. The results indicate that the accuracy of predicting lane-changing 2s in advance is 87.30%. Their classification model is expressed by:

$$ y(x) = \sum_{i=1}^{M} w_i K(x, x_i) $$

where $x$ refers to the RVM input vector based on driver, driving scenario, and vehicle state information. $w_i$ is the weight learned by the model. $K$ is the kernel function. $y(x)$
indicates the behavior category to which the input vector $x$ belongs. $M$ is the number of categories. For each time step $t$, the feature vector for each example $x$ can be represented as:

$$
x(t) = [\text{LateralPos}(t), \ldots, \text{LateralPos}(t - N + 1), \text{SteeringAngle}(t), \ldots, \text{SteeringAngle}(t - N + 1), \text{EyeHistogram}(1), \ldots, \text{EyeHistogram}(9), \text{etc.}]
$$

where $N$ represents the number of past values of each variable that have internally been stored.

Similarly, Morris et al. [48] proposed combining the time series information of a driver’s head pose, surrounding vehicle movements, and vehicle states within a 2s time window into a feature vector input for the RVM classifier to predict lane-changing behaviors. In addition to the simultaneous acquisition of driver and driving scenario information, the three-layer AIO-HMM network proposed by Jian et al. [49] models the impact of driving scenarios on a driver by placing driving scenario information and driver information in different network layers. It utilizes a driver’s head pose, lane information, vehicle speed and GPS to output the probability distribution of driver behaviors.

Although CDI perception algorithms are presented with rich information, shallow networks such as RVM, Random Forest, and HMM have limited feature expression capabilities [78]. As a result, they can only handle time series within a restricted time window and cannot fully leverage the advantages of rich information in CDI perception. In comparison, neural networks can extract features efficiently, leading to a better representation. Particularly, neural networks, such as Long Short-Term Memory (LSTM) networks [79] [80] and Convolutional Neural Networks (CNN) with resistance to forgetting [81], are utilized to address these challenges. Ou et al. [80] proposed a lane-changing prediction model based on deep recurrent neural networks. They integrated the facial movements of a driver, lane markings, and vehicle speed to infer the driver’s steering and lane-changing intentions. The results showed that the model achieved an accuracy of 90.52%. Similarly, both Xing et al. [82] and Jain et al. [83] proposed driver behavior prediction models based on Recurrent Neural Networks (RNN) with LSTM units. Jain et al. [83] utilized a driver’s head position, road conditions, GPS, and vehicle dynamics to infer a driver’s driving intentions. The model achieved a lead time of 3.16 s for predicting the driver’s lane-changing and turning behaviors, with a combined prediction accuracy of 90.5%.

In addition to lane-changing prediction, Yu et al. [84] used driving scenario information, driver’s age, gender, driving experience, and vehicle states as inputs to the Random Forest for predicting a driver’s speeding behavior. The study highlighted that, apart from driving scenarios, driver characteristics such as age and driving experience also significantly affect the performance of speeding prediction.

2) Driver state detection: Many studies on driver state detection have focused solely on driver-related factors [85] [86]. Driver state detection based on CDI simultaneously considers the factors of drivers, driving scenarios, and vehicle states, along with their complex interactions, which enables AVs to understand the correlation between driver states and driving scenarios. Ingre et al. [87] investigated the correlation between blink duration, deviation of the lateral position, and driver fatigue levels. The results suggest that both blink duration and deviation of the lateral position are similarly correlated with driver fatigue, indicating that both the driver and driving scenario information are crucial for driver state prediction.

In addition to fatigue detection, information as such as driver’s head pose, visual gaze, lane markings, and surrounding vehicle information are also commonly used for distraction detection based on CDI. Rezaei et al. [88] utilized a fuzzy fusion algorithm to combine a driver’s head pose with the distance and angle of the detected vehicles relative to the ego vehicle to determine the level of driving danger caused by driver distraction. However, traditional methods, such as fuzzy logic, have a poor ability to learn the optimal amount of contextual information compared to LSTM. Wollmer et al. [51] combined the driver’s head pose with the vehicle’s longitudinal offset and heading in the lane, the steering wheel angle, and speed. The above information and their first and second derivatives were concatenated into a 3x6 matrix and inputted into an LSTM network to determine the driver’s distraction level. The results indicate a prediction accuracy of 96.6% for distraction.

In addition to head pose and visual gaze, physiological signals also serve as a good indicator of the driver’s state. Guo et al. [52] combined electroencephalogram (EEG) signals with road conditions in driving scenarios to assess the driver’s vigilance level.

3) Takeover identification: Currently, it is still a huge challenge for AV systems to navigate safely in complex driving contexts [89]. Therefore, we will remain in the human-vehicle cooperation stage for a considerable time [90]. The stage presents an important safety issue known as the takeover problem [91]. On the one hand, conditional automation allows drivers to engage in NDRT [92], leading to distraction and fatigue, reducing the driver’s SA and worsening takeover performance [93] [94]. On the other hand, when AVs encounter complex driving contexts and exit their ODD, the driver needs to take over control rapidly [95]. The driver must understand the driving scenarios, enter the driving loop, and regain vehicle control. The takeover process affects the vehicle’s speed and trajectory, thus directly impacting driving safety [96].

Many studies utilize driver information, such as eye-tracking data [97] [98] [99], heart rate [100], and driving scenarios [101] to estimate SA or takeover performance. The advantage of CDI in addressing takeover issues lies in the combination of driver and driving scenario information. In order to classify a driver’s takeover readiness, Braunagel et al. [102] proposed a method that considers the driver’s ongoing NDRT, visual attention to the road, traffic flow density, and weather conditions. However, in this method, NDRT, traffic density, and weather are represented as simple numbers, meaning some driver and driving scenario features are missing. To improve the SA estimation system in terms of considering the surrounding situation and treating glance features as combinations, Hayashi et al. [53] proposed a SA prediction model that combines drivers’ visual attention area and the information about ve-
hicles and obstacles in the driving scenarios. Then, a SVM is employed to classify the quality of SA based on these inputs. The results of simulator experiments achieved a prediction accuracy of 83%. Du et al. [54] used a random forest classifier to predict the quality of driver takeover performance based on driver physiological data, including heart rate, electrodermal activity, visual attention, driving scenario types, and traffic flow density. The results achieved an accuracy of 84.3% within a 3s prediction time window.

There is also research on predicting driver takeover intentions by integrating driver and driving scenario information. In [103], a 3D convolution was used to extract driver gaze heatmaps and driving scenario semantic segmentation images. These images were combined with driver physiological signals and vehicle states to predict driver takeover intentions.

Table 1 summarizes the current CDI perception algorithms, including their advantages and disadvantages and the algorithms used. Although various prediction models are applied to different tasks, we find that neural networks based on time-series inputs are more frequently used than other methods in recent CDI perception algorithms. In terms of information collection, driver’s head pose, visual gaze, and lane markings are commonly collected.

Fig. 3. The overall pipeline of CDI perception algorithms. The data collection module utilizes sensors to gather information from drivers, driving scenarios, and vehicle states. The information is then used to extract features, which are inputted into prediction networks to predict driving behaviors, driver states, takeover tasks, and more.

Based on the above analysis, we deduce the general pipeline of CDI perception algorithms, as depicted in Figure 5. The CDI perception algorithms can be divided into data collection, feature extraction, and prediction. Firstly, it is necessary to collect appropriate data from drivers, driving scenarios, and vehicle states based on the purpose of the perception task. Subsequently, data features are automatically extracted using neural networks or manually extracted to form a feature set, which is then used as input for the prediction network.

B. Perception with personalization

The driver plays a crucial role in the human-vehicle-road loop during the human-vehicle cooperation stage. However, different drivers possess various characteristics, such as physiological features, driving preferences, and driving styles. The presence of personalized features presents challenges to driving safety. For example, compared to normal driving styles, aggressive driving is more prone to causing accidents [104]. Current CDI perception approaches typically rely on the concept of an average driver, emphasizing common factors and neglecting the significant role of personalized factors. The lack of consideration for personalized factors renders the driver the weakest link in the human-vehicle-road loop [105].

In examining the relationship between blink duration, deviation of the lateral position, and driver fatigue levels, the study [87] also found significant differences in blink duration among different fatigue levels for various drivers. Therefore, when utilizing driver information, CDI perception systems need to consider the personalized characteristics of drivers. Personalized driver perception typically includes driver state detection and driving style recognition [106]. Personalized driver state detection aims to improve detection accuracy by exploiting the personalized characteristics of drivers. It can be used to improve the detection accuracy of driver information, hence improving the system’s overall detection accuracy. Additionally, driving style recognition can be used to enhance the personalized capabilities of CDI decision-making systems.

1) Driver state detection: Personalized driver state detection improves detection accuracy by identifying differences in physiological signals such as EEG, galvanic skin response (GSR), photoplethysmography (PPG), and external features such as facial characteristics between different drivers. Typically, physiological signals are used to monitor the driver’s vital signs [107] and fatigue level [108]. However, physiological signals from different drivers exhibit variations [109], and the absence of data from a specific driver in the training set may lead to reduced detection accuracy. Thus, it is necessary to obtain labeled driver-specific feature data to train the detection network.

Online SVM [55] utilizes two SVM classifiers. SVM1, trained on a general dataset, is used for initial driver fatigue detection, while SVM2 is trained using feedback data generated by personalized drivers. When the accuracy of SVM2 exceeds that of SVM1, SVM1 is replaced. The method increased the fatigue detection accuracy from 72.05% with the general model to 95.66% for 28 experimental subjects. However, obtaining labeled personalized training data is challenging, and finding an appropriate method to collect personalized driver data poses significant difficulties.

Unlike Online SVM, Choi et al. [56] select a subset of data with the current decision boundary and train a new decision boundary by combining it with new data. The model trained on a general dataset is combined with the personalized model trained on a small amount of personalized data to improve the personalization of the model, enhancing the influence of the personalized model on the decision boundary and ensuring the capability of the general model where the personalized model cannot provide helpful decision boundary. Ye et al. [110] proposed an automatic adaptation model for subjects, where high-confidence unlabeled data from personalized heart rate...
data is used to train the personalized model. The general model trained on the general dataset combines the personalized model to predict driver states. The information from the general and personalized models complements each other, providing personalized support for different subjects and improving the overall detection accuracy.

Additionally, a driver’s external characteristics, such as appearance, exhibit personalized features and require personalized perception. An unsupervised classification algorithm for driver workload detection was proposed in [111]. It utilizes fuzzy C-means (FCM) clustering of a vehicle’s speed and three-axis acceleration and then employs SVM to classify the clustered data and output prediction results. Ming et al. [112] introduced a personalized driver fatigue detection algorithm based on facial features. They proposed a calibration-free perception scheme that adaptively adjusts the threshold for eye closure based on individual drivers, effectively resolving prediction errors caused by small eyes during fatigue recognition.

2) Driving style recognition: Different from personalized driver state detection, driving style recognition emphasizes the identification of characteristics and preferences of the driver in terms of driving behavior, which can support for personalized decision-making, assist in correcting poor driving habits, and promote driving safety.

However, driving style is influenced by various factors, such as the driver’s personality, gender, driving behavior, and environmental characteristics, like weather, driving scenarios, and time of day. Different drivers exhibit distinct driving styles under similar conditions; even the same driver may
demonstrate different driving styles at different times and locations \[113\]. Therefore, a CDI perception system should be capable of perceiving different factors to determine driving style, meeting the drivers’ requirements for future AVs regarding safety and personalization.

The classical method for driving style recognition is based on a predefined threshold and determines the driving style by vehicle states. These methods are also referred to as rule-based (RB) approaches. Murphey et al. \[114\] used the jerk profile, which calculates the jerk ratio \( r \), to determine the driving style. It is categorized as calm driving: \( \gamma < 0.5 \), normal driving: \( 0.5 < \gamma < 1 \), aggressive driving: \( 1 < \gamma \), and others. Although this method is simple and practical, the ambiguous boundaries between different driving styles, such as the boundary between normal and calm driving, may result in lower classification accuracy.

Fuzzy Logic (FL) classifies driving style based on the degree of membership, considering multiple influencing factors and producing more accurate classification results. Guo et al. \[115\] proposed an FL algorithm that utilizes two factors, namely, accelerator pedal position and its rate of change, to classify drivers’ driving styles. Dorr et al. \[116\] selected different combinations from six parameters to determine driving styles based on three road types: urban, rural, and highway. Through simulations of rural and urban driving scenarios, the algorithm achieved a 68% probability of correctly classifying the three driving styles. The RB and FL methods classify driving styles by setting thresholds. However, classification accuracy is closely related to the setting of threshold intervals. Therefore, achieving good classification results requires expert knowledge to determine reasonable threshold intervals. Additionally, the established rules have poor generalization ability when facing new driving scenarios.

Machine learning algorithms can effectively address the issues above by automatically inferring thresholds for various driving styles using a large amount of driving data. The generalization ability of the models can be improved by acquiring data from different driving scenarios. K-means clustering, a commonly used algorithm for driving style recognition without requiring data labels, classifies data by finding those with similar features. Mohammadnazar et al. \[117\] proposed a k-means clustering method for the three-class classification of driving styles, using vehicle speed, lateral, and longitudinal accelerations as input data. They employed three volatility functions to extract six volatility measures for the k-means clustering of driving styles.

Additionally, Wang et al. \[118\] combined K-means clustering with Bayesian networks. They first used Bayesian non-parametric methods to learn the original driving patterns from driving data. Then, they employed K-means clustering to classify driving styles based on the original driving patterns. KNN is another simple yet effective algorithm for driving style recognition. It clusters driving data into K classes based on given driving behavior features and then determines the driving style. In \[57\], seven driving behavior features were used, and data collected from tri-axial accelerometers were used as input for driving style classification based on the KNN algorithm. Experimental data from 110 road sections verified that the accuracy of classifying aggressive and normal driving styles reached 100%. Furthermore, random forests \[119\] and Bayesian networks \[120\] have also been applied in driving style recognition.

In methods such as k-means, KNN, and Bayesian networks, it is necessary to select driving features and classify driving styles based on them. However, manually selecting effective features, especially implicit features, can be challenging. Therefore, CNN’s powerful feature extraction capability has gradually been applied in driving style recognition. Bejani et al. \[58\] employed a CNN to extract explicit and implicit features from smartphone accelerometers, and the result shows that the algorithm achieved a driving style detection accuracy of 95%. A CNN-LSTM driving style recognition network considering temporal information was proposed in \[121\]. The information was first input into the CNN network to extract style features. Then, the LSTM network was utilized to classify driving styles.

Based on the above study, drivers have various physiological characteristics. However, current CDI algorithms utilize large datasets to train perception networks, which may easily capture generic characteristics among different drivers while overlooking the personalized characteristics of a specific driver. The generic prediction network may exhibit limited performance in detecting a specific driver when applying CDI perception systems. As a result, when building CDI perception systems, the impact of driver physiological characteristics such as EEG, ECG, GSR, and eyes-on detection accuracy should be carefully considered. These features are commonly used in CDI perception systems and are most likely to generate personalized characteristics among drivers. Additionally, driving style is frequently used to generate personalized decision-making. Taking the driver’s driving style into account in CDI perception systems allows CDI decision-making systems to make personalized decision-making.

IV. CDI decision-making system

A CDI decision-making system incorporates driving scenarios and driver information into the decision-making loop to govern a vehicle’s motion or assist a driver autonomously. However, heterogeneity is an intrinsic characteristic of drivers. Different drivers’ varying responses to the same stimulus refer to inter-heterogeneity, and the internal stochasticity in the same driver during the driving process is called intra-heterogeneity \[122\]. Heterogeneity makes accurately describing driver behaviors and states challenging, potentially leading to CDI decision-making failures. Therefore, the inter-heterogeneity and intra-heterogeneity issues emphasized in RQ2 and RQ3 are critical considerations in our proposed CDI decision-making systems.

To introduce the CDI decision-making system, we first review the algorithms that can be used for CDI decision-making. To analyze the influence of the inter-heterogeneity of drivers on decision-making systems, we also discuss personalized driver models of car-following, lane-changing, and braking. Furthermore, in the face of stochasticity in driver behaviors, we introduce stochastic models and data-driven approaches.
A. Decision-making algorithms for CDI

CDI decision-making algorithms aim to integrate driver information into the decision-making loop, i.e., comprehensively utilizing driver and driving scenario information. The rich information enables the CDI system to handle driving contexts where the driver and driving scenarios information are tightly coupled, which cannot be effectively handled solely by intelligent cockpits or intelligent driving systems. We reviewed the existing algorithms for CDI decision-making, which can be classified into lane-keeping, collision avoidance, and SA enhancement.

1) Lane keeping: The CDI decision-making system integrates a driver’s state and lane information to generate lane departure warnings or perform corrective steering operations. It aims to intervene in lane deviations and ensure driving safety with minimal driver disturbance. A stochastic model predictive control (SMPC) algorithm for lane keeping was proposed in [123]. This algorithm combines the predicted positions of nearby vehicles and the driver’s steering behavior. The SMPC algorithm acquires driving control to correct the driver’s driving behavior when a possible collision arises.

However, the method above does not consider a driver’s state information. During normal driving, the driver’s visual gaze direction is closely related to a road’s curvature, as the driver’s eyes follow the vehicle’s motion [124]. Therefore, visual information serves as a good indicator of lane departure. A typical approach is utilizing the driver’s visual and lane information to detect unconscious lane deviations. Pohl et al. [59] determine departure interventions based on the driver’s visual distraction and the lane departure. When the rule-based (RB) lane intervention module detects lane departure and the driver is distracted, an electric power-assisted steering gear provides an additional torque offset to intervene in lane departure. A driver’s state decision equation determines whether the driver is distracted:

\[ S(t) = \begin{cases} 1 & D_N(t) > \tau(t) \\ 0 & D_N(t) \leq \tau(t) \end{cases} \]  

(3)

where \( S(t) \) is the driver distraction state; 1 refers to distracted; 0 refers to not distracted; \( D_N(t) \) is driver distraction level; \( \tau(t) \) is adaptive distraction threshold.

Another RB lane departure intervention method proposed in [60] utilizes the driver’s pupil shape, lane markings, distance to the preceding vehicle, and humidity information to determine the driver’s distracted state. Thresholds are defined for different types of information, and the interval values are transformed into linguistic information through linguistic membership functions to determine whether lane departure intervention is necessary.

2) Collision avoidance: Collision avoidance systems detect drivers’ states and driving scenarios to infer safety risks, issue warnings, or activate Automatic Emergency Braking (AEB). A multi-agent collision warning ontology framework, considering the driver, external vehicles, and driving environment, is proposed in [125]. The framework incorporates multi-agent data and establishes a collision warning rule library to assess the danger of driving scenarios and issue collision warnings.

Another multi-agent collision warning ontology framework is presented in [126], which further includes the behavior information of the subject vehicle and sets up a rule-based warning system based on five common hazardous scenarios.

Additionally, Wang et al. [127] proposed a driving safety field theory, which includes a potential field considering road conditions, a behavior field considering personalized driver features, and a kinetic field considering vehicles and pedestrians. Compared to other driver safety models, this model can integrate various information and is suitable for addressing highly complex driver safety issues. Based on this theory, the authors also provided a collision warning system for car-following and merging scenarios [128].

RB decision methods are relatively simple to implement, but the main issue is their limited generality. Rules designed for specific scenarios often lead to decision conflicts or errors when facing new situations. Machine learning methods can adapt to new environments through generalization. Additionally, nonlinear models exhibit higher prediction accuracy compared to RB methods. Armand et al. [61] utilize Bayesian networks to integrate driver intention, driving scenarios, and vehicle states to infer the level of danger at intersections and determine the type of assistance required. After validating their approach with a dataset including 260 runs, the AEB function achieved 100% correct activations, the reminder and warning function had an accuracy of 84%, and the suggestion function had a trigger accuracy of 82% with a 4% false positive rate.

In addition to the aforementioned specific function warning systems, Choi et al. [62] have proposed multi-functional warning systems, including collision warning, lane departure warning, distraction alert, and advanced lane departure alert, based on active appearance model (AAM), multilayer perceptron (MLP), and SVM.

3) Situation awareness enhancement: The CDI decision-making system combines a driver’s gaze behavior and external driving scenario information to remind the driver of important information in driving scenario information and enhance the driver’s SA. Fletcher et al. [63] and Petersso et al. [129] have proposed systems for alerting drivers about missed traffic signs. The system utilizes driver information, including head position and visual gaze direction, vehicle state information, including speed, acceleration, GPS, steering wheel angle, and driving scenario information to alert the driver. The decision-making process involves (1) detecting whether a driver’s visual direction is within the area of the speed limit sign and (2) checking if a vehicle’s status meets the requirements of the speed limit sign. If the driver’s visual direction is not within the area of the speed limit sign and the vehicle’s status does not meet the requirements, the alert system issues a warning. The alert system issues a warning if the driver’s visual direction or the vehicle’s status does not comply. Otherwise, the alert system notifies the driver passively.

Table III summarizes the aforementioned CDI decision-making algorithms, including information and the algorithms used. We found that the decision-making algorithms are mainly RB. These algorithms allow for a clear definition and interpretation of each rule. However, they have limitations when faced with complex problems. The number and complex-
ity of rules may increase with the complexity of the problem, and RB systems built on known data and prior knowledge find it difficult to make accurate decisions in unknown situations.

Based on our research, the overall pipeline of CDI decision-making algorithms is depicted in Figure 4. The CDI decision-making algorithms can be divided into information selection, decision generation, and decision action. Firstly, decision inputs are derived from perception results or raw data from the driver, driving scenarios, and vehicle state. Subsequently, the emphasis is on generating appropriate decisions by utilizing rich information. Finally, particular assistance types are executed based on the condition of the driver and driving scenarios.

### Table II

<table>
<thead>
<tr>
<th>Objective</th>
<th>Ref.</th>
<th>Considerations</th>
<th>Method</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lane-keeping</td>
<td>125</td>
<td>Steering behaviors, surrounding vehicles.</td>
<td>SMPC</td>
</tr>
<tr>
<td></td>
<td>59</td>
<td>Driver head pose, road curvature, lane position, vehicle speed.</td>
<td>Rule-based</td>
</tr>
<tr>
<td></td>
<td>60</td>
<td>Driver head pose, eye information, steering behaviors, vehicle lateral position, lane markings, relative vehicle’s speed, traffic flow, yaw rate.</td>
<td>HSS</td>
</tr>
<tr>
<td>Collision Avoidance</td>
<td>123</td>
<td>Driver head pose, eye information, pedestrians, surrounding vehicles.</td>
<td>Rule-based</td>
</tr>
<tr>
<td></td>
<td>126</td>
<td>Eye information, driver state, pedestrians, surrounding vehicles, vehicle state</td>
<td>Rule-based</td>
</tr>
<tr>
<td></td>
<td>61</td>
<td>Driver behaviors, vehicle speed, pose, brake pedal, throttle pedal.</td>
<td>Bayesian Network work</td>
</tr>
<tr>
<td></td>
<td>62</td>
<td>Eye gaze, facial features, lane markings, surrounding vehicles.</td>
<td>Rule-based</td>
</tr>
<tr>
<td>SA Enhancement</td>
<td>63</td>
<td>Eye gaze, head pose, road signs, vehicle’s speed, acceleration.</td>
<td>Rule-based</td>
</tr>
<tr>
<td></td>
<td>129</td>
<td>Eye information, head pose, road signs, vehicle’s position, speed, GPS.</td>
<td>Rule-based</td>
</tr>
</tbody>
</table>

#### B. Personalized decision-making algorithms

Regarding RQ2, inter-heterogeneity can be reflected in the variations of driver behavior intentions across different spatial and temporal driving contexts. In order to address inter-heterogeneity, it is crucial to analyze individual driver characteristics such as driving styles and driving habits [123]. Therefore, the following section presents car-following, lane-changing, and braking models that consider personalized driver characteristics. The CDI decision-making system achieves personalized decisions by utilizing personalized driver models [130], thereby reducing conflicts during the human-vehicle cooperation phase.

1) Personalized car-following: Car-following (CF) behavior refers to a driver maintaining its current lane and interacting with preceding and following vehicles. In the current stage of human-vehicle cooperation, an important goal is to adapt AVs to the personalized characteristics of drivers. Therefore, establishing a personalized CF model is beneficial for simulating and predicting driver behaviors in CDI decision-making systems, enabling personalized CF behavior and ensuring driving safety. Existing CF models can generally be categorized into two types: mathematical models and data-driven models.

Mathematical models are the earliest models developed to describe driver CF behavior. Personalized mathematical models typically employ a set of equations that incorporate individual driver characteristics. However, the parameters of these models are predefined. It is necessary to adjust these parameters based on different drivers to address inter-heterogeneity. For example, Andersen et al. [64] proposed the Driving-by-Visual-Angle (DVA) CF model, which uses the driver’s visual angle and its rate of change as variables for predicting driver acceleration. To achieve better accuracy, the model requires parameter calibration for specific drivers.

In [65], expected speed is an essential input parameter for driver CF models, reflecting the driver’s anticipated speed for vehicle trajectories. Additionally, Pariota et al. [131] proposed a linear dynamic CF model that utilizes the driver’s anticipated equilibrium spacing and the velocity of the leading vehicle as input parameters to the model’s state space, generating the vehicle’s speed and acceleration. However, these methods are based on pre-defined mathematical formulas, which lack flexibility and make it difficult to adapt to different driving data, such as varying time headways, acceleration rates, and comfort levels caused by various driving styles [132]. It is challenging to address these issues using mathematical models alone, thus requiring alternative methods to tackle the abovementioned problems.

Data-driven models utilize a large volume of driving data to establish personalized driver car-following models by leveraging reinforcement learning and CNN algorithms. These models identify the complex relationship between personalized driver features and driving scenarios. Qin et al. [133] proposed a CNN-LSTM CF model, where driving data with temporal information is inputted into a CNN to extract feature vectors among vehicle parameters. Then, the LSTM is used to predict...
the following speed. The model accurately generates various
driving behaviors and exhibits excellent generalization ability
to adapt to different driving styles.

Unlike neural network models that require labeled data as input, unsupervised learning models such as HMM, Gaussian Mixture Model (GMM), and reinforcement learning can learn without supervision. Wang et al. [134] proposed the bounded generalized GMM-HMM (BGGMM-HMM) personalized CF model, where BGGM extracts personalized driving behavior feature parameters from driving data. The HMM model utilizes these feature parameters to predict driving behaviors based on the state transition probability distribution in different driving scenarios. The reinforcement learning method learns the CF strategy through interactions between drivers and the driving scenarios, optimizing the model using feedback from the reward function. Zhu et al. [135] introduced DDPGvRT, a deep reinforcement learning CF model capable of generating personalized driver behaviors and driving trajectories. It leverages the difference between simulated and observed speed, considering 1 s delay response as the reward function, and learns through trial-and-error interactions using historical driving data.

However, considering transient signals as input to the reinforcement learning model neglects the temporal correlation of motion information. Liao et al. [136] proposed a personalized CF model called LSTM-TD3 that considers driving style to achieve human-like driving memory. Including LSTM memory units allows TD3 to consider the current input parameter values and incorporate historical data. The model establishes a reward function based on different driving styles. Through NGSIM dataset validation, this model exhibits better convergence compared to other reinforcement learning methods and can generate different driving trajectories based on safety, comfort, and other requirements of different driving styles. Unlike traditional mathematical methods that estimate parameters through fitting data, reinforcement learning models learn decision-making mechanisms from training data, allowing them to achieve better generalization capabilities. Moreover, the model can adapt to drivers with different driving styles by incorporating new data to address the issue of heterogeneity.

To leverage the advantages of mathematical and data-driven models, hybrid models have emerged, which utilize driving behavior data generated by mathematical models to improve the training process of data-driven models. Zhang et al. [69] proposed a hybrid model combining the Intelligent Driver Model (IDM) and Neural Processes (NP). Using real driving data, the IDM model with time-varying parameters calibrates the inconsistencies between learned driver behaviors. Simultaneously, the NP-based CF model can generate driving behaviors under specific driving styles. Combining the IDM model with time-varying parameters and the NP-based CF model, the hybrid model considers driver heterogeneity and randomness, simulates observed CF behavior under a specific driving style, and generates CF behavior under unobserved driving styles.

2) Personalized lane-changing: Lane-changing (LC) behavior refers to a vehicle’s voluntary or forced maneuver from its current lane to another lane, describing the lateral interaction between vehicles. Previous research has paid lit-

tle consideration to personalized driver characteristics [137]. Personalized LC models can generally be classified into RB models and learning-based models.

RB personalized LC models consider driving rules, driving experience, traffic regulations, and driver characteristics to establish a decision rule library for LC decisions. An approach considering the influence of different driving styles on LC trajectories was proposed in [138]. It generates different LC trajectories by introducing driving style-related parameters into a sinusoidal LC model.

The above method is not suitable for accommodating diverse driving styles. Butakov et al. [67] proposed a two-layer LC model that generates personalized LC trajectories based on the driver’s real-time characteristics to address this limitation. The lower layer utilizes a sinusoidal LC model to generate vehicle trajectory and acceleration profiles for LC, and the upper layer employs GMM to adjust the personalized parameters of the mathematical model based on the driver’s driving features, such as the desired relative position to surrounding vehicles and the desired gap for LC. Huang et al. [139] introduced a trajectory selection model based on driver preferences. It utilizes the Fuzzy Linguistic Preference Relation (FLPR) method to select personalized LC trajectory clusters that satisfy the driver’s safety, comfort, and stability preferences. Then, the Technique for Order Preference by Similarity to an Ideal Solution (TOPSIS) is applied to rank and select the LC trajectory that best matches the driver’s preferences.

A limitation of RB models is that the defined rules or parameters may be insufficient to describe the heterogeneity of drivers accurately, and the predetermined rules may not cover all driving scenarios, resulting in lower prediction and decision accuracy. Learning-based approaches generate LC trajectories by learning from the driver’s driving data. They establish a mapping from the driving scenarios, vehicle states, and driver’s personalized characteristics to LC decisions using complex network structures. Compared to RB LC methods, learning-based approaches achieve higher accuracy.

An SVM-based LC model was proposed in [140]. By training on drivers’ LC data, SVM determines the timing of LC based on the driver’s preferences and driving scenarios. Meanwhile, Model Predictive Control (MPC) generates LC trajectories suitable for different driving styles based on the input from SVM. Zhang et al. [141] further considered the impact of surrounding and ego vehicles’ driving styles on trajectories and introduced the Driver Operation Profile (DOP) concept. They transformed the trajectories of surrounding vehicles and the ego vehicle into a DOP and utilized a CNN network to extract driving style information. By incorporating environmental factors, the model outputs LC trajectories.

However, the methods above determine the timing or trajectory of LC based on the current vehicle states and the influence of surrounding vehicles, overlooking the influence of the driver’s historical driving data on LC trajectories [142]. Ye et al. [143] proposed a personalized LC model considering temporal information. The vehicle’s lateral acceleration and trajectory, combined with driving style and surrounding vehicles, are used as inputs to an LSTM network to predict
personalized LC decisions. The model achieved a prediction accuracy of 96.5% on the NGSIM dataset. Gao et al. [68] utilized LC temporal information and historical LC data to extract personalized LC parameters through an LSTM network. These parameters were then applied to a Gaussian distribution equation to generate personalized LC trajectories.

The main drawback of neural networks is their poor interpretability as black-box models, making it difficult to understand the reasons behind network decision errors [144]. Zhu et al. [145] proposed a hierarchical learning-based LC model framework based on Gated Recurrent Units (GRU) and a safety field to address this drawback. The GRU neural network’s hierarchical structure was divided into three layers: LC intention recognition layer, trajectory planning layer, and control execution layer. The safety field method was incorporated to supervise driving trajectories and ensure LC safety. The hierarchical approach was employed to tackle the black-box issue of neural networks, and the real-time learning capability enabled the model to adapt to different driving styles.

3) Personalized braking: The impact of the driver’s characteristics on braking behavior mainly includes collision warning threshold and braking control strategy. AVs consider time to collision (TTC) and time headway (THW) as critical collision decision criteria [146] [147]. Traditional collision warning systems emit alerts based on fixed thresholds. However, different driving styles result in variations in TTC and THW among drivers [148]. Non-personalized alerts can lead to driver annoyance and even the deactivation of the collision warning system. Therefore, considering the influence of the driver’s characteristics on braking decisions, it is necessary to develop an adaptive collision warning system.

Muehlfeld et al. [149] proposed a phased adaptive threshold collision warning system that utilizes statistical behavior modeling to analyze the distribution of risk levels in past driving data and calculate personalized distribution warning activation thresholds. Although the method partially achieves adaptive threshold functionality, the phased thresholds reduce the accuracy of adaptation. Therefore, further exploration of the alert mechanism in adaptive warning systems is needed. To accurately obtain adaptive driver collision alert thresholds, Wang et al. [150] proposed a personalized collision warning algorithm that can adjust in real-time based on changes in driver behaviors. The adaptive threshold mechanism consists of a real-time risk assessment model that determines the current level of danger in the driving scenario and an adaptive threshold adjustment that updates the warning threshold in real-time based on the current risk level, collision warning threshold, and driver’s braking behavior. Simulation results show that this method improves prediction accuracy, adapts to driver behavior fluctuations, and accommodates different driving styles.

In addition to parameter updating from driver behaviors, a driver’s states are important influencing factors. In [151], the authors utilized vehicle state information, the driver’s EEG signals, and thermal imaging facial data as inputs to a KNN network to predict the driver’s braking reaction time. The results showed that including the driver’s state information improved the accuracy of the prediction. Iranmanesh et al. [152] considered the impact of a driver’s distracted state on collision warning thresholds. They dynamically adjusted the braking warning thresholds in real-time using the driver’s braking data and fine-tuned them based on the driver’s distracted state.

In addition to collision warning, personalized braking control strategies that align with the driver’s characteristics are equally important. Different driving styles can be reflected in distinct braking pedal signals [69]. Chen et al. [153] and Zhu et al. [70] proposed braking control strategies suitable for different driving styles. For example, Zhu et al. [70] classified drivers into aggressive, normal, and cautious using the GA-SVM algorithm. They designed three personalized braking control strategies based on different brake pedal boost ratios and pressure. Matching the appropriate braking control strategy to the driver’s driving style improved driving comfort and adapted to different driving styles.

C. Driving process uncertainty

For RQ3, The intra-heterogeneity in driver behaviors can lead to the driving process uncertainty, which means that the driver’s current decisions may be influenced by dynamic factors such as individual differences, emotions, attention levels, and driving scenarios, resulting in unpredictable behaviors in specific situations.

Stochastic Model Predictive Control is a widely used approach to address the driving process uncertainty. This method assumes that the system uncertainty is stochastic and utilizes statistical analysis to obtain the probability distribution of the system’s uncertain behaviors [154]. Zhang et al. [71] proposed an electric vehicle-assisted fault-tolerant control system that integrates SMPC and Distribution Model Predictive Control (DMPC) to model the interactions among subsystems such as the driver, automatic steering system, and in-wheel motor within a distributed stochastic model predictive control framework. The concept of cooperative game theory is introduced to describe the conflicts and cooperation between the uncertain driver behaviors and other subsystems, aiming to enhance the stability of electric vehicle operation in the event of motor faults. Li et al. [12] modeled driver behavior in the signal dilemma zone using SMPC. The simulation experiments demonstrated that the proposed driver behavior model framework based on SMPC accurately captures vehicle motion and dynamics in the dilemma zone.

However, driver behavior is characterized by randomness and high variability, making the modeling of driver behavior challenging. One approach to address the high variability of driver behavior is to adopt a data-driven approach, utilizing a large amount of driving data collected under various traffic conditions to model driver behavior [155]. Zeng et al. [156] proposed a stochastic model incorporating road information for driver pedal behavior. They decomposed the driver’s braking behavior into discrete pedal position adjustments. They used the vehicle, road, and current pedal information as inputs to an Input-Output Hidden Markov Model (IOHMM) to estimate the probability distribution of pedal behavior.
Fünfgeld et al. [73] introduced a stochastic framework based on interpretable models and stochastic processes for vehicle dynamics prediction. They utilized road data to predict future vehicle dynamics and employed sequential Monte Carlo simulation to approximate the distribution of the vehicle’s future states. Hubmann et al. [74] proposed introducing a heuristic method to search for an approximate solution to the partially observable Markov decision process (POMDP) problem to improve the convergence of probability predictions. Additionally, considering the uncertainties of specific factors during motion planning also expedited the algorithm’s convergence speed.

V. FRAMEWORK FOR CDI SYSTEMS

This section presents a CDI framework based on lifelong learning. The framework meets the requirements for constructing CDI systems. Additionally, the dynamically expandable neural network and the memory pool enable CDI to continuously expand its network capacity with the arrival of new tasks, achieving lifelong learning for various driving contexts.

A. Lifelong learning-oriented structure

For practical applications, a CDI system is supposed to be capable of handling driving contexts where driver information and driving scenario information are tightly coupled or a takeover is requested. Drivers possess diverse personalized characteristics and exhibit varying distributions of uncertain behaviors. As a result, different drivers may react differently in the same driving scenario, and even the same driver can exhibit varying responses in the same driving scenario at different times. This results in a tremendous number of driving contexts formed between drivers and driving scenarios. Traditional machine learning methods require a priori parameter configuration and data collection for offline training [157]. While maintaining the model’s generalization ability, this approach may limit the model’s capacity to adjust to the personalized characteristics of different drivers.

However, recent research indicates that establishing a unified model and continuously improving its performance through lifelong learning is a more reasonable and promising approach [158]. Unlike traditional learning that adapts to different tasks through offline training, lifelong learning can be incremental, enabling it to acquire new knowledge without the obligation to relearn the already-learned data [159]. Moreover, the mechanisms of forgetting and retention tackle the problems of stability–plasticity, facilitating the perpetual accumulation of the system’s knowledge [160]. Additionally, the assimilated knowledge is transferrable and applicable to future tasks, facilitating the system’s rapid acquisition and adaptation to new tasks, thereby achieving a state of continuous intelligence and adaptability [161]. These features enable lifelong learning-based intelligent systems to emulate the learning process observed in humans [162]. Thus, lifelong learning-based CDI systems can effectively handle a wide range of complex driving contexts, continuously adapt to the personalized characteristics of drivers, and meet the demands of future CDI system development endeavors.

However, the connection between drivers and driving scenarios varies in different driving contexts, meaning the spatial-temporal features of different contexts are distinct. The variability leads to models trained in old driving contexts not generalizing well to unknown driving contexts. The model performance can be improved by learning new driving contexts, which may lead to forgetting old driving contexts, resulting in catastrophic forgetting. The phenomenon is especially prominent, particularly for feature-rich parametric models [163]. Therefore, the key to achieving lifelong learning-based CDI systems is to prevent catastrophic forgetting when learning new driving contexts.

To address the problem of catastrophic forgetting, many researchers [164] [165] [166] have proposed dynamically expandable neural networks. These networks can selectively train the old network based on the training tasks and expand the network capacity by adding neurons or decoders to prevent catastrophic forgetting. Therefore, dynamic neural networks with adjustable parameter capacity can balance stability and plasticity.

This paper employs dynamically expandable neural networks and a memory pool to achieve lifelong learning-based CDI systems, as shown in Figure 5. The additional memory pool can store data from previous tasks, allowing the model to revisit old data to prevent catastrophic forgetting. Moreover, it enhances the model’s learning ability for new tasks based on the knowledge gained from the previous tasks [167], [168]. The dynamically expandable neural networks selectively train the network parameters associated with these variables and output corresponding results. When facing new tasks, the network can utilize previously learned knowledge for assessment. If the loss cannot meet the desired threshold, the network first dynamically expands its capacity. Subsequently, the network is trained by combining old data from the memory pool with new data, achieving lifelong learning.

B. A framework for CDI systems

Finally, based on the requirements for CDI systems and the lifelong learning-oriented structure described in Figure 5, we proposed a framework for CDI systems, as depicted in Figure 6. Unlike the driver assistance-oriented CDI framework proposed in [169], our proposed CDI framework extends beyond mere driver assistance. It seamlessly integrates driver information into the decision-making process, specifically tailored for intricate human-vehicle cooperation scenarios. Furthermore, we emphasize the importance of incorporating perception with personalization and personalized decision-making into the framework to reduce driver uncertainties in the decision-making loop and align with the trend of development in future personalized AVs. Notably, we have incorporated a lifelong learning structure, utilizing dynamically expandable neural networks coupled with a memory pool. This enables a continuous evolution towards more intelligent and effective handling of complex driving contexts. Although lifelong learning has been used in other domains [170] [171] [172], its application in CDI systems marks a pioneering step in our framework. Such innovative integration guarantees that
our CDI system’s capability for personalization can be continuously adapted to specific drivers. We believe our system framework offers valuable insights for researchers striving to craft safe driving models that account for CDI.

VI. DISCUSSION

This paper provides a comprehensive review of the critical research questions for building a CDI system. By addressing these questions, guidance for developing CDI perception and decision-making systems was provided. We reviewed CDI perception and decision-making algorithms and emphasized the need for personalized perception and the capability to handle driver heterogeneity and stochasticity to achieve human-centric AVs. Additionally, we propose a framework for the CDI system based on lifelong learning, which outlines the pipeline of the CDI system. To the authors’ best knowledge, this is the first comprehensive research on CDI systems.

For RQ1, we first surveyed multi-modal perception algorithms for CDI perception systems, which can be categorized into driver behavior prediction, driver state detection, and takeover identification. Based on our research, the pipeline for constructing a multi-modal CDI perception algorithm is presented in Figure 3, which includes various factors that need to be considered. More specifically, the network’s inputs should be selected from a driver, driving scenarios, and vehicle states based on the task’s requirements. Furthermore, it is necessary to analyze the relevance of driver and driving scenario features to the target task. For instance, a driver’s head pose and lane marking information can serve as good indicators of lane-changing behavior. Next, selecting feature representations and prediction networks based on inputs should be carefully evaluated. Currently, time-series features and LSTM networks are widely used. In addition to the pipeline, personalized characteristics such as physiological characteristics and driving styles can impact the overall detection accuracy of CDI perception systems. Therefore, when developing algorithms, it is important to consider the influence of these personalized characteristics.

For RQ2, we first surveyed decision-making algorithms suitable for CDI decision-making systems. As the CDI decision-making loop involves driver information, this poses requirements on the CDI decision-making system to address the heterogeneity of different drivers in spatial and temporal driving contexts. By characterizing the diverse responses of drivers to the same stimuli, personalized decision-making addressed the issue of consistency among drivers in different spatial and temporal driving contexts, which is caused by inter-heterogeneity. It is observed that driving style and individual driving behaviors are frequently utilized as inputs to generate decision-making for specific drivers in personalized CF, LC, and braking models. Similarly, as shown in Figure 4, future CDI decision-making systems can consider driving styles and personalized behaviors in the input and utilize personalized information to generate adaptive decision thresholds or behaviors that can address inter-heterogeneity.

For RQ3, SMPC and data-driven methods are commonly used to address driver uncertainty in the driving process. SMPC utilizes statistical analysis to obtain probability distributions of system uncertainties and can cover some stochastic behaviors. Data-driven methods rely on collecting a large amount of driving data and training models to cover driver stochastic behaviors. Furthermore, with the development of deep learning, data-driven approaches have become a trend. Therefore, in the future, it is possible to construct a comprehensive CDI dataset that encompasses a wide range of driving scenarios to address the driving process uncertainty.

According to the research scope defined in Section II, our review is limited to the issues related to perception and
decision-making in the CDI system. Therefore, this paper does not consider other topics, such as hardware. Additionally, after analyzing the relevant literature on CDI algorithms and considering the current challenges in human-vehicle cooperation, we have carefully proposed considering the driver’s personalized characteristics and driving process uncertainty that a CDI system needs to address. Finally, based on the three problems identified and our considerations for the CDI system, we have proposed a lifelong learning framework based on evolvable neural networks to meet the future development needs of CDI. It should be noted that this technology only applies to a CDI system, and further research is needed for its application in other domains.

In summary, this paper investigates the perception and decision-making algorithms for CDI, driver’s personalized state detection, personalized driver models, and methods to address stochasticity in driving behavior. It also answers three research questions related to CDI systems. In comparison to other related works, our research focuses more specifically on serving the purpose of constructing a CDI system. Finally, based on the research mentioned above, we propose a framework based on dynamically expandable neural networks that can meet the requirements of a CDI system. This investigation is generally useful for constructing a CDI system.

VII. CONCLUSION AND FUTURE WORK

This paper discusses the current research status of perception and decision-making algorithms for CDI and addresses the issues to be considered in building a CDI system. The following findings are obtained: 1) The existing CDI perception algorithms lack consideration of personalized driver states, which can be addressed by introducing personalized detection and recognition. 2) There are few data-driven decision-making algorithms for CDI, and there is a lack of consideration for the driver’s personalized features and the uncertainty of driving processes caused by the driver’s subjective behaviors. Therefore, we suggest using personalized driver models and data-driven approaches to address these issues. 3) There is no guiding framework for building a CDI system, which is crucial for its implementation. We propose a framework to overcome shortcomings.

These findings also indicate future research directions. The application of neural networks in CDI perception and decision-making algorithms is a trend that requires the construction of datasets for training the networks. Additionally, understanding the impact of the driver’s personalized features on the CDI system and selecting appropriate input parameters are also worth investigating. Furthermore, to adapt to the centralized trend of automotive electrical and/or electronic architectures, it is also worth exploring the construction of hardware suitable for CDI systems. In conclusion, this paper provides a research foundation for the future development of CDI systems.

REFERENCES


Zhongpan Zhu is a Research Associate Professor at School of Electronic and Information engineering of Tongji University. He received M.S. degree in Power Machinery and Engineering from School of Automotive Studies of Tongji University in 2012 and the Ph.D. degree in Vehicle Engineering in 2019. He was a postdoctoral research scientist at School of electronic and information engineering of Tongji university from 2019 until 2022. He was the chief of automotive software development section in TJARD institute from 2013 until 2016. His research focuses on the AI algorithms of autonomous vehicles and intelligent robots.

Shuaijie Zhao received the B.E. degree in Engineering from Jilin University, Changchun, China, in 2022. He is currently working toward the M.S. degree with the Department of Automotive Engineering, Tongji University, China. His research interests include intelligent cockpit, human-machine interaction, and image recognition.

Cheng Wang received the B.Sc. degree from the school of automotive engineering at Wuhan University of Technology, Wuhan, China, in 2014 and the M.Sc. degree from the school of automotive studies at Tongji University, Shanghai, China, in 2017 and the Ph.D. degree from the institute of automotive engineering at the Technical University of Darmstadt, Darmstadt, Germany, in 2021. He currently works as a research associate at the University of Edinburgh. His research interest includes explainable AI and safety verification and validation of autonomous vehicles.

Aimin Du received the Ph.D. degree in power and mechanical engineering from Jiangsu University, Zhenjiang, China, in 1999. After finished his postdoctoral research work at Tongji university, Shanghai, China, in 2001, he is currently an Associate Professor with the Department of Automotive Engineering, Tongji University, Shanghai, China. His research interests include the development and control of new energy vehicle power systems, intelligent control of automotive power systems, and server software development. He received the China Machinery Industry Science and Technology Award in 2011.