Deep Learning Algorithms for Smart Cars: A Survey

Mostafa Abotaleb[*1], Ehsaneh khodadadi[2], Nadjem Bailek [3,4]

[1] Department of System Programming, South Ural State University, 454080 Chelyabinsk, Russia

[2]Department of Chemistry and Biochemistry, University of Arkansas, Fayetteville, AR 72701, USA.

[3]Energies and Materials Research Laboratory, Faculty of Sciences and Technology, University of Tamanghasset, Tamanrasset, 10034, Algeria.

[4]Sustainable Development and Computer Science Laboratory, Faculty of Sciences and Technology, Ahmed Draia University of Adrar, Adrar, Algeria
Emails: abotalebmostafa@bk.ru; ekhodada@uark.edu; bailek.nadjem@univ-adrar.edu.dz

Abstract

The rate of progress in autonomous car technology has increased exponentially over the past decade, mostly thanks to advancements in deep learning and artificial intelligence. This work aims to summarize recent progress made in the application of deep learning techniques to the problem of autonomous driving. First, we will go through the deep reinforcement learning paradigm and other AI-based solutions for autonomous driving, such as convolutional and recurrent neural networks. Algorithms for driving scene recognition, path planning, behavior arbitration, and motion control were developed with these techniques in mind. Both the End2End system, which immediately converts sensory input into steering commands, and the modular perception-planning-action pipeline, each module of which is built using deep learning techniques, are the focus of our studies. We also discuss the modern challenges of building AI systems for autonomous driving, such as making sure they are safe to use, finding good places to practice, and creating effective computing hardware. This survey's comparison sheds light on the pros and cons of AI and deep learning approaches to autonomous driving, which aids in making design decisions.1

Keywords: Artificial Intelligence; Deep learning; Self-driving; End2End system.

1. Introduction

The domains of computer vision [1, 2], robotics [2, 3], and Natural Language Processing (NLP) [4, 5] have all seen significant advancements in recent years, and these advancements can largely be attributed to the rise of Deep Learning and AI as two of the primary technologies responsible for these developments. They also have a major bearing on the commercial and academic revolutions in autonomous driving that are currently underway. Over the past few years, we've seen the first steps toward moving the research and development of autonomous vehicles (AVs) and self-driving automobiles out of the laboratory and onto actual streets and highways. If we included them into our current ecological system, we would observe a decrease in traffic accidents and congestion, as well as an increase in people's ability to move around congested urban areas. Although the word "self-driving" may seem self-explanatory, it is actually defined by one of five different SAE Levels. The SAE J3016 standard [4] introduces a scale from 0 to 5 that can be used to rate the degree of automation present in a given vehicle. Vehicles with a lower SAE have the most rudimentary forms of driver assistance, while those with the highest levels of SAE have none. Level 5 autonomous vehicles are completely hands-free, with no controls other than the accelerator and brake.

Doi: https://doi.org/10.54216/JAIM.050202
Received: January 25, 2023 Revised: May 14, 2023 Accepted: September 04, 2023

Traditional methods of vision, path planning, and motion control can address most driving scenarios with relative ease. However, there are some edge circumstances for which these approaches cannot provide a solution.

One of the earliest examples of a fully autonomous vehicle was constructed by Ernst Dickmanns [5] in the 1980s. This laid the groundwork for subsequent research programs, such as PROMETHEUS, whose ultimate goal was to create a completely autonomous vehicle. The VaMP project's autonomous car was able to travel 1,600 kilometers (95 percent of which it drove on its own) in 1994. Similar work was done by the CMU NAVLAB in 1995, when they drove 6,000 kilometers autonomously and completed 98% of the distance without human intervention. Significant progress was made toward the goal of completely autonomous driving with the help of the DARPA Grand Challenges in 2004 and 2005 and the DARPA Urban Challenge in 2007. The goal was to make an autonomous car finish an off-road course as quickly as possible without any human intervention. In that year's race, none of the 15 starting vehicles crossed the finish line. Using principles from machine learning, Stanley won the 2005 tournament by adapting to the unpredictable conditions. As the first time that machine learning and artificial intelligence were recognized as crucial components of autonomous driving, this was a watershed point in the evolution of autonomous vehicles. Since the vast majority of the surveyed works appeared after 2005, this article's focus on that year is also significant.

We provide a comprehensive overview of the different AI and deep learning technologies utilized in autonomous driving, including a discussion of the current state-of-the-art approaches to applying these methods to self-driving automobiles. In addition, we provide extensive coverage of the challenges posed by addressing safety concerns, identifying appropriate training data sources, and acquiring the required computational hardware.

The core of autonomous vehicles are the systems that make decisions on their own. Cameras, radars, LiDARs, ultrasonic sensors, GPS units, and inertial sensors are just some of the on-board sources of data that are processed by these systems. These observations will be used by the car's computer to inform its driving judgments. Basic AI-driven autonomous vehicle block designs are depicted in Figure 1.

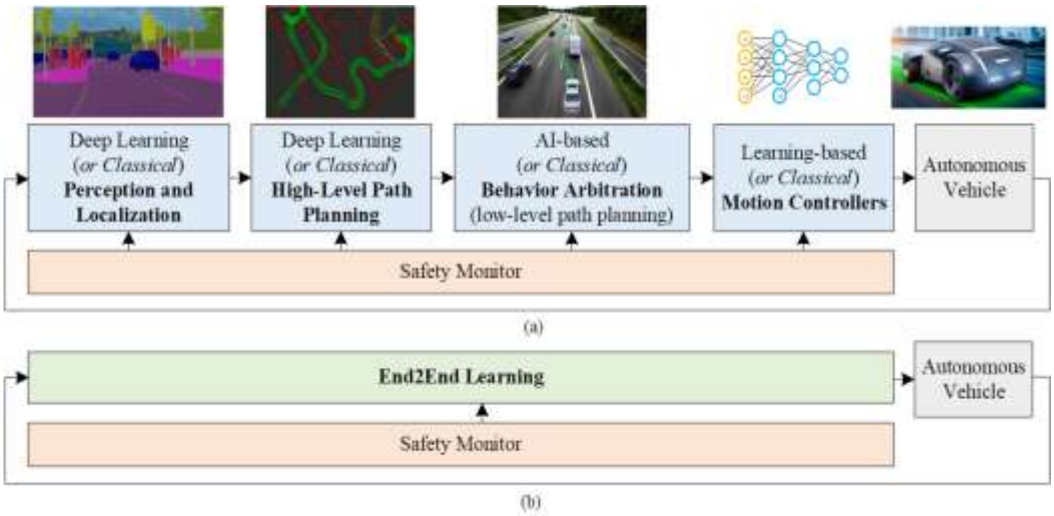

Figure 1: Deep Learning based self-driving car (a) sequential perception- planing-action pipeline (b) End2End system

Both a modular perception-planning-action pipeline (Figure 1(a)) and an End2End learning technique (Figure 1(b)) are employed to compute the driving decisions, with the latter immediately transferring sensory input to control outputs. Models A and B are depicted in Figure 1. The modular pipeline's components can be built using either artificial intelligence (AI) and deep learning techniques or conventional methods that don't rely on learning at all. Combinations of learning and non-learning based components are abundant (a traditional A-star based path planning algorithm, for instance, can take input from a deep learning based object detector). A safety monitor has been created to check the security of each individual module.

The modular pipeline shown in Figure 1(a) has been logically divided into four sub-processes. These parts can be created with either modern AI and deep learning techniques or more conventional

Doi: https://doi.org/10.54216/JAIM.050202
Received: January 25, 2023 Revised: May 14, 2023 Accepted: September 04, 2023

approaches. Here are the parts: "Perception and Localization," "High-Level Path Planning," "Behavior Arbitration," also called "low-Level Path Planning," and "Motion Controllers."

Based on these four criteria, we have organized a large body of literature detailing deep learning methods developed for autonomous vehicles. We have assembled a set of articles that address the numerous hardware, data sources, and safety problems that come up during the process of designing deep learning modules for autonomous vehicles, in addition to the methods that were investigated.

Once an autonomous car has been programmed with a route to follow across a network of roads, its first task is to locate itself in its immediate vicinity. Based on this model, a continuous route is plotted, and the car's next moves are determined by the behavior arbitration system. Finally, a motion control system will correct any problems that develop while the motion is being carried out. These four parts are discussed in [6], along with conventional design methods that do not include AI.

Here we will present a high-level review of the deep learning and AI technologies used in autonomous driving, as well as a survey of the many methodologies used in the development of the aforementioned hierarchical decision making process. We also provide an overview of End2End learning systems, which are employed to compress the hierarchical procedure into a single deep learning architecture that directly maps sensory observations to control outputs.

2. Deep Learning for Path Planning

Autonomous vehicles use a method called path planning to figure out how to get from their current location to a predetermined destination. In order to avoid running into any obstacles, a self-driving car must, in accordance with the path planning methodology, analyze the surrounding area for any potential hazards and then calculate a trajectory along a route that avoids them. According to [7], autonomous driving is a multi-agent scenario in which the host car must utilize complex negotiation skills with other road users while overtaking, giving way, merging, and making left and right turns on unstructured urban freeways. The host car here has to deal with the chaos of urban streets. According to the study's conclusions, serious action needs to be taken to improve driver safety. With a reward function of R(s) = r for a dangerous event to avoid and R(s) [1, 1] for the rest of the trajectories, the goal is to master challenging maneuvers without incident.

Optimal path planning for autonomous vehicles is a relatively new area of study that necessitates fast processing speeds to meet response time goals and optimization constraints. The review available at [8] provides a high-level overview of route planning for autos. There's talk of the route planner's "taxonomic components," like "behavior planners," "motion planners," and "mission planners." Despite the growing interest in utilizing deep learning technology for path planning and behavior arbitration, an assessment of such technologies is missing from [8]. Imitation Learning (IL) [9, 10], and 11] and Deep Reinforcement Learning (DRL) based planning [12], and [13] will be analyzed below. These are two of the most well-known applications of deep learning to the field of path planning.

The goal of Imitation Learning [9], [10], and [11] is to learn how a person normally drives by seeing and analyzing their driving behavior in recorded scenarios [14]. The technique assumes a human demonstration as the starting point for education. As a result, the authors tune in to CNN to absorb the network's planning strategies through proxy. Neuro Trajectory [11] is one such perception-planning deep neural network, and it is trained to forecast the ego-vehicle's trajectory based on the ego's intended state across a given time window. The scientists at Google Brain created this network. The goal of Inverse Reinforcement Learning (IRL) is to learn the reward function of a human driver [15], which is another way of looking at imitation learning. In order to learn reward functions and construct human-like driving trajectories, these techniques use real-world driving data from humans.

Path planning DRL focuses on driving skill acquisition via simulator exercise [7], [16], [12], and [13]. The process of abstracting and then transforming a model of the real-world environment into a virtual one is called a transfer model. It is clear from the text that the target function cannot ensure functional safety without simultaneously creating a serious variance issue, as stated in [7]. Building a policy function with both learnable and non-learnable parts is recommended as a solution to this issue. The objective of the learnable policy is to maximize some sort of reward function, which might be anything from passenger convenience to vehicle safety to the opportunity to overtake another vehicle, etc. The non-learnable policy, meanwhile, meets all the functional safety standards without sacrificing too much convenience.

Doi: https://doi.org/10.54216/JAIM.050202
Received: January 25, 2023 Revised: May 14, 2023 Accepted: September 04, 2023

Both IL and DRL have advantages and disadvantages when it comes to planning routes. One major feature of IL is that it may be trained using data collected in the wild. However, because there is a dearth of data on edge cases (such as lane-hopping, vehicle collisions, and so on), it is challenging to foresee how the trained network would respond to data it has never seen before. However, while DRL systems may explore a wide range of driving scenarios in a virtual environment, the bulk of the time these models display biased behavior when transferred to the real world.

3. Motion Controllers for AI-based Self Driving Cars

The motion controller is in charge of calculating the vehicle's longitudinal and lateral steering commands. Figure 1(a) depicts the use of learning algorithms as a part of Learning Controllers, whereas Figure 1(b) depicts the use of learning algorithms as part of End2End Control Systems, which directly translate sensory input to steering commands. The motion control component contains both of these functions.

A. Learning Controllers

Traditional controllers are dependent on a model with set parameters and a priori knowledge. When robots or other autonomous systems are used in complex scenarios, such as driving, traditional controllers are limited in their ability to prepare for every possible event that may emerge. This makes it difficult for traditional controllers to keep up. In contrast to traditional controllers, learning controllers make use of training data in order to enhance the accuracy of their models over the course of time. The precision of the model's approximation of the real system is improved with the addition of each new batch of training data. This results in higher scalability, more accurate uncertainty estimations, and the capacity to anticipate consequences and disturbances that cannot be simulated in advance of deployment [17]. Take into consideration the state-space nonlinear system that is listed below:

$$(t+1) \quad = ftrue\ (z(t)\ ,\ u(t)\ ) \tag{1}$$

with observable state $z(t) \in Rn$ and control input $u(t)> \in R\ m$, at discrete time t . The true system ft rue is not known exactly and is approximated by the sum of an a-priori model and a learned dynamics model.

In order to function, conventional controllers rely on a model with fixed parameters and assumed background information. Traditional controllers have a limited ability to plan for every possible event that may occur when robots or other autonomous systems are used in complex settings like driving. This makes it difficult for conventional controllers to keep up. In contrast to conventional controllers, learning controllers can refine their models with more training data. Training data is not used by conventional controllers. When more data is added to the training set, the model's approximation of the real system gets closer and closer to the real thing. Scalability is improved, uncertainty estimates are more accurate, and unanticipated impacts and disturbances can be predicted [17]. The scalability is improved as a result as well. Consider the nonlinear system in state space that is given below:

B. End2End Learning Control

In the context of autonomous driving, what is indicated by the term "End2End Learning Control" is a mapping that is direct from the sensory input to the control commands. Inputs often consist of high-dimensional feature spaces, such as images or point clouds. Compare this to the workflow shown in Figure 1(b), in which objects are first recognized in the input image, then a path is planned, and lastly the computed control values are done. This process begins by analyzing the input image to determine what it contains.

It is also feasible to consider End2End learning as a form of back propagation strategy that is designed specifically for use with large-scale and intricate models. In the 1990s, [28] researchers built the Autonomous Land Vehicle in a Neural Network (ALVINN), which was the very first system to use this concept. ALVINN is pre-programmed to follow a predetermined route and to alter the direction in which it travels based on the arc that it detects in the road. DAVE (Darpa Autonomous VEhicle) is believed to have accomplished the next milestone in End2End driving in the middle of the 2000s [29]. DAVE was trained on hours of human driving that were gathered under situations that were comparable to, but not identical to, their own driving conditions. In the recent past, there has been a

Doi: https://doi.org/10.54216/JAIM.050202
Received: January 25, 2023 Revised: May 14, 2023 Accepted: September 04, 2023

steady increase in the processing capacity of computers, which has opened the path for the widespread implementation of End2End learning models. Back-propagation is a technique that may now be effectively implemented on parallel Graphics Processing Units for the purpose of gradient estimation in deep neural networks. This processing is required in order to train large and complex network topologies, and it calls for a significant amount of data to be processed in order to be effective.

Deep Reinforcement Learning (DRL) systems, which are taught and evaluated in simulation [30], as well as deep neural networks [30], which are trained offline using real-world and/or synthetic data, are frequently used in End2End control publications. [30] Both of these types of systems are taught and evaluated in simulation [30]. There have been reports of approaches for translating DRL models learned in simulation to real-world driving [27], in addition to DRL systems that have been trained using real-world picture data [23], [24].

In recent years, NVIDIA®'s PilotNet design has become increasingly popular. This architecture incorporates End2End methodologies. Raw images from a single camera that faces forward are used to train a convolutional neural network (CNN), which then converts those images into directional inputs [25]. The training data consists of pictures and steering inputs taken during actual driving circumstances, which took place in a wide variety of environmental and atmospheric conditions, as well as on a wide variety of surface types. Before training begins, the process of data augmentation is carried out in order to enhance the quality of the data by applying simulated changes to the raw data.

PilotNet has somewhere in the neighborhood of 27 million connections and a quarter of a million characteristics. The testing is carried out in both a virtual environment and a vehicle that exists in the real world. The autonomous performance metric is a measurement that determines the proportion of time that the neural network is in charge of the vehicle. This percentage is determined as follows: (2) Autonomy = 100 (one minus the number of interventions times six seconds)/(time that has elapsed).

Because it takes a human driver at least six seconds to retake control of the car and return it to the appropriate state, an intervention is presumed to have taken place when the virtual vehicle strays more than one meter from the center line. The car was able to drive itself 98% of the time while making the trip from Holmdel to Atlantic Highlands, which is located in New Jersey and is a distance of 20 kilometers. The ability to calculate the steering commands is something that human drivers have to teach PilotNet. The most important step is to identify those features of the input traffic image that have the most influence on the route that is selected by the network. A method is provided for detecting the prominent object regions in the input image, and the low-level features that were learned by PilotNet are equivalent to those that are essential to a human driver.

End2End systems [28, 29] that are analogous to PilotNet and link visual data to steering commands have been disclosed. [30] poses the issue of autonomous driving as a potential ego-motion prediction challenge for the future. The FCN-LSTM technique, which stands for Fully Convolutional Network - Long-Short Term Memory, was developed in order to train pixel-level supervised tasks using a fully convolutional encoder and to predict motion with a temporal encoder. Convolutional Long Short Term Memory, or C-LSTM, is a network that is proposed for directional control in [30]. This network takes into consideration the visual and temporal dependencies of the input data. The article [31] describes how 360-degree cameras were used to accomplish End2End learning. It is generally accepted that human drivers make use of rear-view and side-view mirrors; hence, it is presumed that all accessible environmental data must be gathered and incorporated into the network model before an appropriate control order can be generated and given.

For the purpose of evaluating the Tesla® Autopilot system, [31] suggested using a framework of the End2End Convolutional Neural Net variety. It is designed to compare its own output to that of the Autopilot in order to locate any inconsistencies that may exist. The network was trained using information from more than 420 hours of real-world driving. The capabilities of Autopilot and those of the proposed framework were put to the test in the form of real-time comparisons and contrasts using a Tesla® vehicle as the testing platform. DRL is an alternate way for developing comprehensive driving systems and the evaluation findings showed that both systems were correct 90.4% of the time when discriminating between one another and the handoff of vehicle control to a human operator. This is often carried out in a simulated environment, which shields an autonomous driving agent from any potential dangers that may be present. The TORCS game engine employs a DRL End2End mechanism to perform the computation necessary to determine the steering command. [31] created an asynchronous advantage Actor-Critic (A3C) technique for training a CNN on photos and vehicle velocity input, while taking into consideration a more complex virtual environment. Ideas that are

Doi: https://doi.org/10.54216/JAIM.050202
Received: January 25, 2023 Revised: May 14, 2023 Accepted: September 04, 2023

similar have been improved upon in [26], and the new version allows for faster convergence as well as broader generalization. Both parts are dependent on the same cycle, which consists of receiving the game's current state, making a decision on the next set of commands, and then experiencing the results of that decision during the subsequent iteration. The experimental configuration was made to feel more like the real world with the help of World Rally Championship 6 and other simulations, such as TORCS.

It seems that the next step in the development of DRL-based control will be the incorporation of classic model-based control methods, such as those outlined in Section 6.1. The policy of the neural network is calculated, and this is done on top of the stable and deterministic model that the classical controller provides. In this way, the stringent constraints of the model system are included into the policy of the neural network. A DRL policy that is based on real-world picture data is what the author of [28] suggests using for the task of aggressive driving. In this scenario, a model predictive controller supplies optimal trajectory examples to the CNN, also known as the learner, while it is being trained.

4. Data Sources for Training Autonomous Driving Systems

When it comes to the training and testing of an autonomous driving component, the exploitation of real-world data is without a doubt one of the most significant requirements that must be met. Data gathering on public roadways has shown to be an exceptionally valuable activity for the development stage of such components. This is because such components require a considerable number of data during the development stage. The vehicle that is used for data collection is equipped with a number of sensors, like as radar, LIDAR, GPS, cameras, Inertial Measurement Units (IMU), and ultrasonic sensors, in order to obtain a specific description of the driving scenario. These sensors are used to capture information about the environment around the vehicle. This is done in an effort to collect the greatest amount of information feasible. Because the arrangement of the sensors is dependent on the use that the data is going to be put to, each vehicle has its own distinct set of sensors. A common arrangement for an autonomous vehicle's sensory subsystems is depicted in Figure 2, which may be found here.

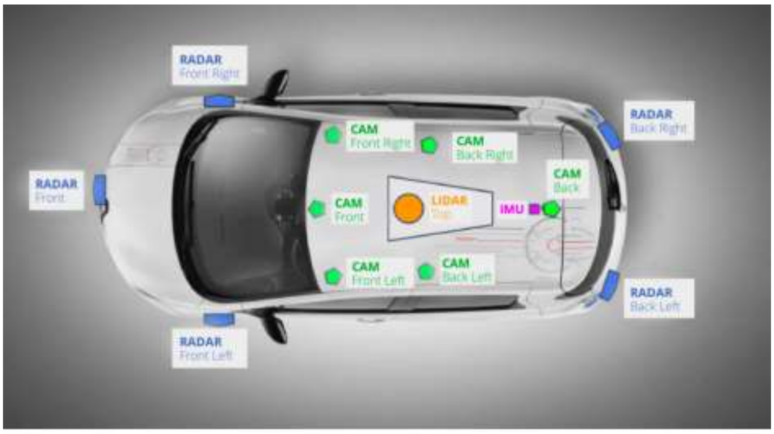

Figure 2: Sensor for self-driving car

The significant and growing interest in the study of autonomous vehicles has led to the release and documentation of numerous driving datasets during the past few years. Size, number of sensors, and file type are only few of the ways in which they diverge. Researchers need only look for the dataset that most closely fits their specific study question. [29] is the number of recently released datasets included in a survey. Only a small fraction of these databases, however, are particularly useful for research into autonomous driving. All areas of computer vision are covered by these datasets.

One of the most comprehensive assessments of open-source datasets for algorithms used in autonomous vehicles can be found in [30]. Twenty-seven datasets with information gathered from public roads are presented here. Different perspectives are used to compare the datasets so that the reader can pick the one that makes the most sense for his or her research.

We have looked everywhere for a master dataset that includes all or most of the available datasets, but to no avail. This could be due to the fact that there are currently no standards for the optimal

Doi: https://doi.org/10.54216/JAIM.050202
Received: January 25, 2023 Revised: May 14, 2023 Accepted: September 04, 2023

configuration of sensors or the presentation of collected data. Each dataset's meaning and interpretation are profoundly affected by the algorithm's intended use. Scale® and nuTonomy®, two IT companies, have just began compiling one of the largest and most comprehensive datasets for autonomous vehicles. The Berkeley Deep Drive dataset, developed by researchers at Berkeley, is one example. The scientific literature contains additional useful datasets that need to be combined7.

In [32], the authors provide a study that aims to assess the current state of the art by collecting and analyzing large-scale naturalistic data of semi-autonomous driving. Their end goal is to more accurately depict the contemporary technological landscape. About 5.5 billion video frames, 99 participants, 29 cars, 405,807 miles, and 5.5 billion miles were analyzed in this study. Unfortunately, the data collected for this study is not released to the public.

In the following sections, we will provide and highlight the unique characteristics of the publicly available datasets that are the most useful in general.

KITTI Vision Benchmark (KITTI) dataset .The Karlsruhe Institute of Technology (KIT) in Germany created this dataset to fulfill the needs of evaluating stereo-vision, optical flow, 3D tracking, 3D object detection, and simultaneous localization and mapping (SLAM) algorithms. It is widely considered to be the most important dataset for research into autonomous cars. Over 2000 references to it can be found in academic journals as of now. The data collection vehicle is equipped with a Velodyne 3D LiDAR, high-precision GPS and IMU sensors, and multiple high-resolution color and grayscale stereo cameras. It features six hours' worth of driving information collected in and around Karlsruhe, Germany, in urban, suburban, and rural settings. The dataset is made accessible under a Creative Commons license that mandates attribution, forbids commercial usage, and necessitates that all users adhere to the same rules.

from the NuScenes data collection. The cities of Boston and Singapore provided the data for this nuTonomy-created dataset, which contains over a thousand individual driving situations. Both cities are well-known for their intense traffic and exceedingly challenging driving conditions. When creating this dataset, the data providers tagged 25 different types of items with precise 3D bounding boxes at a rate of 2 Hz across the board. This was done to facilitate common computer vision tasks like object detection and tracking. Data collecting on automobiles is now underway. About 1.41 million still photos, 400.000 Lidar scans, 1.31 million RADAR scans, 1.11 million object bounding boxes, and 40.000 keyframes will make up the final dataset. Under the rules of the Creative Commons Attribution-Non Commercial Share Alike 3.0 License, the dataset is freely available to users.

For the Automotive Multi-Sensor Dataset the abbreviation "AMUSE" is used. This data, provided by Sweden's Linkoping University, is a compilation of recordings collected in a number of locations with a car furnished with height sensors, an IMU, a velocity sensor, and a GPS. Long data streams from several sensors and cameras, recorded in the proper format, are made available to the public alongside the application programming interface (API) needed to access them. The data collection is released under an unsupported version of the Creative Commons Attribution-NonCommercial-NoDerivs 3.0 License.

The Ford campus's acquired visual and lidar information . The data was collected with a Point Grey Ladybug3 omnidirectional camera system, a Velodyne Lidar scanner, two push-broom forward looking Riegl Lidars, and an Aplanix POS-LV professional inertial measurement unit (IMU). The University of Michigan subsequently supplied it. More than 100 terabytes of information were gathered in and around the Ford Research site in downtown Dearborn, Michigan in 2009. Several simultaneous localization and mapping (SLAM) techniques, utilized in autonomous driving, can be thoroughly tested on this dataset.

Reference to the Udacity data collection at . The vehicle is equipped with a Velodyne 3D Lidar, a Global Positioning System (GPS), and an Inertial Measurement Unit (IMU). Total size of the dataset is 223 GB. Following its recognition, the user is provided with the equivalent steering angle that was recorded by a human driver during the tests.

Oxford's produced data collection . The dataset was gathered at Oxford University in the UK over the course of a year. More than a thousand kilometers of driving data was collected, together with around 20 million pictures from six cameras mounted on the car, plus ground truth data from LIDAR, GPS, and INS. Information was collected in every conceivable setting, including a thunderstorm, the middle of the night, the middle of the day, and in the snow. The vehicle in this dataset repeatedly drove the same route over the course of a year, allowing researchers to investigate long-term localization and

Doi: https://doi.org/10.54216/JAIM.050202
Received: January 25, 2023 Revised: May 14, 2023 Accepted: September 04, 2023

mapping for autonomous vehicles in real-world, dynamic urban environments. This is one of the features that makes this dataset unique.

The term "CamVid" refers to the Cambridge-based Video Dataset with Labels . The University of Cambridge in the United Kingdom released this dataset as the first of its kind to the public. This dataset includes films annotated with object class semantic labels and metadata, and it has been cited more than any other dataset to date. The ground-truth labels assigned to each pixel by the database fall into one of 32 semantic categories. Only one camera, mounted high on the dashboard, is used by the sensors in this setup. Since the car is only driven in urban areas with a moderate quantity of foot traffic and suitable weather circumstances, the scenarios are not very difficult.

Pedestrian data collection used as a standard by Daimler. Daimler AG R&D and the University of Amsterdam have generously shared a dataset useful for studying pedestrian recognition, classification, segmentation, and path prediction. Information on pedestrians is collected from a traffic vehicle using only the mono and stereo cameras installed in the vehicle. This is the first ever collected data set to focus on pedestrians.

Caltech's (Caltech) pedestrian identification dataset. Provided by the California Institute of Technology in the United States, this collection features recordings captured with a camera attached to a moving vehicle, resulting in difficult images of low quality and people who are frequently concealed. Around ten hours' worth of driving scenarios are represented by around two hundred and fifty thousand frames, with three hundred and fifty thousand bounding boxes and two thousand and three hundred individual pedestrian comments. The annotations include not only spatial but also temporal correspondences between the bounding boxes, as well as extensive occlusion label information.

5. Conclusion

Significant progress has been made in autonomous car technology during the past decade, especially as a result of breakthroughs in artificial intelligence and deep learning. At various stages in the process of creating individual parts of a self-driving car, contemporary AI approaches are either directly applied or taken into consideration. Not only have deep learning algorithms affected the structure of traditional perception-planning-action pipelines, but they have also enabled the creation of end-to-end learning systems that can map sensory input to control input without any intermediate steps. Automated vehicles, often known as driverless automobiles, are high-tech conveyances built to safely carry people and goods between locations. There will be many challenging issues to overcome once autonomous vehicles powered by AI begin to appear on public roads. One of the biggest challenges is establishing the functional safety of autonomous vehicles, which is challenging given the current formalism and the ability of neural networks to explain occurrences. Further, deep learning algorithms require large quantities of processing hardware and rely heavily on massive volumes of training data. In this post, we took a look at the many applications of deep learning in autonomous vehicles. The research on performance and computing requirements provides a system level reference for the design of AI-based self-driving vehicles.

Funding: “This research received no external funding”

Conflicts of Interest: “The authors declare no conflict of interest.”

References

[1] A Krizhevsky, I Sutskever, and G. E. Hinton, ImageNet Classification with Deep Convolutional Neural Networks. Advances in Neural Information Processing Systems, 25, 1097–1105, 2012.

[2] M Andrychowicz et al, Learn- ing Dexterous In-Hand Manipulation. CoRR, vol. abs/1808.00177, August 2018. [Online]. Available: https://arxiv.org/abs/1808.00177

[3] Y. Goldberg, Neural Network Methods for Natural Language Processing, ser. Synthesis Lectures on Hu- man Language Technologies. Morgan & Claypool, 37, 2017.

Doi: https://doi.org/10.54216/JAIM.050202
Received: January 25, 2023 Revised: May 14, 2023 Accepted: September 04, 2023

[4] SAE Committee, Taxonomy and Definitions for Terms Related to On-road Motor Vehicle Automated Driving Systems, 2014.

[5] E Dickmanns, V. Graefe, Dynamic Monocular Machine Vision. Machine vision and applications, 1, 223–240, 1988.

[6] B Paden, M Cap, S Z Yong, D S Yershov, E. Frazzoli, A Survey of Motion Planning and Control Techniques for Self-Driving Urban Vehicles. IEEE Trans. Intelligent Vehicles, 1(1), 33– 55, 2016.

[7] S Shalev-Shwartz, S Shammah, A Shashua, Safe, Multi-Agent, Reinforcement Learning for Autonomous Driving, 2016.

[8] S. D. Pendleton, H. Andersen, X. Du, X. Shen, M. Meghjani, Y. H. Eng, D. Rus, and M. H. Ang, Perception, Planning, Control, and Coordination for Autonomous Vehicles. Machines, 5(1), 6, 2017.

[9] E. Rehder, J. Quehl, and C. Stiller, Driving Like a Human: Imitation Learning for Path Planning us- ing Convolutional Neural Networks. Int. Conf. on Robotics and Automation Workshops, 2017.

[10] L. Sun, C. Peng, W. Zhan, and M. Tomizuka, A Fast Integrated Planning and Control Framework for Autonomous Driving via Imitation Learning. ASME 2018 Dynamic Systems and Control Conference, 3, 2018.

[11] S. Grigorescu, B. Trasnea, L. Marina, A. Vasilcoi, and T. Cocias, NeuroTrajectory: A Neuroevolution- ary Approach to Local State Trajectory Learning for Autonomous Vehicles. IEEE Robotics and Automation Letters, 4(4), 3441–3448, 2019.

[12] L. Yu, X. Shao, Y. Wei, and K. Zhou, Intelli- gent Land-Vehicle Model Transfer Trajectory Planning Method Based on Deep Reinforcement Learning. Sensors (Basel, Switzerland), 18, 09, 2018.

[13] C. Paxton, V. Raman, G. D. Hager, and M. Kobilarov, Combining Neural Networks and Tree Search for Task and Motion Planning in Challenging Environments. Int. Conf. on Intelligent Robots and Systems (IROS), 2017.

[14] W. Schwarting, J. Alonso-Mora, and D. Rus, Planning and Decision-Making for Autonomous Vehi- cles," Annual Review of Control, Robotics, and Autonomous Systems, 1, 2018.

[15] T. Gu, J. M. Dolan, and J. Lee, Human-like Planning of Swerve Maneuvers for Autonomous Vehicles. IEEE Intelligent Vehicles Symposium (IV), 716-721, 2016.

[16] A. I. Panov, K. S. Yakovlev, and R. Suvorov, Grid Path Planning with Deep Reinforcement Learning: Preliminary Results. Procedia Computer Science, 123, 347 – 353, 2018.

[17] C. J. Ostafew, J. Collier, A. P. Schoellig, and T. D. Barfoot, Learning-based Nonlinear Model Predic- tive Control to Improve Vision-based Mobile Robot Path Tracking. Journal of Field Robotics, 33(1), 133–152, 2015.

[18] P. J. Nguyen-Tuong D and S. M, Local Gaussian Process Regression for Real Time Online Model Learning. Proceedings of the neural information processing systems Conference, 1193-1200, 2008.

[19] O. Sigaud, C. Salaun, and V. Padois, On-line Regression Algorithms for Learning Mechanical Mod- els of Robots: A Survey. Robotics and Autonomous Systems, 59(12), 1115–1129, 2011.

[20] C. Ostafew, A. Schoellig, and T. D. Barfoot, Visual Teach and Repeat, Repeat, Repeat: Iterative Learning Control to Improve Mobile Robot Path Tracking in Challenging Outdoor Environments, 176–181, 2013.

[21] B. Panomruttanarug, Application of Iterative Learning Control in Tracking a Dubin's Path in Parallel Parking. Int. Journal of Automotive Technology, 18(6), 1099–1107, 2017.

[22] N. R. Kapania and J. C. Gerdes, Path Tracking of Highly Dynamic Autonomous Vehicle Trajectories via Iterative Learning Control. American Control Conference (ACC), 2753– 2758, 2015.

[23] S. Lefvre, A. Carvalho, and F. Borrelli, A Learning- Based Framework for Velocity Control in Autonomous Driving. IEEE Transactions on Automation Science and Engineering, 13(1), 32– 42, 2016.

[24] P. Drews, G. Williams, B. Goldfain, E. A Theodorou, and J. M Rehg, Aggressive Deep Driving: Combining Convolutional Neural Networks and Model Predictive Control, 133–142, 2017.

[25] J. Rawlings and D. Mayne, Model Predictive Control: Theory and Design, Nob Hill Pub., 2009.

Doi: https://doi.org/10.54216/JAIM.050202
Received: January 25, 2023 Revised: May 14, 2023 Accepted: September 04, 2023

[26] M. Kamel, A. Hafez, and X. Yu, A Review on Motion Control of Unmanned Ground and Aerial Vehicles Based on Model Predictive Control Techniques. Engineering Science and Military Technologies, 2, 10–23, 2018.

[27] M. Brunner, U. Rosolia, J. Gonzales, and F. Borrelli, Repetitive Learning Model Predictive Control: An Autonomous Racing Example. IEEE 56th Annual Conference on Decision and Control (CDC), 2017, 2545–2550, 2017.

[28] D. A. Pomerleau, Alvinn: An autonomous Land Vehicle in a Neural Network. in Advances in neural information processing systems, 305–313, 1989.

[29] U. Muller, J. Ben, E. Cosatto, B. Flepp, and Y. L. Cun, Off-road Obstacle Avoidance through End-to- End Learning. Advances in neural information processing systems, 739–746, 2006.

[30] H. Xu, Y. Gao, F. Yu, and T. Darrell, End-to-End Learning of Driving Models from Large-scale Video Datasets. IEEE Conf. on Computer Vision and Pat- tern Recognition (CVPR), 2017.

[31] E. Perot, M. Jaritz, M. Toromanoff, and R. D. Charette, End-to-End Driving in a Realistic Racing Game with Deep Reinforcement Learning, IEEE Conf. on Computer Vision and Pattern Recognition Workshops (CVPRW), 474–475, 2017.

[32] L. Fridman, et al, MIT Autonomous Vehicle Technology Study: Large-Scale Deep Learning Based Analysis of Driver Behavior and Interaction with Automation. IEEE Access 2017, 2017.

Doi: https://doi.org/10.54216/JAIM.050202
Received: January 25, 2023 Revised: May 14, 2023 Accepted: September 04, 2023