IEEE Access
Multidisciplinary ⋮ Rapid Review ⋮ Open Access Journal

Received June 27, 2020, accepted July 13, 2020, date of publication July 21, 2020, date of current version August 3, 2020.

Digital Object Identifier 10.1109/ACCESS.2020.3010940

Insights and Strategies for an Autonomous Vehicle With a Sensor Fusion Innovation: A Fictional Outlook

FARRUKH HAFEEZ[1,2], USMAN ULLAH SHEIKH[1], (Member, IEEE), NASSER ALKHALDI[2], HASSAN ZUHAIR AL GARNI[2], ZEESHAN AHMAD ARFEEN[1,3], (Senior Member, IEEE), AND SAIFULNIZAM A. KHALID[1]

[1]School of Electrical Engineering, Universiti Teknologi Malaysia, Johor Bahru 81310, Malaysia
[2]Electrical and Electronic Engineering Department, Jubail Industrial College, Jubail 35541, Saudi Arabia
[3]Department of Electrical Engineering, The Islamia University of Bahawalpur (IUB), Bahawalpur 63100, Pakistan

Corresponding author: Usman Ullah Sheikh (usman@utm.my)

This work was supported by the Ministry of Education Malaysia and Universiti Teknologi Malaysia (UTM) for their support through the Research University Grant (GUP) under Grant Q.J130000.2523.19H61 and Grant Q.J130000.2523.19H45.

ABSTRACT A few decades ago, the idea of a car driving without human assistance was something inconceivable. With the advent of deep learning-based machine learning in artificial intelligence, this imaginary idea has become part of our life. Like in other fields, these technological revolutions have brought drastic changes to the field of automated driving systems. The autonomous vehicle is in the transition state between level 3 and level 4 of automation, but many mysteries are still waiting to be solved. Understanding the environment as precisely as a human driver is still far in the future. To attain human perception requires the capturing of extensive surrounding information that depends on the onboard sensors installed on the vehicle. Because the recent autonomous vehicle is equipped with several sensors, it captures surrounding information in diverse forms. Combining these multi-domain data with sensor fusion is the open area of research that is considered in this paper. Along with sensor fusion, another area of prime importance that is necessary to be explored is the prediction of pedestrian intentions. Though the study of the prediction of a pedestrian's intentions started approximately fifteen years ago, most of the research is based on detection rather than intention. Furthermore, this paper also discusses related research in the field of prediction of the pedestrian's intentions. At the end of the article, this review paper includes open questions, challenges, and proposed solutions.

INDEX TERMS Advanced driver assistance system, deep learning, pedestrian intention prediction, sensor, sensor fusion.

NOMENCLATURE

ACC	Adaptive cruise control
ADAS	Advanced driver-assistance systems
ADS	Automated driving systems
AEB	Automatic emergency braking
AI	Artificial Intelligence
AP	everywhere Autopilot on everywhere
AP	Highway Autopilot on highway
AV	Autonomous vehicle
CNN	Convolutional neural network
DARPA	Defence advanced research projects agency
DL	Deep learning
DM	Driver monitoring
DMM	Dynamical motion modelling
DMM	Dynamical motion modelling
GNSS	Global navigation satellite system
HOG	Histogram of oriented gradients tutorial
IMU	Inertial measurement unit
IRTAD	International traffic safety data and analysis group
IRTAD	international traffic safety data and analysis group
SAE	Society of automotive engineers
LDWS	Lane departure warning system
LIDAR	Light detection and ranging
LKA	Lane keep assistance
LRR	Long-range radars

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

PA	Park assistance
PBM	Planning-based models
PDM	Balanced Gaussian process dynamical models
PIP	Pedestrian intention predication
RADAR	Radio detection and ranging
SLDS	Switching linear dynamic system
SRR	Short-range radars
SVM	Space vector machine
TJA	Traffic jam assistance
V2I	Vehicle to Infrastructure
V2V	Vehicle to vehicle
VRU	Vulnerable road user

I. INTRODUCTION

According to the 2019 annual report [1] of the International Traffic Safety Data and Analysis Group (IRTAD), over 1.3 million people die annually and ten million people are seriously injured due to accidents. More than fifty percent of those injured are pedestrians, cyclists, and motorcycle users [2]. These statistics indicate that strong measures are required to control such accidents. In the context of autonomous vehicle technology, minimizing these accident rates is one of the prime objectives. However, there are many challenges in making this technology acceptable globally. The impact of technology advancement brought investors and automobile manufacturers into the field of autonomous vehicles. Current investment indicates that by 2050, the autonomous vehicle industry will reach $800 billion. Due to this expected growth, others in addition to automobile manufacturers, government agencies, universities, and academic research centres are devoting their full resources. Kettering University, North Carolina A&T University, Michigan State University, and the University of Toronto are preparing for an upcoming competition that will be held in the coming years [3]. Targets of this competition involve navigation of the automated driving mode in a dense urban environment.

The journey of autonomous vehicles (AVs) has continued for approximately thirty years now. In 1986, a project named PROMETHEUS, considered the first-ever autonomous vehicle project, started. Thirteen automobile vehicle manufacturers as well as nineteen universities and academic research centres were involved in this project [4]. In the US, the first AV-based project was started in 1988 under the name Navlab Thorpe [5] by Carnegie Mellon University. Following this project, in 1996, Japan formed the Advanced Cruise-Assist Highway System Research Association. Among the competitions, the most prominent is the DARPA Grand Challenge starting in 2004. The first competition was held in 2004, and the completion prize money offered was $1 million for the team that first finished a 150-mile route and crossed the California-Nevada border. In 2005, the second round of the DARPA competition was organized. Five vehicles completed the route. After two years, the third competition challenge, popularly known as the Urban Challenge, was held in California. The competition route was 96 km.

Recent autonomous vehicle competitions include the Auto Drive challenge competition, specially targeted for academics, which began in 2018 [3]. Starting in 2018, two successful competitions have been completed. Future competitions are planned for October 2020. The focus will be on urban driving conditions to improve vision and sensing algorithms.

Automobile giants have revised their budgets, trained their employees and formed alliances with software computing companies. More than 40 companies have been listed as developing autonomous vehicles [6]. BMW together with Daimler has allied with Intel Corporation. They are planning to build BMW iNEXT, which will be an open standard-based platform, by 2021. Audi, considered the first company to deploy hands-free autonomous vehicles, already has planned to spend $16B to put autonomous vehicles on the road by 2023.

The autonomous vehicle is also called an intelligent vehicle because of its capability to perceive the surrounding environment and, based on this perception, to take appropriate action. This sensing of environmental conditions is one of the prime steps in the field of automated driving systems that are needed to observe all possible aspects of human brains to reach this perception level. The typical framework of the autonomous vehicle consists of five components: Perception, Localization and Mapping, Path Planning, Decision Making, and Vehicle Controlling. Perception plays a role exactly like human sensing from eyes to monitor the environment and perceive data from the sensor. For these sensing purposes, several sensors are required to collect surrounding data. Based on the data received, localization of vehicles locally and globally is the second step that is designated the Localization and Mapping step. Path Planning is the third step to determine a route based on data received from the sensors. The fourth step (Decision Making) calculates the best possible route based on environmental data, current vehicle conditions, and all possible available paths. In the end, Vehicle Control is responsible for implementing this decision generated from the Decision Making block that can be a change of lane action, slowing down near a pedestrian crosswalk, stopping on a red signal, etc. Figure 1 portrays the framework of an intelligent vehicle portfolio. AV perceives surrounding information, and based on the perceived data and local position, further actions will be determined.

Environmental perception is considered the first and foremost component in the involvement of autonomous vehicles and includes road structure, lane on the road, traffic signs, traffic signals, Vehicle-to-Vehicle (V2V) communication, infrastructure presence, observations of a vulnerable road user, etc.

Figure 2 indicates potential research in these areas [7]. Several research studies and algorithms have been proposed in areas such as localization and mapping and lane and road

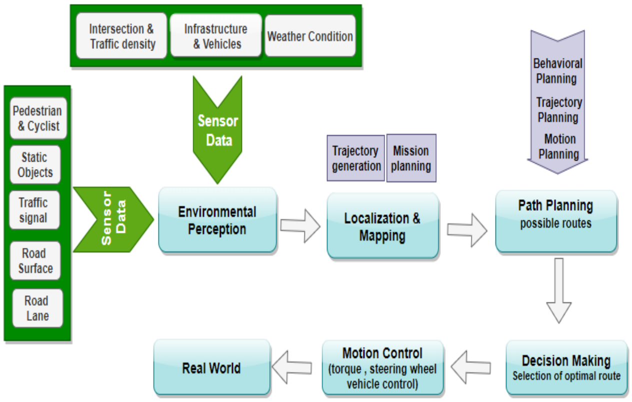

FIGURE 1. A typical framework of an autonomous vehicle process.

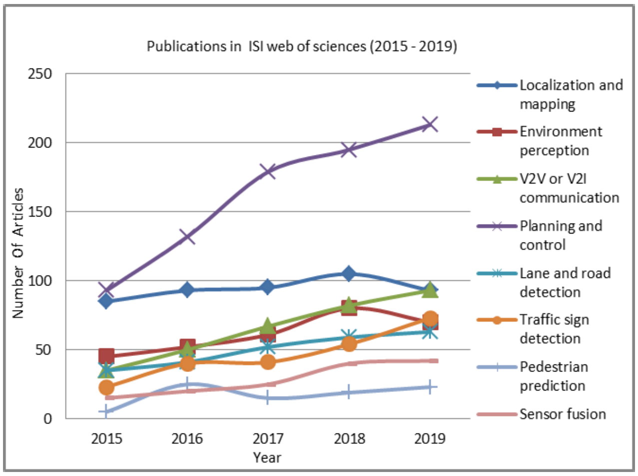

FIGURE 2. The number of publications in the Web of Science database.

detection, but areas such as sensor fusion and prediction of pedestrian intentions are still being explored.

Methods and algorithms that have been proposed previously are still limited to factors such as specific areas or within certain premises and campuses, under uniform weather conditions because algorithms are trained and tested in only sunny weather, etc. Factors such as regional climate variations, the impact of social variation, and cultural norms are not considered or not given enough importance in studies. The implementation of robust algorithms is not possible without considering these elements.

This paper provides a comprehensive literature review in two areas of AV and discusses sensor technology, the output forms, algorithms, and research related to sensor fusion. This paper also covers one of the most important aspects of AV, which is the interaction of AVs with pedestrians. To minimize road accidents by AVs, intelligent algorithms must reach the extent of human understanding capabilities. Special consideration is given to recent studies undertaken in the estimation of pedestrian intentions based on state-of-the-art techniques, especially research work in deep learning (DL). Finally, areas of improvement, unexplored and less explored areas are dug out, and their proposed remedial solution is suggested.

This paper no doubt will provide a comprehensive review with upcoming trends and provide a quick guide to researchers working in the field of sensor fusion and pedestrian intention.

For a fast insight, the paper is segmented into the following sections: Section II covers smart sensors currently used in AVs, their strengths, limitations and data forms. Section III describes sensor fusion techniques. Section IV gives coverage to intelligent algorithms related to pedestrian intentions. Finally, the paper is concluded in Section V.

II. SENSORS IN AV

Sensors function as a source of data collection from the surrounding environment in an AV. These data are then sent for further intensive processing using installed computing devices. If data are representing a true and accurate representation of the surroundings, then correct action is possible [8]. Errors or missing measurements in collecting data may lead to an irremediable loss. Driver assistance facilities depend wholly upon sensors that have been installed on the AV. Table 1 shows the level of automation with a sensor installed and driving assistance facilities. Each sensor extracts surrounding information that can be used for exploring environmental perception in different domains of the study, and this information is summarized in Table 2. Therefore, automobile manufacturers have been applying combinations of sensors. A brief historical application of sensors in commercial and research vehicles can be seen in Table 3. The coming section features elaborate application of these sensors in the field of AVs, as well as their limitations and drawbacks.

A. CAMERA

A camera is one of the basic sensors used in an autonomous vehicle, to accurately identify positions of objects around it. A variety of technologies in cameras can be classified based on coordinate systems or brightness level variation. For the operation of an autonomous vehicle under dynamic conditions, it is required to deploy several types of cameras as shown in Table 2. For example, under low visibility, cameras with dynamic brightness levels work more efficiently than other cameras. In recent AV technology, several different types of cameras are installed. In the subsection, some common types of cameras used in current autonomous vehicles are briefly discussed.

1) MONOCULAR CAMERA/MC

The 2D-camera produces the target object image effectively as a flat two-dimensional plane view. The 2D-image does not provide any height information at all - there are X and Y data, but no Z-axis depth of field data. Images produced with different viewpoints create completely different contours, causing a machine vision to have confined class in applications where information about the shape is critical to performing a task. Currently, in the advanced driver assistance system (ADAS), the monocular camera is used for blind spots, sideways motion, parking assistance,

TABLE 1. Levels of automation in possible sensors, driving facility [9].

Level of Automation	Driving system facility	Year	Sensors	(No. of sensors)	Assistance System
Level 1 Driver Assistance	An advanced driver assistance system by either steering or braking/accelerating (but not both) can assist the driver significantly	2012	Ultrasonic Radar LRR Camera for surroundings	4 1 1	ACC LDWS
Level 2 Partial Assistance	Under some circumstances, both steering and braking/accelerating can be controlled simultaneously by an advanced driver assistance system, but the driver must continue to pay full attention	2015-2019	Ultrasonic Radar LRR Radar SRR Camera for Surround	8 1 4 4	ACC LDWS PA LKA
Level 3 Conditional Assistance	Under all circumstances, all aspects of the driving task can be performed by the driver assistance system, but the driver must be ready to take back control at any time when required	2020-2028	Ultrasonic Radar LRR Radar SRR Long-distance camera Camera for Surround Stereo camera Ubolo Lidar Dead Reckoning	10 2 6 2 5 1 1 1 1	ACC LDWS PA LKA AEB DM TJA
Level 4 High Assistance	Perform all driving activities and monitor the driving environment; in some cases, do all the driving. In these circumstances, the human being does not need to pay attention	2026-2035	Ultra-sonic Radar LRR Radar SRR Long-distance camera Camera for Surround Stereo camera Ubolo Lidar Dead Reckoning	10 2 6 2 5 1 1 1 1	ACC LDWS PA LKA AEB DM TJA Sensor Fusion AP Highway
Level 5 Full Assistance	The human occupants are just passengers and need never be involved in driving	2035 onwards.	Ultrasonic Radar LRR Radar SRR Camera for Surround Long-distance camera Stereo camera Ubolo Lidar Dead Reckoning	10 2 6 5 6 2 1 1 1	ACC LDWS PA LKA AEB DM TJA Sensor Fusion AP Highway AP everywhere

*****ACC** Active cruise control **LDWS** Lane Departure Warning System **PA** Park Assist **LKA** Lane Keep Assist **DM** Driver Monitoring
AEB Automatic Emergency Braking **TJA** Traffic Jam Assist **AP** Autopilot on Highway **AP*** Autopilot on everywhere.

lane recognition for keeping in the lane and crosswalk recognition. 2D-cameras are smaller, cheaper, and easier to install, and calibration can be performed easily. Precise object detection and incorrect vertical distance are among the major problems encountered when using a monocular camera.

2) STEREO CAMERA/SC

Contrary to the 2D-camera, the 3D-camera no longer produces a flat picture. In three-dimensional point clouds of precise coordinates, the position of every pixel in space is known. The 3D-camera simultaneously provides X-, Y- and Z-plane data as well as the respective rotational information. Three-dimensional recognition of the running environment of the vehicle is becoming important to recognize the environment. Therefore, the knowledge of the depth information of the object is required. The 3D-camera can extract in-depth information. Static Object Detection such as a traffic sign, traffic light, and lane detection together with dynamic information can be obtained accurately. A difficult calibration process and computationally complex algorithms are required for object detection and recognition.

TABLE 2. Sensor performance invariant fields [10], [11].

Perception of Target	Camera			Radar		Lidar	Ultrasonic
	2D	3D/IR	Event	SSR	LSR		
Object detection	✓	✓✓	✓✓	-	-	✓✓	✓✓
Object classification	✓✓	✓✓	✓✓	-	-	✓	*
Distance estimation	✓	✓✓	✓✓	✓✓	✓✓	✓✓	✓✓
Edge detection performance	✓✓	✓✓	✓	-	-	✓✓	✓✓
Lane tracking	✓✓	✓✓	✓✓	-	-	*	-
Visibility range	✓	✓✓	✓✓	✓✓	✓✓	<200	-
Poor weather	*	*	✓	✓✓	✓✓	✓	✓
Dark or low illumination performance	*	*	✓✓	✓✓	✓✓	✓✓	✓✓
Range (in metres)	*	few metres	-	2-30	30-150	<200	Up to 2

✓✓ Sensors perfectly suited ✓Sensors with good performance *Sensors may be used with extensive processing
**Sensors may meet the criteria but may have shortcomings

3) INFRARED CAMERA/IR

The infrared camera provides the complete and reliable coverage needed to make AVs safe and functional in any environment at any time in day or night. The IR camera senses radiated signals generated from target objects. As it scans above visible light, the IR camera may detect objects that may not be perceptible to any Camera, Radar, or Lidar. Veoneer, the Swedish autonomous vehicle manufacturer, has been awarded a contract to design the first IR camera for level automation [30].

4) EVENT CAMERA/EC

Event cameras are sensors that are bio-inspired and operate drastically differently from traditional cameras. Rather than capturing images at a fixed rate, they asynchronously measure changes in brightness per pixel, leading to a stream of events encoding the time, location, and sign of the changes in brightness. Event cameras have exceptional properties when compared with traditional cameras: very high dynamic range, low power consumption, and high time resolution. However, since the output consists of a series of asynchronous events rather than individual intensity images, a conventional vision algorithm cannot be applied, so new deep learning-based dynamic algorithms are needed [31].

B. RADIO DETECTION AND RANGING/RADAR

Radar-based radio wave detection systems have been used for decades to accurately calculate the location, speed, and direction of aircraft, warships, and other objects in motion. In the field of AVs, advanced cruise control (ACC) and automatic emergency braking (AEB) are the application areas of radar. Radar works efficiently in dark, rainy, or even foggy weather and is nearly impervious to adverse weather conditions [32]. Figure 3 shows a clear picture of pedestrian recognition while using radio beams [33].

- RADAR GRADING

To fully reach human perception, multiple types of radar are required to be deployed in an autonomous vehicle. The combination of these multiple types of radar provides precious statistics for superior driver assistance structures occurrence range versus long-range radar coverage [34].

- *Short-Range Radar*/SRR

SRRs use the 24-GHz frequency and are used for short-range applications such as blind-spot identification, parking assistance, or detection of obstacles and avoidance of collisions. With an operating range of up to 30 metres, the radar sensor can be used to warn against unidentified threats.

- *Long-Range Radar*/LRR

LRRs that use the 77-GHz (76-81-GHz) band provide better accuracy and resolution in a smaller packet. They are used to measure distance, speed of other vehicles and object detection within a wider field of view such as cross-traffic alert systems. Long-range applications require antennas that

TABLE 3. Emergent progress in AV with intelligent sensors.

Year	Platform	Camera			Radar	Lidar	Ultrasonic	Sonar
		2D	3D	IR				
2008	Boss [12]	✓	-	-	✓	✓	-	-
	Odin [13]	-	-	-	-	✓	-	-
	Junior [14]	-	-	-	✓	✓	-	-
2010	BRAiVE [13]	✓	✓	-	-	✓	-	-
2011	Junior [14]	✓	✓	-	-	✓	-	-
	BRAiVE [15]	-	-	-	-	-	-	-
	BCI [16]	✓	✓	-	✓	✓	-	-
2013	Mercedes-Benz E and S-Class [17]	✓	✓	✓	✓	✓	✓	✓
	Audi's Research Vehicle [18]	✓	✓	✓	✓	✓	✓	✓
2014	Bertha [19]	✓	✓	-	✓	-	-	-
	A1 [20]	✓	✓	✓	-	✓	-	-
	Ford [23]	✓	-	-	✓	-	-	✓
	BMW [21]	✓	-	-	✓	✓	✓	-
	ZMP Robocar HV [22]	✓	-	-	-	✓	-	-
2015	Infiniti Q50S [23]	✓	-	-	✓	-	-	✓
	Lexus RX [24]	✓	✓	-	✓	-	-	✓
	Volvo XC90 [25]	✓	✓	-	✓	-	-	✓
2016	Otto Semi-Trucks [26]	✓	-	-	✓	✓	-	-
2017	Nav [27]	✓	-	✓	-	✓	-	✓
	Robot Car [28]	✓	-	-	✓	-	-	-
2018	Uber car Fusion [29]	✓	✓	-	✓	✓	-	-
	Uber car (XC90) [29]	✓	✓	-	✓	-	-	-
2019	Apollo Auto [29]	✓	✓	-	✓	✓	-	-

provide a higher resolution within a more limited range of scanning. Long-range radar (LRR) systems provide ranges between 80 m and 200 m or greater.

C. LIGHT DETECTION AND RANGING/LIDAR

Lidar uses invisible laser light to determine the distance to objects. In autonomous vehicle technology, Lidar provides

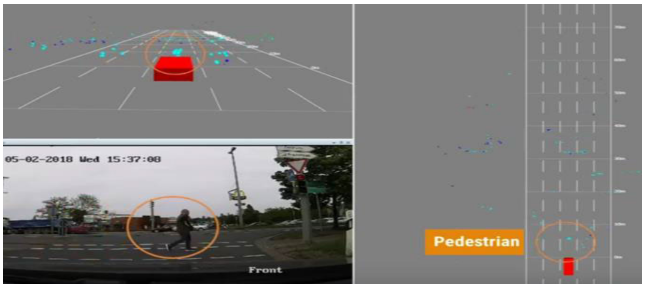

FIGURE 3. Pedestrian Detection using radio wave detection [33].

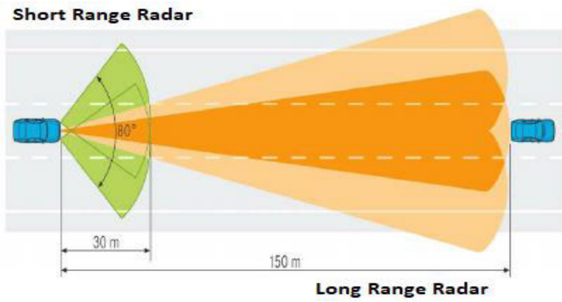

FIGURE 4. Comparison of SSR vs LRR [34].

the highest possible understanding of the traffic, road users and potential hazards surrounding the vehicle. Lidar can measure up to 100 m distance with an accuracy of 2 cm. Lidar is also unaffected by adverse weather conditions such as wind, rain, and snow, and could even be used in heavy snow conditions to map inaccessible areas [35]. For understanding, the images generated by Lidar, particularly for urban areas, are displayed in Figure 5.

FIGURE 5. Typical Urban Area view generated by Lidar Sensor [36].

D. ULTRASONIC SENSORS/US

Ultrasonic sensors emit short high-frequency sound pulses. These propagate at the speed of sound in the air. If they hit an object, they are reflected as echo signals to the sensor, which itself determines the distance to the target based on the timespan between the signal being emitted and the echo being received. In the ADAS, the ultrasonic sensor is commonly used for parking in small parking spaces and emergency braking at the low speed [37]. Figure 6 shows the workings of an ultrasonic sensor in different electromagnetic spectrum bands [38].

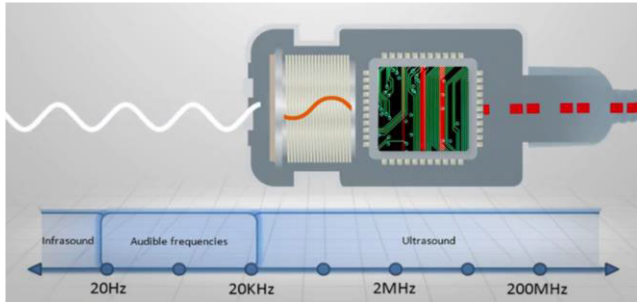

FIGURE 6. Ultrasonic sensor pulse generation [38].

The introduction summarized above specifies the application of these sensors in capturing surrounding data. After obtaining this large amount of data, combining these multiple domain data to extract desired information is the next process, designated as sensor fusion, explained next.

III. SENSOR FUSION

According to a report from the National Transportation Safety Board about an Uber crash [39], the vehicle detected the pedestrian six seconds before the accident. The autonomous driving system classified the pedestrian as an unidentified object, first as a car and then as a bicycle. In other words, the vehicle sensors detected the victim, but the software wrongly determined that it was not in danger and that no evasive action was required. The aforementioned case study reveals that there are still areas of improvement in the perception of the environment under different conditions, road users and surroundings. The potential area includes more accurate and precise sensors for a better understanding of the surroundings and, importantly, combining these different multidomain data obtained from various sensors is the evolution of sensor fusion. Figure 7 portrays multiple domain data generated from sensors on the AV that are processed using AI-based algorithms and perceiving the environment. This whole process is like human perception using the human sensory system.

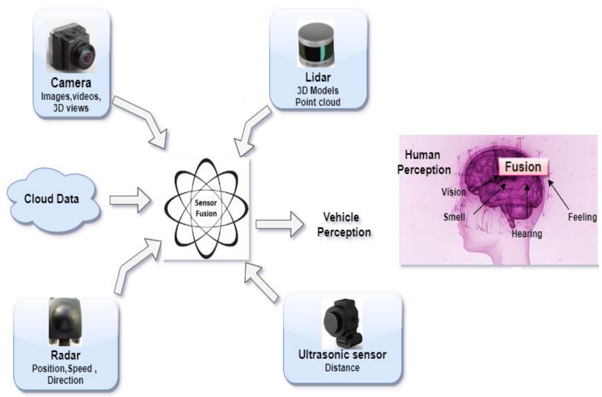

FIGURE 7. Sensor fusion process cycle.

Sensor fusion is a critical element in developing the "brain" of an autonomous vehicle, ensuring intelligent, accurate and timely decisions based on the actions of other participants in traffic.

A. PREVIOUS STUDIES

A remarkable model was presented in the Grand Cooperative Driving Challenge [40]. The vehicle had installed cameras and radar. The fusion of these sensors and the algorithm generates an accurate path and trajectory planning. Strong Vehicle to Vehicle (V2V) communication and collision avoidance was achieved by a sensor fusion algorithm. Li [41] presented a model in which camera and Lidar fusion were used. Fusion results in a real-time drivable region under different road conditions. Lane detection for a structured road was obtained successfully. However, the application was not utilized in the densely urban region. In the paper by Omar *et al.* [42], the authors applied the fusion model, using Cameras, Radar and Lidar. The fusion algorithm was applied in detection, classification, and tracking in urban areas and highways. However, from fusion, the algorithm obtained better results than the conventional operation in classification models, but misclassification was still considerably high. Adverse weather conditions constitute a big challenge even in the transition state of level II and level III automation. Lee [43] presented a simple model that can detect road and lane in adverse weather conditions. The sensing system consists of Cameras and Lidar. The model's weakness includes the detection of curves and sometimes lane detection issues. Chen [44] used a Lidar point cloud and RGB images. The proposed model is based on detecting 3D objects in road scenes. The results were compared with thestate of art 2D algorithm, which obtained accurate 3D locations, size, and orientation of objects.

De Silva [45] discussed issues related to fusion in his research. The proposed method applied a Gaussian process regression algorithm on Lidar with wide-angle camera data, achieving better accuracy and precision in comparison with resolution-matching algorithms. In another research paper, the robust and precise localization algorithm model was presented by Wan *et al.* [46]. They obtained an average accuracy between 5-10 cm in location. The authors used GNSS, Lidar, and IMU in their proposed model, and a Kalman filter was applied to calculate uncertainty estimation. Caltagirone [47] used fully convolutional neural networks in fusing Lidar points of cloud and RGB images. The area of interest was road detection in an urban area. Caltagirone obtained an accuracy of 96.03% in the urban road category. The results were tested and evaluated on the KITTI road benchmark, though the research was limited to specific road types. Shahian [48] presented his model based on Fully Convolutional Neural Network (FCNx), and a traditional Extended Kalman Filter (EKF) used the nonlinear state estimator method. The model worked efficiently in environmental perception areas such as obstacle detection, road segmentation, and tracking. The researchers used a combination of camera, Lidar, and radar. Table 4 summarizes the potential research in the field of sensor fusion.

B. BARRIERS AND IMPEDIMENTS OF SENSOR FUSION

The overall research indicates achievements in the perception of the environment, but reaching human perception is still far away. Some of the challenges in sensor fusion include

- Multimodal sensor nature: Every sensor has its data format, physical units, and differences in spatiotemporal alignment. Processing and extracting desired information from multimodal data is one of the open challenges.
- Data source uncertainty includes challenges such as quantitation errors, noise filtering, calibration errors, or loss of precision, inconsistent data, and missing values and differences in data source reliability.
- Cost reduction: For designing low-cost vehicles, it is necessary to design reduced fusion boxes for easy placement in the vehicle, which applies a limitation on the fusion electronics board and proper cooling design considerations.
- GPS spoofing attacks are also a continuous threat due to the presence of smart sensors.
- The design of less complicated, low computational power and more robust algorithms is still an open challenge in sensor fusion technology.

IV. PEDESTRIAN INTENTION PREDICTION/PIP

Knowing the intention of pedestrians is one of the critical aspects of the autonomous vehicle system. Even after tremendous achievements in deep learning algorithms, pedestrian intention behaviour still has a big space for improvement [51]–[54]. Understanding pedestrian activities and behavioural prediction require that several factors be considered. Beyond human nature, factors such as demographics, environmental conditions, cultural attributes and spatiotemporal factors play an important part in determining intention, as displayed in Figure 8. To study pedestrian behaviour, numerous approaches such as using observation of pedestrians [55], video recording [56], [57], image sensing [58], simulations [59], [60], questionnaires [58], literature surveys [59], [61] and conducting interviews [62] have been adopted. Studies based on pedestrian behaviour can be classified into two broad stages: the first stage can be called general pedestrian studies, and the other stage is in the context of autonomous vehicle studies. General pedestrian studies started in the first twenty years of the 20th century [63]. The research was based on studies of pedestrian factors such as the difference between pedestrian behaviour when walking alone or in a group [64], pedestrian demographics [65], road structure and pedestrian speed [63]. Summarizing, pedestrian research in the early phase can be classified based on pedestrian and environmental factors. However, interaction with vehicles was not extensively addressed during that period. With an increase in the number of vehicles, accident rates began to grow, and researchers and law enforcement analysts started thinking about factors to minimize accident rates [63]. In the second era of pedestrian studies in the context of the autonomous vehicle, researchers highlighted factors that

TABLE 4. Literature survey in sensor fusion.

Related Projects	Sensors procurement	Issues/algorithm designed for	Strengths	Flaws
Grand Cooperative Driving Challenge [40]	Cameras and Radar	Path and trajectory planning collision checking trajectory generation based on fast lattice search	V2V &V2I communication	Overheating of car. The exact position of the location
Qingquan Li [41]	Cameras and Lidar	Real-time optimal-drivable-region for structured or Non-structured roads. Lane detection.	Detected drivable region up to 88% accuracy.	Not suitable for densely occupied vehicles
Felix Kunz [49]	Cameras and Radar	Environmental perception. Motion planning.	Motion planning algorithm updating with a speed of 3 ms.	Needs a highly precise digital map
Ricardo Omar [42]	Cameras, Radar and Lidar	Detection, classification, and tracking in urban areas and highways	Improved detection level on highways	High % of misclassification in urban areas
Unghui Lee [43]	Cameras and Lidar	Road detection, lane detection in adverse weather conditions.	Simple algorithms and can be processed easily.	Lane recognition errors. Vulnerable in sharp curve sections,
Naman Patel [50]	Cameras and Lidar	Autonomous navigation in the indoor environment	Easily recognizes surrounding objects in an indoor environment	Limited for the indoor environment
Xiaozhi Che.[44]	LIDAR point cloud and RGB images	3D object detection in an autonomous driving scenario	Obtained accurate 3D locations, sizes and orientation of objects	The proposed algorithm image detection precision is less than the Mono3D method.
Varuna De Silva [45]	Lidar data with a wide-angle camera data	An issue in fusing heterogeneous sensor data. Developing robust fusion algorithm of heterogeneous sensor data	Obtained better accuracy and precision in comparison with resolution matching algorithms	Trained classifiers on a very small amount of data.
Guowei Wan [46]	GNSS, Lidar and IMU.	The robust and precise localization algorithm	Obtained an average accuracy between 5-10 cm in localization.	Tests performed under normal weather conditions.
Luca Caltagirone [47]	LIDAR point of cloud and RGB images	Road detection using cross fusion method	Obtained accuracy of 96.03 in the urban road category	Needed to test under different road conditions.
Babak Shahian Jahromi [48]	Cameras, Radar and Lidar	Hybrid multisensor fusion for environment perception	Obtained better space segmentation accuracy	Any sensor failure may highly influence detection and tracking.

involve pedestrian intentions as in the general pedestrian era [67]. The scholar Rothenbücher *et al*. [57] highly emphasized the importance of communication between pedestrians and vehicles. Before pedestrian intention prediction, the initial research was based on the detection of pedestrians, which can be assumed to be the primary stage of prediction of intentions. After detecting pedestrians, tracking the pedestrian by its pose or other means is necessary to avoid accidents. Deep learning has been used and has obtained significant results in estimation of pedestrian intentions. Two other approaches, namely, dynamic motion modelling (DMM) and Planning-based Models (PBM), have been used for predicting pedestrian intentions. A discussion of these approaches can be found in the next subsection.

A. DYNAMIC MOTION MODELLING/DMM

DMM is the general approach for the future location of pedestrians based on motion trajectory. Position measurements are obtained using a vision-based pedestrian detector. Schneider [68] used a Bayesian filter, a type of extended Kalman filter for predicting pedestrian trajectory. The results were tested for four different pedestrian dynamics (namely, starting, stopping, bending and crossing). Special care was given to optimization parameters and sensor modelling. Quintero [69] applied Balanced Gaussian Process Dynamic Models (B-GPDMs) and a Naïve-Bayes classifier to predict and pose pedestrian locations and classify intentions within only one second. Both classifiers were combined to increase the accuracy of the action hierarchy. An accurate path

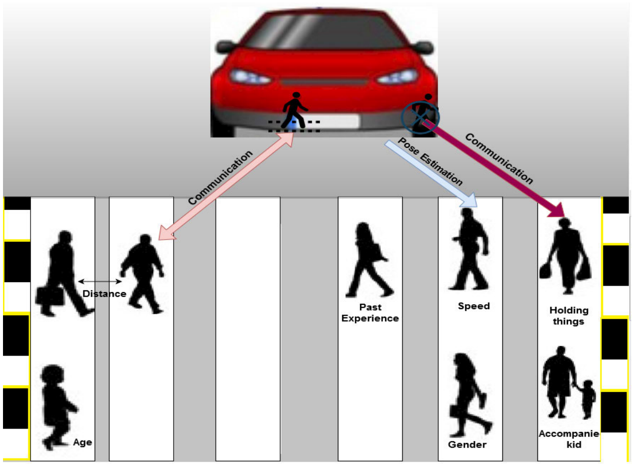

FIGURE 8. Factors involving a pedestrian and pedestrian-to-vehicle communication.

prediction with a mean error of 24.4 cm for walking trajectories was obtained, with 26.67 cm and 37.36 cm for stopping and starting trajectories. Flohr [70] proposed a model that integrates the pedestrian situational awareness, situation criticality and spatial structure of the area as latent states with a switching linear dynamic system (SLDS) to predict changes in pedestrian dynamics. By using pedestrian head orientation, situational awareness is determined. The expected point of closest approach and spatial layout is considered for estimating situation criticality.

B. PLANNING-BASED MODELS/PBMS

In the planning-based model, the pedestrian's future movements are not based explicitly on the intentions of the targets. Instead, they assume that the target (i.e., the pedestrian) has the intention of reaching a specific destination. Kitani [71] proposed future prediction- and forecasting-based application of optimal inverse control over computer vision. The proposed model also considered factors such as continuous activity analysis for improving and obtaining a more accurate result. Ziebart [72] proposed a model. Results were obtained from testing the model in a close environment by using a maximum entropy inverse optimal control method for goal-directed trajectories of pedestrians. This approach proved to be successful in obstacle-sensitive areas. The Karasev [73] model was based on jump-Markov processes to model pedestrian behaviour and predict pedestrian long-term planning trajectories by presuming pedestrian status using a Rao-Blackwellized filter and planning accordingly to a stochastic strategy that represents individual preferences to achieve the same destination. Quintero [74] used three-dimensional body actions and behaviour to predict the destination path of the pedestrian. The authors proposed Gaussian Process Dynamic-based models for this purpose and obtained accuracy up to 7 cm for path prediction, 20 cm in walking trajectories. Angelova [75] applied a deep convolutional neural network (CNN) and used proxy labels during the training process. He found that proxy label-based learning gave more accuracy and stability in results. The output is defined by stop/go logic for pedestrian or cyclist behaviour.

C. DEEP LEARNING APPROACH/DL

A data-driven approach has proved itself a more efficient and best-estimating intention approach for a pedestrian when applied and tested with datasets. The data-driven approach evaluates by using different performance matrices such as bounding box misses, accident rates, and trajectory error. Based on the impressive outcomes, presenting an in-depth literature review of pedestrian intention estimation and predictions in upcoming lines is summarized in Table 5 for a quick overview.

D. PREVIOUS CONTRIBUTIONS

Fang [76] proposed CNN pose-based estimation and a Support Vector Machine (SVM) for classification purposes. The authors proposed that pose estimation can be used to determine the intention of a pedestrian with simple monocular images without using exhaustive algorithms and datasets. A modelled human 2D skeleton structure under different body poses such as bending or turning to stop is understudied. Sebastian [77] proposed a model that generates silhouette-form vehicles equipped with cameras used to detect pedestrian intentions applied to a Motion Contour Image-based HOG-like descriptor (MCHOG) with an SVM classifier for image detection and classification. The authors obtained stopping intentions from 125 - 500 ms before standing with an accuracy range of 80% to 100%. Völz1 [78] proposed Lidar-based images and applied Long-Short-Term-Memory network-based algorithms. The approaches were validated with real-world trajectories and obtained 10-20% accuracy compared with previous methods. Pedestrian prediction analyzed and predicted near crosswalks was noticed. Rehder [79] proposed that instead of complete body part evaluation, even head orientation can be applied to predict intention. Head Detection is performed by the HOG/SVM cascade classifier while orientation is performed by the logistic regression model. Datasets were built for training and testing purposes. Saleh [80] suggested a novel end-to-end data- a data-driven method for the long-term prediction of VRUs, such as pedestrians in urban traffic, based solely on their trajectories.

The problem of intent prediction was conceived as a problem of time series prediction by merely observing a short-window sequence of pedestrian motion trajectory. A forecast of their future lateral positions was made up to 4 seconds ahead. The authors obtained 1.72% more average orientation similarity than other models. Ovidiu [81] proposed a model that was based on Retina net-based detection and classification of pedestrians.

The calculation of timing to cross a road is estimated by a recurrent neural network (LSTM). The JAAD dataset is used for testing and validation purposes. Pedestrian detection performance was to be more accurate than action recognition

TABLE 5. Representative work in pedestrian intention prediction.

Study	Sensor & Data set used	Technique Applied	Outcomes	Demographics/Area Typology and Environmental Considerations
Zhijie Fang [76]	Monocular camera ground truth (GT) [78]	CNN-based posed-based estimation and Support Vector Machine for classification purposes.	Pose estimation can be used to determine the intention of a pedestrian with simple monocular images without using exhaustive algorithms and datasets.	Single Urban environment No considerations
Sebastian [77]	Stereo camera. Daimler Pedestrian Dataset [10]	Motion Contour image-based HOG-like descriptor (MCHOG) with a linear Support Vector Machine (SVM) classifier.	Silhouette-generated form vehicle-equipped cameras are used to detect pedestrian intentions.	Single/Group Urban environment No consideration
Völzl B [79]	Lidar-based 2D images AdaDelta	Long-Short-Term-Memory networks applied to Lidar-based images.	Pedestrian prediction analyzed and predicted near crosswalks. The approaches were validated with real-world trajectories and obtained 10-20% accuracy compared with previous methods	Single Urban environment No considerations
Eike Rehder [80]	Monocular camera Self-constructed dataset	Detection performed by the HOG/SVM cascade classifier. The logistic regression model performs orientation.	Used head orientation in a highly occluded urban environment to predict intention.	Single Urban environment, highly occluded
Khaled Saleh[81]	Daimler Pedestrian Benchmark Data Set	Long-Short-Term Memory networks (LSTMs) architecture	Obtained competent results in comparison to other approaches with a smaller mean lateral error.	Single Urban environment No consideration
Chenchen Zhao [82]	Monocular camera KITTI Data set	A monocular based feed-forward neural network approach	1.72% more average orientation similarity than other models	Single Urban environment No considerations
Schneider [68]	Monocular camera KITTI Data set	Bayesian filters for pedestrian path prediction	Position measurements were obtained using a vision-based pedestrian detector	Single Urban environment No considerations

in comparison with state of art methods. Karam [83] used a depth camera and applied a convolution neural network to classify different pedestrian orientations. Three body landmarks (shoulder, neck, and face) are used to determine orientation. Overall, 85% accuracy was claimed for different pedestrian actions.

Although researchers are still currently applying these three intention prediction models, DMM and PMM have their limitations such as that DMM assumes that all trajectories will have similar dynamics, which is not always the case and finally leads to lower accuracy, especially for long-term prediction. PBM presumes a final destination, which is difficult to predict based on current actions. However, DL proved to be more robust than DMM and PMM techniques. DL algorithms such as RNN + LSTM were found to be more effective and accurate in real-time unseen situations.

E. REMEDIAL ACTIONS WITH FUTURE OUTLOOK

Understanding the intentions of pedestrians exactly like human drivers and responding accordingly is still far from achievement. Reaching this level needs several factors to be improved.

- Comprehensive studies based on different scenarios and conditions are required to face all possible pedestrian actions. Factors such as social norms, pedestrian demographics, group size, and pedestrian distance cannot be neglected for true detection.
- The mode of communication between vehicle and pedestrian plays an important role, although different approaches and mechanisms have been adopted for testing purposes.
- Standardization and exploration of stress-free pedestrian communication must be set by automation organizations.
- Most of the current algorithms are based on the dynamics of pedestrians. To reach human perception, such algorithms should be designed to also consider the surroundings of pedestrians.
- To date, the dataset used for testing purposes is extremely limited. Public testing and validation of large-scale datasets under different weather conditions in different scenarios are required.
- A means of providing high algorithms and data security is required, and failure or partial malfunction should be indicated promptly.

- The application of a robust standard algorithm is still required, and the application of this algorithm should be followed globally.

V. CONCLUSION

Real-time visible data from the environment obtained via sensors is the source for an autonomous vehicle to decide its manoeuvre. Extracting accurate information in every circumstance is possible when the AV is fully equipped with smart sensors. Due to the practical barriers indicated in this paper, two areas of AVs are explored. These areas have been less explored in the previous research and can be verified for further examination from Table 1. This paper addressed the sensor technology: its current deployment, the merits and limitations, and sensor shortcomings in different scenarios, and most importantly, combining these multidomain data are discussed. Many sensor fusion approaches have been designed, but accuracy and intelligent algorithms with less complexity have not been achieved to date. CNN algorithms proved to be the most effective in sensor fusion over two years. Furthermore, the second subject, which is comparatively less investigated, is the estimation of pedestrian intentions. Towards this aim, different practical approaches have been highlighted in this paper.

COMPETING INTERESTS

The scholar's oath that there are no competing interests concerning the journal publication.

REFERENCES

[1] (2019). *International Traffic Safety Data And Analysis Group (IRTAD) Report 2019*. [Online]. Available: https://www.itf-oecd.org/sites/default/files/docs/irtad-road-safety-annual-report-2019.pdf

[2] *Global Status Report on Road Safety 2018*. Accessed: Apr. 9, 2020. [Online]. Available: https://www.who.int/publications-detail/global-status-report-on-road-safety-2018

[3] AutoDrive Challenge. (2020). *SAE International*. [Online]. Available: https://www.sae.org/attend/student-events/autodrive-challenge/

[4] M. Williams, "PROMETHEUS-The European research programme for optimising the road transport system in Europe," *IEE Colloq. Driv. Inf.*, Dec. 1998, pp. 1–9.

[5] C. Thorpe, M. H. Hebert, T. Kanade, and S. A. Shafer, "Vision and navigation for the carnegie-mellon navlab," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 10, no. 3, pp. 362–373, May 1988.

[6] CB Insights. (2019). *Autonomous Vehicles & Car Companies L CB Insights*. [Online]. Available: https://www.cbinsights.com/research/autonomous-driverless-vehicles-corporations-list/

[7] T. Reuters, "Web of science core collection," in *Proc. Web Sci.*, 2020, pp. 1–4.

[8] B. Schoettle, "Sensor fusion: A comparison of sensing capabilities of human drivers and highly automated vehicles," Sustain. Worldw. Transp., Tech. Rep. SWT-2017, 2017, vol. 12, pp. 1–42.

[9] SAE International. (2019). *SAE J3016 Automated-Driving Graphic*. [Online]. Available: https://www.sae.org/news/2019/01/sae-updates-j3016-automated-driving-graphic

[10] E. Yurtsever, J. Lambert, A. Carballo, and K. Takeda, "A survey of autonomous driving: Common practices and emerging technologies," *IEEE Access*, vol. 8, pp. 58443–58469, 2020.

[11] A. D. Kumar, R. Karthika, and K. P. Soman, *Stereo Camera and LIDAR Sensor Fusion-Based Collision Warning System for Autonomous Vehicles*. Singapore: Springer, 2020, pp. 239–252.

[12] C. Urmson *et al.*, "Autonomous driving in urban environments: Boss and the urban challenge," *J. Field Robot.*, vol. 25, no. 8, pp. 425–466, 2008.

[13] P. Grisleri and I. Fedriga, "The BRAiVE autonomous ground vehicle platform," *IFAC Proc. Volumes*, vol. 43, no. 16, pp. 497–502, 2010.

[14] J. Levinson, J. Askeland, J. Becker, J. Dolson, D. Held, S. Kammel, J. Z. Kolter, D. Langer, O. Pink, V. Pratt, M. Sokolsky, G. Stanek, D. Stavens, A. Teichman, M. Werling, and S. Thrun, "Towards fully autonomous driving: Systems and algorithms," in *Proc. IEEE Intell. Vehicles Symp. (IV)*, Jun. 2011, pp. 163–168.

[15] A. Broggi, M. Buzzoni, S. Debattisti, P. Grisleri, M. C. Laghi, P. Medici, and P. Versari, "Extensive tests of autonomous driving technologies," *IEEE Trans. Intell. Transp. Syst.*, vol. 14, no. 3, pp. 1403–1415, Sep. 2013.

[16] D. Göhring, D. Latotzky, M. Wang, and R. Rojas, "Semi-autonomous car control using brain computer interfaces," in *Intelligent Autonomous Systems* (Advances in Intelligent Systems and Computing). Berlin, Germany: Springer, 2013.

[17] L. Ulrich, "Top ten tech cars 2019: Self-driving and electric technologies are infiltrating everyday cars—Slowly," *IEEE Spectrum*, vol. 56, no. 4, pp. 32–43, Apr. 2019.

[18] A. Souppouris. (Dec. 2014). *Riding in Audi's 150 mph Self-Driving RS 7, the Anti-Google Car*. [Online]. Available: https://www.engadget.com/2014/12/18/audi-self-driving-rs-7-concept-

[19] J. Ziegler *et al.*, "Making bertha drive—An autonomous journey on a historic route," *IEEE Intell. Transp. Syst. Mag.*, vol. 6, no. 2, pp. 8–20, Summer 2014.

[20] K. Jo, J. Kim, D. Kim, C. Jang, and M. Sunwoo, "Development of autonomous car—Part II: A case study on the implementation of an autonomous driving system based on distributed architecture," *IEEE Trans. Ind. Electron.*, vol. 62, no. 8, pp. 5119–5132, Aug. 2015.

[21] M. Aeberhard, S. Rauch, M. Bahram, G. Tanzmeister, J. Thomas, Y. Pilat, F. Homm, W. Huber, and N. Kaempchen, "Experience, results and lessons learned from automated driving on Germany's highways," *IEEE Intell. Transp. Syst. Mag.*, vol. 7, no. 1, pp. 42–57, 2015.

[22] S. Kato, E. Takeuchi, Y. Ishiguro, Y. Ninomiya, K. Takeda, and T. Hamada, "An open approach to autonomous vehicles," *IEEE Micro*, vol. 35, no. 6, pp. 60–68, Nov./Dec. 2015.

[23] D. Sherman. (Feb. 2016). *Semi-Autonomous Cars Compared! Tesla Models vs. BMW 750i, Infiniti q50s, and Mercedes-Benz s65 AMG*. [Online]. Available: http://www.caranddriver.com/features/semi-autonomous-cars-compared-tesla-vs-bmw-mercedes-and-infiniti-feature

[24] A. Souppouris. (Dec. 2014). *Riding in Audi's 150 mph Self-Driving RS 7, the Anti-Google Car*. [Online]. Available: https://www.engadget.com/2014/12/18/audi-self-driving-rs-7-concept-

[25] J. M. Gitlin. (May 2016). *From Audi to Volvo, Most 'Self-Driving' Cars Use the Same Hardware*. [Online]. Available: https://arstechnica.com/cars/2016/05/from-audi-to-volvo-most-self-driving-cars-use-the-same-hardware/

[26] J. Stewart. (May 2016). *$30k Retrofit Turns Dumb Semis Into Self-Driving Robots*. [Online]. Available: https://www.wired.com/2016/05/otto-retrofit-autonomous-self-driving-trucks/

[27] Renault. (2017). *Adas: A Range of Technologies Promoting Safety and Easier Driving Experience*. [Online]. Available: https://group.renault.com/en/passion-2/innovation/renault-a-born-innovator/adas-a-range-of-technologies-promoting-safety-and-easier-driving-experience/

[28] W. Maddern, G. Pascoe, C. Linegar, and P. Newman, "1 year, 1000 km: The Oxford RobotCar dataset," Int. J. Rob. Res., 2017.

[29] Baidu. (May 1, 2019). *Apollo Auto*. [Online]. Available: https://github.com/ApolloAuto/apollo

[30] *Veoneer Awarded Thermal Cameras Contract for Level 4 AV*. [Online]. Available: https://www.veoneer.com/en/advanced-driver-assistance-systems

[31] A. I. Maqueda, A. Loquercio, G. Gallego, N. Garcia, and D. Scaramuzza, "Event-based vision meets deep learning on steering prediction for self-driving cars," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit.*, Jun. 2018, pp. 5419–5427.

[32] H. Yamada, T. Kobayashi, Y. Yamaguchi, and Y. Sugiyama, "High-resolution 2D SAR imaging by the millimeter-wave automobile radar," in *Proc. IEEE Conf. Antenna Meas. Appl. (CAMA)*, Dec. 2017, pp. 149–150.

[33] S. Shome, R. Bera, B. Maji, A. K. Bhoi, and P. K. Mallick, *Array Radar Design and Development*. Singapore: Springer, 2020, pp. 539–546.

[34] D. Uttamchandani, *Handbook of Mems for Wireless and Mobile Applications*. Amsterdam, The Netherlands: Elsevier, 2013.

[35] E. Javanmardi, Y. Gu, M. Javanmardi, and S. Kamijo, "Autonomous vehicle self-localization based on abstract map and multi-channel LiDAR in urban area," *IATSS Res.*, vol. 43, no. 1, pp. 1–13, 2019.

[36] S. Casas, W. Luo, and R. Urtasun, "IntentNet: Learning to Predict Intention from Raw Sensor Data," in *Proc. CoRL*, vol. 87, 2018, pp. 947–956.
[37] W. Xu, C. Yan, W. Jia, X. Ji, and J. Liu, "Analyzing and enhancing the security of ultrasonic sensors for autonomous vehicles," *IEEE Internet Things J.*, vol. 5, no. 6, pp. 5015–5029, Dec. 2018.
[38] Z. Li, L. Zhao, Z. Jiang, S. Akhbari, J. Ding, Y. Zhao, Y. Zhao, and L. Lin, "Capacitive micromachined ultrasonic transducer for ultra-low pressure measurement: Theoretical study," *AIP Adv.*, vol. 5, no. 12, Dec. 2015, Art. no. 127231.
[39] *Uber Car Crash March 18*. Accessed: Jul. 23, 2020. [Online]. Available: http://bit.ly/VSD-NTSB
[40] A. Geiger, M. Lauer, F. Moosmann, B. Ranft, H. Rapp, C. Stiller, and J. Ziegler, "Team AnnieWAY's entry to the 2011 grand cooperative driving challenge," *IEEE Trans. Intell. Transp. Syst.*, vol. 13, no. 3, pp. 1008–1017, Sep. 2012.
[41] Q. Li, L. Chen, M. Li, S.-L. Shaw, and A. Nuchter, "A sensor-fusion drivable-region and lane-detection system for autonomous vehicle navigation in challenging road scenarios," *IEEE Trans. Veh. Technol.*, vol. 63, no. 2, pp. 540–555, Feb. 2014.
[42] R. O. Chavez-Garcia and O. Aycard, "Multiple sensor fusion and classification for moving object detection and tracking," *IEEE Trans. Intell. Transp. Syst.*, vol. 17, no. 2, pp. 525–534, Feb. 2016.
[43] U. Lee, J. Jung, S. Shin, Y. Jeong, K. Park, D. H. Shim, and I.-S. Kweon, "EureCar turbo: A self-driving car that can handle adverse weather conditions," *IEEE Int. Conf. Intell. Robot. Syst.*, Oct. 2016, pp. 2301–2306.
[44] C. Xiaozhi, M. Huimin, W. Ji, L. Bo, and X. Tian, "Multi-view 3D object detection network for autonomous driving," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jul. 2017, pp. 1907–1915.
[45] V. De Silva, J. Roche, and A. Kondoz, "Robust fusion of LiDAR and wide-angle camera data for autonomous mobile robots," *Sensors*, vol. 18, no. 8, p. 2730, Aug. 2018.
[46] G. Wan, X. Yang, R. Cai, H. Li, Y. Zhou, H. Wang, and S. Song, "Robust and precise vehicle localization based on multi-sensor fusion in diverse city scenes," in *Proc. IEEE Int. Conf. Robot. Autom.*, May 2018, pp. 4670–4677.
[47] L. Caltagirone, M. Bellone, L. Svensson, and M. Wahde, "LIDAR–camera fusion for road detection using fully convolutional neural networks," *Robot. Auto. Syst.*, vol. 111, pp. 125–131, Jan. 2019.
[48] B. S. Jahromi, T. Tulabandhula, and S. Cetin, "Real-time hybrid multi-sensor fusion framework for perception in autonomous vehicles," *Sensors*, vol. 19, no. 20, pp. 1–23, 2019.
[49] F. Kunz, D. Nuss, J. Wiest, H. Deusch, S. Reuter, F. Gritschneder, A. Scheel, M. Stubler, M. Bach, P. Hatzelmann, C. Wild, and K. Dietmayer, "Autonomous driving at ulm university: A modular, robust, and sensor-independent fusion approach," in *Proc. IEEE Intell. Vehicles Symp. (IV)*, Jun. 2015, pp. 666–673.
[50] N. Patel, A. Choromanska, P. Krishnamurthy, and F. Khorrami, "Sensor modality fusion with CNNs for UGV autonomous driving in indoor environments," in *Proc. IEEE/RSJ Int. Conf. Intell. Robots Syst. (IROS)*, Sep. 2017, pp. 1531–1536.
[51] H. Zhang, Y. Liu, C. Wang, R. Fu, Q. Sun, and Z. Li, "Research on a pedestrian crossing intention recognition model based on natural observation data," *Sensors*, vol. 20, no. 6, p. 1776, Mar. 2020.
[52] S. Zhang, M. Abdel-Aty, J. Yuan, and P. Li, "Prediction of pedestrian crossing intentions at intersections based on long short-term memory recurrent neural network," *Transp. Res. Rec. J. Transp. Res. Board*, vol. 2020, Mar. 2020, Art. no. 036119812091242.
[53] C. Zhao, Y. Qian, and M. Yang, "Monocular pedestrian orientation estimation based on deep 2D-3D feedforward," *Pattern Recognit.*, vol. 100, Apr. 2020, Art. no. 107182.
[54] K. Mangalam, E. Adeli, K.-H. Lee, A. Gaidon, and J. C. Niebles, "Disentangling Human Dynamics for Pedestrian Locomotion Forecasting with Noisy Supervision," in *Proc. IEEE Winter Conf. Appl. Comput. Vis.*, Mar. 2020, pp. 2784–2793.
[55] T. Lagström and V. M. Lundgren, "AVIP-Autonomous vehicleś interaction with pedestrians an investigation of pedestrian-driver communication and development of a vehicle external interface," M.S. thesis, Dept. Product Prod. Develop., Chalmers Univ. Technol., Göteborg, Sweden, 2015, p. 84.
[56] M. Matthews, G. Chowdhary, and E. Kieson, "Intent communication between autonomous vehicles and pedestrians," 2017, *arXiv:1708.07123*. [Online]. Available: https://arxiv.org/abs/1708.07123
[57] D. Rothenbucher, J. Li, D. Sirkin, B. Mok, and W. Ju, "Ghost driver: A field study investigating the interaction between pedestrians and driverless vehicles," in *Proc. 25th IEEE Int. Symp. Robot Hum. Interact. Commun. (RO-MAN)*, Aug. 2016, pp. 795–802.
[58] S. U. Yang, "Driver behavior impact on pedestrians' crossing experience in the conditionally autonomous driving context," M.S. thesis, School Comput. Sci. Commun., KTH Roy. Inst. Technol., Stockholm, Sweden, Dec. 2017, pp. 4–14.
[59] A. Pillai, "Virtual reality based study to analyse pedestrian attitude towards autonomous vehicles," M.S. thesis, School Comput. Sci. Commun., KTH Roy. Inst. Technol., Stockholm, Sweden, 2017.
[60] S. K. Jayaraman, C. Creech, L. P. Robert, Jr., D. M. Tilbury, X. J. Yang, A. K. Pradhan, and K. M. Tsui, "Trust in AV: An uncertainty reduction modelof AV-pedestrian interactions," in *Proc. HRI*, 2018, pp. 133–134.
[61] H. Prakken, "On the problem of making autonomous vehicles conform to traffic law," *Artif. Intell. Law*, vol. 25, no. 3, pp. 341–363, Sep. 2017.
[62] K. Mahadevan, S. Somanath, and E. Sharlin, "Communicating awareness and intent in autonomous vehicle-pedestrian interaction," in *Proc. Conf. Hum. Factors Comput. Syst.*, Apr. 2018, pp. 1–12.
[63] R. L. Moore, "Pedestrian choice judgment," *J. Oper. Res. Soc.*, vol. 4, no. 1, pp. 3–10, 1953.
[64] B. Herwig, "Verhalten von kraftfahrern und fussgangern zebras-treifen," *Zeitschrift Verkehrssicherheit*, vol. 11, pp. 189–202, Nov. 1965.
[65] N. W. Heimstra, J. Nichols, and G. Martin, "An experimental methodology for analysis of child pedestrian behavior," *Pediatrics*, vol. 44, no. 5, pp. 832–838, 1969.
[66] D. Johnston, "Road accident casuality: A critique of the literature and an illustrative case," Grand Rounds, Dept. Psychiatry, Hotel Dieu Hospital, Kingston, ON, Canada, Tech. Rep., 1973.
[67] A. Millard-Ball, "Pedestrians, autonomous vehicles, and cities," *J. Planning Edu. Res.*, vol. 38, no. 1, pp. 6–12, Mar. 2018.
[68] M. Hein and G. Conference, *Pattern Recognition: 35th German Conference, GCPR 2013, Saarbrücken, Germany, September 3–6, 2013, Proceedings*, vol. 8142. Berlin, Germany: Springer, Sep. 2013, p. 10.
[69] R. Quintero, I. Parra, D. F. Llorca, and M. A. Sotelo, "Pedestrian intention and pose prediction through dynamical models and behaviour classification," in *Proc. IEEE 18th Int. Conf. Intell. Transp. Syst.*, Sep. 2015, pp. 83–88.
[70] J. F. P. Kooij, N. Schneider, F. Flohr, and D. M. Gavrila, "Context-based pedestrian path prediction," in *Computer Vision—ECCV* (Lecture Notes in Computer Science), vol. 8694. Cham, Switzerland: Springer, 2014, pp. 618–633.
[71] K. M. Kitani, B. D. Ziebart, J. A. Bagnell, and M. Hebert, "Activity forecasting," in *Computer Vision—ECCV* (Lecture Notes in Computer Science), vol. 7575. Berlin, Germany: Springer, 2012, pp. 1–14.
[72] R. Showcase *et al.*, "Planning-based prediction for pedestrians recommended citation planning-based prediction for pedestrians," School Comput. Sci., Carnegie Mellon Univ., Pittsburgh, PA, USA, Tech. Rep., 2009.
[73] V. Karasev, A. Ayvaci, B. Heisele, and S. Soatto, "Intent-aware long-term prediction of pedestrian motion," in *Proc. IEEE Int. Conf. Robot. Autom.*, Jun. 2016, pp. 2543–2549.
[74] R. Quintero, I. Parra, D. F. Llorca, and M. A. Sotelo, "Pedestrian path prediction based on body language and action classification," in *Proc. 17th Int. IEEE Conf. Intell. Transp. Syst. (ITSC)*, Oct. 2014, pp. 679–684.
[75] J. Cermak and A. Angelova, "Learning with proxy supervision for end-to-end visual learning," in *Proc. IEEE Intell. Vehicles Symp. (IV)*, Jun. 2017, pp. 1–6.
[76] Z. Fang, D. Vázquez, and A. M. López, "On-board detection of pedestrian intentions," *Sensors*, vol. 17, no. 10, pp. 1–14, 2017.
[77] S. Kohler, M. Goldhammer, K. Zindler, K. Doll, and K. Dietmeyer, "Stereo-vision-based pedestrian's intention detection in a moving vehicle," in *Proc. IEEE 18th Int. Conf. Intell. Transp. Syst.*, Sep. 2015, pp. 2317–2322.
[78] B. Volz, K. Behrendt, H. Mielenz, I. Gilitschenski, R. Siegwart, and J. Nieto, "A data-driven approach for pedestrian intention estimation," in *Proc. IEEE 19th Int. Conf. Intell. Transp. Syst. (ITSC)*, Nov. 2016, pp. 2607–2612.
[79] E. Rehder, H. Kloeden, and C. Stiller, "Head detection and orientation estimation for pedestrian safety," in *Proc. 17th Int. IEEE Conf. Intell. Transp. Syst. (ITSC)*, Oct. 2014, pp. 2292–2297.
[80] K. Saleh, M. Hossny, and S. Nahavandi, "Intent prediction of vulnerable road users from motion trajectories using stacked LSTM network," in *Proc. IEEE 20th Int. Conf. Intell. Transp. Syst. (ITSC)*, Oct. 2017, pp. 327–332.

[81] D. O. Pop, A. Rogozan, C. Chatelain, F. Nashashibi, and A. Bensrhair, "Multi-task deep learning for pedestrian detection, action recognition and time to cross prediction," *IEEE Access*, vol. 7, pp. 149318–149327, 2019.

[82] R. Grosse, M. K. Johnson, E. H. Adelson, and W. T. Freeman, "Ground truth dataset and baseline evaluations for intrinsic image algorithms," in *Proc. IEEE 12th Int. Conf. Comput. Vis.*, Sep. 2009, pp. 2335–2342.

[83] K. M. Abughalieh and S. G. Alawneh, "Predicting pedestrian intention to cross the road," *IEEE Access*, vol. 8, pp. 72558–72569, 2020.

FARRUKH HAFEEZ received the B.E. degree in electronics and the M.E. degree in industrial electronics from the NED University of Engineering and Technology, Karachi, Pakistan, in 2005 and 2009, respectively. He is currently pursuing the Ph.D. degree in electrical engineering with Universiti Teknologi Malaysia (UTM). His research interests include computer vision, machine learning, prediction intention of pedestrian, and sensor fusion.

USMAN ULLAH SHEIKH (Member, IEEE) received the B.Eng. degree in electrical and mechatronics engineering, the M.Eng. degree in telecommunications engineering, and the Ph.D. degree in image processing and computer vision from Universiti Teknologi Malaysia, in 2003, 2005, and 2009, respectively. He is currently a Senior Lecturer with Universiti Teknologi Malaysia working on image processing for intelligent surveillance systems. He has published works in many journals and conferences. His research interests include computer vision and embedded systems design. He is a member of IET.

NASSER ALKHALDI received the A.S. degree in instrumentation and control engineering technology from the Jubail Industrial College, the B.E. degree in electrical engineering from Purdue University, the M.E. degree in network engineering and management from DePaul University, and the Ph.D. degree in electrical engineering from Wayne State University. He is currently working as an Assistant Professor of instrumentation and control engineering technology with the Jubail Industrial College, Jubail, Saudi Arabia. His research interests include RF engineering and wireless micro sensors.

HASSAN ZUHAIR AL GARNI received the Master of Engineering degree in electrical and computer engineering and the Ph.D. degree in systems engineering from Concordia University, Canada, in 2013 and 2018, respectively. He is currently working as an Assistant Professor and the Chairman of Electrical and Electronics Engineering Technology Department, Jubail Industrial College. He has published several research articles and conference proceedings in the area of renewable energy. His employment experiences include working in Instrumentation and Control at SABIC-SHARQ for five years, and teaching at the Jubail Industrial College for 11 years. His research interests are in the areas of technical educations, renewable energy, including solar photovoltaic design and optimization including site selection, and panel orientation and tracking systems. He was awarded the Concordia Accelerator Award, in 2018, the Concordia University Conference and Exposition Award in 2017 and 2018, and the IEEE-SMC Student Travel Grant Award.

ZEESHAN AHMAD ARFEEN (Senior Member, IEEE) is currently pursuing the Ph.D. degree with University Technology Malaysia (UTM). He is also an Assistant Professor with the Electrical Power Engineering Department, The Islamia University of Bahawalpur, Pakistan. He is an author and potential reviewer of several high indexed journals. His research interests include electrical transportations, microgrid, energy management systems, and sustainable energy.

SAIFULNIZAM A. KHALID received the B.Eng., M.E.E., and Ph.D. degrees from Universiti Teknologi Malaysia, Johor Bahru, Malaysia, in 1998, 2000, and 2009, respectively. He is currently a Senior Lecturer with the Faculty of Engineering, School of Electrical Engineering, Universiti Teknologi Malaysia. He has published more than 50 articles in H-indexed and Scopus Journal. He is also an Expert Coordinator for Tenaga Nasional Berhad Malaysia. His research interests include deregulated power systems, and application of artificial neural network in power systems and power tracing.