Cognitive Systems Research 83 (2024) 101169

Contents lists available at ScienceDirect

Cognitive Systems Research

journal homepage: www.elsevier.com/locate/cogsys

Human-inspired autonomous driving: A survey

Alice Plebe [a,*], Henrik Svensson [b], Sara Mahmoud [b], Mauro Da Lio [a]

[a] University of Trento, Trento, Italy
[b] University of Skövde, Skövde, Sweden

ARTICLE INFO

Action editor: David Vernon

Keywords:
Autonomous vehicles
Human driving
Human cognition
Imitation learning
Neuromorphic computing

ABSTRACT

Autonomous vehicles promise to revolutionize society and improve the daily life of many, making them a coveted aim for a vast research community. To enable complex reasoning in autonomous vehicles, researchers are exploring new methods beyond traditional engineering approaches, in particular the idea of drawing inspiration from the only existing being able to drive: the human. The mental processes behind the human ability to drive can inspire new approaches with the potential to bridge the gap between artificial drivers and human drivers. In this review, we categorize and evaluate existing work on autonomous driving influenced by cognitive science, neuroscience, and psychology. We propose a taxonomy of the various sources of inspiration and identify the potential advantages with respect to traditional approaches. Although these human-inspired methods have not yet reached widespread adoption, we believe they are critical to the future of fully autonomous vehicles.

1. Introduction

The industry has frequently overlooked the role of human inspiration in the design of autonomous vehicles. One common objection is that humans cause traffic accidents, which is something autonomous vehicles should definitively not draw inspiration from. However, the major human causes of accidents are cognitive impairments such as the influence of alcohol or drugs, tiredness, distraction, and recklessness (Singh, 2015). Otherwise, human drivers cope well with the driving task, efficiently evaluating traffic situations and generating behaviors. In fact, an expert driver in normal conditions is rarely the cause of an accident. Hence, replicating some of those abilities in a vehicle that does not suffer from human imperfections, such as fatigue, might be the way toward fully autonomous vehicles.

In this paper, we review existing work on autonomous vehicles that draw inspiration from cognitive science, neuroscience, and psychology. We analyze what it means to replicate cognition in autonomous vehicles and show how that may differ with respect to how scientists are informed by and use the source model of biological intelligence. There is a multitude of different ways in which the intelligence of biological agents have been and can be used to inspire the development of autonomous vehicles. As will be shown in our review, the inspiration differs with respect to the degree to which the implementation choices are informed by the source model of the biological intelligence. The large variation of where and how the inspiration is used paints a picture that there is no principled or systematic way that the choice of inspirations is made, but rather seem to be opportunistically chosen.

It is our intention to contribute to the field by adding some terminological structure, categorize the main types of inspiration and, to a degree, evaluate the possible advantages of the inspiration. Although our review does not include some of the technical details of the studies and the results, the main focus of our review is the human inspiration and the usefulness of such approaches. The reader is directed to other survey papers for the technical methods for self-driving cars, such as general self-driving cars surveys (Badue et al., 2021; Devi, Malarvezhi, Dayana, & Vadivukkarasi, 2020; Paden, Čáp, Yong, Yershov, & Frazzoli, 2016; Pendleton et al., 2017) or specific such as deep learning for self-driving cars (Grigorescu, Trasnea, Cocias, & Macesanu, 2020; Kuutti, Bowden, Jin, Barber, & Fallah, 2020), or reinforcement learning (Aradi, 2020; Kiran et al., 2021).

The review is structured as follows. In Section 2, we present a taxonomy of sources of inspirations, which identifies four levels of abstraction. Sections 3 to 6 focus on each level of abstraction and describe the related works. Lastly, Section 7 discusses the impact of human inspiration for autonomous vehicles and the potential future directions.

2. Background

To describe the cognitively inspired approaches in autonomous driving, we tease apart two dimensions along which the inspiration from biological agents can be categorized.

* Corresponding author.
E-mail address: alice.plebe@unitn.it (A. Plebe).

https://doi.org/10.1016/j.cogsys.2023.101169
Received 29 August 2022; Received in revised form 21 June 2023; Accepted 31 August 2023
Available online 7 September 2023
1389-0417/

	(A) BEHAVIORAL INSPIRATION	(B) FUNCTIONAL INSPIRATION	(C) ARCHITECTURAL INSPIRATION	(D) CELLULAR INSPIRATION
(1) DATA STAGE	Behavioral cloning, imitation learning {17}	Human-like system parametrization {2}	Visual cortex {2}	Retinomorphic cameras {5}
(2) IMPLEMENT. STAGE	Human-like constraints of decision space {2}	Attention, social collaboration, emotions {14}	Cerebellum, serotonin {7}	Place cells, grid cells {2}
(3) THEORY STAGE	Observation of external agents {2}	ACT-R, consciousness, microdecisions {7}	Basal ganglia {4}	Spiking neural networks {4}

Fig. 1. Taxonomy of inspirations. The columns represent the four levels of abstraction of the inspiration: behavioral (A), functional (B), architectural (C), and cellular (D). The rows indicate the stages at which the inspiration is modeled: data (1), implementation (2), and theory (3). The table reports a selection of examples of inspiration from the collected papers. In blue, the total number of analyzed papers in each class.

The first dimension describes the level of abstraction at which the biological organism is taken as source of inspiration. The inspiration is differentiated into four levels of abstraction: behavioral (A), functional (B), architectural (C), and cellular (D). The behavioral level (A) refers to complete agent-environment interactions, i.e., observations of humans performing driving tasks. The functional level (B) takes us "inside" the human and places the inspiration at the level of cognitive abilities, such as those found in standard textbooks of cognitive psychology (including but not limited to attention, perception, memory, problem solving and decision making (Smith & Kosslyn, 2013)). The architectural level (C) refers to global brain systems, prototypically the human or mammalian brain. The cellular level (D) refers to approaches that takes inspiration from the behavior of a single neuron or group of neurons.

The second dimension describes at which level of explanation the inspiration is formulated. The description is loosely based on the *level of explanation* in cognitive and psychological sciences (Guest & Martin, 2021), which translates into three stages at which the inspiration is modeled: data (1), implementation (2), and theory (3). Data (1) is the lowest stage and refers to the structure of the data perceived from the environment, or how it flows and processes in the brain. The implementation stage (2) focuses on how to specify a descriptive formula that can be translated into computational implementation. Theory (3) is the highest stage and refers to more general theories of cognition and behavior often formulated in natural language. An intrinsic feature of this layered model is that each level of explanation is influenced by the level above it. Thus, there often exists an explicit or implicit *theory* of the behavior (e.g. cognition consists of a uni-directional flow of information from sensing to acting) which influences the way it is modeled at the *implementation* level (e.g. separate modules handle each function). This makes the categorization on this dimension somewhat difficult as the approach may encompass more than one level, but also the scientist taking inspiration from the implementation level may be unaware of the influence from the theory level. With that in mind, we note that the review is only based on the explicit claims made in the reviewed papers, not the *implicit* assumptions that may be inferred. It is possible, for both dimensions, that inspiration are drawn from more than a single level, but in the review (see Table 2), we point to a single category that we see as the primary influence for each reviewed paper. Fig. 1 shows the proposed taxonomy with the twelve classes in the two dimensions. The table provides a prototypical example for each category in the taxonomy, along with the number of papers identified for the category.

In addition, the analyzed works adopt well-established methods that are transversal to the categories of human imitation. We categorize and code these methods in Table 1. While some works may combine more than one of the listed methods, in Table 2 we indicate only the primary adopted method.

Table 1
List of the main methodologies adopted in the collected papers, along with the corresponding number of papers that employed each methodology and a unique code that is referenced in Table 2.

Code	#	Methodology
AB	5	Agent-based modeling
AE	3	Autoencoder
AL	3	Adversarial learning
BC	3	Behavioral cloning
CI	4	Conditional imitation learning
CC	3	Computational cognitive architecture
CN	6	Convolutional neural network
CP	3	Classical sense-think-act pipeline
FN	3	Fuzzy neural networks
IL	6	Imitation learning
OA	5	Other algorithms
PG	8	Probabilistic graphical model
RL	6	Reinforcement learning
RN	3	Recurrent neural network
SP	7	Spiking neural network

3. Inspiration at behavioral level

We identify a first category of works aiming to learn driving behaviors by observing humans. This category encompasses the papers of class A in Table 2. The following works do not focus on theories explaining how humans drive; rather, they focus on imitating human behaviors at a high abstraction level. The papers collected here vary in the way imitation is realized. Fig. 2 gives a general overview of the different methods to learn behaviors through imitation. In addition, some papers also distinguish themselves in that they choose to imitate a particular human driver rather than a generalized version derived from a large dataset consisting of many human drivers.

The idea of learning through imitation has, in fact, a strong meta-cognitive foundation. Imitation learning is a compelling research topic in neuroscience, especially since the discovery of mirror neurons during the end of the last century; in humans and other primates, mirror neurons have a key role in the ability to imitate the actions of another agent (Iacoboni et al., 1999; Rizzolatti, Fogassi, & Gallese, 2001; Rizzolatti & Sinigaglia, 2008). Imitation learning has been spreading also in artificial intelligence (Schaal, 1999), robotics, and autonomous

Table 2
Selection of analyzed papers. For each paper, we indicate the purpose in the context of autonomous driving; the class of inspiration as in Fig. 1 (A = behavioral; B = functional; C = architectural; D = cellular; 1 = data; 2 = implementation; 3 = theory) ; the main method applied in the paper using the codes from Table 1 ; and the benefit provided by the inspiration.

Paper	Application/Purpose	Class	Method	Benefit of inspiration
Chan, Partouche, and Pasquier (2007)	Curve anticipation and negotiation	A1	FN	Use human driving data to extract rules for better negotiation of new unseen roads
Gu, Hashimoto, Hsu, Iryo-asano, and Kamijo (2017)	Left turn at intersections with multiple pedestrians	A1	PG	Use human data to better recognize pedestrians' behaviors and act accordingly
Li, Chang, and Chen (2003)	Limited set of driving tasks, using a miniature robotic vehicle	A1	PG	Use human experience to design fuzzy rule based control for human-like driving skills
Codevilla, Müller, López, Koltun, and Dosovitskiy (2018)	Driving policy for intersections following human directives	A1	CI	Take into account human's internal states in the driving policy
Hawke et al. (2020)	Driving policy in urban routes with simple traffic	A1	CI	Take into account human's route commands in urban driving
Eraqi, Moustafa, and Honer (2022)	Driving policy in lane keeping or switching	A1	CI	Take into account dynamic route commands in driving
Teng et al. (2023)	Driving policy in CARLA simulated environment	A1	CI	Take care of interpretability
Hang, Lv, Xing, Huang, and Hu (2021)	Decision making in merging and overtaking	A1	OA	Use human driving data to extract metrics defining human-like driving styles
Omeiza, Anjomshoae, Webb, Jirotka, and Kunze (2022)	Generate driving explanations	A1	AB	Use human driving commentary to learn automatic explanations
Xu et al. (2021)	Motion planner in highway scenarios	A1	OA	Use human driving data to improve comfort, efficiency, and safety
Markelić et al. (2011)	Advanced driver-assistance system, using a real vehicle	A1	IL	Use expert demonstrations to detect unusual human behaviors
Sharma, Tewolde, and Kwon (2018)	End-to-end steering control	A1	BC	Reduce the amount of data needed to effectively training the model
Kumaar, Navaneethkrishnan, Hegde, Raja, and Vishwanath (2019)	End-to-end steering control	A1	BC	Clone human behaviors to improve flexibility in response to long-term dynamics
Hecker, Dai, Liniger, Hahner, and Gool (2020)	End-to-end driving agent	A1	AL	Discriminate between human driving data and generated output to produce more human-like actions
Koeberle et al. (2021)	Autonomous driving in roundabouts and lane merging	A1	AL	Discriminate between human driving data and generated output to produce more human-like actions
Kuefler, Morton, Wheeler, and Kochenderfer (2017)	Autonomous driving in highway scenarios	A1	AL	Discriminate between generated actions and human data to produce human-like emergent behaviors
Rhinehart, McAllister, and Levine (2019)	Autonomous agent	A1	IL	Generate expert-like behaviors without reward function crafting
Li, Ota, and Dong (2018)	Decision making in scenarios with multiple external human drivers	A2	BC	Learn safer driving actions in trafficked scenarios
Liang, Wang, Yang, and Xing (2018)	Long-term driving strategies in urban environments	A2	RL	Exploit expert human demonstrations to initialize the action exploration in a reasonable space
Schulz, Mattar, Hehn, and Kooij (2021)	Object detection in blind corners	A3	IL	Use acoustic data to detect approaching vehicles behind blind corners before they enter in line-of-sight
Zhang and Ohn-Bar (2021)	Driving policy	A3	IL	Observe other human drivers to learn new maneuvers
Bagdatli and Dokuz (2021)	Discretionary lane change	B1	PG	Use human driving data to identify the factors that prompt drivers to change lane
Suresh and Manivannan (2017)	Limited set of driving tasks, using a real vehicle	B1	PG	Use human driving data to identify logical states of scenarios
Ouyang, Cui, Dong, Li, and Niu (2022)	Multi-sensor target detection	B2	CN	Real-time efficiency and high detection accuracy using visual attention
Buckman, Pierson, Schwarting, Karaman, and Rus (2019)	Intersection negotiation	B2	AB	Decrease the individual waiting times and the overall group congestion by measuring a social psychology metric

(continued on next page)

Table 2 (*continued*).

Qiao, Schneider, and Dolan (2021)	Behavior planning at urban intersections	B2	RL	Better convergence to an optimal policy with respect to traditional reinforcement learning
Sun, Zhan, Tomizuka, and Dragan (2018)	Decision making in merging and overtaking	B2	IL	Use "courtesy" to produce trajectories well accepted by other human drivers
Riaz and Niazi (2018)	Rear-end collision avoidance	B2	AB	Simulate levels of fear to reduce the number of driving rules
Butt, Riaz, Mehmood, and Akram (2021)	Rear-end collision avoidance	B2	AB	Simulate levels of fear to reduce the number of driving rules derived from expert drivers
Pal, Mohandoss, and Mitra (2019)	End-to-end driving agent	B2	CN	Reduce computation time with visual attention
Wei et al. (2021)	End-to-end driving agent	B2	CN	Use visual attention to reduce computation time and focus on salient areas for safety improvement
Kim, Rohrbach, Darrell, Canny, and Akata (2018)	Generate driving explanations	B2	CN	Use visual attention to synthesize appropriate driving commentaries
Da Lio, Donà, Rosati Papini, and Plebe (2022)	Autonomous agent accepting human directives	B2	OA	Mutual understanding of intentions for efficient collaboration with humans
Gonsalves and Upadhyay (2021)	Autonomous agent, using a miniature robotic vehicle	B2	RL	Use the human driving cycle as a blueprint for implementation
Ha and Schmidhuber (2018)	Autonomous agent, using a simplified simulated vehicle	B2	RN	Combine prediction with memory to learn in "hallucinated" scenarios
Czubenko, Kowalczuk, and Ordys (2015)	Autonomous agent	B2	AB	Emotions bias the decision making towards specific reactions for specific situations
Zhang, Zhou, Liu, and Hussain (2019)	Autonomous agent	B2	CP	Context-aware module adapting the driving strategy to different scenarios
Chen, Sun, Zhao, Li, and Liu (2021)	Lane keeping	B3	CC	Improve transition smoothness in driver-vehicle cooperation using the HSIC cognitive architecture
Saeed et al. (2019)	Optimal route computation	B3	FN	Use cognitive memory to store route experiences for improved decision making
Salvucci (2006)	Computational model of driver behavior for distraction recognition	B3	CC	Better understanding of driver behavior using the ACT-R cognitive architecture
Xie, Chen, Tomizuka, Zheng, and Wang (2020)	Computational model of driver behavior (no evaluation)	B3	CP	Potential future adoption in human-like autonomous vehicles
Sprenger (2020)	Decision making (no implementation)	B3	OA	Use "microdecisions" to potentially compensate the lack of proper consciousnesses in intelligent vehicles
Wiedermann and Leeuwen (2021)	Autonomous agent (no implementation)	B3	CP	Require "minimal machine consciousness" to achieve intelligent vehicles
Zakaria (2021)	Autonomous agent with imperfect sensors	B3	PG	Improve response to surprising events using "neuronal units of thoughts"
Plebe, Kooij, Rosati Papini, and Da Lio (2021)	Perception of driving space	C1	AE	Use cortical magnification to jointly exploit egocentric and allocentric spatial representations
Kashyap, Fowlkes, Krichmar, and Member (2021)	Object- and ego-motion estimation	C1	AE	Estimate object motion from the ego-motion field using sparse representations similar to the visual cortex
Plebe and Da Lio (2020)	Learning representations of driving concepts	C2	AE	Achieve interpretable representations inspired by convergence–divergence zones in the brain
Nezhadalinaei, Zhang, Mahdizadeh, and Jamshidi (2021)	Object detection and tracking	C2	SP	Better efficiency combining spiking neural networks, CNNs, and conditional random fields
Pasquier and Oentaryo (2008)	Decision making for limited set of tasks	C2	FN	Emergence of tactical driving skills using cerebellum-like control mechanisms
Kim and Langari (2013)	Adaptive cruise control	C2	CC	Better inter-vehicle distance tracking performance using brain limbic system based control
Xing, Zou, and Krichmar (2020)	Autonomous navigation	C2	RL	More flexibility in time-critical navigation by imitating the function of serotonin
Ahmedov, Yi, and Sui (2021)	Autonomous agent	C2	IL	Improve interpretability and robustness using two "neural circuit policies" to mimic the asymmetric hemispheres
Chen, Zhang, Shang, Chen, and Zheng (2017)	Autonomous agent	C2	RN	Bridge the gap from complex perception to control taking inspiration from visual and motor cortices

(*continued on next page*)

Table 2 (*continued*).

Tenison et al. (2019)	Detection of traversable paths for navigation	C3	OA	Reduce computational burden using "Feynman machine" dynamic system
Hussain, Abdullah, Yang, and Gurney (2012)	Low-level motor control	C3	PG	Smoother and more efficient motor control using a controller switch inspired by basal ganglia
Yang, Hussain, and Gurney (2013)	Low-level motor control	C3	PG	Smoother and more efficient motor control using two controller switches inspired by basal ganglia
Lechner et al. (2020)	End-to-end lane keeping	C3	RN	Improve interpretability and robustness of the system using "neural circuit policies" inspired by *C. Elegans*
Kashyap (2020)	Stereo depth scene reconstruction	D1	SP	Adopt retinomorphic camera combined with TrueNorth neuromorphic processor[a]
López-Randulfe, Duswald, Bing, and Knoll (2021)	Radar-based object detection	D1	SP	Adopt spiking neural networks with weights derived analytically from Fourier transform[a]
Maqueda, Loquercio, Gallego, García, and Scaramuzza (2018)	Prediction of steering angle	D1	CN	Achieve better performance with retinomorphic cameras compared to standard cameras[a]
Fischl et al. (2017)	Autonomous driving with a miniature robotic vehicle	D1	CN	Adopt retinomorphic camera combined with TrueNorth neuromorphic processor[a]
Patton et al. (2021)	Autonomous driving with F1Tenth vehicle	D1	SP	Adopt μCaspian neuromorphic processor[a]
Taniguchi, Fukawa, and Yamakawa (2021)	Simultaneous localization and mapping	D2	PG	Extend existing SLAM models with hippocampus-like structures (to be evaluated in future work)
Kaiser et al. (2019)	End-to-end lane following	D2	RL	Adopt spiking neural networks trained with probabilistic sampling of synaptic values[a]
Shalumov, Halaly, and Tsur (2021)	LiDAR-driven collision avoidance	D3	SP	Adopt spiking neural networks organized into the Neural Engineering Framework[a]
Liu, Lu, Luo, and Yang (2021)	Obstacle avoidance and target tracking	D3	SP	Adopt spiking neural networks trained in analogy with spike-timing-dependent plasticity[a]
Bing, Meschede, Chen, Knoll, and Huang (2020)	End-to-end lane keeping and obstacle avoidance	D3	SP	Adopt Spiking neural networks trained in analogy with spike-timing-dependent plasticity[a]
Banino et al. (2018)	End-to-end autonomous navigation	D3	RL	Learn complex control policies from a sparse reward using grid-cell representations

[a] Papers do not claim other benefits besides the strong biological plausibility (Sections 6 and 7 further discuss this point).

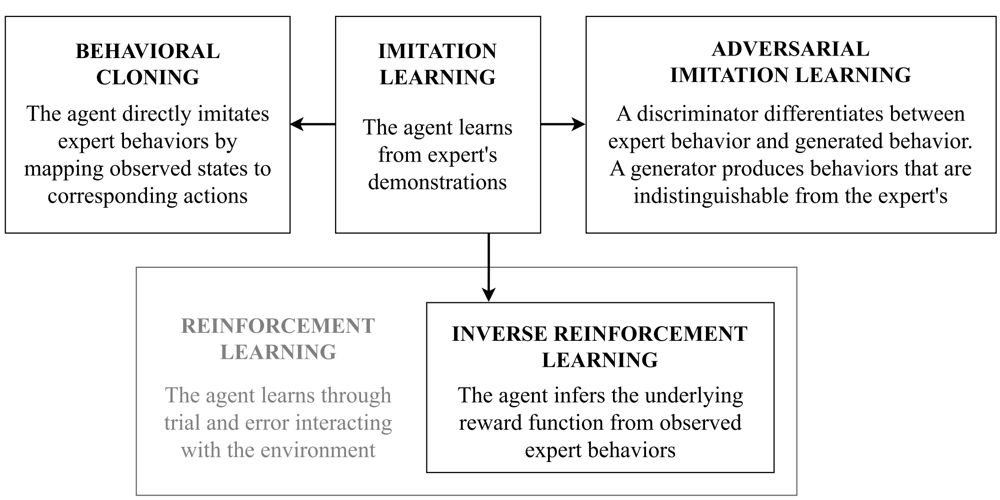

Fig. 2. Classification of imitation learning approaches. This classification is not intended to be exhaustive but rather gives a general overview of the topic in the context of autonomous driving.

driving. However, while in robotics we find several implementations of imitation learning based on mirror neurons (Demiris, Aziz-Zadeh, & Bonaiuto, 2014; Lopes, Melo, Montesano, & Santos-Victor, 2010; Tani & Ito, 2004), imitation learning approaches in autonomous driving leverage different concepts.

3.1. *Reinforcement learning and its variations*

The predominant method for integrating imitation learning into autonomous driving is *reinforcement learning* (RL). RL is a well-established framework for machine learning that has improved greatly in combination with deep neural networks. The crucial hindrance to using RL for imitation learning is that formulating an appropriate reward function is tricky. Nonetheless, several strategies have been proposed—we refer the reader to a recent survey on the topic (Ding, 2020). A popular strategy is *inverse reinforcement learning* (IRL) (Ng & Russell, 2000), which can extract a reward function from observed expert behaviors. The extracted reward function can be used in a standard RL model to derive the policy. This expert imitation strategy is also called *apprenticeship learning* (Abbeel & Ng, 2004). Unfortunately, IRL has significantly intensive computational cost. Moreover, since the problem is typically ill posed, the same observations can lead to different reward functions. Liang et al. (2018) propose a variation of the approach, called *controllable imitative RL*, and apply it to learn long-term driving strategies in urban environments. They exploit the expert human demonstrations to initialize the action exploration of the RL algorithm in a reasonable space.

3.2. Behavioral cloning

Another approach for imitation learning is *behavioral cloning* (Torabi, Warnell, & Stone, 2018). In this case, it is not necessary to formulate any reward function: the discrepancy between expert actions and model policy is minimized by supervised training. A simple behavioral cloning strategy is implemented in Sharma et al. (2018) with a deep convolutional network trained by observing humans driving in the TORCS simulator. Even if not explicitly mentioned, also (Li et al., 2018) use behavioral cloning to predict steering angle and speed in simulated trafficked scenarios. A notable shortcoming of simple end-to-end behavioral cloning is that it is difficult for the target model to derive driving policies from demonstrations of human driving, since the perceptual input given by cameras is too ambiguous with respect to the human driving decisions. Several attempts have been proposed to address this problem. Kumaar et al. (2019) propose a more sophisticated architecture made by long-term recurrent convolutional networks, and they apply it in end-to-end system for steering control. Similarly, Codevilla et al. (2018) find necessary to add information to help the model learning the driving policy. They take into account the expert's internal states representing the long-term driving intentions. In Hawke et al. (2020) behavioral cloning is enforced with the explicit representation of driver's route commands, with three possible values for *go-straight*, *turn-left* or *turn-right*. The target scenario of this work is urban driving with simple traffic conditions. The same conditioning command is used by Eraqi et al. (2022) in a more sophisticated approach that fuse LIDAR and camera input. In addition, route commands are generated automatically using CARLA simulator autopilot features. Behavioral cloning conditioned by driver's route commands is also the choice in Teng et al. (2023), where the novelty is in taking care of interpretability, which has become one of the hot topics for deep learning in general, and particularly in AV applications (Kamath & Liu, 2021). A certain degree of interpretability is achieved with two submodels, the first one does not use conditional imitation learning and produces a semantic interpretabile top view of the road scene, which is fed into a second submodel that predicts lateral and longitudinal commands, exploiting human imitation. Explainability is pursued more explicitly by Omeiza et al. (2022), proposing a system that synthesize spoken driving commentary, for informin the end-user about driving operations. Their system is based on decision trees, where planned commentary are derived by behavioral cloning from recorded commentaries of human driving instructors.

A distinct approach is proposed by Pal et al. (2019), where the cloned behavior is the eye gaze of human drivers. The information is exploited to improve the perception computation time in the model. Not just vision but also hearing can be imitated. In some cases, humans can detect vehicles behind blind corners by using auditory cues. The work of Schulz et al. (2021) show that it is possible to imitate a similar sensing modality to detect approaching vehicles before they enter the line-of-sight.

3.3. Adversarial imitation learning

A further method of learning by imitation is *adversarial imitation learning* (Ho & Ermon, 2016). Here, the imitating model acts as the generator network, while the discriminator network evaluates how well the resulting action-value function matches the expert behavior. Adversarial imitation learning is adopted by Kuefler et al. (2017) to drive autonomously in highway scenarios, and by Koeberle et al. (2021) for lane merging and roundabout scenarios. In Hecker et al. (2020), adversarial imitation learning is used to better refine an autonomous driving agent; as a result, the agent performs more human-like behaviors. In Rhinehart et al. (2019), model-based RL and imitation learning are combined in a model that accomplishes tasks never seen before, such as navigating to target regions and avoiding potholes. Similarly, Xu et al. (2021) implement a highway motion planner that combines traditional trajectory planning with an evaluation method based on expert demonstration that leads to more human-like trajectories. Adversarial imitation learning and its variants allow the imitation of expert drivers under several complex traffic scenarios. Their typical shortcoming, however, is the limited robustness of the driving policy to situations diverging from the training distribution, that prevents their application in everyday traffic situations.

3.4. Direct derivation of human driving styles

There are works that do not leverage any imitation learning strategy—they manually derive descriptive parameters from human driving data. In an earlier attempt to perform basic autonomous maneuvers, Li et al. (2003) develop a complex control based on fuzzy rules, including hard-coded human driving skills derived from expert drivers. Chan et al. (2007) implement a neuro-fuzzy architecture to extract rules from human experts for better curve anticipation and negotiation. Gu et al. (2017) adopt a dynamic Bayesian network to assess parameters describing pedestrian behaviors at crowded intersections. The network decides whether is safe to cross the intersection based on the gaps between pedestrians and the crosswalk. Hang et al. (2021) propose to use two non-cooperative game approaches—Nash equilibrium and Stackelberg games—to implement a human-like decision making algorithm for merging and overtaking. They model three human driving styles (aggressive, conservative, normal) based on metrics extracted from the NGSIM dataset. The direct derivation of human decision making rules for driving is the best way to ensure the inclusion of human driving styles in the artificial system, but at the price of not integrating deep learning components, thus losing the efficiency of today's most efficient components in AV systems.

3.5. Which humans to imitate

When learning through imitation, a key question is the following: who do you want to imitate? In the papers reviewed so far, the sources of imitation are anonymous human drivers, usually collected in large-scale datasets. However, there are exceptions. In Markelić et al. (2011), the imitation comes from the specific human that drives the car that the automated assistant system is running. The system learns the driver's behavior and becomes able to issue warning signals when the driver deviates from their usual behavior. On the opposite side, Zhang and Ohn-Bar (2021) propose to imitate all human drivers except for the one riding the ego-car. Given a scene, the model transforms the other vehicles' observations into the ego point of view and infers the expert actions. In this way, the authors aim to replicate the human ability to observe others and learn new behaviors never performed before. An interesting advantage of this approach is that it is possible to exploit a dataset to the full by analyzing all the surrounding vehicles.

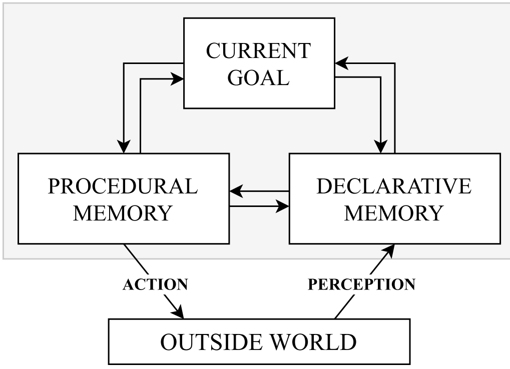

Fig. 3. General overview of the flow of information among the main modules of ACT-R.

4. Inspiration at functional level

In the second category, we collect works that take human cognitive abilities as a reference to develop autonomous driving systems. This category encompasses the papers of class B in Table 2. The works collected here draw inspiration from various aspects of human cognition, including attention, memory, consciousness, social cognition, and decision making. Furthermore, some papers attempt to explore how certain general theories of cognition can be applied in the context of autonomous driving.

4.1. General theories of cognition

We identify a first group of papers that are grounded in various general theories of cognition. The work of Salvucci (2006) commits to one of the most famous and comprehensive cognitive architectures, the *Adaptive Control of Thought-Rational* (ACT-R) by Anderson and Lebiere (1998), represented in a summary form in Fig. 3. The ACT-R architecture was never intended for engineering applications, least of all autonomous vehicles: the performance of models based on ACT-R is far from the real-life requirements of the automotive industry. However, Salvucci shows that a cognitive architecture can account for human driving behavior. In Chen et al. (2021), another less known cognitive architecture called *Human-Simulated Intelligent Control* (HSIC) (Zhou & Bai, 1983) is used. The HSIC is not a general cognitive architecture like ACT-R, but rather a control method based on human "kinesthetic intelligence". A key feature of kinesthetic intelligence is the intermittency of acting and waiting, typically exhibited by humans in a target tracking task. Using this idea, Chen et al. realize a system for lane keeping. The limitation of this system is in assuming a perfect knowledge of the lane geometry, hence there is no integration with a realistic sensory system.

In Gonsalves and Upadhyay (2021), no specific cognitive theory is cited. Nonetheless, their idea of cognition is spelled out clearly, and it corresponds to a traditional view of cognition as consisting of a pipeline of four stages: perception, scene generation, planning, and action, sometimes referred to as the sandwich model of cognition (Hurley, 2001). They implement a system based on this pipeline and test it on a miniature robot car equipped with two cameras. A traditional view of cognition is also endorsed in Zhang et al. (2019), where the pipeline is defined as perception, scene cognition, decision making, and control. Although a traditional view of cognition may be a valid starting point, current cognitive science have revealed an intimate relationship between perception and action, that cast doubt on a strict interpretation of the traditional pipeline as a model of human cognition.

At first glance, the pipeline in Ha and Schmidhuber (2018) also appears to be tied to a traditional view of cognition, with the modules visual (V), memory (M), and control (C). However, the novelty of the work lies in the memory module, where the representations are in terms of predictions of future sensory data. Since the model can generate future predictions, it can also hallucinate novel road scenarios and train itself on them. This ability reflects an intriguing function dreams have in humans: sophisticated exercises of simulations of various aspects of life. This phenomenological aspect of dreams is central in cognitive theories of mental simulation (Windridge, Svensson, & Thill, 2021). The idea of dreaming is also exploited in Da Lio, Donà, Rosati Papini, Biral and Svensson (2020) and Plebe, Rosati Papini, Donà, and Da Lio (2019) as a means of self-training a co-driver agent. The agent itself departs from the traditional sense-think-act organization and instead uses a layered control sensorimotor architecture modeled by the affordance competition hypothesis (Da Lio, Donà, Rosati Papini and Gurney, 2020; Plebe, Rosati Papini, Cherubini, & Da Lio, 2022).

There are works that attempt to build cognitive theories specifically for the driving task. From several theoretical ideas, Xie et al. (2020) extract three rules on the cognition of driving concerning the role of stored knowledge, the goal-oriented nature of driving, and the flexibility of behavior. Following these rules, they implement a model call DRIVE, described in detail but without experimental results yet. Zakaria (2021) proposes a new concept called *neuronal unit of thoughts* (NUTs). NUTs is a mixed symbolic-subsymbolic architecture where nodes are, at the same time, similar to neurons in a neural network and containers of high-level semantic meaning. For example, connecting two NUTs *speed* and *sharp curve* activates a third NUT *slowdown*. According to the author, the model naturally improves the response time to surprising events, and it is evaluated in the TORCS driving simulator. It seems that the major limitation of NUT approach is in scaling to realistic road conditions, no experiment seems to have been attempted other than the simple TORCS simulator.

4.2. Decision making

The next group of papers focus on specific cognitive mechanisms, rather than general cognitive theories as in the works presented so far. One of the most inspiring cognitive processes is *decision making*. In Suresh and Manivannan (2017), the decision-making system consists of binary decision trees based on a set of human decision rules derived from the literature. Bagdatli and Dokuz (2021), on the other hand, conducted a psychological study to identify the factors that induce humans to change lane. They design decision rules based on these factors and implement the decision-making system using a fuzzy cognitive map (Kosko, 1986). The work of Qiao et al. (2021) focuses on the hierarchical nature of human decision-making. The higher level of decision is prone to rational choices, whereas the lower level is more oriented to control. The authors build a hierarchical reinforcement learning to plan vehicle behavior at urban intersections, which shows better convergence than traditional algorithms.

Emotions also play a key role in the decision process (Damasio, 1994). The work of Czubenko et al. (2015) adopts emotions to modify the current set of possible decisions in a situation. For example, an emotion of surprise enables the decision of emergency braking, which is not available in a neutral emotional state. The emotion system is implemented using fuzzy rules. In addition, Riaz and Niazi (2018) use fuzzy logic to implement a system to avoid rear-end collisions. The system takes into consideration only the emotion of fear, and is based on a model of human fear popular in cognitive science (Ortony, Clore, & Collins, 1990). The system is improved by Butt et al. (2021) by designing the set of possible decisions according to a survey of experts. A shortcoming of this group of papers is the lack of integration between the decision process and a system with the efficiency given by deep learning modules, necessary to scale to full driving conditions.

4.3. Attention and memory

Another crucial cognitive mechanism is *attention* (Lindsay, 2020), which is investigated especially in the visual context (Braun, Koch, & Davis, 2001; Carrasco, 2011). The work of Ouyang et al. (2022)

expressly takes inspiration from saccades and implements a model called *SaccadeFork* for target detection, achieving real-time efficiency and high detection accuracy. The model uses a previous network (Law, Teng, Russakovsky, & Deng, 2020), which adopts a sort of attention mechanism to detect object centers. A more elaborate approach that combines saliency and convolutions is implemented in Wei et al. (2021), where the attention module explicitly targets motion planning (instead of generic perception tasks) and produces better results in terms of safety. Attention through saliency in convolutions is also exploited in Kim et al. (2018), but with different purposes: the automatic synthesis of textual explanations of the behavior of a self-driving vehicle, the same task pursued by Omeiza et al. (2022) (see Section 3.2). The work of Saeed et al. (2019) is inspired by another primary aspect of cognition: *memory organization* (Baddeley, 1992). The authors propose a system for vehicle route planning with a detailed organization that includes semantic, episodic, associative, short-term, and working memory of route information.

4.4. Consciousness and social cognition

An intriguing mental property—the most elusive in cognitive science—is *consciousness* (Chalmers, 1996; Dennett, 1992; Tye, 1995). Sprenger (2020) argues that conscious decisions are the result of micro-temporal processes below the threshold of consciousness. Since autonomous driving agents cannot make conscious decision, they can better rely on what Sprenger calls *microdecisions*. However, his proposal is at the moment only speculative, with no demonstrative implementation. On the opposite side, Wiedermann and Leeuwen (2021) argue that fully autonomous systems need a certain level of consciousness. The advocated minimal machine consciousness covers capabilities like self-knowledge, self-monitoring, self-awareness, and self-informing. Both papers are conceptual proposals that fall short of experimental evidence.

Lastly, there are works that take inspiration from social cognition (Augoustinos, Walker, & Donaghue, 1995) and rely on the idea that driving is essentially a social activity. Buckman et al. (2019) tackle the problem of intersection negotiation by applying a metric called *social value orientation*, which measures the willingness to help another vehicle at one's expense. By including this metric in the coordination policy, individual waiting times decrease as well as the overall group congestion. Prosocial behavior is also pursued by Sun et al. (2018), who propose an artificial form of "courtesy". A courteous autonomous agent balances its selfish objective with the potential inconvenience it brings to the other drivers. For example, a selfish agent overtakes first and forces the other driver to brake. By adding a courtesy term to the cost computation, the agent leaves more space when it merges, or it even yields to the other driver. An additional crucial aspect of social cognition is how humans grasp a mutual understanding of their intentions. In Da Lio et al. (2022), a co-driver is able to interpret suggestions from the human driver in terms of what is affordable in the current road situation. Depending on the situation, the system either executes the desired behavior or vetoes the request if considered dangerous.

5. Inspiration at architectural level

The next category of works draws on the mechanisms that take place in the human brain when driving. This category encompasses the papers of class C in Table 2. In several papers, inspiration comes mainly from specific brain areas, including the cerebellum and the basal ganglia. There are papers that draw inspiration from the way representations are organized in the brain cortices. In addition, we include a few works that look for inspiration in non-human brains.

5.1. Subcortical areas

The work of Pasquier and Oentaryo (2008) leverages the principles of cerebellum motion control (Ito, 1984) and realizes one of the earliest attempts to learn a limited set of driving skills, such as parking and U turns. Kim and Langari (2013) take inspiration from two main components of the brain emotion circuits, i.e., the amygdala and the orbito-frontal cortex (LeDoux, 2000). They simulate the two components to realize a system for adaptive cruise control. The artificial amygdala learns associations between sensory input and the experienced situations—either negative or positive—while the artificial orbito-frontal cortex prevents any inappropriate actions with inhibitory signals.

The basal ganglia are the source of inspiration for a number of works. Hussain, Gurney, Abdullah, and Chambers (2008) refer specifically to the cortico-basal ganglia loops, which play a fundamental role in reward and action selection (Haber, 2011), to implement a vehicle controller. The controller is loosely based on an earlier computational model of action selection in the basal ganglia (Gurney, Prescott, & Redgrave, 2001). The system switches between a set of possible controllers—initially two, later extended to three (Yang, Hussain, & Gurney, 2012)—depending of the demands of the current task. The model is further refined in Yang et al. (2013), where two basal ganglia deal separately with the longitudinal and lateral control. While in the cortico-basal ganglia network the key neuromodulator for action selection is dopamine, serotonin is the neuromodulator linked to impulsivity control, punishment prediction, and harm aversion (Avery & Krichmar, 2017)—although its precise role is still rather elusive (Hu, 2016). Xing et al. (2020) exploit the putative role of serotonin in promoting patience for navigation and road following. They implement an artificial serotonergic system on a small robot vehicle and show how the system is more flexible under time-critical constraints. Despite the undoubted interest of this study, it is clear that patience is only one of the desirable characteristics of leadership policy, and it would therefore be necessary to see how the method fits into an overall system.

5.2. Cortical areas

In the following works, inspiration comes from the way representations are organized in the brain cortices. Chen et al. (2017) simulate representations in the visual and motor cortices using a combination of convolutional and recurrent neural networks, and they apply an attention mechanism over an allocentric mapping of targets. Their strategy allows the model to effectively bridge the gap from complex perception problems to simple control commands. Concerning the brain organization of representations, *convergence–divergence zones* (CDZs) (Damasio, 1989; Meyer & Damasio, 2009) are an architecture of (biological) neurons that combine representation learning and mental simulation. The convergent component builds high-level conceptual representations from sensory stimuli, while the divergent component can reenact the sensorial experiences linked to the representations; Fig. 4 provides an example. The simulative nature of representations is key to making concepts emerge in the brain (Olier, Barakova, Regazzoni, & Rauterberg, 2017). This idea is already exploited in robotics (Droniou, Ivaldi, & Sigaud, 2015), and is applied to autonomous vehicles in Plebe and Da Lio (2020). They implement a CDZ-inspired convolutional autoencoder to learn conceptual representations of "vehicle" and "lane", where the representations maintain semantic organization and temporal coherence.

Another way of organizing neural representations comes from the dorsal visual pathway (Beyeler, Rounds, Carlson, Dutt, & Krichmar, 2019). Here, the representations are sparse and linked to ego-motion. There is ample evidence that the sparseness of the neuronal representation of stimuli is a highly efficient coding scheme in the cortex (Rolls & Tovee, 1995). This principle is applied by Kashyap et al. (2021) for object- and ego-motion estimation in driving scenes. They obtain a

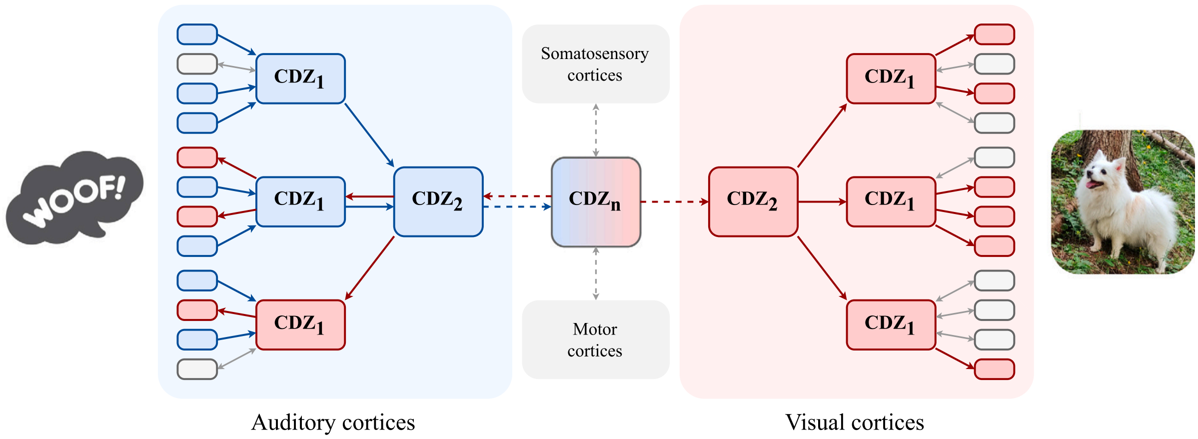

Fig. 4. Example of flow of activation in convergence–divergence zones (CDZs). An auditory stimulus elicits an activity pattern in neuron ensembles in early auditory cortices, which send converging forward projections (blue arrows) to higher-order CDZs. The CDZ in higher-order association cortices sends divergent back projections (red arrows) to CDZs in various sensorimotor areas, including the visual cortices. The CDZs in early visual cortices reconstruct the activity patterns previously associated with the barking sound, for example the appearance of a fluffy dog.

similar efficiency by introducing a sparsity measure in the loss function. The result is that the model needs only about 5% of the neurons to maintain accuracy in trajectory prediction. In addition to the representation structure, the primary visual cortex possesses another interesting feature called *cortical magnification* (Duncan & Boynton, 2003). It consists of a distortion of the retinal image with the purpose of focusing the neural resources on the crucial part of the scene (Born, Trott, & Hartmann, 2015). A similar idea is exploited by Plebe et al. (2021) for perception of the road scene. They implement a warped mapping of occupancy grids, obtaining a natural transition from egocentric to allocentric spatial representations of road objects.

5.3. Non-human brain

This review primarily focuses on methods for autonomous vehicles that draw inspiration from human cognition and brain function. However, while the ability to drive is crucial at higher levels of inspiration (behavioral and functional), it does not appear to be a requirement at lower levels, i.e., architectural and cellular inspirations. While all the papers reviewed in this Section assume a human brain, it is important to note that this is not a necessary condition in reality. Brain components like the cerebellum or basal ganglia, mechanisms such as neuromodulation, and neural organizations like convergence–divergence zones are common among non-human animals, particularly mammals. There is no specific driving-related development in the human brain that is specifically exploited in these papers. Nevertheless, the majority of these papers still assume the human brain as the source of imitation, with only a few exceptions.

The neural system of the nematode *Caenorhabditis Elegans* is used as a blueprint to design small artificial networks of 19 neurons called *neural circuit policies* (NCP) (Lechner et al., 2020; Milford, 2020). The small system of the *C. Elegans* is organized in a four-layer hierarchy made of sensory neurons, interneurons, command neurons, and motor neurons. A network of 19 neurons may seem ridiculous compared to the usual number of parameters in deep neural networks. However, the NCP neurons dramatically surpass the performance of traditional artificial neurons, because the neural dynamics in NCPs come from biologically plausible continuous-time ordinary differential equations. For this reason, NCPs can be successfully employed in complex systems, such as the one used by Lechner et al. (2020) for lane keeping control. Also, Ahmedov et al. (2021) use NCPs but in combination with a traditional convolutional neural network. They implement two CNN-NCP stacks to mimic the arrangement of the brain into two slightly asymmetric hemispheres.

6. Inspiration at cellular level

In the last category of works, inspiration traces back to the cellular level, i.e., the functioning of neurons in the human brain. This category encompasses the papers of class D in Table 2. The papers collected here draw inspiration from the mechanism of neuronal spikes, or action potentials, and from how the timing and pattern of spikes across neurons encode and convey information in the nervous system.

6.1. Spiking neural networks

Spiking neural networks (SNNs) (Izhikevich, 2003; Maass & Bishop, 1999) are the most consolidated framework for artificial neurons. They behave more similarly to biological neurons than deep neural networks and, at the same time, are simple enough to be deployed in engineering applications. The neurons in SNNs encode more information than traditional static neurons, combining spatial and temporal relations (Sougné, 2001). This advantage, however, is counterbalanced by the fact that SNNs are difficult to train because the neural transfer function is usually non-differentiable, and thus stochastic gradient descent cannot be used. For this reason, it is not easy to determine if SNNs are beneficial in any case (Pfeiffer & Pfeil, 2018). Adopting SNNs is frequently justified only by the fact that the networks are biologically plausible. Undivided attention to biological plausibility is reasonable when the SNN application also yields a better understanding of behaviors and phenomena occurring in the brain. This is not the case for autonomous driving, which requires more practical advantages.

Different approaches have been proposed to overcome the issue of training SNNs. In Kaiser et al. (2016), the authors design small networks with manually tuned connections. With this architecture, they realize a simple system for lane following evaluated in a simulated environment. Based on this seminal paper, the same research team (Kaiser et al., 2019) develops a learning rule using probabilistic sampling over a set of possible synaptic values (Kappel, Habenschuss, Legenstein, & Maass, 2015). The learning rule is implemented in an end-to-end system for lane following.

A popular solution to SNN learning derives from a feature of the human brain called *spike-timing-dependent plasticity* (STDP) (Markram, Gerstner, & Sjöström, 2011). This process adjusts the strength of connections between neurons in the brain on the basis of the relative timing of the neurons' spikes. Izhikevich (2007) proposes an artificial implementation of STDP modulated by reinforcement learning. The proposal is the basis for another learning strategy called *reward-modulated* STDP (R-STDP), adopted by Bing et al. (2018) in a system

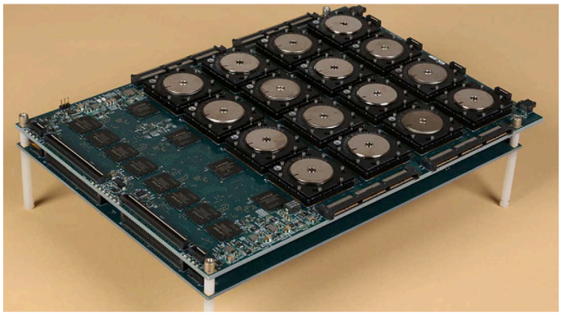

Fig. 5. A circuit board with 16 TrueNorth chips, designed by researchers at IBM under DARPA's Systems of Neuromorphic Adaptive Plastic Scalable Electronics (SyNAPSE) program (Srinivasa & Cruz-Albrecht, 2012).

for lane keeping and obstacle avoidance on a two-wheeled robot. A refined implementation is in Bing et al. (2020), which adopts a standard deep Q-learning algorithm for lane keeping and transfers the learned policy to the SNN. Liu et al. (2021) present another variant of R-STDP to train a system for target tracking and obstacle avoidance on a mobile robot. They first train a standard deep neural network, then convert the network into spiking neurons. In this way, however, the theoretical advantage of temporal coding is lost. Similarly, Nezhadalinaei et al. (2021) address object detection and tracking with a spiking convolutional network combined with conditional random fields and probabilistic particle filters. López-Randulfe et al. (2021) design a derivation of synaptic weights in SNN based on the discrete Fourier transform, for radar-based object detection.

An elegant way to escape from the issue of training SNNs is to avoid learning at all. A well-established neural framework for neurocognitive studies is *Neural Engineering Framework* (NEF) (Eliasmith, 2013; Eliasmith & Anderson, 2003), where groups of SNNs are organized in higher units called semantic pointers (Blouw, Solodkin, Thagard, & Eliasmith, 2015) that can be used without any training. Shalumov et al. (2021) adopt NEF for speed and steering control using LiDAR data.

Besides SNNs, there are other—less known—approaches coming from the domain of dynamical systems. For example, the *Feynman machine* (Laukien, Crowder, & Byrne, 2018) is a hierarchical temporal memory-based stack of spiking autoencoders that captures features of the cortex like incremental learning. It is used in Tenison et al. (2019) for path detection, with processor specifications so low that computation can be carried out on computers as small as a Raspberry Pi. Another source of inspiration comes from the famous *place cells* and *grid cells* in hippocampal formation (O'Keefe & Nadel, 1978; O'Keefe & Recce, 1993), which act as cognitive representations of specific locations in space. The works of Banino et al. (2018) and Taniguchi et al. (2021) adopt similar grid-cell representations for navigation and mapping applications.

6.2. Neuromorphic hardware

The main benefit of the papers presented in this category is to pave the way for the integration of neuromorphic hardware—the ideal complement to SNNs—into autonomous driving systems. Some works that experiment with neuromorphic hardware are already available. A relatively affordable neuromorphic chip is IBM's TrueNorth (Fig. 5), which features 1 million spiking neurons with 256 million synapses, for as low as 20 milliwatts per square centimeter power density. Fischl et al. (2017) employs the TrueNorth with a NEF model to realize a complete driving system for a 6-wheel small mobile robot. TrueNorth is also the choice of Kashyap (2020) for stereo depth scene reconstruction. Another small and low-cost neuromorphic chip is the μCaspian, with 256 neurons and 4096 synapses (Mitchell, Schuman, & Potok, 2020). Patton et al. (2021) use the μCaspian for autonomous driving in an F1/10 vehicle, a small-scale robotic car.

In addition to neuromorphic chips, neuromorphic cameras (also called *event cameras* or *Dynamic Vision Sensors*) are attracting increasing attention (Brandli, Berner, Yang, Liu, & Delbruck, 2014). The cameras are inspired by biological retinas and, instead of acquiring standard images at a fixed rate, they output pixel-wise changes of intensity asynchronously. The works of Fischl et al. (2017) and Kashyap (2020) mentioned above make use of neuromorphic cameras. Moreover, Maqueda et al. (2018) adopt event cameras to predict the steering angle of the ego vehicle. The goal of this work is not to develop a framework to actually control a vehicle; the work aims to show the advantages offered by neuromorphic cameras in the context of autonomous driving.

7. Discussion

Today, research on autonomous driving is carried out in large part using traditional methods from the fields of mechanical, electronic, and computer engineering—disciplines that are distant from the topics covered by neuroscience and cognitive science. However, it is a fact that fully autonomous vehicles are still far from being achieved (Jain, Del Pero, Grimmett, & Ondruska, 2021). This suggests that investigating alternative paradigms could be a promising direction, especially paradigms inspired by the only existing agents capable of driving—the human beings. Scientometric analyses (da Silva et al., 2020; Faisal, Yigitcanlar, Kamruzzaman, & Paz, 2021; Gandia et al., 2019) show how active research on autonomous driving is, especially since the last decade. At the same time, it is evident that cognitive- and brain-inspired studies are still marginal, accounting to about 1% of the overall research on autonomous driving.

The recent leap forward made by autonomous vehicles is due to advances in artificial intelligence, above all deep learning. Artificial intelligence is founded on the principle of imitating human intelligence. Artificial neural networks are also historically inspired by brain circuits—although this distant origin is irrelevant in most current applications. In fact, the majority of works reviewed here adopt deep neural networks, but this is no indication of the brain inspiration of the approach. There are, indeed, papers that present deep learning approaches that claim to be brain-inspired, although they do not refer to cognitive or neuroscientific principles (Chen, Chen, Zhang, & Hu, 2019; Huh & Hossain, 2021; Li & Gao, 2018; Wang & Xia, 2021).

When committing to the idea of taking cognitive or neurobiological inspiration, one has to face the problem of integrating the inspiration with the available state-of-the-art techniques, such as deep learning. Because of this, it is often challenging to compete on standard benchmarks and measure the benefit of the approach. Inspiration at the behavioral level, presented in Section 3, manifests a more tangible benefit than the other categories, despite being the category with less theoretical commitment. Methods inspired by external behaviors of human drivers merge efficiently with well-established processing pipelines, and they obtain competitive results compared to traditional engineering algorithms. However, imitation at the behavioral level hardly leads to radical innovations. On the other hand, inspiration at the cellular level, presented in Section 6, promotes pioneering approaches that are strongly committed to neurological facts. Since these approaches explore new research paths, they pay a price in terms of being competitive with traditional methods. In fact, many papers do not have comparative evaluations, and the most common benefit is the promise of drastically reducing energy consumption.

There have been some criticism that building an AI-system based on how a human *is* behaving is not the same (and not enough) as to build an AI-system that behaves as how the human *ought* to behave. According to Kim, Hooker, and Donaldson (2021), this may lead to unethical behavior and other unwanted consequences in AI-systems. This criticism is at least applicable to deep-learning approaches and the behaviorally inspired approaches at the data stage, as they are basically building the behavior of the autonomous vehicle from observations of how humans behave. To what extent it also applies to the approaches that take inspiration from cognitive abilities, brains, or neurons is less obvious and beyond the scope of this paper. It could also be noted that there are other criticisms such that biological systems have a kind of genuine understanding that is impossible to replicate in non-biological artificial systems which might also apply to all levels of inspiration (Ziemke & Sharkey, 2001). Another criticism is that current state-of-the-art AI techniques will have difficulties replicating some aspects of human cognition, in particular the ability to attribute to and understand mental states of other humans—the so called *Theory of Mind* (Aru, Labash, Corcoll, & Vicente, 2023). Indeed, the focus on humans and animals and their behavior might be misleading altogether. As noted by Forbus (2010): “Airplanes were created by a careful study of how aerodynamics worked, not by studying the details of birds. The deepest insights on how birds fly came ultimately from applying aerodynamic principles discovered while trying to create airplanes … ”. We cannot exclude the possibility that when/if we finally develop a fully autonomous vehicle, it might not have taken inspiration from humans.

The works analyzed here do not definitively establish unequivocal benefits from human inspirations. However, this review reveals a wide spectrum of promising sources of inspiration. Almost every cognitive function and area of the brain has captured the attention of research efforts. In this analysis, we have examined numerous components and mechanisms, including attention, memory organization, decision-making, emotions, consciousness, social cognition, the hippocampus, neuromodulation, the cerebellum, the amygdala, the orbito-frontal cortex, the basal ganglia, convergence–divergence zones, sparse representations, cortical magnification, the neural circuitry of the *C. Elegans*, spiking neural networks, and spike-timing-dependent plasticity.

7.1. Future directions

Given the current scenario surveyed in this paper and discussed in the previous Section, it is challenging to outline the future directions that human-inspired research on autonomous driving may take. It is even more arduous to predict whether research driven by cognitive science will ultimately yield long-term benefits. This review does not attempt to engage in excessive speculation; instead, we offer some observations.

On one hand, the current competition for a successful self-driving system compels manufacturers to focus on immediate improvements that can be achieved more easily through non-technological changes, such as increasing datasets. From this short-term perspective, a less demanding form of human imitation, such as behavioral cloning, can be included. On the other hand, it appears that major paradigm breakthroughs are necessary to achieve full autonomous driving, and the human-inspired approach is undoubtedly a strong contender since humans are currently the only agents driving in full autonomy.

There are indications that the level of inspiration that seems to perform the least—the cellular level—may have the best long-term prospects. Observers have pointed to neuromorphic computing as the future of various applications, including autonomous vehicles (Roy, Jaiswal, & Panda, 2019; Schuman et al., 2022). Moreover, inspiration at the functional and architectural levels, while still limited in number, has the advantage of drawing inspiration from a wide spectrum of sources, as discussed in the previous Section. The breadth of exploration at these levels is a promising factor for the emergence of a research direction that garners a critical mass of attention. It is by reaching this critical volume that medium-term success becomes possible. In conclusion, it is worthwhile to pursue research in the direction of autonomous vehicles endowed with high-level cognitive capabilities directly inspired by humans.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

No data was used for the research described in the article.

Acknowledgments

During the preparation of this work the authors used ChatGPT in order to grammar-check part of the manuscript. After using this tool/service, the authors reviewed and edited the content as needed and take full responsibility for the content of the publication.

References

Abbeel, P., & Ng, A. Y. (2004). Apprenticeship learning via inverse reinforcement learning. In *International conference on machine learning*.

Ahmedov, H. B., Yi, D., & Sui, J. (2021). Application of a brain-inspired deep imitation learning algorithm in autonomous driving[Formula presented]. *Software Impacts*, *10*(October), Article 100165.

Anderson, J. R., & Lebiere, C. J. (1998). *The atomic components of thought*. East Sussex (UK): Psychology Press.

Aradi, S. (2020). Survey of deep reinforcement learning for motion planning of autonomous vehicles. *IEEE Transactions on Intelligent Transportation Systems*.

Aru, J., Labash, A., Corcoll, O., & Vicente, R. (2023). Mind the gap: Challenges of deep learning approaches to theory of mind. *Artificial Intelligence Review*, 1–16.

Augoustinos, M., Walker, I., & Donaghue, N. (1995). *Social cognition: An integrated introduction*. Thousand Oaks (CA): SAGE Publications.

Avery, M. C., & Krichmar, J. L. (2017). Neuromodulatory systems and their interactions: A review of models, theories, and experiments. *Frontiers in Neural Circuits*, *11*, 108.

Baddeley, A. (1992). Working memory. *Science*, *255*, 556–559.

Badue, C., Guidolini, R., Carneiro, R. V., Azevedo, P., Cardoso, V. B., Forechi, A., et al. (2021). Self-driving cars: A survey. *Expert Systems with Applications*, *165*, Article 113816.

Bagdatli, M. E. C., & Dokuz, A. S. (2021). Modeling discretionary lane-changing decisions using an improved fuzzy cognitive map with association rule mining. *Transportation Letters*, *13*(8), 623–633.

Banino, A., Barry, C., Uria, B., Blundell, C., Lillicrap, T., Mirowski, P., et al. (2018). Vector-based navigation using grid-like representations in artificial agents. *Nature*, *557*, 429–433.

Beyeler, M., Rounds, E. L., Carlson, K. D., Dutt, N., & Krichmar, J. L. (2019). Neural correlates of sparse coding and dimensionality reduction. *PLoS Computational Biology*, *15*, Article e1006908.

Bing, Z., Meschede, C., Chen, G., Knoll, A., & Huang, K. (2020). Indirect and direct training of spiking neural networks for end-to-end control of a lane-keeping vehicle. *Neural Networks*, *121*, 21–36.

Bing, Z., Meschede, C., Huang, K., Chen, G., Röhrbein, F., Akl, M., et al. (2018). End to End Learning of Spiking Neural Network based on R-STDP for a Lane Keeping Vehicle. In *IEEE international conference on robotics and automation (ICRA)* (pp. 4725–4732).

Blouw, P., Solodkin, E., Thagard, P., & Eliasmith, C. (2015). Concepts as semantic pointers: A framework and computational model. *Cognitive Science*, 1–35.

Born, R., Trott, A. R., & Hartmann, T. S. (2015). Cortical magnification plus cortical plasticity equals vision? *Vision Research*, *111*, 161–169.

Brandli, C., Berner, R., Yang, M., Liu, S.-C., & Delbruck, T. (2014). A 240 × 180 130 dB 3 μs latency global shutter spatiotemporal vision sensor. *IEEE Journal of Solid-state Circuits*, *49*, 2333–2341.

Braun, J., Koch, C., & Davis, J. L. (2001). *Visual attention and cortical circuits*. Cambridge (MA): MIT Press.

Buckman, N., Pierson, A., Schwarting, W., Karaman, S., & Rus, D. (2019). Sharing is Caring: Socially-Compliant Autonomous Intersection Negotiation. *IEEE international conference on intelligent robots and systems*, 6136–6143.

Butt, M. A., Riaz, F., Mehmood, Y., & Akram, S. (2021). REEEC-AGENT: human driver cognition and emotions-inspired rear-end collision avoidance method for autonomous vehicles. *Simulation*, *97*(9), 601–617.

Carrasco, M. (2011). Visual attention: The past 25 years. *Vision Research*, *51*, 1484–1525.

Chalmers, D. (1996). *The conscious mind: In search of a fundamental theory*. Oxford (UK): Oxford University Press.

Chan, M., Partouche, D., & Pasquier, M. (2007). An intelligent driving system for automatically anticipating and negotiating road curves. In *IEEE international conference on intelligent robots and systems* (pp. 117–122).

Chen, J., Chen, J., Zhang, R., & Hu, X. (2019). Toward a brain-inspired system: Deep recurrent reinforcement learning for a simulated self-driving agent. *Frontiers in Neurorobotics*, *13*(June), 1–9, arXiv:1903.12517.

Chen, J., Sun, D., Zhao, M., Li, Y., & Liu, Z. (2021). A New Lane Keeping Method Based on Human-Simulated Intelligent Control. *IEEE Transactions on Intelligent Transportation Systems*, 1–12.

Chen, S., Zhang, S., Shang, J., Chen, B., & Zheng, N. (2017). Brain-inspired cognitive model with attention for self-driving cars. *IEEE Transactions on Cognitive and Developmental Systems*, *11*(1), 13–25.

Codevilla, F., Müller, M., López, A., Koltun, V., & Dosovitskiy, A. (2018). End-to-end Driving via Conditional Imitation Learning. In *IEEE international conference on robotics and automation (ICRA)* (pp. 4693–4700).

Czubenko, M., Kowalczuk, Z., & Ordys, A. (2015). Autonomous Driver Based on an Intelligent System of Decision-Making. *Cognitive Computation*, *7*(5), 569–581.

Da Lio, M., Donà, R., Rosati Papini, G. P., Biral, F., & Svensson, H. (2020). A Mental Simulation Approach for Learning Neural-Network Predictive Control (in Self-Driving Cars). *IEEE Access*, *8*, 192041–192064. http://dx.doi.org/10.1109/ACCESS.2020.3032780.

Da Lio, M., Donà, R., Rosati Papini, G. P., & Gurney, K. (2020). Agent architecture for adaptive behaviors in autonomous driving. *IEEE Access*, *8*, 154906–154923.

Da Lio, M., Donà, R., Rosati Papini, G. P., & Plebe, A. (2022). The biasing of action selection produces emergent human-robot interactions in autonomous driving. *IEEE Robotics and Automation Letters*, *7*(2), 1254–1261.

da Silva, J. P. N., Gandia, R. M., Sugano, J. Y., Oliveira, C. C., Antonialli, F., Pedrosa, G., et al. (2020). From technology to market: a bibliometric and integrative review on autonomous vehicles. *World Review of Intermodal Transportation Research*, *9*, 1–26.

Damasio, A. (1989). The brain binds entities and events by multiregional activation from convergence zones. *Neural Computation*, *1*, 123–132.

Damasio, A. (1994). *Descartes' error: Emotion, reason and the human brain*. New York: Avon Books.

Demiris, Y., Aziz-Zadeh, L., & Bonaiuto, J. (2014). Information processing in the mirror neuron system in primates and machines. *Neuroinformatics*, *12*, 63–91.

Dennett, D. C. (1992). *Consciousness explained*. New York: Back Bay Books.

Devi, S., Malarvezhi, P., Dayana, R., & Vadivukkarasi, K. (2020). A comprehensive survey on autonomous driving cars: A perspective view. *Wireless Personal Communications*, *114*(3), 2121–2133.

Ding, Z. (2020). Imitation learning. In H. Dong, Z. Ding, & S. Zhang (Eds.), *Deep reinforcement learning: Fundamentals, research and applications* (pp. 273–675). Berlin: Springer-Verlag.

Droniou, A., Ivaldi, S., & Sigaud, O. (2015). Deep unsupervised network for multimodal perception, representation and classification. *Robotics and Autonomous Systems*, *71*, 83–98.

Duncan, R. O., & Boynton, G. M. (2003). Cortical magnification within human primary visual cortex correlates with acuity thresholds. *Neuron*, *38*, 659–671.

Eliasmith, C. (2013). *How to build a brain: A neural architecture for biological cognition*. Oxford (UK): Oxford University Press.

Eliasmith, C., & Anderson, C. H. (2003). *Neural engineering computation, representation, and dynamics in neurobiological systems*. Cambridge (MA): MIT Press.

Eraqi, H. M., Moustafa, M. N., & Honer, J. (2022). Dynamic conditional imitation learning for autonomous driving. *IEEE Transactions on Intelligent Transportation Systems*, *23*, 22988–23001.

Faisal, A., Yigitcanlar, T., Kamruzzaman, M., & Paz, A. (2021). Mapping two decades of autonomous vehicle research: A systematic scientometric analysis. *Journal of Urban Technology*, *28*, 45–74.

Fischl, K. D., Tognetti, G., Mendat, D. R., Orchard, G., Rattray, J., Sapsanis, C., et al. (2017). Neuromorphic Self-Driving Robot with Retinomorphic Vision and Spike-Based Processing / Closed-Loop Control. In *Annual conference on information sciences and systems (CISS)*.

Forbus, K. D. (2010). AI and cognitive science: The past and next 30 years. *Topics in Cognitive Science*, *2*(3), 345–356.

Gandia, R. M., Antonialli, F., Cavazza, B. H., Neto, A. M., de Lima, D. A., Sugano, J. Y., et al. (2019). Autonomous vehicles: scientometric and bibliometric review. *Transport Reviews*, *39*, 9–28.

Gonsalves, T., & Upadhyay, J. (2021). Integrated deep learning for self-driving robotic cars. *Artificial Intelligence for Future Generation Robotics*, 93–118.

Grigorescu, S., Trasnea, B., Cocias, T., & Macesanu, G. (2020). A survey of deep learning techniques for autonomous driving. *Journal of Field Robotics*, *37*(3), 362–386.

Gu, Y., Hashimoto, Y., Hsu, L.-t., Iryo-asano, M., & Kamijo, S. (2017). Human-like motion planning model for driving in signalized intersections. *IATSS Research*, *41*(3), 129–139.

Guest, O., & Martin, A. E. (2021). How computational modeling can force theory building in psychological science. *Perspectives on Psychological Science*, *16*(4), 789–802.

Gurney, K., Prescott, T. J., & Redgrave, P. (2001). A computational model of action selection in the basal ganglia. II analysis and simulation of behaviour. *Biological Cybernetics*, *84*, 411–423.

Ha, D., & Schmidhuber, J. (2018). World Models. arxiv, arXiv:arXiv:1803.10122v4.

Haber, S. N. (2011). Neural circuits of reward and decision making: Integrative networks across corticobasal ganglia loops. In R. B. Mars, J. Sallet, M. F. S. Rushworth, & N. Yeung (Eds.), *Neural basis of motivational and cognitive control* (pp. 22–35). Cambridge (MA): MIT Press.

Hang, P., Lv, C., Xing, Y., Huang, C., & Hu, Z. (2021). Human-Like Decision Making for Autonomous Driving: A Noncooperative Game Theoretic Approach. *IEEE Transactions on Intelligent Transportation Systems*, *22*, 2076–2087, arXiv:2005.11064.

Hawke, J., Shen, R., Gurau, C., Sharma, S., Reda, D., Mazur, N. N. P., et al. (2020). Urban driving with conditional imitation learning. In *international conference on robotics and automation* (pp. 251–257).

Hecker, S., Dai, D., Liniger, A., Hahner, M., & Gool, L. V. (2020). Learning Accurate and Human-Like Driving using Semantic Maps and Attention. In *IEEE/RSJ international conference on intelligent robots and systems (IROS)* (pp. 2346–2353).

Ho, J., & Ermon, S. (2016). Generative adversarial imitation learning. In D. D. Lee, M. Sugiyama, U. V. Luxburg, I. Guyon, & R. Garnett (Eds.), *Advances in neural information processing systems*. Curran Associates, Inc.

Hu, H. (2016). Reward and aversion. *Annual Review of Neuroscience*, *39*, 297–324.

Huh, E. N., & Hossain, M. I. (2021). Brainware computing: Concepts, scopes and challenges. *Applied Sciences*, *11*(11).

Hurley, S. (2001). Perception and action: Alternative views. *Synthese*, *129*, 3–40.

Hussain, A., Abdullah, R., Yang, E., & Gurney, K. (2012). An Intelligent Multiple-Controller Framework for the Integrated Control of Autonomous Vehicles. In *International conference on brain inspired cognitive systems* (pp. 92–101).

Hussain, A., Gurney, K., Abdullah, R., & Chambers, J. (2008). Emergent common functional principles in control theory and the vertebrate brain: A case study with autonomous vehicle control. In *International conference on artificial neural networks* (pp. 949–958). Springer.

Iacoboni, M., Woods, R. P., Brass, M., Bekkering, H., Mazziotta, J. C., & Rizzolatti, G. (1999). Cortical mechanisms of human imitation. *Science*, *286*, 2526–2528.

Ito, M. (1984). *The cerebellum and neural control*. New York: Raven Press.

Izhikevich, E. M. (2003). Simple model of spiking neurons. *IEEE Transactions on Neural Networks*, *14*, 1569–1572.

Izhikevich, E. M. (2007). Solving the distal reward problem through linkage of STDP and dopamine signaling. *Cerebral Cortex*, *17*, 2443–2452.

Jain, A., Del Pero, L., Grimmett, H., & Ondruska, P. (2021). Autonomy 2.0: Why is self-driving always 5 years away? arXiv preprint arXiv:2107.08142.

Kaiser, J., Hoff, M., Konle, A., Tieck, J. C. V., Kappel, D., Reichard, D., et al. (2019). Embodied Synaptic Plasticity With Online Reinforcement Learning. *Frontiers in Neurorobotics*, *13*(October), 1–11.

Kaiser, J., Tieck, J. C. V., Hubschneider, C., Wolf, P., Weber, M., Hoff, M., et al. (2016). Towards a framework for end-to-end control of a simulated vehicle with spiking neural networks. In *IEEE international conference on simulation, modeling, and programming for autonomous robots* (pp. 127–134).

Kamath, U., & Liu, J. (2021). *Explainable artificial intelligence: An introduction to interpretable machine learning*. Berlin: Springer-Verlag.

Kappel, D., Habenschuss, S., Legenstein, R., & Maass, W. (2015). Network plasticity as Bayesian inference. *PLoS Computational Biology*, *11*, Article e1004485.

Kashyap, H. J. (2020). *Brain inspired neural network models of visual motion perception and tracking in dynamic scenes* (Ph.D. thesis), UC Irvine.

Kashyap, H. J., Fowlkes, C. C., Krichmar, J. L., & Member, S. (2021). Sparse Representations for Object- and Ego-Motion Estimations in Dynamic Scenes. *IEEE Transactions on Neural Networks and Learning Systems*, *32*(6), 2521–2534.

Kim, T. W., Hooker, J., & Donaldson, T. (2021). Taking principles seriously: A hybrid approach to value alignment in artificial intelligence. *Journal of Artificial Intelligence Research*, *70*, 871–890.

Kim, C., & Langari, R. (2013). Application of brain limbic system to adaptive cruise control. *International Journal of Vehicle Autonomous Systems*, *11*(1), 22–41.

Kim, J., Rohrbach, A., Darrell, T., Canny, J., & Akata, Z. (2018). Textual explanations for self-driving vehicles. In *Proc. of european conference on computer vision* (pp. 563–578).

Kiran, B. R., Sobh, I., Talpaert, V., Mannion, P., Al Sallab, A. A., Yogamani, S., et al. (2021). Deep reinforcement learning for autonomous driving: A survey. *IEEE Transactions on Intelligent Transportation Systems*.

Koeberle, Y., Sabatini, S., Tsishkou, D., Sabourin, C., Koeberle, Y., Sabatini, S., et al. (2021). Learning human like driving policies from real interactive driving scenes. *HAL Open Science*.

Kosko, B. (1986). Fuzzy cognitive maps. *International Journal of Man–Machine Studies*, *24*, 65–75.

Kuefler, A., Morton, J., Wheeler, T., & Kochenderfer, M. (2017). Imitating Driver Behavior with Generative Adversarial Networks. In *IEEE intelligent vehicles symposium (IV)*. IEEE.

Kumaar, S., Navaneethkrishnan, B., Hegde, S., Raja, P., & Vishwanath, R. M. (2019). Towards Behavioural Cloning for Autonomous Driving. In *IEEE international conference on robotic computing (IRC)* (pp. 560–567).

Kuutti, S., Bowden, R., Jin, Y., Barber, P., & Fallah, S. (2020). A survey of deep learning applications to autonomous vehicle control. *IEEE Transactions on Intelligent Transportation Systems*, *22*(2), 712–733.

Laukien, E., Crowder, R., & Byrne, F. (2018). Feynman machine: A novel neural architecture for cortical and machine intelligence. In *AAAI conference on artificial intelligence* (pp. ss–17–07).

Law, H., Teng, Y., Russakovsky, O., & Deng, J. (2020). CornerNet-Lite: Efficient Keypoint Based Object Detection. arXiv:1904.08900,

Lechner, M., Hasani, R., Amini, A., Henzinger, T. A., Rus, D., & Grosu, R. (2020). Neural circuit policies enabling auditable autonomy. *Nature Machine Intelligence*, *2*(10), 642–652.

LeDoux, J. E. (2000). Emotion circuits in the brain. *Annual Review of Neuroscience*, *23*, 155–184.

Li, T.-H. S., Chang, S.-J., & Chen, Y.-X. (2003). Implementation of Human-Like Driving Skills by Autonomous Fuzzy Behavior Control on an FPGA-Based Car-Like Mobile Robot. *IEEE Transactions on Industrial Electronics*, *50*(5), 867–880.

Li, D., & Gao, H. (2018). A Hardware Platform Framework for an Intelligent Vehicle Based on a Driving Brain. *Engineering*, *4*(4), 464–470.

Li, L., Ota, K., & Dong, M. (2018). Humanlike Driving: Empirical Decision-Making System for Autonomous Vehicles. *IEEE Transactions on Vehicular Technology*, *67*(8), 6814–6823.

Liang, X., Wang, T., Yang, L., & Xing, E. (2018). CIRL : Controllable Imitative Reinforcement Learning for Vision-based Self-driving. In *European conference on computer vision (ECCV)* (pp. 584–599).

Lindsay, G. W. (2020). Attention in psychology, neuroscience, and machine learning. *Frontiers in Computational Neuroscience*, *14*, Article 29.

Liu, J., Lu, H., Luo, Y., & Yang, S. (2021). Spiking neural network-based multi-task autonomous learning for mobile robots. *Engineering Applications of Artificial Intelligence*, *104*(June), Article 104362.

Lopes, M., Melo, F., Montesano, L., & Santos-Victor, J. (2010). Abstraction levels for robotic imitation: Overview and computational approaches. In O. Sigaud, & J. Peters (Eds.), *From motor learning to interaction learning in robots* (pp. 313–355). Berlin: Springer-Verlag.

López-Randulfe, J., Duswald, T., Bing, Z., & Knoll, A. (2021). Spiking Neural Network for Fourier Transform and Object Detection for Automotive Radar. *Frontiers in Neurorobotics*, *15*(June), 1–7.

Maass, W., & Bishop, C. M. (Eds.), (1999). Pulsed neural networks. *Pulsed neural networks*. Cambridge (MA): MIT Press.

Maqueda, A. I., Loquercio, A., Gallego, G., García, N., & Scaramuzza, D. (2018). Event-based Vision meets Deep Learning on Steering Prediction for Self-driving Cars. In *IEEE conference on computer vision and pattern recognition (CVPR)* (pp. 5419–5427).

Markelić, I., Kjær-Nielsen, A., Pauwels, K., Baunegaard With Jensen, L., Chumerin, N., Vidugiriene, A., et al. (2011). The driving school system: Learning basic driving skills from a teacher in a real car. *IEEE Transactions on Intelligent Transportation Systems*, *12*(4), 1135–1146.

Markram, H., Gerstner, W., & Sjöström, P. J. (2011). A history of spike-timing-dependent plasticity. *Frontiers in Synaptic Neuroscience*, *3*, 4.

Meyer, K., & Damasio, A. (2009). Convergence and divergence in a neural architecture for recognition and memory. *Trends in Neuroscience*, *32*, 376–382.

Milford, M. (2020). C. Elegans inspires self-driving cars. *Nature Machine Intelligence*, *2*(11), 661–662.

Mitchell, J. P., Schuman, C. D., & Potok, T. E. (2020). A small, low cost event-driven architecture for spiking neural networks on FPGAs. In *International conference on neuromorphic systems*.

Nezhadalinaei, F., Zhang, L., Mahdizadeh, M., & Jamshidi, F. (2021). Motion object detection and tracking optimization in autonomous vehicles in specific range with optimized deep neural network. In *2021 7th international conference on web research (ICWR)* (pp. 53–63). IEEE.

Ng, A. Y., & Russell, S. J. (2000). Algorithms for inverse reinforcement learning. In *International conference on machine learning*.

O'Keefe, J., & Nadel, L. (1978). *The hippocampus as a cognitive map*. Oxford (UK): Oxford University Press.

O'Keefe, J., & Recce, M. (1993). Phase relationship between hippocampal place units and the EEG theta rhytm. *Hippocampus*, *3*, 317–330.

Olier, J. S., Barakova, E., Regazzoni, C., & Rauterberg, M. (2017). Re-framing the characteristics of concepts and their relation to learning and cognition in artificial agents. *Cognitive Systems Research*, *44*, 50–68.

Omeiza, D., Anjomshoae, S., Webb, H., Jirotka, M., & Kunze, L. (2022). From spoken thoughts to automated driving commentary: Predicting and explaining intelligent vehicles' actions. In *IEEE intelligent vehicles symposium* (pp. 1040–1047).

Ortony, A., Clore, G. L., & Collins, A. (1990). *The cognitive structure of emotions*. Cambridge (UK): Cambridge University Press.

Ouyang, Z., Cui, J., Dong, X., Li, Y., & Niu, J. (2022). SaccadeFork: A lightweight multi-sensor fusion-based target detector. *Information Fusion*, *77*(August 2021), 172–183.

Paden, B., Čáp, M., Yong, S. Z., Yershov, D., & Frazzoli, E. (2016). A survey of motion planning and control techniques for self-driving urban vehicles. *IEEE Transactions on intelligent vehicles*, *1*(1), 33–55.

Pal, S., Mohandoss, T., & Mitra, P. (2019). Visual Attention for Behavioral Cloning in Autonomous Driving. In *International conference on machine vision (ICMV)*.

Pasquier, M., & Oentaryo, R. J. (2008). Learning to drive the human way: a step towards intelligent vehicles. *International Journal of Vehicle Autonomous Systems*, *6*(1–2), 24–47.

Patton, R., Schuman, C., Kulkarni, S., Parsa, M., Mitchell, J. P., Haas, N. Q., et al. (2021). Neuromorphic computing for autonomous racing. In *International conference on neuromorphic systems 2021* (pp. 1–5).

Pendleton, S. D., Andersen, H., Du, X., Shen, X., Meghjani, M., Eng, Y. H., et al. (2017). Perception, planning, control, and coordination for autonomous vehicles. *Machines*, *5*(1), 6.

Pfeiffer, M., & Pfeil, T. (2018). Deep learning with spiking neurons: Opportunities and challenges. *Frontiers in Neuroscience*, *12*, 774.

Plebe, A., & Da Lio, M. (2020). On the road with 16 neurons: Towards interpretable and manipulable latent representations for visual predictions in driving scenarios. *IEEE Access*, *8*, 179716–179734.

Plebe, A., Kooij, J. F., Rosati Papini, G. P., & Da Lio, M. (2021). Occupancy grid mapping with cognitive plausibility for autonomous driving applications. In *Proceedings of the IEEE/CVF international conference on computer vision (ICCV)* (pp. 2934–2941).

Plebe, A., Rosati Papini, G. P., Cherubini, A., & Da Lio, M. (2022). Distributed cognition for collaboration between human drivers and self-driving cars. *Frontiers in Artificial Intelligence*, *5:910801*.

Plebe, A., Rosati Papini, G. P., Donà, R., & Da Lio, M. (2019). Dreaming mechanism for training bio-inspired driving agents. In *Proceedings of the 2nd international conference on intelligent human systems integration (IHSI)* (pp. 429–434). Springer, Cham.

Qiao, Z., Schneider, J., & Dolan, J. M. (2021). Behavior Planning at Urban Intersections through Hierarchical Reinforcement Learning. In *IEEE international conference on robotics and automation (ICRA)* (pp. 2667–2673). IEEE, arXiv:2011.04697.

Rhinehart, N., McAllister, R., & Levine, S. (2019). Deep imitative models for flexible inference, planning, and control. In *International conference on learning representations*. arXiv:arXiv:1810.06544v4.

Riaz, F., & Niazi, M. A. (2018). Enhanced emotion enabled cognitive agent-based rear-end collision avoidance controller for autonomous vehicles. *Simulation*, *94*(11), 957–977.

Rizzolatti, G., Fogassi, L., & Gallese, V. (2001). Neurophysiological mechanisms underlying the understanding and imitation of action. *Nature Reviews Neuroscience*, *2*, 661–670.

Rizzolatti, G., & Sinigaglia, C. (2008). *Mirrors in the brain: How our minds share actions and emotions*. Oxford (UK): Oxford University Press.

Rolls, E., & Tovee, M. J. (1995). Sparseness of the neuronal representation of stimuli in the primate temporal visual cortex. *Journal of Neurophysiology*, *73*, 713–726.

Roy, K., Jaiswal, A., & Panda, P. (2019). Towards spike-based machine intelligence with neuromorphic computing. *Nature*, *575*, 607–617.

Saeed, Y., Ahmed, K., Zareei, M., Zeb, A., Vargas-Rosales, C., & Awan, K. M. (2019). In-Vehicle Cognitive Route Decision Using Fuzzy Modeling and Artificial Neural Network. *IEEE Access*, *7*, 20262–20272.

Salvucci, D. D. (2006). Modeling Driver Behavior in a Cognitive Architecture. *Human Factors*, *48*(2), 362–380.

Schaal, S. (1999). Is imitation learning the route to humanoid robots? *Trends in Cognitive Sciences*, *3*, 233–242.

Schulz, Y., Mattar, A. K., Hehn, T. M., & Kooij, J. F. (2021). Hearing What You Cannot See: Acoustic Vehicle Detection around Corners. *IEEE Robotics and Automation Letters*, *6*(2), 2587–2594, arXiv:2007.15739.

Schuman, C. D., Kulkarni, S. R., Parsa, M., Mitchell, J. P., Date, P., & Kay, B. (2022). Opportunities for neuromorphic computing algorithms and applications. *Nature Computational Science*, *2*, 10–19.

Shalumov, A., Halaly, R., & Tsur, E. E. (2021). LiDAR-driven spiking neural network for collision avoidance in autonomous driving. *Bioinspiration and Biomimetics*, *16*(6).

Sharma, S., Tewolde, G., & Kwon, J. (2018). Behavioral Cloning for Lateral Motion Control of Autonomous Vehicles using Deep Learning. In *IEEE international conference on electro/information technology (EIT)* (pp. 228–233). IEEE.

Singh, S. (2015). *Critical Reasons for Crashes Investigated in the National Motor Vehicle Crash Causation Survey*: *Technical Report DOT HS 812 115*, Washington (DC): National Highway Traffic Safety Administration.

Smith, E. E., & Kosslyn, S. M. (2013). *Cognitive psychology: Mind and brain: Pearson new international edition*. Pearson Higher Ed.

Sougné, J. (2001). Binding and multiple instantiation in a distributed network of spiking nodes. *Connection Science*, *13*, 99–126.

Sprenger, F. (2020). Microdecisions and autonomy in self-driving cars: virtual probabilities. *AI & SOCIETY*, 1–16.

Srinivasa, N., & Cruz-Albrecht, J. M. (2012). Neuromorphic adaptive plastic scalable electronics: Analog learning systems. *IEEE Pulse*, *3*(1), 51–56. http://dx.doi.org/10.1109/MPUL.2011.2175639.

Sun, L., Zhan, W., Tomizuka, M., & Dragan, A. D. (2018). Courteous Autonomous Cars. In *IEEE/RSJ international conference on intelligent robots and systems (IROS)*.

Suresh, P., & Manivannan, P. V. (2017). Human driver emulation and cognitive decision making for autonomous cars. In *Proceedings of 2016 international conference on robotics: Current trends and future challenges (RCTFC)* (pp. 1–6). IEEE.

Tani, J., & Ito, M. (2004). Self-organization of distributedly represented multiple behavior schemata in a mirror system: Reviews of robot experiments using RNNPB. *Neural Networks*, *17*, 1273–1289.

Taniguchi, A., Fukawa, A., & Yamakawa, H. (2021). Hippocampal formation-inspired probabilistic generative model. arXiv:2103.07356,

Teng, S., Chen, L., Ai, Y., Zhou, Y., Xuanyuan, Z., & Hu, X. (2023). Hierarchical interpretable imitation learning for end-to-end autonomous driving. *IEEE Transactions on Intelligent Vehicles*, *8*, 673–683.

Tenison, I., Bharathan, R. P., Kurian, D., Rajan, J. S., Imthias, A. T., & Muhammedali, S. K. (2019). Feynman Machine : A Cortical Machine Intelligence for Path Detection in Advanced Driver-Assistance Systems. *IEEE Region 10 Conference (TENCON)*, 246–250.

Torabi, F., Warnell, G., & Stone, P. (2018). Behavioral cloning from observation. In *International joint conferences on artificial intelligence* (pp. 4950–4957).

Tye, M. (1995). *Ten problems of consciousness*. Cambridge (MA): MIT Press.

Wang, L., & Xia, Y. (2021). Artificial intelligence brain. In *International conference on computer engineering and artificial intelligence* (pp. 266–270).

Wei, B., Ren, M., Zeng, W., Liang, M., Yang, B., & Urtasun, R. (2021). Perceive, Attend, and Drive: Learning Spatial Attention for Safe Self-Driving. In *IEEE international conference on robotics and automation (ICRA)* (pp. 4875–4881). IEEE, arXiv:2011.01153.

Wiedermann, J., & Leeuwen, J. V. (2021). Autonomous Vehicles that Cooperate and Understand: Intelligent Algorithms under the Hood. *Technical Report January*, Utrecht University, Utrecht, The Netherlands.

Windridge, D., Svensson, H., & Thill, S. (2021). On the utility of dreaming: A general model for how learning in artificial agents can benefit from data hallucination. *Adaptive Behavior*, *29*(3), 267–280, arXiv:arXiv:1803.10122.

Xie, S., Chen, S., Tomizuka, M., Zheng, N., & Wang, J. (2020). To Develop Human-like Automated Driving Strategy Based on Cognitive Construction: Appraisal and Perspective. In *IEEE international conference on intelligent transportation systems (ITSC)*.

Xing, J., Zou, X., & Krichmar, J. L. (2020). Neuromodulated Patience for Robot and Self-Driving Vehicle Navigation. In *International joint conference on neural networks (IJCNN)*.

Xu, D., Ding, Z., He, X., Zhao, H., Moze, M., Aioun, F., et al. (2021). Learning from naturalistic driving data for human-like autonomous highway driving. *IEEE Transactions on Intelligent Transportation Systems*, *22*(12), 7341–7354.

Yang, E., Hussain, A., & Gurney, K. (2012). Neurobiologically-inspired soft switching control of autonomous vehicles. *7366 LNAI*, In *International conference on brain inspired cognitive systems* (pp. 82–91).

Yang, E., Hussain, A., & Gurney, K. (2013). A Basal Ganglia Inspired Soft Switching Approach to the Motion Control of a Car-Like Autonomous Vehicle. In *International conference on brain inspired cognitive systems* (pp. 245–254).

Zakaria, N. (2021). Neuronal Unit of Thoughts (NUTs); A Probabilistic Formalism for Higher-Order Cognition. *Lecture Notes in Networks and Systems*, *204*, 855–871.

Zhang, J., & Ohn-Bar, E. (2021). Learning by Watching. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition* (pp. 12706–12716). arXiv:2106.05966.

Zhang, X., Zhou, M., Liu, H., & Hussain, A. (2019). A Cognitively Inspired System Architecture for the Mengshi Cognitive Vehicle. *Cognitive Computation*.

Zhou, Q., & Bai, J. (1983). An intelligent controller of novel design. In *Proceedings of a multi-national instrument conference* (pp. 137–149). Shanghai (China).

Ziemke, T., & Sharkey, N. E. (2001). A stroll through the worlds of robots and animals: Applying jakob von uexkülls theory of meaning to adaptive robots and artificial life. *Semiotica*.