

Received August 27, 2021, accepted September 26, 2021, date of publication September 29, 2021, date of current version October 7, 2021.

Digital Object Identifier 10.1109/ACCESS.2021.3116303

# Behavior Prediction of Traffic Actors for Intelligent Vehicle Using Artificial Intelligence Techniques: A Review

SURESH KOLEKAR<sup>1</sup>, SHILPA GITE<sup>1,2</sup>, BISWAJEET PRADHAN<sup>3,4,5</sup>, (Senior Member, IEEE), AND KETAN KOTECHA<sup>1,2</sup>

<sup>1</sup>Computer Science and Information Technology Department, Symbiosis Institute of Technology, Symbiosis International (Deemed) University, Pune 412115, India

<sup>2</sup>Symbiosis Centre of Applied AI (SCAAI), Symbiosis International (Deemed) University, Pune 412115, India

<sup>3</sup>Centre for Advanced Modelling and Geospatial Information Systems (CAMGIS), Faculty of Engineering and IT, University of Technology Sydney, Ultimo, NSW 2007, Australia

<sup>4</sup>Center of Excellence for Climate Change Research, King Abdulaziz University, Jeddah 21589, Saudi Arabia

<sup>5</sup>Earth Observation Centre, Institute of Climate Change, Universiti Kebangsaan Malaysia (UKM), Bangi, Selangor 43600, Malaysia

Corresponding authors: Shilpa Gite (shilpa.gite@sitpune.edu.in), Biswajeet Pradhan (biswajeet.pradhan@uts.edu.au), and Ketan Kotecha (head@scaai.siu.edu.in)

This work was supported by the Centre for Advanced Modelling and Geospatial Information Systems (CAMGIS), Faculty of Engineering and IT, University of Technology Sydney (UTS).

**ABSTRACT** Intelligent vehicle technology has made tremendous progress due to Artificial Intelligence (AI) techniques. Accurate behavior prediction of surrounding traffic actors is essential for the safe and secure navigation of the intelligent vehicle. Minor misbehavior of these vehicles on the busy roads may lead to an accident. Due to this, there is a need for vehicle behavior research work in today's era. This research article reviews traffic actors' behavior prediction techniques for intelligent vehicles to perceive, infer, and anticipate other vehicles' intentions and future actions. It identifies the key strategies and methods for AI, emerging trends, datasets, and ongoing research issues in these fields. As per the authors' knowledge, this is the first systematic literature review dedicated to the vehicle behavior study examining existing academic literature published by peer review venues between 2011 and 2021. A systematic review was undertaken to examine these papers, and five primary research questions have been addressed. The findings show that using sophisticated input representation that includes traffic rules and road geometry, artificial intelligence-based solutions applied to behavior prediction of traffic actors for intelligent vehicles have shown promising success, particularly in complex driving scenarios. Finally, the paper summarizes the most widely used approaches in behavior prediction of traffic actors for intelligent vehicles, which the authors believe serves as a foundation for future research in behavior prediction of surrounding traffic actors for secure and accurate intelligent vehicle navigation.

**INDEX TERMS** Intelligent driving, deep learning, intelligent vehicles, vehicle behavior prediction, pedestrian behavior prediction.

## I. INTRODUCTION

Many companies, such as Waymo and Lyft, are now working on intelligent vehicle technology for various vehicles. Although intelligent vehicles are still in the early stages of development, partially automated systems have been used in the automobile industry for a few years. Since the middle of

The associate editor coordinating the review of this manuscript and approving it for publication was Razi Iqbal<sup>1</sup>.

the 1980s, several universities, research centers, and automobile manufacturers have researched and built intelligent vehicles. Furthermore, for efficient and secure navigation on the road with mixed traffic actors, the intelligent vehicle should understand the current state of surrounding traffic actors and predict their future behavior [1].

This general problem of behavior prediction of traffic actors for the intelligent vehicle is mainly categorized into two parts. One part of this problem is to predict

the behavior of pedestrians, and another part is to predict the behavior of surrounding vehicles. Surrounding vehicles generally are of diverse types, driverless vehicles or human-driven vehicles [2]. Furthermore, an intelligent vehicle system should have the ability to perceive, understand, and predict human behavior to interact with the human environment safely. Human behavior comes in various forms: full-body motion, gestures, facial expression, or movement through space by walking, using a mobility device [3]. Accurately predicting human behavior is a very challenging task due to their complex, dynamic and random behavior. Although the motion behavior of vehicles (such as cars, trucks, and buses) has well-defined rules and environmental conditions, it is not an easy task due to several challenges.

There is interdependence amongst vehicles' behavior, where the behavior of one vehicle affects the behavior of another vehicle. Therefore, predicting vehicle motion behavior requires an understanding of the behaviors of the surrounding vehicles [4].

Intelligent vehicle technology has made tremendous progress due to the advancement of artificial intelligence (AI) techniques. Machine learning, deep learning, and artificial intelligence are the leading backbone technologies behind the progress of the intelligent vehicle [5]. Therefore, the behavior prediction of surrounding traffic actors begins to migrate from the classical to the AI-based method. Recurrent neural networks (RNN) are widely used in predicting pedestrians' behavior and nearby vehicles (e.g., car, truck, bus, etc.). One variant of RNN, the long short-term memory (LSTM) model, has been popularly used to predict the behavior of pedestrians and surrounding vehicles [6], [7]. Another variant of RNN, the gated recurrent network, predicts vehicle trajectory by combining conditional variational autoencoders [8].

Accurate behavior prediction of pedestrians and vehicles is crucial for intelligent vehicles' safe and secure navigation. Minor misbehavior of intelligent vehicles on crowded roads leads to an accident. Due to this, much research still needs to be done to see an intelligent vehicle on the road. Big companies and players in the automobile industry and researchers make intelligent vehicles as smart and reliable as human-driven vehicles. Hence, it is crucial to identify the existing research related to pedestrian behavior prediction and vehicle behavior prediction concerning an intelligent vehicle. It is necessary to conduct a systematic literature review to identify research trends and gaps concerning the behavior prediction of traffic actors. Towards this goal, existing studies and work on behavior prediction of pedestrian and surrounding vehicles concerning the intelligent vehicle examine critically and use these insights to develop new directions.

## A. PRIOR RESEARCH

Specifically, in the field of behavior prediction of traffic actors for an intelligent vehicle, as per our knowledge, there are very few Systematic Literature Reviews (SLRs) papers are available. One of the most recent review papers on vehicle

behavior prediction for intelligent driving using a deep learning approach was Mozaffari *et al.* [2]. In their work, the authors discussed challenges and problems associated with predicting future vehicle trajectories during complex driving scenarios. They provided a comprehensive review of the different approaches used to solve vehicle behavior prediction, i.e., physics-based, maneuver-based, and interaction-aware models. Based on input representation, output type, and prediction model, various researchers have used different approaches. In our view, this work gives a valuable start to researchers who might be interested in vehicle behavior prediction for the safe navigation of intelligent vehicles.

Ridel *et al.* [9] conducted a review in 2018 to predict the behavior of pedestrians in urban scenarios for intelligent vehicles. In this work, the authors discussed the state-of-the-art research developments and challenges to overcome towards finding solutions closer to the human ability to predict and interpret the behavior of pedestrians. This task requires high response time, accuracy, and precision in the real world. However, a lot more research still needs to be done to develop an intelligent vehicle that can ensure the safety of pedestrians on the road.

In a very recent work in 2020, Dunne *et al.* [10] conducted SLR to present the computational model for predicting human behavior in an intelligent environment. The authors have provided the frequently used dataset in human behavior prediction and prediction accuracy, which lies in the range of 43.9% to 100%. This research focuses on human behavior prediction in smart homes, offices, vehicles, and healthcare, which is the basis of research in this paper. This research focuses on pedestrian behavior prediction and vehicle behavior prediction in the area of intelligent vehicles. Table 1 shows the prior research in the field of behavior prediction of surrounding actors for intelligent vehicles.

As seen in the literature, as mentioned earlier, no existing systematic review focuses on the challenges and problems related to input representation for the prediction model of behavior prediction of traffic actors. In addition, the existing systematic literature review lacks a comprehensive review focused on the publicly available datasets. The literature also lacks an exhaustive study on methods or tools used for behavior prediction of traffic actors in the context of intelligent vehicles.

This systematic literature review (SLR) aims to critically examine existing research articles and their outcomes in the formulated research issue. Table 2 lists the research questions created to help focus on this SLR. As per the author's knowledge, this is the first SLR to cover the behavior of two significant traffic actors, namely pedestrians and surrounding vehicles, which affect the creation of intelligent vehicles.

Contributions of this SLR are summarized as follows:

- Eighty-three primary research studies were identified on behavior prediction of traffic actors for the intelligent vehicle. Other researchers can use these studies to advance their work in this area.

TABLE 1. Prior research related to this work.

Reference	Objectives and Topics	Domain	Type
[2]	It discusses challenges and difficulties in the prediction of possible vehicle trajectories in dynamic driving situations. It carefully examines the different methods employed to address vehicle behavior prediction problems, such as physical, maneuvering, and interaction-conscious models.	Vehicle behavior prediction for an intelligent vehicle	Review
[9]	It discusses emerging cutting-edge technology and the obstacles that must be overcome to find a better solution similar to the human level performance to anticipate and interpret pedestrian behavior. This challenge requires high response time, accuracy, and precision in the real world.	Pedestrian behavior prediction in the urban scenario for intelligent vehicle	Review
[10]	It includes information on a widely used human behavior prediction dataset and prediction accuracy ranging from 43.9% to 100%. It specializes in human behavior prediction in intelligent houses, offices, vehicles, and healthcare.	Human behavior prediction in smart houses, offices, vehicles, and healthcare	Systematic Literature Review

- The significant challenges and issues regarding input representation for the behavior prediction model are discussed.
- A comprehensive review of the availability and quality of publicly available datasets is performed.
- A summary of the existing artificial intelligence techniques available for behavior prediction of traffic actors is presented.
- The research gaps and future research directions were identified, which will help researchers and business organizations choose the proper method for behavior prediction of traffic actors in intelligent vehicles.

The following is an outline of the article. The proposed methodology, as well as our research questions, is detailed in section 2. Section 3 contains the findings and answers to the proposed research questions. In section 4, the main findings from existing literature are discussed. Section 5 includes the future research directions. Finally, in section 6, conclusions are presented.

## II. RESEARCH METHODOLOGY

The Preferred Reporting Items for Systematic Reviews and Meta-Analysis (PRISMA) method is used to perform a systematic review [88]. PRISMA is a collection of guidelines for the composition and structure of systematic reviews and other meta-analyses based on data.

The reviewing process in our paper consists of three steps: the formulation of research questions, the search phase, and the criteria for inclusion and exclusion of research articles. The steps for our research analysis are detailed as follows.

This systematic review is organized to cover the breadth of the study under consideration by categorizing and evaluating existing publications. The first step is to define the research questions so that the coverage rate of current works can be accurately described and draw some perspectives that can

help researchers generate new ideas by analyzing similar works. Table 2 lists the research questions that were used in SLR.

The second step in our SLR is to identify information sources. Related manuscripts were found using Scopus and Web of Science. Table 3 lists the primary and secondary keywords to form search queries for the identification of research articles. The third step is to develop procedures for reviewing the technical and scientific articles that these searches produced to identify relevant papers to our situation.

The proposed approach is divided into two steps: (i) use of Boolean operators AND/OR to identify search terms from research questions to prepare a list of keywords; (ii) use Boolean operators AND/OR to select queries that are used to search for and collect all relevant data. Table 4 displays the search queries used in this article.

### A. INCLUSION AND EXCLUSION CRITERIA

A list of inclusion criteria for selection and exclusion criteria for rejection of research papers were developed to select relevant research studies for systematic review (Table 5).

In the screening procedure, three inclusion criteria steps are used:

(i) Abstract-based screening: Irrelevant research papers based on information and keywords in the abstracts of research articles were rejected. Abstracts of research articles that met at least 40 percent of the IC were retained for the next steps.

(ii) Full-text screening: Research papers that do not discuss or relate to the search query given in Table 3, i.e., papers with abstracts that only represent minor aspects of the search query, were removed.

(iii) Quality-analysis step: Quality analysis on the remaining research papers were performed and excluded those that do not meet any of the following criteria:

**TABLE 2. Research questions used in this SLR.**

Number	Research Questions	Motivations
RQ 1	What is the distribution of published papers predicting pedestrian and surrounding vehicle activity for intelligent vehicles by year, publication, and domain?	It helps to figure out when, where, and who performed the research studies.
RQ 2	What are the main challenges and problems facing the prediction of traffic actor’s behavior?	It aids in exploring the fundamental problems that occur when predicting the behaviors of traffic actors and the benefits and limitations of current solutions.
RQ 3	What are the different datasets available for the behavior prediction of traffic actors?	It aids in identifying a dataset that contains appropriate training and testing data for successful research results.
RQ 4	What are the main methods used related to artificial intelligence for the behavior prediction of traffic actors?	It aids in identifying appropriate artificial intelligence approaches used for today’s intelligent applications.
RQ 5	What are the future directions for predicting pedestrian and vehicle behavior early and accurately?	It aids in finding important research avenues that have yet to be explored.

**TABLE 3. List of primary and secondary keywords used.**

Fundamental Keyword	“Behavior prediction”
Primary Keywords	“Intelligent vehicle,” “Artificial Intelligence,” “Vehicle behavior prediction,” “Pedestrian behavior prediction,”
Secondary Keywords	“Intelligent driving,” “Autonomous vehicle,” “Trajectory prediction,” “Motion prediction,” “Machine Learning,” “Deep learning,”

**TABLE 4. Search queries.**

Database	Query	# Initial result
Scopus	(( machine AND learning ) OR ( artificial AND intelligence ) OR ( deep AND learning )) AND (( human AND behavior AND prediction ) OR ( pedestrian AND behavior AND prediction ) OR ( vehicle AND behavior AND prediction ) OR ( trajectory AND prediction ) OR ( motion AND prediction )) AND (( autonomous AND vehicle ) OR ( autonomous AND driving ) OR ( intelligent AND vehicle ))	362
Web of Science		241

- C1: Research articles must provide findings and results.
- C2: Research articles must provide empirical proof of finding.
- C3: Research objectives and outcomes must be well presented.
- C4: Research articles must use appropriate and sufficient references.

**B. CONDUCTION OF SLR**

The following main steps were used to select appropriate papers for this review: Search strategy began with 603 papers, 362 obtained from Scopus databases, and 241 from the Web of Science database. The total number

of discoveries was reduced to 450 after duplicates were removed. The 450 were then examined using criteria for inclusion and exclusion based on the title and abstract, which reduced our results to 189; additional eligibility criteria based on the quality-analysis step and full-text availability yielded 83 results. These 83 papers are carefully reviewed to arrive at the conclusions discussed in the following section.

**III. RESULTS**

This section summarizes the findings of systematic analysis. It answers the research questions posed above based on the results of this review process.

TABLE 5. List of inclusion and exclusion criteria.

Inclusion criteria
Research articles must be original research articles instead of review/survey articles.
Research articles that are published between 2011 and 2021.
Research articles that answer at least one research question.
Research articles must contain search keywords either in the title, abstract, or full text.
The proposed solution in the research article must be evaluated.
The solution developed must be aimed to solve problems on behavior prediction of traffic actors.
Exclusion criteria
Non-English Research articles
Duplicate research articles
Research articles with non-availability of full text
Research articles that are not relevant to pedestrians and vehicle behavior prediction concerning intelligent vehicle

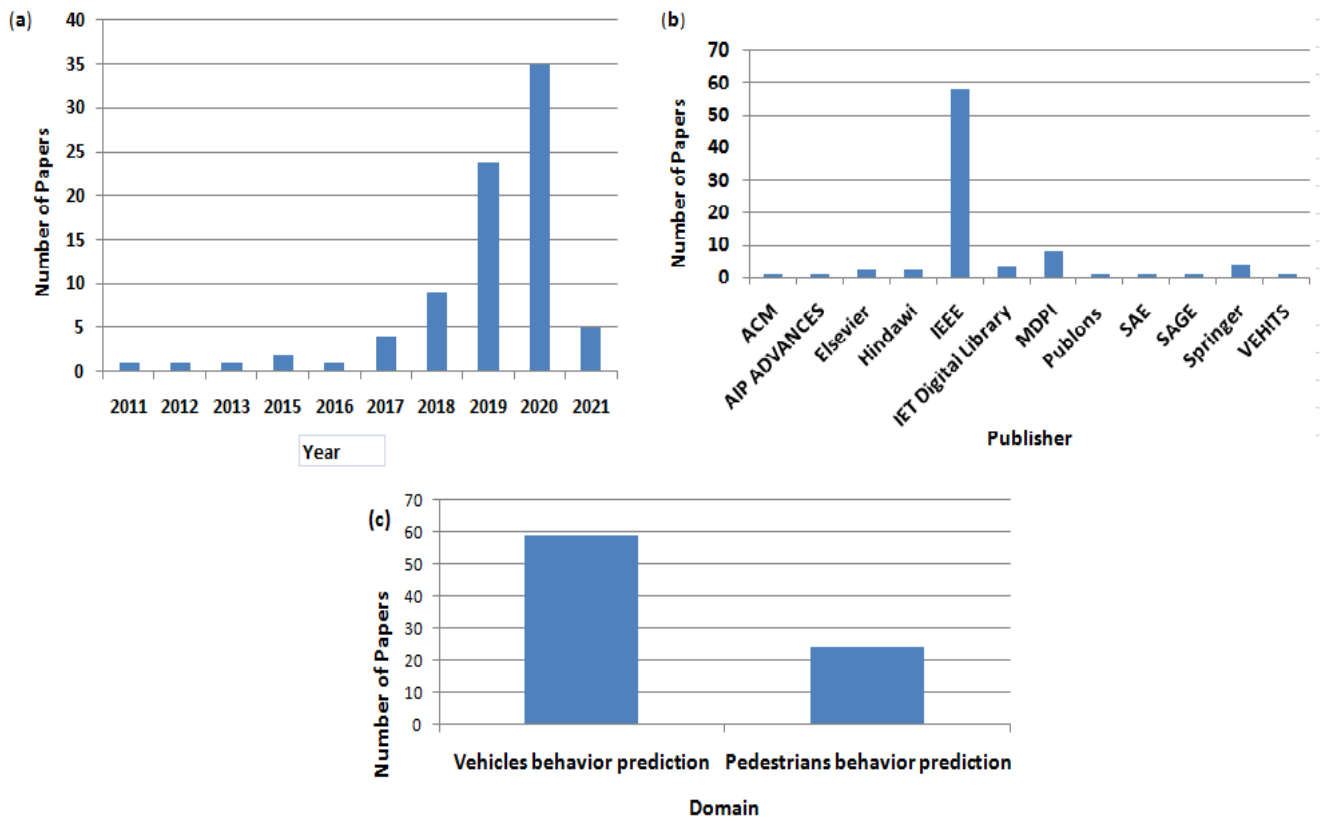


FIGURE 1. Distribution of selected papers by (a) publication year, (b) publisher, and (c) domains.

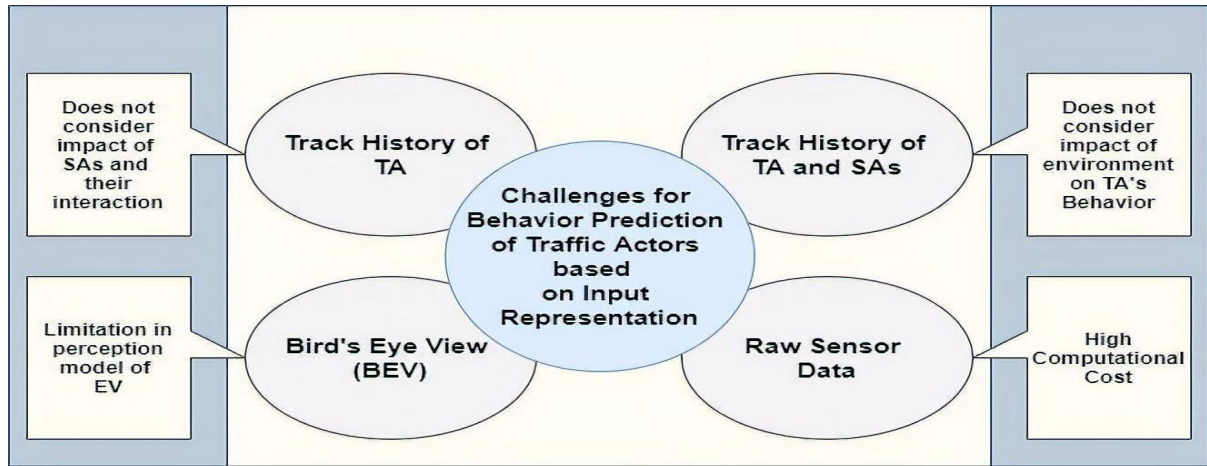
**RQ 1. What is the distribution of published papers predicting pedestrian and surrounding vehicle activity for intelligent vehicles by year, publication, and domain?**

Authors have selected only papers on pedestrians’ behavior prediction and vehicle behavior prediction related to an intelligent vehicle. There are 83 publications in total. Figure 1 shows the distribution of selected research articles by (a) publication year, (b) publisher, and (c) domain. The graph in Fig. 1(a) shows that the trend of intelligent vehicles predicting pedestrian and surrounding vehicle activity is growing, and predicting traffic actor behavior had not been proposed before 2011. After 2011, however, this trend began

to rise, with 24 essential papers published in 2019, 35 in 2020, and 5 in 2021. In addition, the majority of selected papers are published in IEEE Explore, including 58 articles, followed up by eight papers of MDPI, as shown in Fig. 1(b). Fig. 1(c) shows that the majority of selected papers (59 papers) are related to vehicle behavior prediction, followed by pedestrian behavior prediction (24 papers).

**RQ 2. What are the main challenges and problems facing the prediction of traffic actor’s behavior?**

Behavior prediction for an intelligent vehicle is not a trivial task because of several problems and challenges. First, traffic actors’ behavior is interdependent, where the behavior of



**FIGURE 2.** Challenges for behavior prediction of traffic actors based on input representation.

traffic actors affects the behavior of other traffic actors. Second, traffic rules and road geometry can change the behavior of traffic actors. Third, the future behavior is multimodal, forgiven history of behavior of traffic actors, there may exist more than one possible future behavior for it.

To define the problem of behavior prediction of traffic actors, authors adopt the following terminology according to [2]:

- **Target Actors (TAs)** are the actors whose behavior needs to predict the safe navigation of the intelligent vehicle.
- **Ego Vehicle (EV)** is the intelligent vehicle that observes the behavior of surrounding traffic actors to predict their future behavior.
- **Surrounding Actors (SAs)** are the traffic actors whose behavior needs to be predicted as their behavior influences the navigation of the intelligent vehicle.
- **Non-Effective Actors (NAs)** are the remaining actors in a driving environment that have no impact on the TA's behavior.
- **Bird's-Eye View (BEV)** is an elevated view of an object from above, with a bird's perspective.
- **Raw Sensor Data** is unprocessed sensor data obtained from EV's sensor.

Based on input data to the prediction model for behavior prediction of traffic actors, authors reviewed and classified selected articles into four categories: track history of the TA only, track history of the TA and SAs, simplified bird's-eye view, and raw sensor data.

- **Track history of the TA only:** The Traditional approach for predicting the behavior of traffic actors is to consider its current state and track the history of its states throughout a given time without taking into account the behavior of SAs. In [12], [13], information about different gestures (hand gesture, looking, node), actions like walking or standing, a reaction like moving slow,

speed up, and crossing action is used to understand the behavior of pedestrians. In [11], there is a switching linear dynamical system (SLDS) model and multi-layered bidirectional LSTM model that compared to predict the behavior of a pedestrian. SLDS and LSTM model use dataset which provides bounding boxes, disparity, and X and Z coordinates of target actor only. The proposed novel objective feature uses the periodicity of human walking (gait), mirror symmetry of the human body, and changes in ground reaction forces in a human gait cycle to predict poses and global positions for several pedestrians simultaneously [14].

Track history of the target actor only is instrumental in predicting short-term behavior. Still, it is not sufficient to predict the long-term behavior of the target actor, as long-term behavior is highly dependent on surrounding actors and the environment.

- **Track history of the TA and SAs:** In this approach, interaction among traffic actors also considers the track history of TA. State and track history of SAs, similar to TA, can be used as input for the prediction model to improve performance in predicting the behavior of TA. In [15], the lane-changing behavior of drivers is recorded using time-series input data from vehicular time-dependent dynamic states (positions, velocities, velocity differences, and position differences of nearby traffic actors, for example). Behavior prediction of the pedestrian is challenging because it requires reasoning about the traffic actor's past movement, social interactions among varying numbers and kinds of actors, constraints from scene context, and the stochastic nature of pedestrians. Novel Multi-Agent Tensor Fusion (MATF) network encodes multiple agents past trajectory, and the scene context into Multi-Agent Tensor then applies convolution fusion to capture multi-agent interaction [16].

This input representation assumes that the state and track history of all SVs are always observable. But it is challenging due to sensor impairments like noise and occlusion. The performance of prediction models for behavior prediction of traffic actors depends not only on the track history of the TA and SAs, but it is also very much dependent on environmental conditions and traffic rules.

- **Simplified bird's-eye view:** Simplified Bird's-eye View (BEV) of the environment is used to consider the interaction among traffic actors to predict the behavior of TA for an intelligent vehicle. In this representation, a series of polygons and lines depict static and dynamic objects, road lanes, and other environmental elements in a Bird's-Eye View (BEV) picture. BEV of the world is used as input to a prediction model for traffic actor activity prediction for intelligent vehicles in [17]–[19], and [31]. In [20], a novel approach to perform vehicle trajectory prediction employing graphic representation is described. The vehicles are represented using Gaussian distribution into bird eye view (BEV). The deep learning model has been trained using a highD dataset collected from aerial imagery. Occupancy Grid Map (OGM) was used by Mohajerin and Rohani [21] for multi-step prediction of drivable space for intelligent vehicles. An OGM divides the space around the ego vehicle (EV) into equivalent cells that reflect the state of occupation of nearby regions, i.e., free or occupied. The simplified BEV is versatile in terms of representation complexity, and it also allows data fusion from various types of sensors into a single BEV image. However, it inherits the flaws of the perception module, which is used to estimate the states of static and dynamic traffic actors in the driving context.
- **Raw sensor data:** The raw sensor data obtained from the EV's sensors is fed to the prediction model in this input representation. As a result, the input data includes all possible information about the surrounding area. It helps the model learn from all available sensory data to extract useful features. In [5], [23], [24], and [32], raw sensory data collected from EV sensors are used and extract valuable features as input to a prediction model for behavior prediction of traffic actors for an intelligent vehicle. For intelligent vehicles, [22] used a lane shifting decision model. An intelligent vehicle's monitoring and sensor system collects parameters (e.g., position, velocity, acceleration, etc.) from surrounding traffic actors. The identification and decision of vehicle changing lanes were studied by mining information from the vehicle's historical motion data.

The raw sensor data used as input for the prediction model has a high dimension, requiring much computing power to process it. The various challenges and problems for predicting traffic actors for intelligent vehicles based on input representation to the prediction model can be summarized as depicted in Figure 2.

### RQ 3. What are the different datasets available for the behavior prediction of traffic actors?

The cornerstone of a model based on artificial intelligence is data. Obtaining specific, impartial data from the appropriate source would aid in the development of a more accurate and reliable model. The most commonly used publicly accessible datasets for behavior prediction of different traffic actors for intelligent vehicles are covered in this section.

#### • Public dataset for pedestrian behavior prediction

The authors looked at existing research articles about predicting pedestrian behavior in the sense of intelligent vehicles. Researchers appear to have used various datasets to train an AI-based model for predicting pedestrian behavior to aid in the safe navigation of intelligent vehicles. It is critical to collect a suitable dataset with adequate quantity and quality for AI-based models to achieve good results. Deep learning models, in particular, necessitate a large amount of data for training to achieve output close to that of humans.

The Joint Attention in Autonomous Driving (JAAD) dataset [110] was used in the research studies [12], [51], and [71]. This dataset focuses on pedestrian and driver activity at crossings and the factors that affect it. It contains 346 richly annotated short video clips culled from over 240 hours of driving footage shot in different weather conditions across North America and Eastern Europe. There are multiple tags (weather, places, etc.) for each video, as well as time-stamped action marks from a setlist (e.g., stopped, walking, looking, etc.). A list of demographic characteristics for each pedestrian is also given, and a list of identifiable traffic scene elements for each frame.

Driving in dynamic urban environments is one of the most challenging challenges for intelligent vehicles. Understanding motion, actions, intent, and pedestrian pose dramatically enhances our ability to operate intelligent vehicles safely and efficiently in a crowded area. The PedX dataset [89] is used to understand pedestrian motion, behavior, intent, and pose [89]. Over 5,000 high-resolution stereo image pairs and 2,500 frames of 3D LIDAR (Light Detection and Ranging) point clouds [117] are used. The LIDAR sensors and cameras have been calibrated and synchronized in real-time. Three four-way stops with significant pedestrian-vehicle contact have been selected. Cameras are mounted on the car's windshield to capture driver-perspective pictures. Two sets of stereo cameras were used to capture photographs of all four crosswalks at an intersection, one facing forward and the other facing the incoming road from the left. It provides accurate 2D and 3D labels for each instance, including over 14000 pedestrian models with a distance of 5-45m from the cameras.

The Daimler Pedestrian Path Prediction Benchmark Dataset (GCPR13) [111] is used to infer pedestrian activity from motion tracking. It contains 68 pedestrian sequences representing four distinct forms of pedestrian behavior: crossing, stopping, beginning to walk, and bending [11].

The ETH [112] and UCY [113] datasets are the most widely used literature predicting pedestrian trajectories. The

ETH dataset includes two scenes (named ETH and Hotel) taken from a bird's eye view (named after the ETH Zurich University). There are 750 separate pedestrian trajectories in total. One frame per 0.4 seconds is annotated with pedestrian locations. The UCY dataset (named after the University of Cyprus) includes three scenes (Zara1, Zara2, and Univ) captured from a bird's eye view. It contains over 900 separate pedestrian trajectories in total. One frame per 0.4 seconds is annotated with pedestrian locations. The two datasets are often combined since they include five scenes (ETH, Hotel, Univ, Zara1, and Zara2) and over 1600 pedestrian trajectories.

#### • Public datasets for vehicle behavior prediction

For intelligent vehicle navigation to be a secure, accurate, and early prediction of adjacent vehicle activity is critical. The quantity and consistency of the dataset determine the performance of an AI-based model. Various research groups and industries working in intelligent vehicles have made datasets freely available to assist researchers working on this project worldwide. It is noticed from the selected research articles on vehicle behavior prediction that the researchers used different datasets to predict adjacent vehicle behavior.

The most widely used dataset in the literature is the Next Generation Simulation (NGSIM) public dataset [114], which contains actual driving data for predicting the future of a vehicle next to it. The Next Generation Simulation (NGSIM) public dataset was used in the research studies [5], [15], [16], [22], [58]. In Los Angeles, California, the Next Generation Simulation (NGSIM) software collected comprehensive vehicle trajectory data on southbound US 101 and Lankershim Boulevard, eastbound I-80 in Emeryville, California, and Peachtree Street in Atlanta, Georgia. A network of synchronized digital video cameras was used to collect data. The vehicle trajectory data was transcribed from the video using NGVIDEO, a specialized software application created for the NGSIM program. Every one-tenth of a second, this vehicle trajectory data provided the exact location of each vehicle within the study area, resulting in detailed lane positions and locations relative to other vehicles.

A highD dataset [115] for predicting vehicle trajectories using graphic representations is proposed as a BEV representation of intelligent vehicles, surrounding traffic actors, and the environment via aerial imagery [20]. The highD dataset is a brand-new collection of naturalistic vehicle trajectories gathered from German highways. The aerial viewpoint overcomes traditional limitations of existing traffic data collection methods, such as occlusions when using a drone. More than 110500 vehicles were counted in traffic at six separate locations. The trajectory of each vehicle is automatically extracted, including vehicle type, scale, and maneuvers. The positioning error is usually less than ten centimeters when using cutting-edge computer vision algorithms. Although the dataset was developed for the safety validation of highly automated vehicles, it can also be used for other tasks such as traffic pattern analysis and driver model parameterization.

An intelligent driving platform is being used to create a new dataset of demanding benchmarks for stereo, optical flow, visual odometry, and 3D object detection. The KITTI dataset [116] was collected on a moving platform while traveling in and around Karlsruhe, Germany. A hybrid GPS/IMU system includes camera images, laser scans, high-precision GPS measurements, and IMU accelerations. The primary goal of this dataset is to accelerate the development of computer vision and robotic algorithms for self-driving vehicles. The KITTI dataset was used in the research studies [21], [30], and [39].

The recently announced Waymo Open Dataset [72] platforms for crowd-sourced intelligent vehicles, such as 3D detection and tracking, address some of the most fundamental challenges. Although the dataset contains a wealth of high-quality, multi-source driving data, academics are more interested in the underlying driving strategy implemented in Waymo self-driving cars that are unavailable due to AV manufacturers' proprietary rights. As a result, academic researchers must make different assumptions to use intelligent vehicle components in their models or simulations, which may not accurately represent real-world traffic interactions. It contains radar, LIDAR, and camera data from one thousand 20-second segments obtained from WaymoLevel-5 intelligent vehicles in different traffic conditions. Table 6 provides a list of the data sets, applications, and their functionality.

#### **RQ 4. What are the main methods used related to artificial intelligence for the behavior prediction of traffic actors?**

The prediction of traffic-activity behaviors, which helps to prevent, estimate and anticipate the intentions of the pedestrians and other nearby vehicles, is an essential component of reliable, safe, and efficient, intelligent driving. In intelligent driving situations, scenes seldom have a single target. Several objects, some of which may be moving relative to the vehicle and one another, must typically be identified and monitored simultaneously. As a result, the most applicable literature approaches to dealing with multiple objects aim to resolve multiple object tracking issues [2].

Many findings from related literature show that artificial intelligence-based models are commonly used for predicting traffic actor behavior in the context of intelligent vehicles.

Based on the type of artificial intelligence approaches used, findings were divided into several subsections, as shown in figure 3. The behavior prediction problem in intelligent vehicles is relatively new as most research papers are in 2019, 2020, and 2021. However, due to advancements in artificial intelligence techniques, researchers are working to solve this complex problem. Chronology of the different artificial intelligence techniques used in this field is given in figure 4. The probabilistic model is used widely in the problem of behavior prediction of traffic actors for intelligent vehicles followed by recurrent neural network and generative adversarial network, as shown in figure 5.



**TABLE 6.** List of publicly available datasets, applications, and their functionality.

Reference	Dataset Name	Application and Purpose	Feature
[12], [51], [71]	Joint Attention in Autonomous Driving (JAAD) dataset	The JAAD Dataset was used to propose a method for analyzing pedestrian activity using pose estimation.	Approximately 650 samples of pedestrian activity in various environments and weather conditions are present.
[14], [15],[51]	PedX dataset	The PedX dataset is used to consider, infer, and predict pedestrian intentions and future behavior.	It is a large-scale pedestrian dataset collected from crowded intersections.
[11]	Daimler Pedestrian Path Prediction Benchmark Dataset (GCPR13)	From motion tracks, a dataset was created to infer pedestrians.	It consists of 68 pedestrian sequences involving four distinct forms of pedestrian behavior: crossing, stopping, beginning to walk, and bending in.
[48],[80],[83],[85]	ETH and UCY datasets	The datasets ETH and UCY were used to correctly understand and forecast potential pedestrian trajectories to prevent collisions and prepare a safe route.	The ETH dataset contains two scenes (ETH and Hotel) taken from a bird's eye view. In total, it comprises 750 different pedestrian trajectories. The UCY dataset contains three scenes (Zara1, Zara2, and Univ) taken from a bird's eye view. In total, it comprises more than 900 different pedestrian trajectories.
[68], [77], [78], [84]	Next-generation Simulation (NGSIM) public dataset	The NGSIM Dataset contains real-world driving data that can be used to forecast the behavior of nearby vehicles in the future.	A series of synchronous digital video cameras were used to collect data. By using NGVIDEO, data for vehicle trajectories were collected from videos.
[20], [69]	highD dataset	highD dataset, which was used to predict vehicle trajectory using a graphic representation.	It is a real dataset for vehicle trajectories collected from German highways. It is an aerial dataset that overcomes the limitation of the existing dataset, such as occlusions.
[21], [30], [39]	KITTI dataset	The dataset was created for stereo, visual odometry, 3D object detection, and optical flow task.	It was created from a moving platform in Germany. Hybrid GPU/IMU system used for collecting data.
[72]	Waymo Open Dataset	Waymo Dataset is a crowdsourcing forum for 3D detection and tracking task.	This data collection includes data from the 1000 20-second camera, radar, and LIDAR segments with autonomous labels collected under different traffic situations.
[18]	NuScenes dataset	The nuScenes dataset contains 3d object annotations and is a large-scale intelligent driving dataset.	It comes with a complete sensor suite and 1000 scenes, each lasting 20 seconds.
[60], [86]	SHRP2 NDS dataset	It is an extensive collection of the real Driving dataset available.	In SHARP2, there are four driving situations. A crash, a non-subject conflict, a near-crash, and a balanced baseline.

*Probabilistic Model (PM)*: To make optimal behavior decisions or safely perform actions, intelligent vehicles need a detailed representation of their environment. Since most environmental states are not measurable and must be assumed, gathering the requisite data is hampered. For example, no sensor can detect traffic participants' plans. Sensory data is frequently limited to noisy pose measurements, velocity, and simple geometric features [38]. Furthermore, machine tracking algorithms are typically restricted to physical models and the most straightforward heuristics.

In contrast, human drivers automatically put themselves in the shoes of other traffic participants to think about their behavior. It is insufficient for accurate long-term forecasting and, as a result, anticipatory driving. When it comes to forward-looking driver assistance and intelligent vehicles, incorporating semantics and background knowledge is crucial [44]. The use of probabilistic approaches helps link the

symbolic and metric action representations while also offering a reasonable way to deal with semantic formulation ambiguity and inaccuracy. Since they are enriched with a dynamic perception of circumstances and their meaning, combining both degrees of abstraction allows for a more precise estimation of the state-of-the-art on the one side. On the other hand, the method obtains a symbolic situation description and forecasts it in the future, which is the foundation for probabilistic decision-making [81] and [92].

Schulz *et al.* [43] used a deep neural network to propose Markovian behavioral models that were probabilistic and interactional and were dependent on the driver's path. The following models can all be beneficial to advance motion planning algorithms (e.g., Monte-Carlo tree search (MCTS), partially measurable Markov decision-making processes (POMDP), or algorithms of intent estimates and trajectory predictions (e.g., Bayesian Dynamic Networks (DBNs).

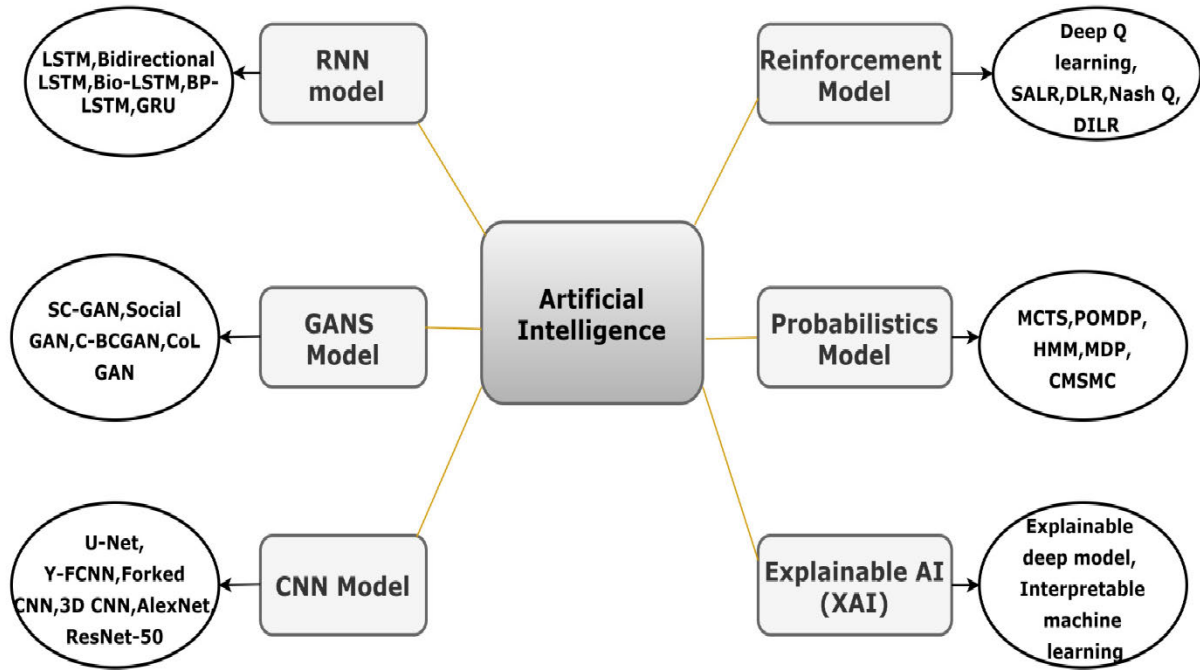


FIGURE 3. Behavior prediction using different artificial intelligence-based models.

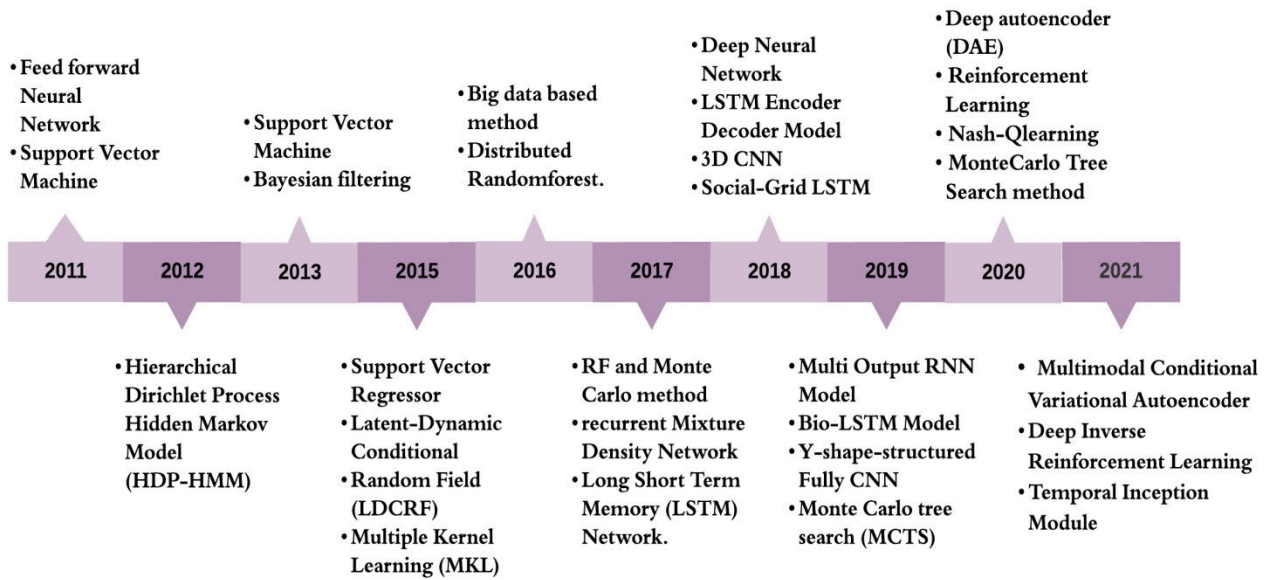


FIGURE 4. Timeline diagram of different artificial intelligence techniques.

While Driver-based vehicle threat evaluation algorithms predict future paths and compute the degree of danger, stochastic approaches define future paths with probability density functions (PDFs) determined with Monte Carlo simulations using statistical methods (MC). Since stochastic processes account for uncertainty, they are safer and more reliable than deterministic methods. The Kalman filter (KF), originally

based on a linear system model [46], is the most well-known and widely used stochastic motion prediction tool.

Driving vehicles through dynamically changing traffic scenarios, particularly on city streets, is a difficult task. In intelligent vehicles, predicting the driving activities of nearby vehicles is essential. Most conventional driving behavior prediction models are limited to a single traffic situation and

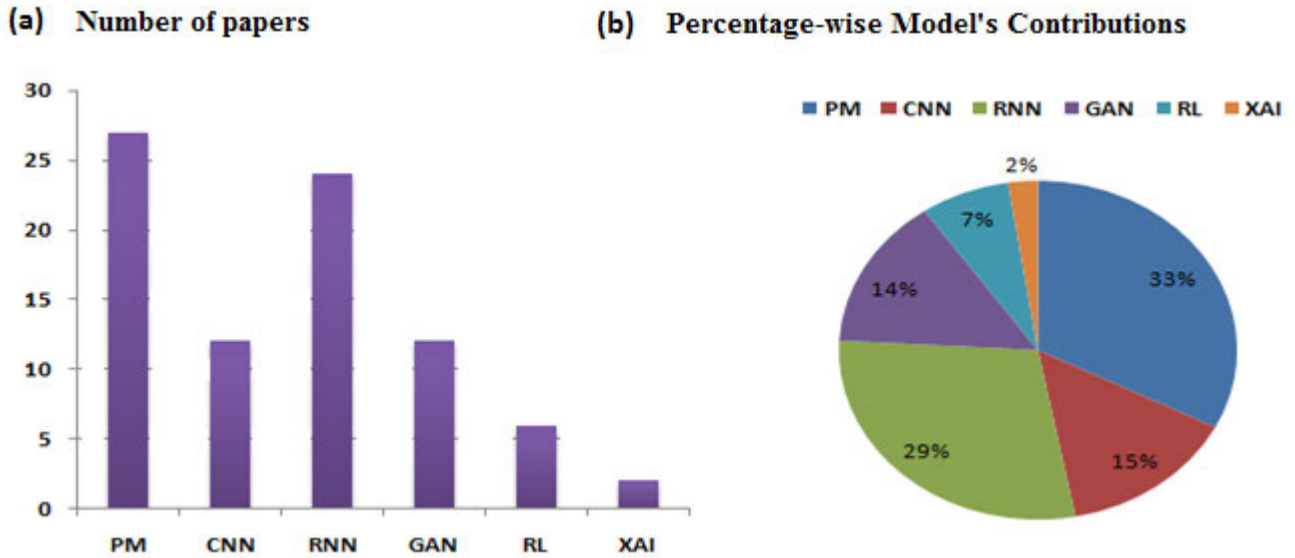


FIGURE 5. AI-based model's contributions based on: (a) number of papers, and (b) percentage-wise.

cannot be adapted to other scenarios. Furthermore, previous experience of driving was never deemed necessary. A novel ontology model was developed to model traffic scenarios. Hidden Markov Models (HMMs) were used to learn continuous driving behavior characteristics [61] and [74]. Based on the characteristics of the situation, a knowledge base was developed to identify model adaptation techniques and to store previous chances. Finally, the potential action of the target vehicle was predicted using both a posteriori and a priori probability. The proposed method was thoroughly tested using an actual intelligent vehicle.

**Convolutional Neural Network (CNN) Model:** Convolutional neural networks include convolution layers, which convolve a filter with learnable weights over the data, pooling layers that minimize input by sub-sampling, and fully-connected layers that map their input to the desired output dimension. To extract features from image data, CNNs are frequently used. They've had great success in the field of computer vision. This success encourages researchers in other fields to portray their data as an image to apply CNNs. However, one-dimensional CNNs have recently become famous for extracting features from one-dimensional signals [2].

Physical models, maneuvering models, and interaction-conscious models are used for vehicle trajectory predictions. The first form is focused on mathematical models tailored to vehicle dynamics [25]. The second attempts to anticipate the driver's intentions and produces a trajectory that matches the expected maneuver. The third form predicts trajectories by modeling interdependencies between traffic agents in some way. The fundamental Bird's-Eye View (BEV) representation and cutting-edge CNN trace forecasts in crowded road situations [20].

As the U-net model [24], the prediction core was selected. This model was used for image regression. The scene with sizes of  $h * w$  has a BEV. On the input side, a d-channel image is generated by stacking representations of previous samples. An image of potential samples on the output side is the network objective. The network core learns the fundamental behavior of the vehicles in the input block and then produces the exact representation of the vehicles. Rather than trajectories or numeric locations, this approach generates an estimation of the potential existence of the input scene. Intelligent vehicles must consider several potential future trajectories of the surrounding actors due to the inherent complexity of traffic activity to ensure a safe and efficient journey. Multimodality of vehicle movement prediction model is proposed to resolve this crucial aspect of the intelligent driving issue. The method produces a raster image encoding each vehicle actor's surrounding background before using a CNN model to produce many potential trajectories and probabilities.

In addition, the driver's focus of attention (FoA) is essential in collecting knowledge about the world, which is essential and makes the driver's car more human. The learning and forecasting of the focus of attention (FOA) are suggested in a Y-shape organized completely revolutionary neural network (Y-FCNN). This network applies to the RGB and Optical Flow layers for the first time to obtain low-level feature maps and then combine the two encoded low-level feature maps. Dilated Convolution allows for a broader receptive area with high-resolution information [35]. It then uses the final forecast. Vehicles and other traffic participants are linked in the future and fitted with various sensors, allowing for communication on various levels, such as situation prediction and intention detection. In their cooperative approach for

cyclist beginning action identification, Bieshaar *et al.* [62] use a boosted stacking ensemble system to realize feature and judgment level cooperation. A novel technique based on 3D Convolutional Neural Network (CNN) is proposed to detect beginning motions on image sequences by studying spatial and temporal characteristics. A mobile device-based beginning activity monitoring scheme that uses the cyclist's smart devices supplements the CNN. Both model outputs are coupled in a stacking ensemble solution using an extreme gradient boosting classifier, resulting in a fast and robust cooperative starting movement detector.

However, it is crucial to accurately estimate and monitor human posture in many applications, as the estimated poses are crucial for inferring their particular behavior. A forked CNN architecture was used to predict the real-world location of the skeletal joints in 3-D space using the radar-to-image representation. The proposed approach was evaluated in a single human scenario for four key motions: (I) walking, (II) swinging left arm, (III) swinging right arm, and (IV) swinging both arms [45].

Convolutional neural networks help predict vehicle behavior because they can take image-like data, generate image-like output, and maintain the spatial relationship of the input data while processing it. These capabilities allow for the simulation of vehicle interaction and driving scene background and the generation of occupancy maps. However, CNN lacks a mechanism to model data series, which is essential for modeling temporal dependencies among vehicle states over time in vehicle behavior prediction.

*Recurrent Neural Network (RNN) Model:* The simplest recurrent neural network (also known as Vanilla RNN) can be considered a two-layer fully-connected neural network with a feedback hidden layer. This small change allows for more effective modeling of sequential data. The Vanilla RNN processes the current step's input data and the memory of previous steps stored in previously hidden neurons at every input sequence. In theory, a basic RNN with several hidden units can learn to approximate some sequence-to-sequence mapping. However, gated RNNs are used because it is impossible to train this network to learn long sequences in practice due to the gradient vanishing or exploding. Instead of a simple, completely connected hidden layer, each cell of these networks employs a gated architecture. The most popular gated RNNs are the long short-term memory (LSTM) and gated recurrent unit (GRU). The most commonly used deep models for forecasting vehicle operation are LSTMs [40], [52], [75], and [79].

Understanding how other vehicles behave is critical to improving the protection and mobility of intelligent vehicles. Onboard sensors such as Radar, LIDAR, and Cameras can detect the motion of nearby vehicles and provide information such as location, velocity, and yaw. Benterki *et al.* [5] suggested a hybrid approach to get the future positions of neighboring vehicles by combining maneuver classification with neural networks and trajectory prediction with Long Short-term Memory (LSTM) networks. Furthermore, given

3D poses and locations measured with inaccuracy in prior frames, a biomechanically influenced recurrent neural network [14] is used to predict the pedestrian orientation and 3D articulated body pose in a global coordinate frame. The proposed network forecasts numerous pedestrian poses and global locations simultaneously from up to 45 meters away from the cameras (urban intersection scale). As outputs of the proposed network, full-body 3D meshes with Skinned Multi-Person Linear (SMPL) model parameters are created.

Map-mask patches were used by Palli-Thazha *et al.* [18] to boost the estimation of trajectories for various groups of interacting traffic agents. 3D LIDAR points and maps in the form of binary masks are used for this. In drivable and non-drivable regions, LSTM encoder-decoder architecture is proposed that uses Map-Mask patches to render trajectory predictions for different groups of traffic agents. Furthermore, a hierarchical multi-sequence learning network is used to predict long-term interactive trajectories for surrounding vehicles using a structural-LSTM (long short-term memory) network [59]. For each interacting vehicle, Structural-LSTM first assigns one LSTM. Then, using a radial relation, these LSTMs share their cell states and hidden states with their spatially adjacent LSTMs and recurrently examine the output state and the other LSTMs in a deeper layer. Finally, the network forecasts trajectories for nearby vehicles based on all output states.

While recurrent neural networks (RNNs) are among the most commonly used neural networks for data analysis and prediction, such as trajectory prediction, they have limitations in modeling spatial relationships such as vehicle interactions and image-like data such as driving scene background.

*Generative Adversarial Networks (GANs) Model:* Generative adversarial networks (GAN) were introduced by Goodfellow in 2014 [90]. It's an unsupervised learning technique based on the mini-max principle, in which the generator and discriminator networks compete to see who can outperform the other. The training is split between the two networks. The discriminator learns to differentiate between the produced image and the actual image in the original article's dataset, while the creator learns to create images similar to real images. In a steady-state setting, the discriminator should predict whether or not an image from the generator network is present with 50% precision.

On the other hand, the original GAN algorithm is unreliable and difficult to implement since it employs Jensen-Shannon (JS) divergence as its loss function. Since JS is a ratio of two odds that may or may not align at first, it produces zero or infinity, resulting in vanishing gradients in the discriminator network. The Earth Mover Distance (EMD), which is continuous almost everywhere, replaces the JS distance in WGAN. According to the author, it reduces the need to strike a careful balance between training the discriminator and the generator. Since the discriminator network in WGAN does not produce a probability and does not distinguish between synthetic and actual input, the author renamed it the critical network [13].

Motion synthesis, augmented reality, defense, and intelligent vehicles are just a few of the applications for predicting and understanding human motion dynamics. Because of the recent success of generative adversarial networks (GAN), there has been much interest in using deep neural network architectures and learning algorithms to perform probabilistic estimation and synthetic data generation. Human pose prediction GAN (HPGAN) [13] is proposed for predicting and interpreting human motion dynamics. HP-GAN incorporates features from Wasserstein GAN with gradient penalty (WGAN-GP), GAN, and sequence-specific optimization, to produce a realistic human motion sequence, and at the same time to measure the consistency of the created sequence. In addition, Gilitschenski *et al.* [17] suggested architecture for studying background maps and trajectory prediction at the same time. The prediction network, which adds to the context map after the trajectory embedding, is built using a modified version of the Social GAN architecture. On randomly sampled map patches and the corresponding image patches, image description, map explanation, and label explanation losses provide additional control for learning the context map. Finally, to enforce the map gradient's sparsity, the map is regularized with a norm penalization loss and a gradient norm penalization loss.

In addition, a Conditional Generative Neural System (CGNS) [28] and a Coordination-Bayesian Conditional Generative Adversarial Network (C-BCGAN) [42] have been used in the literature to predict future trajectories of nearby vehicles for safe, intelligent vehicle navigation. However, while GANs are an elegant data generation mechanism, they are challenging to train and produce output due to unstable training and unsupervised learning methods.

*Reinforcement Learning (RL) Model:* Reinforcement learning (RL) has become a robust learning system capable of learning complex policies in high-dimensional environments due to deep representation learning. In the RL model, an intelligent agent interacts with its environment to enhance its output at a given task. An agent is described as something that uses sensors to perceive its environment and actuators to act in that environment. An expert does not tell RL agents how to act; instead, a reward function  $R$  evaluates an agent's output. The agent selects an action for each state encountered and receives an occasional reward from its environment based on the utility of its decision. The agent aims to maximize the total rewards earned throughout its lifetime. Using information gained about the potential utility (i.e., discounted amount of expected future rewards) of various state-action pairs, the agent gradually increases its long-term reward. Managing the trade-off between discovery and exploitation is one of the most challenging aspects of reinforcement learning. An agent must use its information to choose acts proven to result in high rewards to maximize the rewards it receives. On the other hand, it must take the risk of attempting new actions that may result in higher rewards than the current best-valued actions for each system state to discover those beneficial actions. Put another way, the learning agent must use what it already

knows to achieve incentives, but it must also discover the unknown to make better action choices in the future [91].

Reinforcement learning is used in various areas, including video games, robotics, and intelligent vehicles. Under a reinforcement learning system, driving in congested environments can be conceived as a decision-making challenge. SARL-SGAN-KCE was proposed by Li *et al.* [34], which combines a deep socially conscious attentive value network with a human multimodal trajectory prediction model to identify the optimal driving strategy. The proposed algorithm combines multimodal pedestrian trajectory forecasting and vehicle kinematic constraints to ensure smooth pedestrian-vehicle interactions, efficient operation, and safety. Deep Q-learning networks (DQNs) [26] are also proposed to learn policies that optimize intelligent vehicle intersection handling. They have two goals in mind. The first goal is to learn an adaptive standoff that aims to increase the safety margin while maintaining the ability to make the turn within a set time frame. The second goal is to cause the least amount of damage to other vehicles when negotiating the intersection in the allotted time.

Reinforcement learning is a technique for solving highly complex problems that cannot be solved using traditional methods. This learning paradigm is remarkably similar to human learning. As a result, it is on the verge of achieving excellence. The model can correct mistakes made during the training phase. On the other hand, reinforcement learning assumes that the universe is Markovian, which it is not. The Markovian model describes a series of future events in which the previous event's condition solely determines the probability of each occurrence. Furthermore, the curse of dimensionality severely restricts reinforcement learning in real-world physical systems.

*Explainable Artificial Intelligence (XAI) Model:* The remarkable advances in Deep Learning (DL) algorithms have sparked excitement for using Artificial Intelligence (AI) technologies in almost every domain; however, the algorithms' opaqueness has raised concerns about their use in safety-critical systems. The 'explainability' dimension is critical because it explains the inner workings of black-box algorithms and introduces accountability and transparency dimensions critical for regulators, customers, and service providers. Explainable Artificial Intelligence (XAI) is a series of techniques and methods for converting so-called black-box AI algorithms to white-box algorithms, in which the results obtained by these algorithms, as well as the variables, parameters, and steps taken by the algorithm to reach those results, are straightforward and explainable [93].

There are three dimensions to consider when evaluating the comprehensiveness of AI models, as discussed below.

- **Explainability:** This is an active function of a learning model that allows the model's processes to be clearly explained. The aim is to make the inner workings of the learning model clearer. It's worth noting that sensitive applications necessitate explainability for scientific curiosity's sake and because the risk factor takes

precedence over all other considerations when human lives are at stake.

- **Interpretability:** Unlike explainability, interpretability is a function of a learning model that allows users to comprehend and make sense of it.
- **Transparency:** Transparency is often linked to understandability, with a learning model being deemed transparent if it is understandable without using an interface. The term “transparent” refers to a learning paradigm that is implicitly understandable without any additional components.

There are, in general, two approaches in making models explainable: developing models to be explainable by necessity or implementing techniques for an explanation after the performance (post-hoc). The explanations can be classified into two categories. The first is processing explanations, in which one follows inputs to outputs, for example, by addressing the question, “Why does this particular input lead to that particular output?” It is a black-box approach, as it does not require access to the AI’s internals. The second category includes representation explanations, such as responses to the question, “What knowledge does the network contain?” The latter approach, which requires access to the AI’s internals, is white or grey. Figure 8 shows the basic architecture of Explainable Artificial Intelligence (XAI).

Furthermore, deep Convolutional Neural Networks (CNNs) have emerged as front-runners in the field of driver observation; however, due to their end-to-end nature, they are frequently perceived as black boxes. The interpretability of such models is critical for establishing trust and is a significant concern for integrating CNNs into real-world systems. In a detailed analysis, Roitberg *et al.* [35] suggested a diagnostic system for internally evaluating such models and clarifying the learned spatiotemporal representations. The authors look at standard driver monitoring models from three perspectives: (1) visually explaining the prediction by combining the gradient concerning intermediate features and the corresponding activation maps, (2) looking at what the network has learned by clustering the internal representations and discovering how individual classes relate at the feature-level, and (3) conducting statistical analysis (e.g., common versus rare behaviors).

However, discussing the need for clarity and clarification of AI-related solutions in intelligent vehicles is essential. For a variety of reasons, it’s essential to understand why an intelligent car made an unexpected decision (such as driving in the wrong direction, turning in the wrong direction, applying sudden brakes, having trouble identifying objects, colliding with other objects, or failing to apply brakes, for example). The first purpose is to address the issue and increase user experience and confidence in intelligent vehicle technology. The second explanation is that performing forensics and determining the cause of the accident is crucial in the event of an accident. However, that is only possible if the intelligent car’s decisions are straightforward and explainable [93].

Table 7 summarizes the reviewed literature, techniques, purpose, datasets, results, and future directions.

#### **RQ 5. What are the future directions for predicting pedestrian and vehicle behavior early and accurately?**

Following a review of papers on traffic actor behavior prediction for intelligent vehicles, the authors identified the core open issues and numerous ideas for potential studies in these fields. The concepts and future research areas can be classified into six major categories, which are as follows:

*Input Data Representation:* Based on input data representation to the prediction model for behavior prediction of traffic actors, all reviewed articles consider four categories of data representation: track history of the TA only [11], [12] and [49], track history of the TA and SAs [15], and [16], simplified BEV [17] to [20], and raw sensor data [22], [23], [36], and [70]. The majority of current literature uses a complete view of the ambient environment and the states of different traffic actors as an input to the prediction model, which is not realistic in real-world scenarios due to sensor impairments (e.g., occlusion, noise) [2]. However, covering all road parts with all sensors mounted on the ego vehicle is impractical, reducing the efficiency of the behavior prediction model in the intelligent vehicle [76]. Possible solutions are as follows:

- 1) Use of noise reduction techniques to improve the efficiency of sensor input.
- 2) Aerial imaging techniques may be used with existing sensor data to provide a complete view of the ego vehicle’s surroundings.
- 3) High-definition imaging technology may be used to precisely locate an ego car.

HD maps for intelligent driving integrate and view data from various sources, including vehicle sensors, LIDAR, onboard cameras, satellite imaging, and GPS, in real-time. The combination of this data reflects the car’s precise position concerning all landmarks, offering detailed, real-time information on road gradients and limits, traffic signaling, lane orientation, predicted curves, and safety conditions. 4) The use of connected, intelligent vehicles may be beneficial [64]. Connected intelligent vehicles may provide additional environmental knowledge to help the behavior prediction model perform better overall.

*Adaptive Mobility:* All infrastructure development for the transportation system and driving rules are designed for human drivers. In practice, intelligent vehicles can share the road with human drivers. Adaptive mobility addresses the issue of intelligent vehicles entering our human world. In reality, all transportation infrastructure and driving rules are designed with human-driven vehicles in mind [82]. Some previous studies [37], [41], and [96] used a real driving behavior dataset to train an intelligent vehicle to mimic human driving behavior. As a result, it is fair to expect an intelligent vehicle to model and replicate human driver behavior and reasoning.

*Use of Enhanced Contextual Cues:* Intelligent systems should have a deep semantic scene understanding to analyze and forecast human motion and prepare and maneuver

TABLE 7. Summary of literature review.

Reference	Technique	Purpose	Dataset	Hyperparameters	Result
[43]	Monte Carlo Tree Search (MCTS), Dynamic Bayesian Network (DBN)	Accurate behavior prediction of complex human driver behavior for an intelligent vehicle.	Real-time driving data is collected using a measuring vehicle equipped with LIDAR and radar sensors and GPS/INS-based localization.	Activation function: ELU Number of layers: 05 Number of neurons per layer:400 Dropout: 0.5	The linear model provides the best results for long-term predictions.
[74]	Hidden Markov Models (HMM)	Driving intelligent vehicles through dynamically changing traffic scenarios, particularly on city streets, is a difficult task. A novel ontology model was developed to model traffic scenarios.	Data was collected through various sensors equipped on the vehicle.	Parameters for HMM: Observation window size Ow: Number between 5 and 40, Number of hidden states N: Number between 2 and 20, The number of Mixtures M: Number between 2 and 10.	On average, the prediction time horizon was increased by 56% (0.76 seconds) for lane shifting behaviors.
[73]	Monte Carlo Tree Search method	Predict the different tasks of connected and intelligent vehicles.	Dataset collected from cameras installed on an intelligent vehicle.	Step sizes: 0.1 s and 0.5 s, 5 s look-ahead.	MCTS and state vector with look-ahead information improves performance in the massive control task.
[56]	Traj-clusiVAT	In an intelligent vehicle, the trajectory prediction problem is critical.	T-Drive taxi trajectory data, Singapore taxi trajectory data.	For T-drive data: k': 150, a=0.05 and n=500. For Singapore data: k':300, a: 0.06, and n: 1000.	Traj-clusiVAT outperforms the current algorithm for short-term and long-term prediction.
[57]	PolyMLP	Use sensor-based early detection of sensitive circumstances to avoid collisions with vulnerable road users (VRUs).	Detailed Pedestrian Dataset, Cyclist Dataset	Mean squares error (MSE), Euclidean error.	Time series data of 2D pedestrian's velocity as input serve as best for trajectory prediction.
[60]	SVM,KNN, DT,RF,DNN, ELM	Driver behavior profiling is gaining popularity due to its utility in a variety of applications.	SHRP2 naturalistic driving study (NDS) dataset	KNN: Chebyshev distance, the Euclidean distance, and the Minkowski distance.	RF regressor outperforms the DNN regressor.
[20]	CNN	To model vehicle trajectory in congested highway scenarios.	highD dataset	Epoch: 1, learning rate= $10^{-6}$ Momentum=0.9, Loss function: MSE.	When using the position extraction process, big vehicles such as trucks have produced dual detections.
[33]	Y-shaped structured Fully Convolutional Neural Network (Y-FCNN)	Predict the driver's focus of Attention for a real-time prediction model.	Dr (eye) ve dataset	Optimization: ADAM, Learning rate: 0.01, batch normalization Loss function: K L divergence, Mini batch size: 24 GPU: GTX950M.	Precision is at a rate of about 22 frames per second.
[62]	3D Convolutional Neural Network (CNN)	Vehicles and other traffic participants will be linked in the future with different sensors for communication.	Real-world data is with 49 test subjects and 84 starting motions.	Activation function: ReLU, Batch normalization.	F1-score : 90% a mean detection time : 0.02 seconds.
[66]	AlexNet , VGG-19 , ResNet-50 , and MobileNet-v2 (MNv2)	To forecast future actor movement while accounting for and capturing the inherent complexity of the prediction assignment.	Operating intelligent vehicles manually in Pittsburgh and Phoenix collected 240 hours of data in different traffic conditions.	Batch size: 64, Optimizer: Adam, Learning rate: 0.0001	Extensive testing of the system indicates that it has functional advantages, and after successful testing, the framework was applied to a fleet of SDVs.

TABLE 7. (Continued.) Summary of literature review.

[67]	Structural analysis (SCNN) model	Vision-based vehicle behavior analysis.	Photoshop (PS) software has been used to annotate data. The annotation procedure is broken down into five steps.	Activation function: ReLU, Loss functions: BCEWithLogitsLoss for SCNN and Cross-EntropyLoss for CNN.	Accuracy: 91.1%
[5]	LSTM	The prediction of vehicle behavior is based on a hybrid approach that classifies maneuvers and forecasts their trajectory.	NGSIM dataset	Activation : softmax Loss function: Categorical Cross-entropy, Optimizer: ADAM, Learning rate: 0.001, Epoch: 50.	Root Mean Squares (RMS) errors are 0.30 m and 3.1 m, respectively.
[14]	Biomechanically inspired recurrent neural network. (Bio-LSTM)	To consider, infer, and foresee pedestrian intentions and potential action.	PedX dataset	Framework: Keras, CPU: Intel i7 3.60GHz, GPU: 2 NVIDIA TITAN X, RMSE.	With a global translation range of about 45 meters, this method can achieve about 85 mm in outdoor intersection data and 73 mm in mocap data.
[18]	LSTM Encoder-Decoder architecture	To predict the motion of interacting traffic agents.	NuScenes dataset	Optimizer: Adam , epoch :100, Learning rate: 1e-4.	The proposed model outperforms the constant curvature model in terms of overall efficiency.
[59]	Structural-LSTM	A long-term interactive trajectory prediction system predicts trajectories for nearby vehicles using a hierarchical multi-sequence learning network.	NGSIM dataset	Batch size: 256, Optimizer: Adam, Epoch: 200, Learning rate: 0.0001, GPU: TITAN-Xp.	The longitudinal RMS errors: It reduced to less than 1.93 m. The Lateral RMS errors: It reduced to less than 0.31 m.
[65]	BP-LSTM	Vehicle speed prediction is very critical for various applications of ITS.	A real naturalistic driving dataset is used.	Input= 4 units, hidden layer 1= 50 units, hidden layer 2= 50 units, output =1, number of training = 100.	The proposed vehicle speed prediction system outperforms others.
[13]	WGAN-GP	Predicting and understanding probabilistic human motion dynamics.	NTURGB-D and Human 3.6 M datasets.	Generator loss: WGAN-GP adversary loss, pose gradient loss, and bone loss. Critic loss: WGAN-GP.	The result indicates that there is a 97.8% chance that the series is an authentic human sequence.
[17]	Social GAN	Motion prediction in cluttered environments.	ETH and UCY dataset	Batch Size: 32, epoch: 200, Optimizer: Adam, learning rate: 5e-3	When compared to using only different SGAN versions, using newly proposed context maps improves performance significantly.
[28]	CGNS	To jointly forecast potential trajectories of several highly interconnected actors, considering static background knowledge and interactions between multiple entities.	ETH and UCY dataset	Optimizer: Adam, learning rate: 0.002,	Due to recurrent neural networks, the P-LSTM and S-LSTM provide an increase of comparable accuracy.
[34]	SARL-SGAN-KCE	Help to identify the optimal driving policy with the human multimodal trajectory prediction model.	ETH dataset	Acceleration: 6.4 m/s-2, Maximum turning angle: 120 <sup>0</sup> .	The proposed method was compared to the current state-of-the-art method for crowd navigation, and the analysis found that the proposed method is more straightforward and more human-like.



TABLE 7. (Continued.) Summary of literature review.

[63]	Nash-Q learning	To increase the performance and comfort of intelligent driving vehicles while maintaining their protection.	The standard lane merge scenario shows the dynamic game phase between the ego and cut-in vehicles.	$Tp=0.1$ s, $Np=20$ , $Lo=3.5$ m, $Sr=1$ , $M=0.5$ , $N=0.5$ , $\gamma = 0.9$ , $Vc=20$ m/s, $Vs= 15$ m/s.	Increase efficiency and comfort by 15.75% and 20.71%, respectively.
[87]	Embedding pedestrian dynamics in a D-IRL.	crowd behavior prediction	Stanford Drone, SAIVT Multi-Spectral Trajectory datasets.	First layer: total 32 1 * 1 convolution kernel, Activation function: ReLu, Output layer: one 1*1 convolution kernel.	The proposed approach is resilient to longer-term forecasts into the future.
[104]	Conditional variational autoencoder	Trajectory prediction uses a multimodal generative model.	ETH and UCY dataset	Loss function: Kernel Density Estimate-based Negative Log-Likelihood (KDE NLL).	CAVE technique is better to simulate future human trajectory prediction. It obtains BoN ADE=0.21, BoN FDE=0.41, and KDE NLL= -1.14, which are much better.
[105]	Gated Recurrent Unit Neural Network	Highway speed prediction	Dynamic tachometer graph (DTG) data	Loss function: mean absolute error (MAE), mean absolute percentage error (MAPE) and mean squared error (MSE).	For the bus lane, GRU had an MAE of 3.42, LSTM had an MAE of 4.06, and ARIMA had an MAE of 4.24. For the standard lane, the GRU MAE was 7.36, the LSTM MAE was 8.37, and the ARIMA MAE was 10.61.
[107]	Temporal Inception Module	Human motion prediction	Human3.6M and CMU motion capture dataset.	Epoch: 50, Learning rate decay: 0.96 every 2 epoch, Batch size:16.	The Temporal Inception Module achieves state-of-the-art performance by encoding varying length input subsequences at different temporal scales.
[108]	Long Short Term Memory (LSTM)	Evaluating the risk of driving behavior and collision of vehicles.	BSM dataset	Optimizer: Adam, Learning rate: 0.001, Batch size: 20, Epoch: 500.	The LSTM model beats other models in forecasting vehicle acceleration or deceleration over short time intervals. Still, the vehicle dynamic model has the weakest prediction effort, especially over large time horizons of 1.0 seconds.

alongside them. Context comprehension for better trajectory prediction in static environment features and semantics is still a relatively unexplored field. In situations where the target agent is not behaving alone, socially conscious methods [34] improve over socially unaware methods. Long-term motion trajectory prediction is one activity where contextual cues become especially relevant. Although context-agnostic motion and behavioral trends are helpful for short-term predictions, long-term predictions should consider intentions dependent on the context and surrounding climate.

*Driving Scenario:* Most current projects are restricted to a single driving situation, such as a roundabout, highway, intersection, or T junction [53]–[55]. However, a vehicle behavior prediction module should predict behavior in any driving situation in an intelligent vehicle. Future research should focus on developing a model that can be used in a variety of driving situations. End-to-end learning for intelligent vehicles can be done with reinforcement learning [47]. Additionally, the process of driver behavior cloning can be used to predict behavior in any driving situation [109].

*Domain Adaptation:* Domain adaptation is the process of applying an algorithm trained in one domain to a different target domain. Typically, intelligent driving systems necessitate

gathering and annotating a large amount of training data. On the other hand, using simulated environments makes data collection much more straightforward, but models trained in simulated environments often struggle to generalize to real-world situations. Using domain adaptation, a machine learning algorithm trained on samples from a source domain generalizes to a target domain. A GAN-based pixel-level domain adaptation technique may be used in the future. The adaptation process provides logical samples and generalizes well to object classes that were not seen during testing [94].

Figure 6 shows the domain adaptation for intelligent vehicles [100].

Moreover, Reinforcement Learning for an intelligent driving model can be used in a real-world setting after being trained in a virtual environment.

*Explainable Artificial Intelligence:* Explainable AI (XAI) is artificial intelligence (AI) that allows people to understand the solution’s findings. It contrasts with the “black box” nature of machine learning, in which even the creators of the AI are unable to explain why it made a particular decision. However, users are increasingly delegating more tasks to computers as automation becomes more prevalent. Such complex systems are typically built using “black box” Artificial

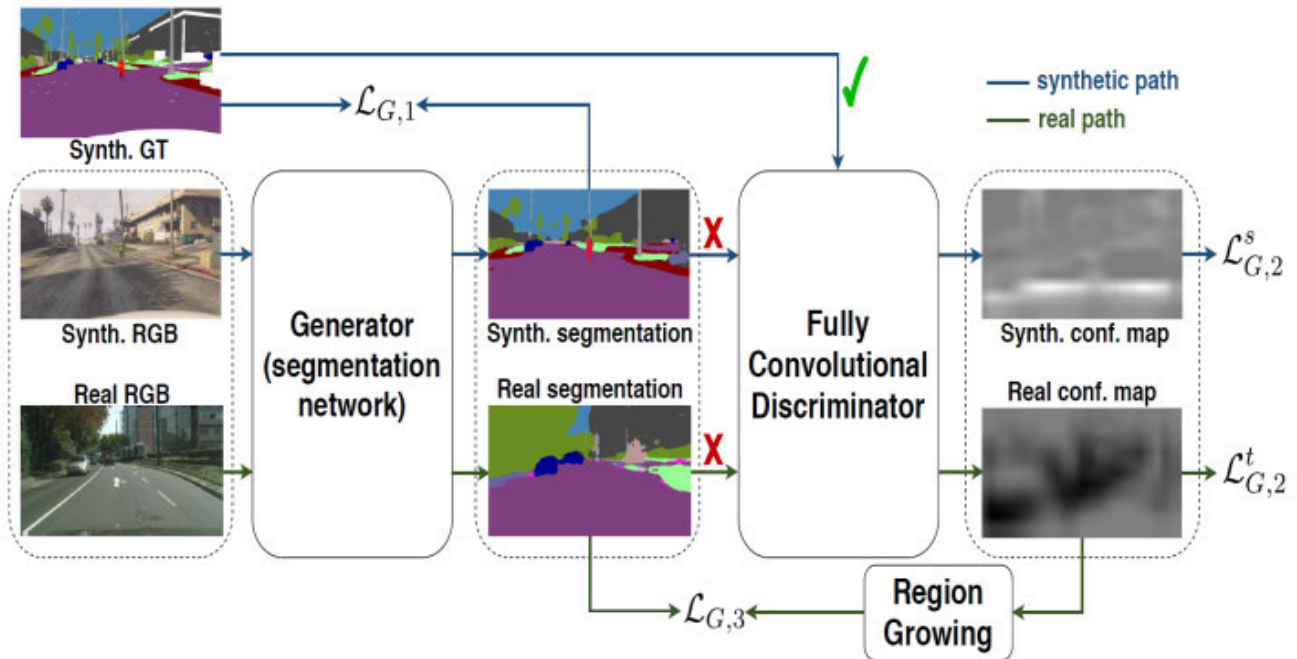


FIGURE 6. Domain adaptation for the intelligent vehicle.

Intelligence (AI), making them difficult to comprehend for users. It is particularly true in the field of intelligent driving, where the level of automation is continually growing due to the use of cutting-edge AI solutions [35].

Since interpretability and clarity are key factors for increasing confidence and protection, future research into Explainable AI (XAI) in the context of intelligent driving is relevant. Figure 7 shows the future direction to improve behavior predictions of traffic actors for intelligent vehicles.

#### IV. DISCUSSIONS

In this Systematic Review, various academic research articles on behavior prediction of traffic actors for accurate and safe navigation of intelligent vehicles were examined. There are two important surrounding traffic actors for intelligent vehicles: one is pedestrian, and the other is vehicles (e.g., car, bus, truck, etc.). The behavior of pedestrians and surrounding vehicles is highly stochastic and dynamic. Therefore, understating behavior and predicting their future trajectory is very important for intelligent vehicles [95].

By conducting this systematic review using the PRISMA protocol, authors answered research questions about various challenges and problems associated with input representation to behavior prediction models, datasets, artificial intelligence-based approaches used for behavior prediction models, and future directions as follows:

- Input representation for prediction models is categorized into four: track history of TA, TA and SA, BEV, and raw sensors. Each input representation has its advantages and disadvantages. The track history of TA

and TA and SA requires less computing time and less response time, which is a crucial point to consider. But the performance of the prediction model is less compared to BEV and raw sensor data input representation. BEV and raw sensor input representation require more computing time and have more response time. But the prediction model's performance is more than the track history of TA and the track history of TA and SA input representation.

- A high-quality publicly available real-world dataset for behavior prediction of surrounding traffic actors for intelligent vehicles is a need of time.
- Behavior prediction problems use various artificial intelligence-based approaches. Due to advancements in AI-based approaches, predicting future trajectories of surrounding traffic actors becomes possible. By examining the literature, the authors found that six different approaches of AI-based models are used. i.e., Probabilistic model, convolutional neural network, recurrent neural network, generative adversarial network, reinforcement learning, and Explainable Artificial Intelligence.
- In the future, to improve the performance of the behavior prediction model, researchers need to select proper input data representation, which includes high definition map (HD map). The majority of the existing research articles are restricted to a single driving situation. However, a vehicle behavior prediction module should predict behavior in any driving situation in an intelligent vehicle. Furthermore, gathering a large amount of training data from the real world is a challenging task. Domain

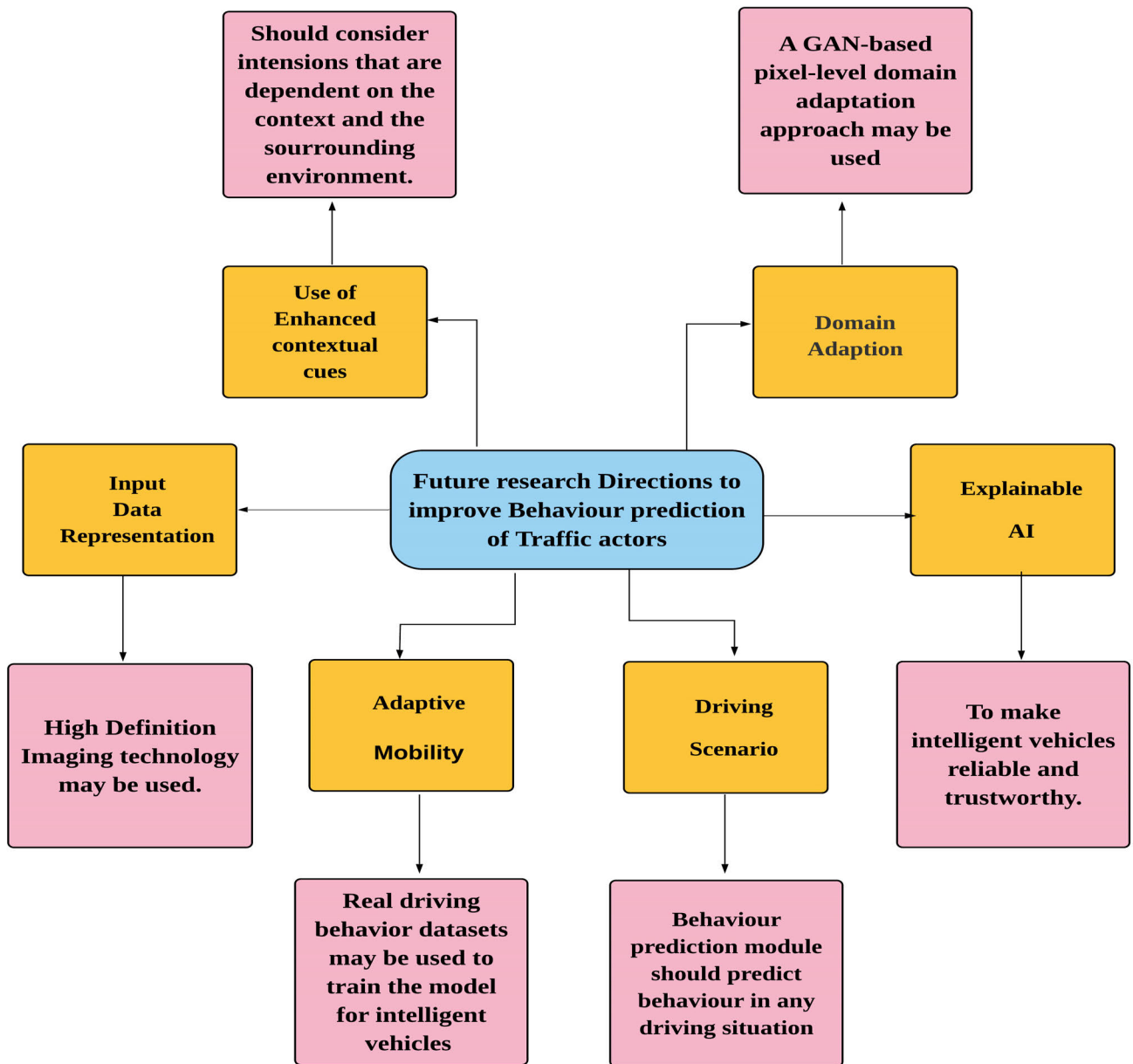


FIGURE 7. Future directions to improve behavior prediction of traffic actors for intelligent vehicles.

adaptation can be used to exploit training data from a simulated environment.

## V. THE FUTURE AHEAD

Apart from the above-mentioned future research work, the authors would like to put forth a few more research directions in predicting the behavior of traffic actors for intelligent vehicles.

### A. REINFORCEMENT LEARNING

Reinforcement learning is a machine learning algorithm in which an actor learns to perform a given task by repeatedly acting in a complex environment. Intelligent vehicles

should have real-time decision-making ability for safe and secure navigation in complex driving scenarios. A significant issue for accurate prediction of traffic actors in the complex driving environment; through experience, humans have mastered this process of driving in a complex environment. Therefore, reinforcement learning can be used for traffic actors' behavior prediction to give intelligent vehicles real-time decision-making ability [47]. Figure 8 shows the architecture of reinforcement learning used for intelligent driving [101].

Conventional reinforcement learning is, by nature black-box machine learning algorithm. Due to their inability to interpret outcomes given by the model, it is difficult to rely

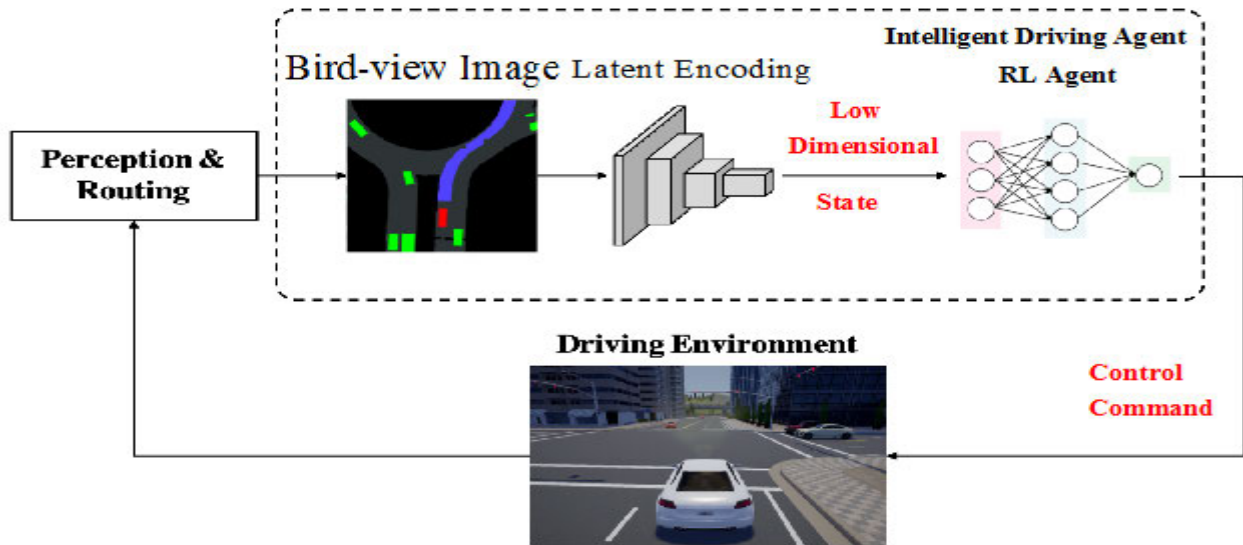


FIGURE 8. The architecture of reinforcement learning for intelligent driving.

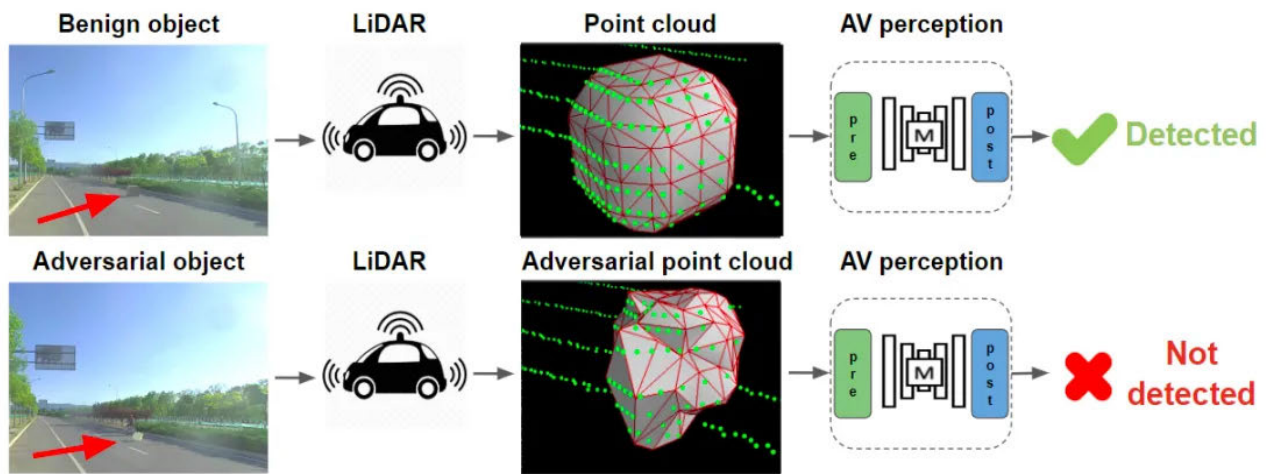


FIGURE 9. Adversarial attacks on the intelligent vehicle.

on intelligent vehicle training using reinforcement learning. Explainable reinforcement learning can be used to make intelligent vehicles reliable and trustworthy.

**B. ADVERSARIAL MACHINE LEARNING**

Adversarial machine learning is an approach that tries to mislead models by providing false input. The most typical reason is to cause a machine learning model to malfunction. An adversarial attack can be used to attack the artificial intelligence-based model. Most artificial intelligence algorithms work on given training and testing datasets. When those models are used in the real world, adversarial attackers may supply data that attack statistical assumptions to fool models [97]. Figure 9 shows the adversarial attack on intelligent vehicles [102].

Adversarial attack on the behavior prediction model of an intelligent vehicle causes a severe impact on predicting the correct behavior of surrounding traffic actors. Therefore, little misbehavior of the intelligent vehicle due to adversarial attack causes a severe impact on human life. To deal with this attack, adversarial machine learning can predict the behavior of traffic actors for secure and reliable intelligent vehicles.

**C. FEDERATED LEARNING**

Federated learning is a collaborative machine learning algorithm without using centralized training data. It is a decentralized form of machine learning. The conventional machine learning approach uses the centralized system to train the model, which is used for real-time prediction. This architecture collects data from local devices and sensors and is sent

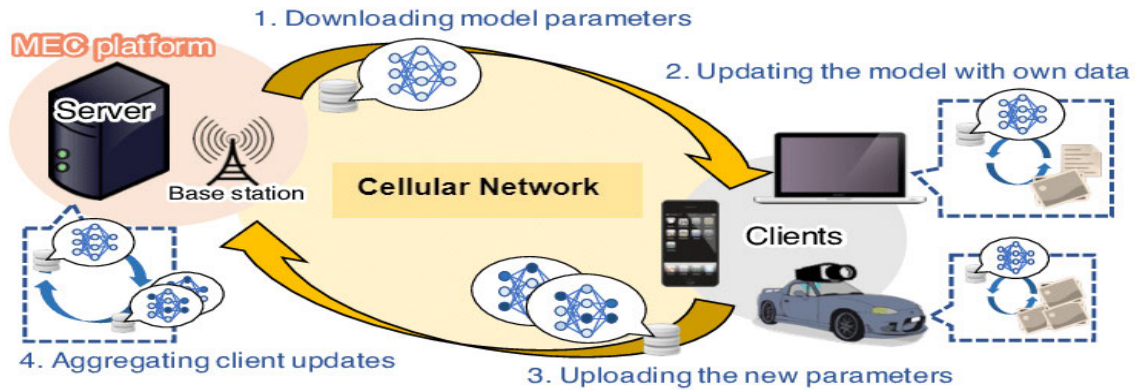


FIGURE 10. Federated learning for the intelligent vehicle.

back to a centralized system, and the outcome subsequently returns to the local device. This entire process takes a reasonable amount of time.

In contrast, the federated machine approach downloads the current model from a centralized system and updates the model locally using their local data. These locally trained models are then returned to a centralized system and aggregated. Then the single improved global model is sent to a local device for making real-time predictions [98]. Figure 10 shows the federated learning for intelligent vehicles [103].

The behavior prediction model of an intelligent vehicle needs to respond quickly to predict the behavior of traffic actors in complex real-world situations to avoid accidents. A Federated machine learning approach can be used to train machine learning models for behavior prediction of traffic actors for quick response.

#### D. ETHICAL ARTIFICIAL INTELLIGENCE

Artificial intelligence ethics is a subset of technology ethics that deals with artificially intelligent systems. It is separated into two categories: human ethics, which is concerned with how people design, create, utilize, and handle artificially intelligent systems, and machine ethics, which is concerned with how robots behave. Ethical artificial intelligence is used to develop an artificially intelligent system that can respond using a code of conduct in critical situations. It provides ethical guidelines and best practice documents to help researchers develop ethically sound artificial intelligence algorithms [99].

The behavior prediction model of an intelligent model needs to consider the social aspect of the surrounding environment. Intelligent vehicles interact with humans as well as surrounding vehicles. Therefore, intelligent vehicle systems should act ethically to predict behavior to minimize the risk of human life.

#### VI. CONCLUSION

There has been a lot of ongoing research on intelligent vehicles and the behavior prediction of surrounding traffic actors for the safety of intelligent vehicles. This article reviewed various behavior prediction techniques using the latest input

representation, AI-based solutions for intelligent vehicles in complex and real-time driving scenarios with their insightful analysis. The main findings of this study are the identifying challenges in input representation that affect the performance of the behavior prediction model. The study also identifies a high-quality publicly available real-world dataset that is the backbone for developing a behavior prediction model. It presented challenges and future research directions in the field of behavior prediction problems. However, several unsolved issues must be resolved before they can be used in intelligent vehicle applications. Although most current solutions consider vehicle interaction, variables such as the atmosphere, a collection of traffic rules are not explicitly inputted into the prediction model. Domain adaptation and explainable artificial intelligence should also be discussed in the real-world adoption of intelligent vehicles. These findings could be a foundation for future research in behavior prediction of surrounding traffic actors for secure and accurate intelligent vehicle navigation. The presented study would help the automobile industry to design secure, efficient, and reliable intelligent vehicles.

#### AUTHOR CONTRIBUTIONS

Conceptualization: Suresh Kolekar and Biswajeet Pradhan; methodology: Suresh Kolekar, Shilpa Gite, and Biswajeet Pradhan; formal analysis and data curation: Suresh Kolekar; writing—original draft preparation: Suresh Kolekar; writing—review and editing: Shilpa Gite, Biswajeet Pradhan, and Ketan Kotecha; supervision: Shilpa Gite and Biswajeet Pradhan; and funding: Biswajeet Pradhan. All authors have read and agreed to the published version of the manuscript.

#### CONFLICT OF INTEREST

The authors declare no conflict of interest.

#### REFERENCES

- [1] F. Leon and M. Gavrilescu, "A review of tracking, prediction and decision making methods for autonomous driving," 2019, *arXiv:1909.07707*. [Online]. Available: <https://arxiv.org/abs/1909.07707>
- [2] S. Mozaffari, O. Y. Al-Jarrah, M. Dianati, P. Jennings, and A. Mouzakitis, "Deep learning-based vehicle behavior prediction for autonomous driving applications: A review," *IEEE Trans. Intell. Transp. Syst.*, early access, Aug. 4, 2020, doi: [10.1109/TITS.2020.3012034](https://doi.org/10.1109/TITS.2020.3012034).

- [3] A. Rudenko, L. Palmieri, M. Herman, K. M. Kitani, D. M. Gavrila, and K. O. Arras, "Human motion trajectory prediction: A survey," *Int. J. Robot. Res.*, vol. 39, no. 8, pp. 895–935, Jul. 2020.
- [4] L. Claussmann, M. Revilloud, D. Gruyer, and S. Glaser, "A review of motion planning for highway autonomous driving," *IEEE Trans. Intell. Transp. Syst.*, vol. 21, no. 5, pp. 1826–1848, May 2019, doi: [10.1109/TITS.2019.2913998](https://doi.org/10.1109/TITS.2019.2913998).
- [5] A. Benterki, M. Boukhfir, V. Judalet, and C. Maaoui, "Artificial intelligence for vehicle behavior anticipation: Hybrid approach based on maneuver classification and trajectory prediction," *IEEE Access*, vol. 8, pp. 56992–57002, 2020, doi: [10.1109/ACCESS.2020.2982170](https://doi.org/10.1109/ACCESS.2020.2982170).
- [6] K. Goel and A. Robicquet. (2016). *Neighborhood LSTM?: Human Trajectory Prediction in Crowded Spaces*. [Online]. Available: [http://openaccess.thecvf.com/content\\_cvpr\\_2016/html/Alahi\\_Social\\_LSTM\\_Human\\_CVPR\\_2016\\_paper.html](http://openaccess.thecvf.com/content_cvpr_2016/html/Alahi_Social_LSTM_Human_CVPR_2016_paper.html)
- [7] A. Gupta, J. Johnson, L. Fei-Fei, S. Savarese, and A. Alahi, "Social GAN: Socially acceptable trajectories with generative adversarial networks," 2018, *arXiv:1803.10892*. [Online]. Available: <http://arxiv.org/abs/1803.10892>
- [8] N. Lee, W. Choi, P. Vernaza, C. B. Choy, P. H. S. Torr, and M. Chandraker, "DESIRE: Distant future prediction in dynamic scenes with interacting agents," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jul. 2017, pp. 2165–2174, doi: [10.1109/CVPR.2017.233](https://doi.org/10.1109/CVPR.2017.233).
- [9] D. Ridet, E. Rehder, M. Lauer, C. Stiller, and D. Wolf, "A literature review on the prediction of pedestrian behavior in urban scenarios," in *Proc. 21st Int. Conf. Intell. Transp. Syst. (ITSC)*, Nov. 2018, pp. 3105–3112, doi: [10.1109/ITSC.2018.8569415](https://doi.org/10.1109/ITSC.2018.8569415).
- [10] R. Dunne, O. Matthews, J. Vega, S. Harper, and T. Morris, "Computational methods for predicting human behaviour in smart environments," Univ. Manchester, Manchester, U.K., Tech. Rep. 94406, 2020, pp. 1–21, doi: [10.13140/RG.2.2.25330.94406](https://doi.org/10.13140/RG.2.2.25330.94406).
- [11] J. J. Dabrowski, J. P. de Villiers, A. Rahman, and C. Beyers, "Deep learning and statistical models for time-critical pedestrian behaviour prediction," in *Proc. Int. Conf. Neural Inf. Process.*, 2020, pp. 458–465, doi: [10.1007/978-3-030-36808-1\\_50](https://doi.org/10.1007/978-3-030-36808-1_50).
- [12] A. Marginean, R. Brehar, and M. Negru, "Understanding pedestrian behaviour with pose estimation and recurrent networks," in *Proc. 6th Int. Symp. Elect. Electron. Eng. (ISEEE)*, Oct. 2019, pp. 1–6, doi: [10.1109/ISEEE48094.2019.9136126](https://doi.org/10.1109/ISEEE48094.2019.9136126).
- [13] E. Barsoum, J. Kender, and Z. Liu, "HP-GAN: Probabilistic 3D human motion prediction via GAN," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. Workshops (CVPRW)*, Jun. 2018, pp. 1499–149909, doi: [10.1109/CVPRW.2018.00191](https://doi.org/10.1109/CVPRW.2018.00191).
- [14] X. Du, R. Vasudevan, and M. Johnson-Roberson, "Bio-LSTM: A biomechanically inspired recurrent neural network for 3-D pedestrian pose and gait prediction," *IEEE Robot. Autom. Lett.*, vol. 4, no. 2, pp. 1501–1508, Apr. 2019, doi: [10.1109/LRA.2019.2895266](https://doi.org/10.1109/LRA.2019.2895266).
- [15] L. Huang, H. Guo, R. Zhang, H. Wang, and J. Wu, "Capturing drivers' lane changing behaviors on operational level by data driven methods," *IEEE Access*, vol. 6, pp. 57497–57506, 2018, doi: [10.1109/ACCESS.2018.2873942](https://doi.org/10.1109/ACCESS.2018.2873942).
- [16] T. Zhao, Y. Xu, M. Monfort, W. Choi, C. Baker, Y. Zhao, Y. Wang, and Y. N. Wu, "Multi-agent tensor fusion for contextual trajectory prediction," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2019, pp. 12118–12126, doi: [10.1109/CVPR.2019.01240](https://doi.org/10.1109/CVPR.2019.01240).
- [17] I. Gilitschenski, G. Rosman, A. Gupta, S. Karaman, and D. Rus, "Deep context maps: Agent trajectory prediction using location-specific latent maps," *IEEE Robot. Autom. Lett.*, vol. 5, no. 4, pp. 5097–5104, Oct. 2020, doi: [10.1109/LRA.2020.3004800](https://doi.org/10.1109/LRA.2020.3004800).
- [18] V. Palli-Thazha, D. Filliat, and J. Ibanez-Guzman, "Trajectory prediction of traffic agents: Incorporating context into machine learning approaches," in *Proc. IEEE 91st Veh. Technol. Conf. (VTC-Spring)*, May 2020, pp. 1–6, doi: [10.1109/VTC2020-Spring48590.2020.9128848](https://doi.org/10.1109/VTC2020-Spring48590.2020.9128848).
- [19] E. Wang, H. Cui, S. Yalamanchi, M. Moorthy, and N. Djuric, "Improving movement predictions of traffic actors in bird's-eye view models using GANs and differentiable trajectory rasterization," in *Proc. 26th ACM SIGKDD Int. Conf. Knowl. Discovery Data Mining*, Aug. 2020, pp. 2340–2348, doi: [10.1145/3394486.3403283](https://doi.org/10.1145/3394486.3403283).
- [20] R. Izquierdo, A. Quintanar, I. Parra, D. Fernandez-Llorca, and M. A. Sotelo, "Vehicle trajectory prediction in crowded highway scenarios using bird eye view representations and CNNs," in *Proc. IEEE 23rd Int. Conf. Intell. Transp. Syst. (ITSC)*, Sep. 2020, pp. 1–6, doi: [10.1109/ITSC45102.2020.9294732](https://doi.org/10.1109/ITSC45102.2020.9294732).
- [21] N. Mohajerin and M. Rohani, "Multi-step prediction of occupancy grid maps with recurrent neural networks," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2019, pp. 10592–10600, doi: [10.1109/CVPR.2019.01085](https://doi.org/10.1109/CVPR.2019.01085).
- [22] X. Gu, Y. Han, and J. Yu, "A novel lane-changing decision model for autonomous vehicles based on deep autoencoder network and XGBoost," *IEEE Access*, vol. 8, pp. 9846–9863, 2020, doi: [10.1109/ACCESS.2020.2964294](https://doi.org/10.1109/ACCESS.2020.2964294).
- [23] Y. Jeong and K. Yi, "Bidirectional long shot-term memory-based interactive motion prediction of cut-in vehicles in urban environments," *IEEE Access*, vol. 8, pp. 106183–106197, 2020, doi: [10.1109/ACCESS.2020.2994929](https://doi.org/10.1109/ACCESS.2020.2994929).
- [24] H. Cui, V. Radosavljevic, F.-C. Chou, T.-H. Lin, T. Nguyen, T.-K. Huang, J. Schneider, and N. Djuric, "Multimodal trajectory predictions for autonomous driving using deep convolutional networks," in *Proc. Int. Conf. Robot. Autom. (ICRA)*, May 2019, pp. 2090–2096, doi: [10.1109/ICRA.2019.8793868](https://doi.org/10.1109/ICRA.2019.8793868).
- [25] H. Cui, T. Nguyen, F.-C. Chou, T.-H. Lin, J. Schneider, D. Bradley, and N. Djuric, "Deep kinematic models for kinematically feasible vehicle trajectory predictions," in *Proc. IEEE Int. Conf. Robot. Autom. (ICRA)*, May 2020, pp. 10563–10569, doi: [10.1109/ICRA40945.2020.9197560](https://doi.org/10.1109/ICRA40945.2020.9197560).
- [26] D. Isele, A. Nakhaei, and K. Fujimura, "Safe reinforcement learning on autonomous vehicles," in *Proc. IEEE/RSJ Int. Conf. Intell. Robots Syst. (IROS)*, Oct. 2018, pp. 6162–6167, doi: [10.1109/IROS.2018.8593420](https://doi.org/10.1109/IROS.2018.8593420).
- [27] K. Cho, T. Ha, G. Lee, and S. Oh, "Deep predictive autonomous driving using multi-agent joint trajectory prediction and traffic rules," in *Proc. IEEE/RSJ Int. Conf. Intell. Robots Syst. (IROS)*, Nov. 2019, pp. 2076–2081, doi: [10.1109/IROS40897.2019.8967708](https://doi.org/10.1109/IROS40897.2019.8967708).
- [28] J. Li, H. Ma, and M. Tomizuka, "Conditional generative neural system for probabilistic trajectory prediction," in *Proc. IEEE/RSJ Int. Conf. Intell. Robots Syst. (IROS)*, Nov. 2019, pp. 6150–6156, doi: [10.1109/IROS40897.2019.8967822](https://doi.org/10.1109/IROS40897.2019.8967822).
- [29] C. Anderson, X. Du, R. Vasudevan, and M. Johnson-Roberson, "Stochastic sampling simulation for pedestrian trajectory prediction," in *Proc. IEEE/RSJ Int. Conf. Intell. Robots Syst. (IROS)*, Nov. 2019, pp. 4236–4243, doi: [10.1109/IROS40897.2019.8967857](https://doi.org/10.1109/IROS40897.2019.8967857).
- [30] N. Zaghari, M. Fathy, S. M. Jameii, M. Sabokrou, and M. Shahverdy, "Improving the learning of self-driving vehicles based on real driving behavior using deep neural network techniques," *J. Supercomput.*, vol. 77, no. 4, pp. 3752–3794, Apr. 2021.
- [31] S. J. Kamble and M. R. Kounte, "On road intelligent vehicle path predication and clustering using machine learning approach," in *Proc. 3rd Int. Conf. I-SMAC (IoT Social, Mobile, Analytics Cloud) (I-SMAC)*, Dec. 2019, pp. 501–505, doi: [10.1109/I-SMAC47947.2019.9032648](https://doi.org/10.1109/I-SMAC47947.2019.9032648).
- [32] S. C. Jugade, A. C. Victorino, V. B. Cherfaoui, and S. Kanarachos, "Sensor based prediction of human driving decisions using feed forward neural networks for intelligent vehicles," in *Proc. 21st Int. Conf. Intell. Transp. Syst. (ITSC)*, Nov. 2018, pp. 691–696, doi: [10.1109/ITSC.2018.8569441](https://doi.org/10.1109/ITSC.2018.8569441).
- [33] M. Ning, C. Lu, and J. Gong, "An efficient model for driving focus of attention prediction using deep learning," in *Proc. IEEE Intell. Transp. Syst. Conf. (ITSC)*, Oct. 2019, pp. 1192–1197, doi: [10.1109/ITSC.2019.8917337](https://doi.org/10.1109/ITSC.2019.8917337).
- [34] K. Li, M. Shan, K. Narula, S. Worrall, and E. Nebot, "Socially aware crowd navigation with multimodal pedestrian trajectory prediction for autonomous vehicles," in *Proc. IEEE 23rd Int. Conf. Intell. Transp. Syst. (ITSC)*, Sep. 2020, doi: [10.1109/ITSC45102.2020.9294304](https://doi.org/10.1109/ITSC45102.2020.9294304).
- [35] A. Roitberg, M. Haurilet, S. Reis, and R. Stiefelhagen, "CNN-based driver activity understanding: Shedding light on deep spatiotemporal representations," in *Proc. IEEE 23rd Int. Conf. Intell. Transp. Syst. (ITSC)*, Sep. 2020, pp. 1–6, doi: [10.1109/ITSC45102.2020.9294731](https://doi.org/10.1109/ITSC45102.2020.9294731).
- [36] E. Ohn-Bar, A. Tawari, S. Martin, and M. M. Trivedi, "On surveillance for safety critical events: In-vehicle video networks for predictive driver assistance systems," *Comput. Vis. Image Understand.*, vol. 134, pp. 130–140, May 2015, doi: [10.1016/j.cviu.2014.10.003](https://doi.org/10.1016/j.cviu.2014.10.003).
- [37] T. Taniguchi, S. Nagasaka, K. Hitomi, N. P. Chandrasiri, T. Bando, and K. Takenaka, "Sequence prediction of driving behavior using double articulation analyzer," *IEEE Trans. Syst., Man, Cybern., Syst.*, vol. 46, no. 9, pp. 1300–1313, Sep. 2016, doi: [10.1109/TSMC.2015.2465933](https://doi.org/10.1109/TSMC.2015.2465933).
- [38] P. Kumar, M. Perrollaz, S. Lefevre, and C. Laugier, "Learning-based approach for online lane change intention prediction," in *Proc. IEEE Intell. Vehicles Symp. (IV)*, Jun. 2013, pp. 797–802, doi: [10.1109/IVS.2013.6629564](https://doi.org/10.1109/IVS.2013.6629564).

- [39] J. Wu, J. Ruenz, and M. Althoff, "Probabilistic map-based pedestrian motion prediction taking traffic participants into consideration," in *Proc. IEEE Intell. Vehicles Symp. (IV)*, Jun. 2018, pp. 1285–1292, doi: [10.1109/IVS.2018.8500562](https://doi.org/10.1109/IVS.2018.8500562).
- [40] S. H. Park, B. Kim, C. M. Kang, C. C. Chung, and J. W. Choi, "Sequence-to-sequence prediction of vehicle trajectory via LSTM encoder-decoder architecture," in *Proc. IEEE Intell. Vehicles Symp. (IV)*, Jun. 2018, pp. 1672–1678, doi: [10.1109/IVS.2018.8500658](https://doi.org/10.1109/IVS.2018.8500658).
- [41] Z. Li, C. Gong, and C. Lu, "Transferable driver behavior learning via distribution adaptation in the lane change scenario," in *Proc. IEEE Intell. Vehicles Symp. (IV)*, Jun. 2019, pp. 193–200, doi: [10.1109/IVS.2019.8813781](https://doi.org/10.1109/IVS.2019.8813781).
- [42] J. Li, H. Ma, W. Zhan, and M. Tomizuka, "Coordination and trajectory prediction for vehicle interactions via Bayesian generative modeling," in *Proc. IEEE Intell. Vehicles Symp. (IV)*, Jun. 2019, pp. 2496–2503, doi: [10.1109/IVS.2019.8813821](https://doi.org/10.1109/IVS.2019.8813821).
- [43] J. Schulz, C. Hubmann, N. Morin, J. Lochner, and D. Burschka, "Learning interaction-aware probabilistic driver behavior models from urban scenarios," in *Proc. IEEE Intell. Vehicles Symp. (IV)*, Jun. 2019, pp. 1326–1333, doi: [10.1109/IVS.2019.8814080](https://doi.org/10.1109/IVS.2019.8814080).
- [44] J. Manttari and J. Folkesson, "Incorporating uncertainty in predicting vehicle maneuvers at intersections with complex interactions," in *Proc. IEEE Intell. Vehicles Symp. (IV)*, Jun. 2019, pp. 2156–2163, doi: [10.1109/IVS.2019.8814159](https://doi.org/10.1109/IVS.2019.8814159).
- [45] A. Sengupta, F. Jin, R. Zhang, and S. Cao, "mm-Pose: Real-time human skeletal posture estimation using mmwave radars and CNNs," *IEEE Sensors J.*, vol. 20, no. 17, pp. 10032–10044, Sep. 2020, doi: [10.1109/JSEN.2020.2991741](https://doi.org/10.1109/JSEN.2020.2991741).
- [46] K. Okamoto, K. Berntorp, and S. Di Cairano, "Driver intention-based vehicle threat assessment using random forests and particle filtering," *IFAC-PapersOnLine*, vol. 50, no. 1, pp. 13860–13865, Jul. 2017, doi: [10.1016/j.ifacol.2017.08.2231](https://doi.org/10.1016/j.ifacol.2017.08.2231).
- [47] S. M. Grigorescu, B. Trasnea, L. Marina, A. Vasilcoi, and T. Cocias, "NeuroTrajectory: A neuroevolutionary approach to local state trajectory learning for autonomous vehicles," *IEEE Robot. Autom. Lett.*, vol. 4, no. 4, pp. 3441–3448, Oct. 2019, doi: [10.1109/LRA.2019.2926224](https://doi.org/10.1109/LRA.2019.2926224).
- [48] C. Scholler, V. Aravantinos, F. Lay, and A. Knoll, "What the constant velocity model can teach us about pedestrian motion prediction," *IEEE Robot. Autom. Lett.*, vol. 5, no. 2, pp. 1696–1703, Apr. 2020, doi: [10.1109/LRA.2020.2969925](https://doi.org/10.1109/LRA.2020.2969925).
- [49] E. Rehder, F. Wirth, M. Lauer, and C. Stiller, "Pedestrian prediction by planning using deep neural networks," in *Proc. IEEE Int. Conf. Robot. Autom. (ICRA)*, May 2018, pp. 5903–5908, doi: [10.1109/ICRA.2018.8460203](https://doi.org/10.1109/ICRA.2018.8460203).
- [50] C. Anderson, R. Vasudevan, and M. Johnson-Roberson, "Off the beaten sidewalk: Pedestrian prediction in shared spaces for autonomous vehicles," *IEEE Robot. Autom. Lett.*, vol. 5, no. 4, pp. 6892–6899, Oct. 2020, doi: [10.1109/LRA.2020.3023713](https://doi.org/10.1109/LRA.2020.3023713).
- [51] X. Du, R. Vasudevan, and M. Johnson-Roberson, "Unsupervised pedestrian pose prediction: A deep predictive coding network-based approach for autonomous vehicle perception," *IEEE Robot. Autom. Mag.*, vol. 27, no. 2, pp. 129–138, Jun. 2020, doi: [10.1109/MRA.2020.2976313](https://doi.org/10.1109/MRA.2020.2976313).
- [52] U. Dogan, J. Edelbrunner, and I. Iossifidis, "Autonomous driving: A comparison of machine learning techniques by means of the prediction of lane change behavior," in *Proc. IEEE Int. Conf. Robot. Biomimetics*, Dec. 2011, pp. 1837–1843, doi: [10.1109/ROBIO.2011.6181557](https://doi.org/10.1109/ROBIO.2011.6181557).
- [53] D. Yi, J. Su, C. Liu, and W.-H. Chen, "Trajectory clustering aided personalized driver intention prediction for intelligent vehicles," *IEEE Trans. Ind. Informat.*, vol. 15, no. 6, pp. 3693–3702, Jun. 2018, doi: [10.1109/TII.2018.2890141](https://doi.org/10.1109/TII.2018.2890141).
- [54] B. Völz, H. Mielenz, I. Gilitschenski, R. Siegart, and J. Nieto, "Inferring pedestrian motions at urban crosswalks," *IEEE Trans. Intell. Transp. Syst.*, vol. 20, no. 2, pp. 544–555, Feb. 2019, doi: [10.1109/TITS.2018.2827956](https://doi.org/10.1109/TITS.2018.2827956).
- [55] W. Qian, A. Beshah, and W. Thomas, "Predictive maneuver planning for an autonomous vehicle in public highway traffic," *IEEE Trans. Intell. Transp. Syst.*, vol. 20, no. 4, pp. 1303–1315, Apr. 2019, doi: [10.1109/TITS.2018.2848472](https://doi.org/10.1109/TITS.2018.2848472).
- [56] P. Rathore, D. Kumar, S. Rajasegarar, M. Palaniswami, and J. C. Bezdek, "A scalable framework for trajectory prediction," *IEEE Trans. Intell. Transp. Syst.*, vol. 20, no. 10, pp. 3860–3874, Oct. 2019, doi: [10.1109/TITS.2019.2899179](https://doi.org/10.1109/TITS.2019.2899179).
- [57] M. Goldhammer, S. Köhler, S. Zernetsch, K. Doll, B. Sick, and K. Dietmayer, "Intentions of vulnerable road users-detection and forecasting by means of machine learning," *IEEE Trans. Intell. Transp. Syst.*, vol. 21, no. 7, pp. 3035–3045, Jul. 2020, doi: [10.1109/TITS.2019.2923319](https://doi.org/10.1109/TITS.2019.2923319).
- [58] J. Li, W. Zhan, Y. Hu, and M. Tomizuka, "Generic tracking and probabilistic prediction framework and its application in autonomous driving," 2019, *arXiv:1908.09031*. [Online]. Available: <http://arxiv.org/abs/1908.09031>
- [59] L. Hou, L. Xin, S. E. Li, B. Cheng, and W. Wang, "Interactive trajectory prediction of surrounding road users for autonomous driving using structural-LSTM network," *IEEE Trans. Intell. Transp. Syst.*, vol. 21, no. 11, pp. 4615–4625, Nov. 2020, doi: [10.1109/TITS.2019.2942089](https://doi.org/10.1109/TITS.2019.2942089).
- [60] A. E. Abdelrahman, H. S. Hassanein, and N. Abu-Ali, "Robust data-driven framework for driver behavior profiling using supervised machine learning," *IEEE Trans. Intell. Transp. Syst.*, early access, Nov. 17, 2020, doi: [10.1109/TITS.2020.3035700](https://doi.org/10.1109/TITS.2020.3035700).
- [61] Y. Zheng and J. H. L. Hansen, "Lane-change detection from steering signal using spectral segmentation and learning-based classification," *IEEE Trans. Intell. Veh.*, vol. 2, no. 1, pp. 14–24, Mar. 2017, doi: [10.1109/TIV.2017.2708600](https://doi.org/10.1109/TIV.2017.2708600).
- [62] M. Bieshaar, S. Zernetsch, A. Hubert, B. Sick, and K. Doll, "Cooperative starting movement detection of cyclists using convolutional neural networks and a boosted stacking ensemble," *IEEE Trans. Intell. Vehicles*, vol. 3, no. 4, pp. 534–544, Dec. 2018, doi: [10.1109/TIV.2018.2873900](https://doi.org/10.1109/TIV.2018.2873900).
- [63] C. Xu, W. Zhao, L. Li, Q. Chen, D. Kuang, and J. Zhou, "A nash Q-learning based motion decision algorithm with considering interaction to traffic participants," *IEEE Trans. Veh. Technol.*, vol. 69, no. 11, pp. 12621–12634, Nov. 2020, doi: [10.1109/TVT.2020.3027352](https://doi.org/10.1109/TVT.2020.3027352).
- [64] H. A. Najada and I. Mahgoub, "Autonomous vehicles safe-optimal trajectory selection based on big data analysis and predefined user preferences," in *Proc. IEEE 7th Annu. Ubiquitous Comput., Electron. Mobile Commun. Conf. (UEMCON)*, Oct. 2016, pp. 1–6, doi: [10.1109/UEMCON.2016.7777922](https://doi.org/10.1109/UEMCON.2016.7777922).
- [65] L. Yufang, C. Mingnuo, and Z. Wanzhong, "Investigating long-term vehicle speed prediction based on BP-LSTM algorithms," *IET Intell. Transp. Syst.*, vol. 13, no. 8, pp. 1281–1290, 2019, doi: [10.1049/iet-its.2018.5593](https://doi.org/10.1049/iet-its.2018.5593).
- [66] N. Djuric, V. Radosavljevic, H. Cui, T. Nguyen, F.-C. Chou, T.-H. Lin, N. Singh, and J. Schneider, "Uncertainty-aware short-term motion prediction of traffic actors for autonomous driving," in *Proc. IEEE Winter Conf. Appl. Comput. Vis. (WACV)*, Mar. 2020, pp. 2084–2093, doi: [10.1109/WACV45572.2020.9093332](https://doi.org/10.1109/WACV45572.2020.9093332).
- [67] L. Mou, H. Xie, S. Mao, P. Zhao, and Y. Chen, "Vision-based vehicle behaviour analysis: A structured learning approach via convolutional neural networks," *IET Intell. Transp. Syst.*, vol. 14, no. 7, pp. 792–801, 2020, doi: [10.1049/iet-its.2019.0419](https://doi.org/10.1049/iet-its.2019.0419).
- [68] B. Lu, S. Ni, and S. S. Washburn, "A support vector regression approach for investigating multianticipative driving behavior," *Math. Problems Eng.*, vol. 2015, Sep. 2015, Art. no. 701926.
- [69] V. Mahajan, C. Katrakazas, and C. Antoniou, "Prediction of lane-changing maneuvers with automatic labeling and deep learning," *Transp. Res. Rec., J. Transp. Res. Board.*, vol. 2674, no. 7, pp. 336–347, Jul. 2020, doi: [10.1177/0361198120922210](https://doi.org/10.1177/0361198120922210).
- [70] H. Wang, M. Gu, S. Wu, and C. Wang, "A driver's car-following behavior prediction model based on multi-sensors data," *EURASIP J. Wireless Commun. Netw.*, vol. 2020, no. 1, p. 10, 2020, doi: [10.1186/s13638-020-1639-2](https://doi.org/10.1186/s13638-020-1639-2).
- [71] J. Gesnoui, S. Pechberti, G. Bresson, B. Stanculescu, and F. Moutarde, "Predicting intentions of pedestrians from 2D skeletal pose sequences with a representation-focused multi-branch deep learning network," *Algorithms*, vol. 13, no. 12, p. 331, 2020, doi: [10.3390/a13120331](https://doi.org/10.3390/a13120331).
- [72] Z. Gu, Z. Li, X. Di, and R. Shi, "An LSTM-based autonomous driving model using a waymo open dataset," *Appl. Sci.*, vol. 10, no. 6, p. 2046, 2020, doi: [10.3390/app10062046](https://doi.org/10.3390/app10062046).
- [73] B. Kövári, F. Hegedüs, and T. Bécsi, "Design of a reinforcement learning-based lane keeping planning agent for automated vehicles," *Appl. Sci.*, vol. 10, no. 20, p. 7171, 2020, doi: [10.3390/app10207171](https://doi.org/10.3390/app10207171).
- [74] X. Geng, H. Liang, B. Yu, P. Zhao, L. He, and R. Huang, "A scenario-adaptive driving behavior prediction approach to urban autonomous driving," *Appl. Sci.*, vol. 7, no. 4, p. 426, Apr. 2017, doi: [10.3390/app7040426](https://doi.org/10.3390/app7040426).
- [75] Y. Xiao and Q. Nian, "Vehicle location prediction based on spatiotemporal feature transformation and hybrid LSTM neural network," *Information*, vol. 11, no. 2, p. 84, Feb. 2020, doi: [10.3390/info11020084](https://doi.org/10.3390/info11020084).

- [76] Y. Yoon, T. Kim, H. Lee, and J. Park, "Road-aware trajectory prediction for autonomous driving on highways," *Sensors*, vol. 20, no. 17, p. 4703, 2020, doi: [10.3390/s20174703](https://doi.org/10.3390/s20174703).
- [77] H. Zhang and R. Fu, "A hybrid approach for turning intention prediction based on time series forecasting and deep learning," *Sensors*, vol. 20, no. 17, p. 4887, Aug. 2020.
- [78] H. Zhang, Y. Wang, J. Liu, C. Li, T. Ma, and C. Yin, "A multimodal states based vehicle descriptor and dilated convolutional social pooling for vehicle trajectory prediction," SAE Tech. Paper 2020-01-5113, Jan. 2020, doi: [10.4271/2020-01-5113](https://doi.org/10.4271/2020-01-5113).
- [79] J. Bock, T. Beemelmans, M. Klösges, and J. Kotte, "Self-learning trajectory prediction with recurrent neural networks at intelligent intersections," in *Proc. 3rd Int. Conf. Vehicle Technol. Intell. Transp. Syst.*, 2017, pp. 346–351, doi: [10.5220/0006374003460351](https://doi.org/10.5220/0006374003460351).
- [80] B. Cheng, X. Xu, Y. Zeng, J. Ren, and S. Jung, "Pedestrian trajectory prediction via the social-grid LSTM model," *J. Eng.*, vol. 2018, no. 16, pp. 1468–1474, 2018, doi: [10.1049/joe.2018.8316](https://doi.org/10.1049/joe.2018.8316).
- [81] F. Tang and C. Li, "Intelligent vehicle lateral tracking control based on multiple model prediction," *AIP Adv.*, vol. 10, no. 7, Jul. 2020, Art. no. 075107, doi: [10.1063/1.5141506](https://doi.org/10.1063/1.5141506).
- [82] A. Li, H. Jiang, J. Zhou, and X. Zhou, "Learning human-like trajectory planning on urban two-lane curved roads from experienced drivers," *IEEE Access*, vol. 7, pp. 65828–65838, 2019, doi: [10.1109/ACCESS.2019.2918728](https://doi.org/10.1109/ACCESS.2019.2918728).
- [83] S. Liu, H. Liu, H. Bi, and T. Mao, "CoL-GAN: Plausible and collision-less trajectory prediction by attention-based GAN," *IEEE Access*, vol. 8, pp. 101662–101671, 2020, doi: [10.1109/ACCESS.2020.2987072](https://doi.org/10.1109/ACCESS.2020.2987072).
- [84] L. Tang, H. Wang, W. Zhang, Z. Mei, and L. Li, "Driver lane change intention recognition of intelligent vehicle based on long short-term memory network," *IEEE Access*, vol. 8, pp. 136898–136905, 2020, doi: [10.1109/ACCESS.2020.3011550](https://doi.org/10.1109/ACCESS.2020.3011550).
- [85] W. Wu, B. Yang, D. Wang, and W. Zhang, "A novel trajectory generator based on a constrained GAN and a latent variables predictor," *IEEE Access*, vol. 8, pp. 212529–212540, 2020, doi: [10.1109/ACCESS.2020.3039801](https://doi.org/10.1109/ACCESS.2020.3039801).
- [86] A. Abdelrahman, H. S. Hassanein, and N. Abu-Ali, "A cloud-based environment-aware driver profiling framework using ensemble supervised learning," in *Proc. IEEE Int. Conf. Commun. (ICC)*, May 2019, pp. 1–6, doi: [10.1109/ICC.2019.8761675](https://doi.org/10.1109/ICC.2019.8761675).
- [87] T. Fernando, S. Denman, S. Sridharan, and C. Fookes, "Neighbourhood context embeddings in deep inverse reinforcement learning for predicting pedestrian motion over long time horizons," in *Proc. IEEE/CVF Int. Conf. Comput. Vis. Workshop (ICCVW)*, Oct. 2019, pp. 1179–1187, doi: [10.1109/ICCVW.2019.00149](https://doi.org/10.1109/ICCVW.2019.00149).
- [88] D. Moher, A. Liberati, J. Tetzlaff, and D. G. Altman, "Preferred reporting items for systematic reviews and meta-analyses: The PRISMA statement," *Ann. Internal Med.*, vol. 339, no. 7716, pp. 332–336, 2009, doi: [10.1136/bmj.b2535](https://doi.org/10.1136/bmj.b2535).
- [89] W. Kim, M. S. Ramanagopal, C. Barto, M.-Y. Yu, K. Rosaen, N. Goumas, R. Vasudevan, and M. Johnson-Roberson, "PedX: Benchmark dataset for metric 3-D pose estimation of pedestrians in complex urban intersections," *IEEE Robot. Autom. Lett.*, vol. 4, no. 2, pp. 1940–1947, Apr. 2019, doi: [10.1109/LRA.2019.2896705](https://doi.org/10.1109/LRA.2019.2896705).
- [90] I. Goodfellow, J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley, S. Ozair, A. Courville, and Y. Bengio, "Generative adversarial networks," *Commun. ACM*, vol. 63, no. 11, pp. 139–144, 2020, doi: [10.1145/3422622](https://doi.org/10.1145/3422622).
- [91] B. R. Kiran, I. Sobh, V. Talpaert, P. Mannion, A. A. A. Sallab, S. Yogamani, and P. Perez, "Deep reinforcement learning for autonomous driving: A survey," *IEEE Trans. Intell. Transp. Syst.*, early access, Feb. 9, 2021, doi: [10.1109/ITITS.2021.3054625](https://doi.org/10.1109/ITITS.2021.3054625).
- [92] T. Gindele, S. Brechtel, and R. Dillmann, "A probabilistic model for estimating driver behaviors and vehicle trajectories in traffic environments," in *Proc. 13th Int. IEEE Conf. Intell. Transp. Syst.*, Sep. 2010, pp. 1625–1631, doi: [10.1109/ITSC.2010.5625262](https://doi.org/10.1109/ITSC.2010.5625262).
- [93] F. Hussain, R. Hussain, and E. Hossain, "Explainable artificial intelligence (XAI): An engineering perspective," 2021, *arXiv:2101.03613*. [Online]. Available: <http://arxiv.org/abs/2101.03613>
- [94] M. Uricar, P. Krizek, D. Hurych, I. Sobh, S. Yogamani, and P. Denny, "Yes, we GAN: Applying adversarial techniques for autonomous driving," in *Proc. Int. Symp. Electron. Imag. Sci. Technol.*, 2019, vol. 2019, no. 15, pp. 1–48, doi: [10.2352/ISSN.2470-1173.2019.15.AVM-048](https://doi.org/10.2352/ISSN.2470-1173.2019.15.AVM-048).
- [95] S. Kolekar, S. Gite, and B. Pradhan, "Demystifying artificial intelligence based behavior prediction of traffic actors for autonomous vehicle—A bibliometric analysis of trends and techniques," *Library Philosophy Pract.*, vol. 5132, pp. 1–24, Mar. 2021.
- [96] S. Gite and H. Agrawal, "Early prediction of driver's action using deep neural networks," *Int. J. Inf. Retr. Res.*, vol. 9, no. 2, pp. 11–27, Apr. 2019, doi: [10.4018/IJRR.2019040102](https://doi.org/10.4018/IJRR.2019040102).
- [97] P. Sharma, D. Austin, and H. Liu, "Attacks on machine learning: Adversarial examples in connected and autonomous vehicles," in *Proc. IEEE Int. Symp. Technol. Homeland Secur. (HST)*, Nov. 2019, pp. 1–7, doi: [10.1109/HST47167.2019.9032989](https://doi.org/10.1109/HST47167.2019.9032989).
- [98] S. R. Pokhrel and J. Choi, "A decentralized federated learning approach for connected autonomous vehicles," in *Proc. IEEE Wireless Commun. Netw. Conf. Workshops (WCNCW)*, Apr. 2020, pp. 1–6, doi: [10.1109/WCNCW48565.2020.9124733](https://doi.org/10.1109/WCNCW48565.2020.9124733).
- [99] M. Ryan, "The future of transportation: Ethical, legal, social and economic impacts of self-driving vehicles in the year 2025," *Sci. Eng. Ethics*, vol. 26, pp. 1185–1208, Sep. 2020.
- [100] M. Bassetton, U. Michieli, G. Agresti, and P. Zanuttigh, "Unsupervised domain adaptation for semantic segmentation of urban scenes," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. Workshops (CVPRW)*, Jun. 2019, pp. 1211–1220, doi: [10.1109/CVPRW.2019.00160](https://doi.org/10.1109/CVPRW.2019.00160).
- [101] J. Chen, B. Yuan, and M. Tomizuka, "Model-free deep reinforcement learning for urban autonomous driving," in *Proc. IEEE Intell. Transp. Syst. Conf. (ITSC)*, Oct. 2019, pp. 2765–2771, doi: [10.1109/ITSC.2019.8917306](https://doi.org/10.1109/ITSC.2019.8917306).
- [102] Y. Cao, C. Xiao, D. Yang, J. Fang, R. Yang, M. Liu, and B. Li, "Adversarial objects against LiDAR-based autonomous driving systems," 2019, *arXiv:1907.05418*. [Online]. Available: <http://arxiv.org/abs/1907.05418>
- [103] T. Nishio and R. Yonetani, "Client selection for federated learning with heterogeneous resources in mobile edge," in *Proc. IEEE Int. Conf. Commun. (ICC)*, May 2019, pp. 1–7, doi: [10.1109/ICC.2019.8761315](https://doi.org/10.1109/ICC.2019.8761315).
- [104] B. Ivanovic, K. Leung, E. Schmerling, and M. Pavone, "Multimodal deep generative models for trajectory prediction: A conditional variational autoencoder approach," *IEEE Robot. Automat. Lett.*, vol. 6, no. 2, pp. 295–302, Apr. 2021, doi: [10.1109/LRA.2020.3043163](https://doi.org/10.1109/LRA.2020.3043163).
- [105] M.-H. Jeong, T.-Y. Lee, S.-B. Jeon, and M. Youm, "Highway speed prediction using gated recurrent unit neural networks," *Appl. Sci.*, vol. 11, no. 7, p. 3059, Mar. 2021, doi: [10.3390/app11073059](https://doi.org/10.3390/app11073059).
- [106] T. Fernando, S. Denman, S. Sridharan, and C. Fookes, "Deep inverse reinforcement learning for behavior prediction in autonomous driving: Accurate forecasts of vehicle motion," *IEEE Signal Process. Mag.*, vol. 38, no. 1, pp. 87–96, Jan. 2021, doi: [10.1109/MSP.2020.2988287](https://doi.org/10.1109/MSP.2020.2988287).
- [107] T. Lebailly, S. Kiciroglu, M. Salzmann, P. Fua, and W. Wang, "Motion prediction using temporal inception module," in *Proc. Asian Conf. Comput. Vis.*, in Lecture Notes in Computer Science, vol. 12623. New York, NY, USA: Springer, 2021, pp. 651–665.
- [108] L. Peng, Y. Wang, F. Zhang, J. Zhang, and Z. Li, "Evaluation of emergency driving behaviour and vehicle collision risk in connected vehicle environment: A deep learning approach," *IET Intell. Transp. Syst.*, vol. 15, no. 4, pp. 584–594, 2021.
- [109] S. Gite, B. Pradhan, A. Alamri, and K. Kotecha, "ADMT: Advanced driver's movement tracking system using spatio-temporal interest points and maneuver anticipation using deep neural networks," *IEEE Access*, vol. 9, pp. 99312–99326, 2021, doi: [10.1109/ACCESS.2021.3096032](https://doi.org/10.1109/ACCESS.2021.3096032).
- [110] I. Kotseruba, A. Rasouli, and J. K. Tsotsos, "Joint attention in autonomous driving (JAAD)," 2016, *arXiv:1609.04741*. [Online]. Available: <http://arxiv.org/abs/1609.04741>
- [111] N. Schneider and D. M. Gavrila, "Pedestrian path prediction with recursive Bayesian filters: A comparative study," in *Proc. German Conf. Pattern Recognit. (GCPR)*, in Lecture Notes in Computer Science, vol. 8142. New York, NY, USA: Springer, 2013, pp. 174–183.
- [112] S. Pellegrini, A. Ess, K. Schindler, and L. van Gool, "You'll never walk alone: Modeling social behavior for multi-target tracking," in *Proc. IEEE 12th Int. Conf. Comput. Vis.*, Oct. 2009, pp. 261–268, doi: [10.1109/ICCV.2009.5459260](https://doi.org/10.1109/ICCV.2009.5459260).
- [113] A. Lerner, Y. Chrysanthou, and D. Lischinski, "Crowds by example," *Comput. Graph. Forum*, vol. 26, no. 3, pp. 655–664, 2007.
- [114] V. Alexiadis, J. Colyar, J. Halkias, R. Hranac, and G. McHale, "The next generation simulation program," *ITE J.*, vol. 74, no. 8, pp. 22–26, 2004.
- [115] R. Krajewski, J. Bock, L. Kloeker, and L. Eckstein, "The highD dataset: A drone dataset of naturalistic vehicle trajectories on German highways for validation of highly automated driving systems," in *Proc. 21st Int. Conf. Intell. Transp. Syst. (ITSC)*, Nov. 2018, pp. 2118–2125.



- [116] A. Geiger, P. Lenz, and R. Urtasun, "Are we ready for autonomous driving? The KITTI vision benchmark suite," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2012, pp. 3354–3361.
- [117] D. Lv, X. Ying, Y. Cui, J. Song, K. Qian, and M. Li, "Research on the technology of LIDAR data processing," in *Proc. 1st Int. Conf. Electron. Instrum. Syst. (EIS)*, Jun. 2017, pp. 1–5, doi: 10.1109/EIS.2017.8298694.



**SURESH KOLEKAR** received the master's degree in computer engineering from the College of Engineering, Pune (COEP). He is currently pursuing the Ph.D. degree with the Symbiosis Institute of Technology, Symbiosis International (Deemed University), Pune. His research interests include machine learning, deep learning, explainable artificial intelligence, and intelligent vehicle.



**SHILPA GITE** received the Ph.D. degree in deep learning for assistive driving in semi-autonomous vehicles from Symbiosis International (Deemed) University, Pune, India, in 2019. She is currently working as an Associate Professor with the Computer Science Department, Symbiosis Institute of Technology, Pune. She is also working as an Associate Faculty with the Symbiosis Centre of Applied AI (SCAAD). She has around 13 years of teaching experience. She is guiding Ph.D. students in biomedical imaging, self-driving cars, and natural language processing areas. She has published more than 60 research papers in international journals and 20 Scopus indexed international conferences. Her research interests include deep learning, machine learning, medical imaging, and computer vision. She was a recipient of the Best Paper Award at the 11th IEMERA Conference held virtually at Imperial College, London, in October 2020.



**BISWAJEET PRADHAN** (Senior Member, IEEE) received the Habilitation degree in remote sensing from Dresden University of Technology, Germany, in 2011. He is currently the Director of the Centre for Advanced Modelling and Geospatial Information Systems (CAMGIS), Faculty of Engineering and IT. He is also a Distinguished Professor with the University of Technology Sydney. He is also an internationally established Scientist in geospatial information systems (GIS), remote sensing and image processing, complex modeling/geo-computing, machine learning, soft-computing applications, natural hazards, and environmental modeling. Since 2015, he has been serving as the Ambassador Scientist for Alexander Humboldt Foundation, Germany. Out of his more than 700 articles, more than 650 have been published in Science Citation Index (SCI/SCIE) technical journals. He has authored eight books and 13 book chapters. He has been receiving 55 awards in recognition of his excellence in teaching, service, and research, since 2006. He was a recipient of Alexander von Humboldt Fellowship from Germany and Alexander von Humboldt Research Fellowship from Germany. From 2016 to 2020, he was listed as the World's Most Highly Cited Researcher by Clarivate Analytics Report as one of the world's most influential minds. From 2018 to 2020, he was awarded the World Class Professor by the Ministry of Research, Technology and Higher Education, Indonesia. He is an associate editor and an editorial member of more than eight ISI journals. He has widely traveled abroad, visiting more than 55 countries to present his research findings.



**KETAN KOTECHA** is currently an Administrator and a Teacher of deep learning. His research interests include artificial intelligence, computer algorithms, machine learning, and deep learning. He has expertise and experience in cutting-edge research and projects in AI and deep learning for the last 25 years. He has published more than 100 widely in several excellent peer-reviewed journals on various topics ranging from cutting edge AI, education policies, teaching-learning practices, and AI for all. He was a recipient of the two SPARC projects worth INR 166 lakhs from MHRD Government of India in AI in collaboration with Arizona State University, USA, and The University of Queensland, Australia. He was also a recipient of numerous prestigious awards, like the Erasmus+ Faculty Mobility Grant to Poland, the DUO-India Professors Fellowship for research in responsible AI in collaboration with Brunel University, U.K., the LEAP Grant at Cambridge University, U.K., the UKIERI Grant with Aston University, U.K., and a Grant from the Royal Academy of Engineering, U.K., under Newton Bhabha Fund. He has published three patents and delivered keynote speeches at various national and international forums, including at the Machine Intelligence Laboratory, USA, IIT Bombay under the World Bank Project, the International Indian Science Festival organized by the Department of Science and Technology, Government of India, and many more. He is an Associate Editor of IEEE Access journal.

...