Abstract

Autonomous vehicles introduce dramatic changes to the way in which we travel. This technology has the potential to impact the transportation across a wide array of categories including safety, congestion, and travel behavior. In this report we review the main issues involved in autonomous driving as discussed in the literature and shed light on topics that we believe require further development.

1 Introduction

The challenges in autonomous driving require a combination of many capabilities, among them: localization, motion planning, vehicle’s systems control (steering, acceleration/deceleration, signaling, etc.), road perception, prediction of other road-users behavior, awareness of dangers, etc. These capabilities have a various level of importance for various levels of driving automation.

Even though the first autonomous vehicle (AV) was experimented in 1926 [104], a real-modern autonomous vehicle was first presented in 1986 by a team from Carnegie Mellon University [58]. Since 2010, many major automotive manufacturers, such as GM, Mercedes Benz, Ford, Toyota, and many more are developing AVs [90]. A demonstration of AV equipped with numerous sensors and communication systems was presented by Toyota in 2013 [10]. In 2016, Google’s AV has passed over one million kilometers. These events present a glimpse to the landmarks in the development of AV that, due to the complexity of the task, progresses slowly with considerable amount of attention to safety and reliability.

AVs need a large amount of data for reliable decision making. This data comes from a variety of onboard sensors and algorithms that perform data fusion and estimation on one hand, and from outer sources like other AVs (V2V) environmental devices (I2V) and a combination of them (X2V) on the other hand. Figure 1 illustrates some data transfer architectures for AV.

In SAE’s (Society of Automotive Engineers) automation level definitions, "driving mode" means "a type of driving scenario with characteristic dynamic driving task requirements (e.g., expressway merging, high speed cruising, low speed traffic jam, closed-campus operations, etc.)"[93]. Table 1 presents the 5 levels of automation according to SAE (6 levels, including the “no-automation” level).
Figure 1: Plane.
<table>
<thead>
<tr>
<th>Automation Level</th>
<th>Definition</th>
<th>Examples</th>
</tr>
</thead>
<tbody>
<tr>
<td>0- No automation</td>
<td>All aspects of the dynamic driving are performed by the human driver.</td>
<td>Adaptive cruise control</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Parking assistance</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Lane keeping</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Collision avoidance system</td>
</tr>
<tr>
<td>1- Driver assistance (Hands on)</td>
<td>Shared performances by the automation system and the human driver.</td>
<td></td>
</tr>
<tr>
<td>2- Partial automation (Hands off)</td>
<td>All aspects of the dynamic driving are automated, but the human driver supervises the automation system.</td>
<td>Most AVs that travel today.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Uber, Tesla, Waymo, etc.</td>
</tr>
<tr>
<td>3- Conditional automation (Eyes off)</td>
<td>All aspects of the dynamic driving are automated, but, the human driver may be requested to intervene.</td>
<td>Uber, Tesla, Waymo, etc.</td>
</tr>
<tr>
<td>4- High automation (Mind off)</td>
<td>Like the above but, the automated driving handles even if the human driver does not respond to a request to intervene.</td>
<td>Uber, Tesla, Waymo, etc.</td>
</tr>
<tr>
<td>5- Full automation (Steering wheel optional)</td>
<td>Human intervene never requested.</td>
<td>Tesla</td>
</tr>
</tbody>
</table>

*Table 1: AV levels of automation (according to SAE)*
Autonomous vehicles are believed to bring beneficial change to the way in which we travel that will impact an array of categories such as safety, congestion, and travel behavior. Crash savings, travel time reduction, fuel efficiency and parking benefits are believed to save thousands of dollars per year per a single AV [35]. Although its great vision, the implementation and mass-market penetration of AVs will most likely take time. As for today, initial costs are commonly unaffordable for the common user.

2 Autonomous Driving

These days, the concept of AVs is based on the vision of replacing the driver with an autonomous system that will drive in the traditional roads with other road users (other vehicles and pedestrians). The path to this autonomous goal passes through the five levels described in Table 1. Since the mission is to replace the driver, high-end sensors must be integrated in order to achieve a safe and efficient drive to the target.

2.1 Lane-Detection

One of the most important qualifications for AV and advanced driver assistant systems is the lane-detection. Lack of clarity of lane markings, poor visibility due to bad weather, illumination and light reflection, shadows, and dense road-based instructions may cause the lane detection algorithm to fail. Most lane detection approaches are based on analyzing the 2-D image captured from a camera (usually mounted behind the front windshield to retrieve lane information). These vision-based approaches can be categorized into two methods: feature-based and model-based. The model-based method commonly uses a mathematical model with parameters to describe the lane structure [81,131]. For example, researchers in [2] presented real-time lane marker detection in urban streets. Their method generates a top view of the road, uses Random sample consensus (RANSAC) line fitting for the initial assumed line model, and a fast RANSAC algorithm for Bézier Splines fitting. The feature-based methods, though known for their robustness against noise, are difficult to implement since they require some prior-known geometric parameters and heavy computation. The feature-based methods analyze images and detect the gradients of pixel information or the color of patterns to recognize the lane markings. For example, the researchers in [72] presented a robust and real-time vision-based lane detection algorithm by reducing the image to an efficient region of interest in order to reduce the high noise level and the calculation time. Their method also removes any false lane markings and tracks the real lane markings using the accumulated statistical data. Their experimental results show that the algorithm gives accurate lane estimate and fulfill the real-time operational requirements on embedded systems with low computing power. For more examples see: [16,109,108,30,95].

2.2 Path Planning

The path planning task for AV has been researched for the last decades. It is mostly common to divide the path planning problem into global and local planning. The planning techniques can be classified as (i) Graph search algorithms, such as Dijkstra [17,23,97,5] or A-star [112,36,76,83] which assume a set of known configurations with the goal to find the path from two configurations passing through the
known configuration set; (ii) Sampling based planners, such as RRT (Rapidly-exploring Random Trees) [69,19,101,59] relaxes the approach of grid sampling the configuration space by sampling it in the region of interest with the desired density. Interpolating curve planners are used for “smoothing” the paths given by the path planners.

2.3 Motion Planning

As AV is aimed to have fully automated driving functionality available in a variety set of scenarios, it raises the need for universally applicable environment sensing and understanding. The sensors input is then used by the motion planning algorithms. For example, Brown et al. [20] introduce a control framework that integrates local path planning together with path tracking using model-predictive-control. The controller first plans a trajectory that considers the vehicle state (position and velocity) and a desired path to the target position. Then, two safe envelopes are considered: one for stability and the other for obstacle avoidance. Moriwaki [84] presents an optimal steering control for electric autonomous vehicle based on $H^\infty$. The objective of the scheme chosen is to be a reference trajectory following while keeping good damping and a certain stability margin. da Silva and de Sousa [29] use dynamic programming for AV motion control. Their objective is following a desired path while keeping the cross-error under some predefined threshold. Kessler et al. [60] introduced two novel approaches for extracting a topological road-graph with possible intersection options from sensor data along with a geometric representation of the available maneuvering space. Also, a search and optimization-based path planning method for guiding the vehicle along a selected track in the road-graph and within the free-space is presented. They compared the methods presented in simulation and showed results of a test drive with a research vehicle. Their evaluations show the applicability in low speed maneuvering scenarios and the stability of the algorithms even for low quality input data. For more schemes for trajectory tracking and path following see [121,128,63].

The field of motion planning which simultaneously considers safety and comfort is yet to be fully considered. Magdici et al. [79] presented a fail-safe motion planner for autonomous vehicles, which simultaneously guarantees safety and comfort by combining an optimal trajectory with an emergency maneuver. Solea et al. [105] presented a path-following control using sliding mode path-following control, together with a smooth velocity planner, imposing the comfort of the human body. An important field of study is the feeling-of-safety (sometimes considered as trusted autonomy) as in automated vehicles while it performs trajectory tracking in urban environments. In these cases, the path must be smoothed previously in a planning stage before the trajectory tracking task. The researchers in [71] implemented the 4th and 5th degree Bézier curves in their path planning generation. They focused on urban scenarios (intersections, roundabouts, and lane changing) and speed planning for comfortable and safe behaviors.

Since AVs act in human environment, ethical issues should be considered in the motion control. The authors in [28,115] deal with such ethical realism to the framing of AVs decisions and control. The use of Machine-Learning (ML) seems to be the most attractive technique for AVs perception (see for example [103,1]). In [74] the authors present a neural network model to analyze the data captured by the sensors. Then, a decision-making system calculates suitable control signals for the vehicle based on the observations. Isele et al. [54] solve intersection problems of AVs by Deep Reinforcement Learning (DRL). The system learns active sensing behaviors and enables safe maneuvers in the case of occlusions.

The complexity and high variety scenarios of road driving make motion planning a challenging task. Banerjee et al. [7] investigated all disengagement and accident reports obtained from public DMV
databases between 2014-2017, and found that ML-based systems are the primary cause of 64% of all disengagements. Koopman and Wagner [65] state in their paper that "...there does not seem to be a way to make ultra-dependable levels of guarantees as to how such a system will behave when it encounters data not in the training set nor test data set". Researchers address this challenge by applying additional algorithms to block unsafe maneuvers. Mirchevska et al. [82] use Reinforcement Learning for lane changing of AVs. They address the uncertainty issues by machine learning combined with safety validations that ensures only safe actions are taken.

Figure 2: An illustration of the approximated free-space (red) and the road-graph with direction options (blue). Based on the directional choice to go straight, a smooth path can be planned (green) [60].

2.4 Sensors

One of the first works on sensory systems for AVs was presented by Waxman [119], where the authors use a camera for the control of the vehicle. The hardware back-days was inefficient such that the frame-rate was smaller than that of the controller. The researchers maintained continuous motion by what they called 'looking ahead' and then, until they took another frame, "driving blind" for a short distance. The use of cameras as an input for closing the loop of AVs gained momentum at the early 90's of the last century. Authors in [57] use an improved processor to control the vehicle. The vision system estimates the lateral position and deviation of the vehicle relatively to the white lines in the frame. All companies dealing with AV uses data-fusion. This enables overlap data in region of interest for safety (see e.f. [25] explaining Uber reduced LiDAR array in their self-driving cars from 7 units to 1 in 2016, creating pedestrian blind spots).

2.4.1 Camera sensors

Cameras became the most common modality sensor due to its high information content, lower cost and operating power, and the ability to integrate it with ultra-sonic sensors or radar as auxiliary sensors if necessary. Autonomous cars often have video cameras in order to see and interpret the objects in the road. By equipping cars with multiple cameras, the vehicles are capable of maintaining a 360° view of their external environment, thereby providing a broader picture of the traffic conditions around them.

Today, 3D cameras are available and utilized for displaying highly detailed and realistic images. Using computer vision algorithms (such as OpenCV) may automatically detect objects, classify them,
and determine the distances between them and the vehicle. For example, the cameras can easily identify other vehicles, pedestrians, cyclists, traffic signs and signals, road markings, bridges, and guardrails. It should be noted that poor weather conditions such as rain, fog, or snow can prevent cameras from clearly seeing the obstacles in the roadway. Additionally, in situations where the colors of objects are similar to the background, or where the contrast between them is low, the detection algorithms may fail [12, 68].

The examples of the use of cameras as the main or exclusive sensor are many. Heng et al. [48] present the AutoVision project, which aims to develop a localization and 3D scene perception capabilities for an AVs, that based on vision sensors only. Kuramoto et al. [67] developed a scheme for computing 3D positions of far detected vehicles based on mono-camera observations.

NODAR [92] accelerates mass-market autonomy through the development of high-performance stereo vision systems. According to their website, whether passenger vehicles, UAVs, or fork lifts, autonomous vehicles require high-density 3D information to understand their surroundings. These markets are extremely price sensitive, yet require ultra-fast, highly accurate, long-range 3D sensing.

### 2.4.2 Radar Sensors

Radar (Radio Detection and Ranging) sensors make up a crucial contribution to the overall function of autonomous driving by sending out radio waves that detect objects and gauge their distance and velocity in relation to the vehicle in real time [13]. Radar is a key technology for AVs and driver assistance systems due to its robustness to environmental variations such as inclement weather (fog, rain, etc.), lighting extremes and long range [37]. Due to these advantages, the radar serves as complementary technology to the cameras.

Radar sensors may be categorized into two groups: Short range sensors (24 GHz) are usually used to enable blind spot monitoring, lane-keeping and parking assistance. The long-range sensors (77 GHz) are used for maintaining safe distance and brake assistance [106].

Radar sensors may be used to identify vulnerable road users. Stolz et al. [107] have researched the ability of using radar sensor for identifying cyclists in auto emergency braking (AEB) systems. Nabati and Qi [87] proposed the RRPN (Radar Region Proposal Network) algorithm for object detection. Based on the radar observation they mapped objects in to the image coordinates.

The impact of the pedestrian’s direction of movement, occlusion, antenna beam elevation angle, linear vehicle movement, and other factors are investigated and discussed in [8]. Their experiments show that although over 95% of pedestrians can be correctly detected in optimal conditions, under real life conditions, due to insufficient Doppler frequency and spatial resolution as well as antenna side lobe effects, radar only based pedestrian recognition will not perform the required tasks. An emerge research work that deals with dual function radar-communications (DFRC), which enables dual functionality to the radar as a sensor and as a communication device, is presented in [77].

### 2.4.3 LiDAR Sensors

LiDAR (Light Detection And Ranging) is currently the most commonly used sensor capable of delivering the most accurate real time 3D data due to its use of laser based point cloud. The real-time LiDAR sensor technology is being used in a variety of commercial applications including autonomous vehicles, vehicle safety systems, 3D aerial mapping and security. Though the benefits of 3D point-cloud data are clear, most AVs requires multiple LiDAR, which make the AV’s sensory system expensive [44].
LiDAR provides 3D images of the detected objects and map of the environment. Moreover, LiDAR can be configured to create a full 360-degree map around the vehicle rather than relying on a narrow field of view in relation to many other radar sensors. These two advantages make autonomous vehicle manufacturers such as Google, Uber, and Toyota use LiDAR systems in their AV's sensors suit. For researches on LiDAR and pedestrian recognition by AI see [114] [89].

LiDAR–radar sensor fusion is more robust to environmental change than camera since it uses a synergy laser and radio frequency signal (see [38,14,52,45]). For example, Kwon et al. [70] proposed a new detection scheme for occluded pedestrian detection by means of LiDAR–radar sensor fusion. The object within the occlusion region of interest is detected by the radar measurement information and the occluded object is estimated as a pedestrian based on human Doppler distribution.

As for 2020, LiDAR sensors are much more expensive than radars for use in AVs. The systems required for autonomous driving can cost well beyond $10,000, while the top sensor being used by Google and Uber costs up to $80,000. In addition, bad weather condition such as snow or fog may block the LiDAR sensors and affect their ability to detect objects in the road. Researchers suggest that
it is possible to overcome the LiDAR’s cones by using stereoscopic camera system (see e.g. [118]). By taking the inner workings of convolutional neural networks into consideration, researchers propose to convert image-based depth maps to pseudo-lidar representations — essentially mimicking the LiDAR signal.

3 Autonomous vehicles as a cooperative system

Traditionally, the term autonomous vehicle refers to the technology that enables automatic operation of the vehicle’s control functions (e.g. steering, throttle, braking, etc.). As such, the vehicle may be equipped with an array of sensors and actuators that is needed for the loop-closure. A complementary approach to the operation of AVs is the concept of multi-agent-systems or cooperative autonomous vehicles. In this approach, the control of each vehicle, apart from the environment interception, is related to the operations of all other AVs in its vicinity [49].

Individual vehicles may benefit from information obtained from other vehicles in the vicinity, especially information relating to traffic congestion and safety hazards. Vehicular communication systems use vehicles and roadside units as the communicating nodes in a peer-to-peer network, providing each other with information. As a cooperative approach, vehicular communication systems can allow all cooperating vehicles to be more effective. According to a 2010 study by the US National Highway Traffic Safety Administration, vehicular communication systems could help avoid up to 79% of all traffic accidents [88].

The cooperative operation of AVs has many significant advantages. Researchers define three critical technologies that increase the benefits of AVs [78]: Cooperative driving, Big-Data and Autonomous-capabilities. Hult et al. [51] discuss the safety improvement that is involved with autonomous driving in the cooperative driving scheme. According to the “Traffic Safety Facts Annual Report of 2017” [94] more than 80% of the accident with known and reported factors caused by error done by drivers. For example, at 8% of those accidents, the factor was the failure to remain in the proper lane, 12% of the drives involved with those accidents were driving under the influence of alcohol, drugs, or medication. We believe that autonomous driving in general, and cooperative driving in particular will significantly increase safety.

Figure 5: Vehicle coordination relies on tight interaction between control, communication, and sensing [51].
Moreover, Zhou et al. [130] show that cooperative control may significantly improve the smoothness of the vehicles' flow and increase the road capacity in comparison to human-driven vehicles. They formulate a control law for the vehicles' acceleration as a response to the actions of other vehicles in the vicinity. A discussion about ecological and fuel consumption can be found in [6]. As part of the capacity effectiveness of cooperative AVs, the fuel consumption is reduced as well as the air pollution (see also [4]).

The degree of centralization of the controller was discussed in the literature. The researchers in [55] present an algorithm for the concept of planning for multi agent systems and use it for AVs control. The authors formulate a multi-layer reservation policy to manage the interactions between vehicles. The aim of the policies is to reduce energy consumption, increase the smoothness of the traffic flow and increase intersection throughput. In [85] researchers present an algorithm for controlling the speed of an AV to track a speed-profile generated for cooperative AVs. The algorithm monitors the control messages that include the current desired speed, as well as the future speed-profile to overcome anomalous communications. In [91] the authors present a three layers algorithm for cooperative lane-changing decision-making. Although the vehicles share information, the main decision making was done independently.

4 Cooperative sensing

A cooperative sensing enables information sharing by all road-users and stationary devices located in the same environment. Such an approach enables to get observations beyond the line-of-sight and beyond the field-of-view. However, it introduces new challenges such as how to locate all observation in the same map, identification of vehicles, high volume of communication, privacy, etc. The authors in [61] present a framework to deal with those challenges. They demonstrate it on a group of AVs equipped with a single LIDAR and a single camera. The communication and control was conducted by open source liberties of Robotics Operation System (ROS). In [62] the authors present a scheme for cooperative sensing which apply to the vehicle see-through, lifted-seat or satellite-view. They investigate the improvements such as the safety and smoothness of the driving. Based on the literature review, it seems that this important subject should be further investigated.

4.1 Infrastructure to Vehicle communication (i2V)

Infrastructure to Vehicle (i2V) communication gives the ability to send and receive data from stationary stations in the AV's vicinity. The ability to monitor the pedestrians is essential if we desire to achieve autonomous driving. Today, many streets are 24/7 camera-viewed for traffic monitoring and for improving citizens’ safety. The big data from these camera network may be used as sensory data for the AV. The most important concept in the future of AVs is safety. Today, most AVs control is based on on-vehicle sensors alone. These are mostly important but, additional external data may be mostly efficient for pedestrian tracking and accident avoiding. An example for using street-view cameras was demonstrated by Kristoffersen et al. [66] where they used thermal cameras to overcome the challenges due to changing lighting conditions and the complexity of scenes with many people occluding one another. They introduced the use of a stereo thermal camera setup for pedestrian counting and investigated the reconstruction of 3D points in a pedestrian street with two thermal cameras. Then, they propose an algorithm for pedestrian counting based on clustering and tracking of the 3D point clouds.
I2V may also improve the decision making by AVs. Perumal et al. [99] present an algorithm for AVs motion planning based on observations for localization and moving obstacles (road users) locations. Grembek et al. [39] present an algorithm for intelligent intersection based on I2V observations that reduce the lack in the information they need to avoid wrong decisions that may cause an accident.

4.2 Vehicle to vehicle communication (V2V)

Vehicle networking may overcome some difficulties with computer vision being able to recognize brake lights, turn signals, etc. Vehicle to vehicle (V2V) communication is more difficult to realize due to its decentralized structure. This feature aims at organizing the interaction among vehicles and possibly developing collaborations among them. In vehicle networking, the decision is based on the interchanged information between a group of vehicles in the same vicinity. This obviously requires communication technology and protocols agreements (see CAR2CAR Consortium [27]). In this concept, the vehicles also serve as routers and allow communication over multi-hop to distant vehicles and roadside stations. Delays, partial measurements, safety, etc. must be considered. For example, in [53] the authors present an approach for AVs’ collision warning system that is robust to communication uncertainty. The communication between AVs may improve significantly its performances, yet, it holds risks as well. Cyber security is a crucial factor in V2V. Amoozadeh et al. [3] show by simulations that an insider attack may cause significant damage to the AV’s control and suggest some principals to improve AVs’ security.

4.3 Communication between AVs and pedestrians

For safe autonomous driving in urban environment, communication of AVs with other road users is required [100]. Habibovic et al. [41] state that communication between AVs and other road users that enables negotiation is essential. They examine some external devices and negotiable schemes. Their conclusion is that more research has to be done in order to formulate agreed standards, or language, for such communication. Dey and Terken [33] study the importance of eye contact and gestures between pedestrians and drivers. They found that motion patterns of the vehicle are more effective than eye contact between drivers and pedestrians for efficient traffic negotiations. These surprising findings open an opportunity for efficient communications between pedestrians and AVs. Researchers also present optional devices to support such communication, for example in [32] where visual interfaces is presented. Bazilial et al. [9] provide a survey on the external human-machine interfaces (eHMIs). They found that textual eHMIs are clearer for a pedestrian than other methods. Moreover, they investigate how the text color and perspective of the textual message affect the comprehension of the message and found that egocentric eHMIs are clearer.

In addition to the sensory structure on the AV, communicating awareness and intent in autonomous vehicle-pedestrian interaction was considered in [80]. In the paper, the researchers investigate the usefulness of interfaces that explicitly communicate awareness and intent of autonomous vehicles to pedestrians, focusing on crosswalk scenarios. Based on the study outcomes they found that interfaces communicating vehicle awareness and intent can help pedestrians attempting to cross. According to the research, the communication method should use a combination of visual, auditory, and physical means (e.g. a phone held by a participant vibrates when it is safe to cross).
5 Localization

The localization of autonomous vehicles is an indispensable task. Although localization schemes are well discussed in the literature, the localization of AVs is a bit unique in terms of the required accuracy, data availability, frequency of the environment updates etc. Some researchers even claim that AVs localization is currently an unsolved problem [102].

The accurate position of AVs is in the base of the decision making for the path planning, safety of movement, etc. Consequently, researchers and commercial companies developing new localization schemes to face this challenge. Wisely plane may consider the estimation error apart from the estimation itself. Wong et al. [120] present a scheme to estimate the localization error based on 2D geographic information alone. They estimate the localization error with 87.4% of predictions within 5cm.

Many research papers and patents were written to improve the localization of AVs using passive images [24,21,22]. Such sensors are low-cost and, together with smart image processing, they provide high accuracy of the AV position. The common ground of such schemes is the use of visual observations applied by an on-board camera to improve the localization accepted from a GNSS.

On the other hand, the use of LiDARs is becoming more and more customary for many observations involved with the AVs operations, especially for collision avoidance. As a result, a map-based localization may use LiDARs as favorite sensor since it has a high resolution and high accuracy. Wang et al. [116] present three steps for map-based localization using LiDAR measurements. First, point clouds from a single frame curbed based on the vehicle dynamic are inserted into the current vehicle’s coordinate system. Then, a contour lines of these points is conducted. The last step is the matching between the map and the contour lines. Mukhtar et al. [86] use sparse 3D LiDAR scan data for map-based localization in order to reduce the sensor cost.

The control of AVs involves the use of a high number of sensors, include IMU, wheels’ odometry measurements, GNSS, LiDAR, cameras, etc. As a result, methods for data fusion for localization are very common. For example, DeBitetto et al. [31] use inertial sensors as well as RADAR data to improve GPS localization of AVs. Yu et al. [125] present a localization scheme for AVs in urban area. They used a prior point cloud of the environment, but since the environment is changed frequently, this prior data may be irrelevant. The authors developed a novel data fusion algorithm that estimates the reliability of each point from the prior map based on the new observations.

Because of the complexity of the programming and data mining, researchers prefer the use of machine-learning and neural networks [126]. In [56] researchers investigate a map of nodes and edges they call hybrid-map. The hybrid-map enables to implement different types of machine-learning methods. The authors demonstrate this concept for an autonomous vehicle equipped with two LiDARs as input for the scheme. In the conclusions the authors state that real applications need further verification and improvements to ensure a robust system.

6 Non-controlled road users’ behavior

One of the biggest challenges in operating an autonomous traffic is to predict the motion of non-controlled road users, e.g. human-driven-vehicles, pedestrians, cyclists, pets etc. Twaddle et al. [110] focus on the increasing need for bicycle behavior models in urban areas. Yao et al. [123] proposed a behavior model for conflicts in vehicles-bicycles mixed flow. Li et al. [75] proposes a cellular automaton model to analyzes the behavioral characteristics of bicycles’ illegal lane-changing behavior. A resent research dealing with AV and pedestrian safety explored the potentials and limits of pedestrian...
The research analyzed nearly 5000 pedestrian fatalities recordings in 2015 in the Fatality Analysis Reporting System, and virtually reconstructed them under a hypothetical scenario that replaces the involved vehicles with AV equipped with state-of-the-art technology. They concluded that although technologies are being developed to successfully detect pedestrians, the current costs and operating conditions substantially decrease the potential for reducing pedestrian fatalities in the short term. The behavior of humans as crowds (see [111,64]) compared with pedestrians that cross a busy road are different. The high heterogeneity between pedestrians and vehicles in terms of maneuverability, speeds, field-of-view etc. makes the prediction of the pedestrian behavior much more complicated than that of the crowd behavior. Crowd, on the other hand, is typically considered as homogeneous, and the high density enables to assume continues interactions between individuals ([47]). So, the prediction of pedestrians requires deeper understanding of pedestrian and driver behavior. Since psychological considerations are convoluted with the trajectory planned by the pedestrians, researches on the pedestrian’s attitude may have high impact on the behavioral models. For example, Zhou et al. [129] used structural equation modeling to predict pedestrian crossing behavior. The authors present a questionnaire with a scenario involved with violating road-crossing rules and asked pedestrians about their attitude to such violating. The results of such research may be of importance in understanding the pedestrian’s behavior. Ye et al. [124] study the pedestrian behavior where road-crossing is done in groups. They analyze the interactions between groups of pedestrians and vehicles at unsignalized intersections. A model that based on the multidimensional “dirty faces” game was used to simulate the scenarios. Pawar et al. [98] analyze and evaluate the dilemma-zone for crossing pedestrians at mid-block crossings.

Researchers developed pedestrians’ models that consider local behavior of the individual pedestrian. Hoogendoorn and Bovy [50] consider pedestrians as autonomous controllers, which minimize a cost function while ongoing toward a target destination. Blue and Adler [15] show that a simple set of rules can effectively percept the pedestrians’ behavior at the micro level. They modeled bi-directional pedestrian motion using a Cellular Automata micro-simulation to confirm this claim.

Pedestrian crossing models may involve high number of factors. Duives et al. [34] evaluate pedestrian behavior models by considering eight distinct motion-based cases and six phenomena of crowd movement. The researchers show that models of pedestrian crossing must fit the specific

![Figure 6: Road sign detection procedure.](image-url)
scenario. Even though pedestrian shares the same motivation, e.g. trying to cross the road safer and faster as possible, each pedestrian adopts his/her own target location, level of urgency, physical capabilities, etc. [96,11,47]. Guo et al. [40] confirms that the behavior of crossing pedestrians depends mainly on the waiting time. Hacohen et.al [43] present a statistical algorithm for pedestrian crossing behavior model that depends on the pedestrian urgency as a single factor to be tuned.

Due the complexity of such predictions, researchers focus their attention on specific interactions between pedestrians and vehicles. Hashimoto et al. [46] developed a particle filter-based model of pedestrian behavior. The authors focused on the scenario of left-turning vehicles at signalized intersections when crossing at signalized crosswalks. Wang et al. [117] developed a pedestrian model for scenarios of midblock crosswalks and intersections. Lee and Lam [73] presented a model which estimates the walking speed of pedestrians at crowded crosswalks. Bonnin et al. [18] consider zebra-crossings.

An innovative strategy for developing pedestrian models is by implementing common robotic motion planning algorithms to construct the predicted trajectory of the pedestrians. Such methods refer to the vehicles as obstacles that should be avoided, and the targets considered as the other sides of the crosswalk. Zeng et al. [127] implement artificial-potential-field algorithm. Hacohen et al. [42] use Probability-Navigation-Function to predict artificial-potential-field algorithm. Hacohen et al. [122] introduced a Voronoi diagram-based algorithm for pedestrian flow model, and Van Waizman et al. [113] developed their method based on Velocity-Obstacle.

7 Conclusions

Autonomous vehicles are the future of the transportation. The intense research of the big players in the industry such as Google and Tesla towards this technology will make transportation safer, more comfortable and more efficient. Nevertheless, most of the research today is focused on designing an AV that will be able to drive in today’s roads today alongside regular vehicles. In addition to this important development, research on cooperative sensing and driving still need to be done. According to the cited research papers and the authors opinion, some important points are still in the development stage and still need much more work. In this list we like to consider gaps that are unique to the AV field and not considers by other disciplines.

1. Environmental sensors - the usage of street cameras and collecting the data to locate cars and other road users.

2. Communication methods with pedestrians - more research need to be done in order to fully understand the impact of AV on pedestrians during street sharing. In some places, streets are designed to be pedestrian-friendly but still allow vehicles through, In such cases, Vehicle to Pedestrians (V2P) communication is essential.

3. Consideration of non-controlled road users (pedestrians, cyclists, etc.) - development of communication methods and prediction algorithms for their behavior.

4. Formulating ethics considerations for the AV decision making and data sharing. For example, preemption, give way for emergency vehicles or for decision making in cases where an accident is inevitable.
Apart from the gaps in the above list, we identify some more issues that should be investigated in order to improve and enable autonomous driving:

1. The price tag of the sensory suit for the AV is high, such that researchers point that this issue delays the AV success. More algorithms using alternatives sensory input data to the LiDAR will enable lowering the costs of the technology and moving forward to the solution faster. For example, 3D observation using mono camera.

2. Sensory data fusion - these methods improve the robustness to weather or other environmental conditions.

3. Artificial intelligence is an emerging field in many disciplines. The advantages of AI should be considered in the operation of AVs as a tool for data analysis and decision making. AI in an aspect of AV that requires more research.

References


*Autonomous Vehicle, Sensing and Communication Survey*


