iJRASET
International Journal For Research in Applied Science and Engineering Technology

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 13 Issue: VI Month of publication: June 2025

DOI: https://doi.org/10.22214/ijraset.2025.72240

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 13 Issue VI June 2025- Available at www.ijraset.com

Multi-Sensor Fusion Using Deep Neural Networks for Robust Object Detection in Autonomous Driving

Mr. Chayan Bhattacharjee[1], Ms. Sayali Parab[2]
[1]*Department of Information Technology, Chikitsak Samuha's Patkar Varde College, Mumbai, India*
[2]*Department of Information Technology, SES's L. S. Raheja College of Arts & Commerce, Mumbai, India*

Abstract: Nowadays, self-directed learning and automation are not restricted to human beings only. If you stare out at the automotive horizon, you can see a new, exciting era coming into the limelight: the age of self-driving cars. An age when humans will no longer need to keep their eyes on the road. No more concerns about distraction while driving or those stressful rush hour commutes, vehicles will whisk us where we want to go, blazingly fast and efficiently. Deep learning is one potential solution for object detection and scene perception problems, which can enable algorithm-driven and data-driven cars. This research paper presents a comprehensive survey of deep learning applications for object detection and scene perception in autonomous vehicles. In this research paper, the theory underlying self-driving vehicles from a deep learning perspective and current implementations of self-driving cars, followed by their critical evaluations, are examined. Thus, in this paper, the gap between deep learning and self-driving cars is bridged through a comprehensive survey.

Keywords: Self-driving car, Machine learning, Deep learning, Convolutional neural networks, Object detection, Multimodal sensor fusion, LiDAR, Computer vision, Autonomous driving initiatives.

I. INTRODUCTION

With recent advances in artificial intelligence (AI), machine learning (ML) and deep learning (DL), various applications of these techniques have gained prominence and come to fore. One such application is self-driving cars, which is anticipated to have a profound and revolutionary impact on society and the way people commute. Although the acceptance and domestication of technology can face initial or prolonged reluctance, yet these cars will mark the first far-reaching integration of personal robots into human society. However, for self-driving cars to become a functional reality, they need to be equipped with perception and cognition to tackle high-pressure real-life scenarios, arrive at suitable decisions, and take appropriate and safe action at all times.

Embedded in the self-driving vehicles' AI are visual recognition systems (VRS) that encompass image classification, object detection, segmentation, and localization for basic ocular performance. Object detection is emerging as a subdomain of computer vision (CV) that benefits from DL, especially convolutional neural networks (CNNs). This paper discusses the self-driving cars' vision systems, the role of DL to interpret complex vision, enhance perception, and actuate kinematic manoeuvres in self-driving cars. It surveys methods that tailor DL to perform object detection and scene perception in self-driving cars.

The rest of the article is structured as follows: Section II provides a brief introduction to the evolution of self-driving cars and discusses the levels of automation to gradually and progressively achieve fully autonomous vehicles. Section III introduces big data, the role of big data in autonomous vehicles and collecting driving data using various cameras. Processing driving data captured using various sensors in real time is a significant challenge, and one of the promising solutions, such as multimodal sensor fusion, is discussed in this section. Section IV introduces deep learning (DL) and the factors that make DL a powerful technique in computer vision. Section IV delves deeper into CNNs, RNNs, DBNs, and other widely used DL techniques in CV. Lastly, we enlist some promising future directions to achieve next-generation autonomous vehicles based on the survey and conclude the paper.

II. EVOLUTION OF SELF-DRIVING CARS

A. Brief history of self-driving vehicles

The concept of self-driving cars has been around for almost 80 years, first reported in 1939 World's Fair in New York by General Motors' (GM) Futurama. Contemporary developments in communication networks and wireless connectivity, the arrival of accurate and robust sensors that continuously miniaturize in size and cost, coupled with AI have been the cornerstone for autonomous driving systems. Embedded in these self-driving systems are human-machine interface applications, network-enabled controls, multiple sensor data fusion, 3D drive scene analysis, and software-defined signal processing to transport materials, payloads, goods, and people. The accuracy of autonomous navigation depends significantly on attaining precise localization, unobtrusive data collection, fused data-set generation, and uninterrupted high-level communication with other vehicles and surrounding smart infrastructure. In the last decade, Carnegie Mellon University and the Defence Advanced Research Projects Agency (DARPA) self-driving cars have contributed to autonomous vehicles advancement. Tesla Motors implemented an autopilot technology in its electric vehicles, where the cameras and sensors predicted collisions with up to 76% accuracy, leading to a collision prevention rate of over 90%. Google, Tesla Motors, General Motors, Waymo, Uber, nuTonomy, and other automobile companies envision a future with autonomous vehicles in approximately 15–20 years time.

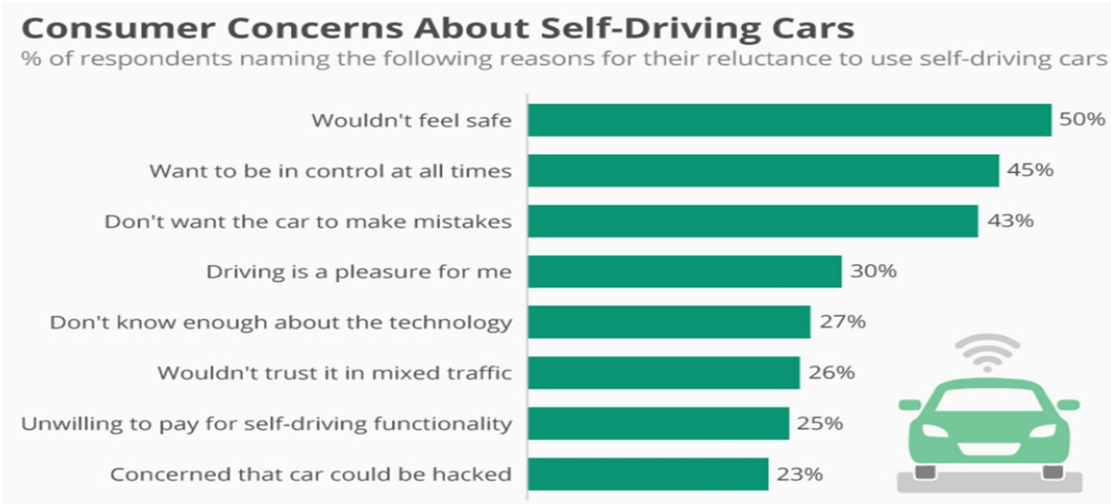

Fig 1: Based on 1260 consumers from 10 countries (Source: Statista.com)

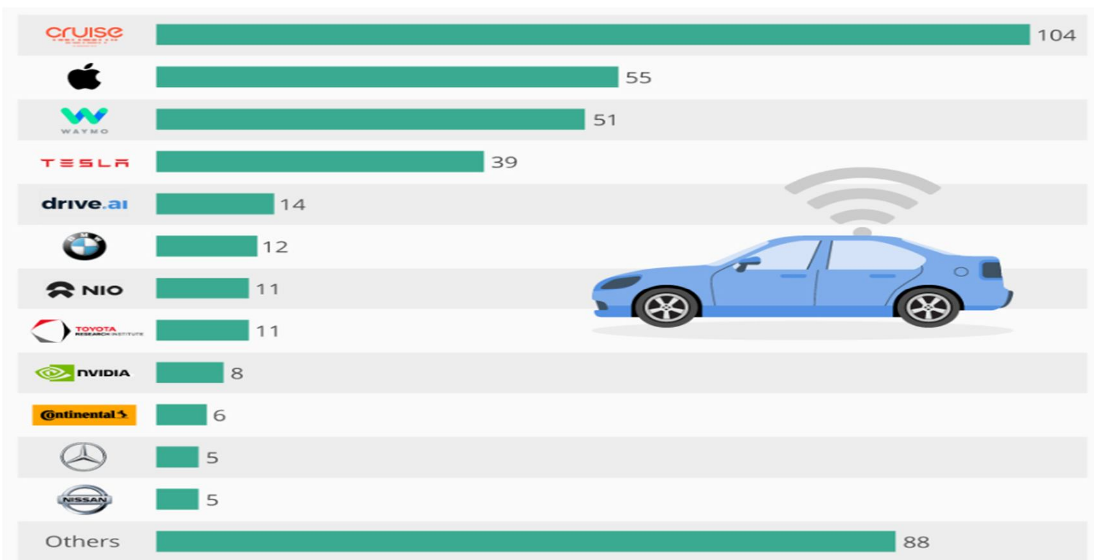

Fig 2: Companies testing self-driving Car in California as of 9th May 2018 (Statista)

Several infrastructure upgrades such as an automated highway system, robotic vehicle cruising management systems, 6G cell-free mobile communication systems with real-time video processing and near-zero latency, are parallel research areas that would contribute to realizing full-fledged autonomous vehicles kick-starting a greener future through autonomous electric vehicles. It is worth mentioning that while the self-driving cars have gained intense attention in the last decade, driver-less transportation has been in existence for over a decade.

Trains are a prominent example of widespread use of self-driving technology. Some of such train examples include the SkyTrain in Vancouver, Canada, Docklands Light Railway (DLR) in London, United Kingdom, Yurikamome in Tokyo, Japan, London Heathrow airport's ultra-pods and many more. These autonomous rail systems transport thousands of passengers on a daily basis. The majority of passengers commuting through self-driving trains are not worried about using those trains. However, the aforementioned trains and autonomous pods operate on enclosed tracks, isolated from the public roads, and bypass the need to interact with other vehicles or pedestrians. In contrast, self-driving cars are set to encounter various users, thereby resulting in complex interactions and the possibility of collisions. Whether people will be as accepting of self-driving cars as they appear to be of existing autonomous transport is an active area of research.

B. Advantages of self-driving cars

- The advances in wireless networking, software-defined networking, and information and communication technology (ICT) have found applications in intelligent transportation systems (ITS) to reduce collisions, reduce pollution, ameliorate mobility issues, provide newer ways of public transportation, and share resources, materials, and space. According to studies, there are 1.3 million deaths every year due to drunk, drugged, distracted, and drowsy driving, which can potentially be saved with the help of autonomous AI systems by eliminating some of the human follies. The following advantages motivate the current research in self-driving cars: For users, the advantages may be reduced stress, faster commutes, reduced travel times, enhanced user productivity, optimum fuel consumption, and reduced carbon emissions. These cars can be programmed to drive defensively, stay clear of blind spots, and follow speed limits.
- For Governments, self-driving cars would assist in traffic enforcement, enhance roadway capacity, reduce road casualties and the number of on-road driving-related accidents, and lead to better observance of speed limits.
- Self-driving cars are envisioned to eliminate drunk driving issues, eliminate issues related to distracted driving, texting and other cell phone use, less braking and accelerating, and less gridlock on highways. Reduced accidents are expected to be beneficial for children and the elderly, encouraging people to feel comfortable and amiable towards self-driving cars.
- Autonomous electric vehicles would introduce a greener mode of transport, leading to less greenhouse and noise pollution, along with increased mobility for the elderly and disabled people. In the current driving landscape, cars are parked for a long time. With self-driving cars, parking lots can be converted to parks and other green infrastructure.
- Self-driving cars would be equipped to improve scheduling and routing, and provide the best routes to improve travel times, while also lowering travel costs. This would also reduce or even eliminate car ownership, they would expand shared access, keep transportation personalized, efficient, and reliable.

C. Disadvantages and drawbacks of self-driving cars

Cars are one of the most widespread and readily available modes of transportation, and while technology has developed safer cars, driving is still a dangerous activity. Self-driving cars formulate a scenario where a few lines of source code, coupled with AI get to decide the life of human beings. Some disadvantages of self-driving cars are outlined as follows: · The foremost catastrophic consequence of self-driving cars would be the elimination of jobs in the transportation industry.

- Although the role of AI in our society is consistently evolving, an AI system making critical decisions need to respect societal values and conform to social norms to gain acceptance. It is argued that in case autonomous vehicles and AI systems malfunction, a person would not die or suffer injuries if they were in control of the system.
- Driving at intersections without traffic lights, malfunctioning traffic lights, uncontrolled intersections, busy intersections, and regions with humans in close proximity are a challenge for self-driving cars. As self-driving cars use the global positioning system (GPS) for localization, they are deemed unsuitable to drive in non-mapped areas.
- The scope of the car's connectivity, the car being online at all times, makes it susceptible to hacking. The safety and convenience offered by self-driving cars might compromise the privacy of passengers as their movements will be tracked and logged.

D. Communication between different entities in self-driving cars

Two well-known collisions mentioned below, involving vehicles operating with a certain degree of autonomous technology, emphasize the benefits of vehicle-to-vehicle (V2V), vehicle-to-infrastructure (V2I), and vehicle-to-everything (V2X) communication in self-driving cars:

- A fatal accident involving a semi tractor-trailer that turned in front of a Tesla car operating on its autopilot program in Florida, caused due to sensors failing to detect a turning vehicle.
- A fatal Uber crash in Arizona.

Investigation and analysis of these accidents indicates that these accidents could have been avoided if the involved vehicles were communicating with each other. The V2V and V2X broadcast a vehicle's current location to nearby traffic, alert traffic to upcoming manoeuvres, traffic jams, accidents and road constructions. A crash, three cars ahead, is too far to be detected by sensors but can easily be communicated over longer distances using V2V. The V2I technology consists of sending traffic light information (TLI) to self-driving cars' acceleration and braking systems, which can assist in planning routes based on the frequency of traffic light changes. The V2V can provide 360-degree road-situation awareness to enhance safety.

Although the user of these techniques requires all vehicles to operate on a standard mode of communication such as dedicated short range communication (DSRC) to relay critical information, a formal policy to mandate DSRC in vehicles is still under development. Scientists, researchers, and experts have historically viewed the lack of computational infrastructure as a major bottleneck that prevents achieving reliable V2V, V2I, and V2X communication. Deploying roadside infrastructure partly mitigated the problem by providing uninterrupted wireless coverage while also improving handover and coverage.

E. Levels of automation: semi-automated, automated and self-driving cars

Autonomy in self-driving cars is based on progression from human-centred autonomy to complete autonomy where all the driving tasks are governed and controlled by the vehicle's AI system, and human interaction is summoned only when necessary. To investigate the capabilities of the present AI systems, Fig. 1 briefly outlines the six levels of vehicular automation defined by the Society of Automotive Engineers (SAE). Level "0" represents the old-fashioned car that is thoroughly controlled by a human driver. Most of the cars today are in this level. With rising SAE level, the share of automated driving functions is increasing. Level "1" and "2" include driving assistance supporting the car driver in certain traffic situations, e.g. automated cruise control, lane assistant, highway assistant. In these technology levels, the car driver has to keep an eye on the driving situation continuously to be able to perform interventions. From level "3" on, the car takes over control in certain traffic conditions, e.g. highway or traffic jam assistant. Here, the driver is able to do other things in the car, e.g. reading, writing or relaxing. But an important issue of highly automated driving functions in level "3" is that the driver has to take over control, if the automated functions are disturbed or not able to handle complex traffic situations. In level "4", fully automated functions provide all required operations of driving control, so that no human interventions are required. The driver has the possibility to take over control and to steer the car by hand in case of system fail functions or personal desire, but in general the car acts as a self-driving device. Finally, level "5" represents the full autonomous vehicle that does not need human interaction anymore. In this level, the cars are robots that transport passengers and/or goods independently.

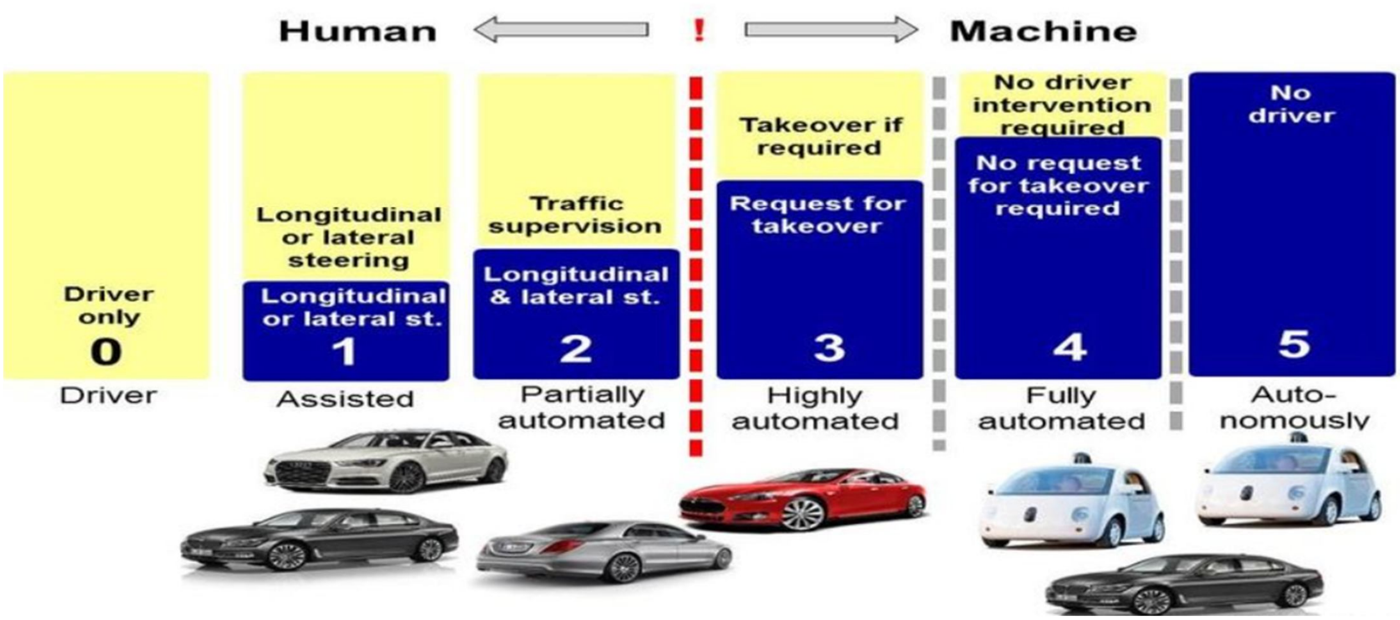

Fig 3: Levels of autonomous cars according to SAE.

III. BIG DATA AND BIG-SENSED DATA FOR SELF-DRIVING CARS

The availability of big data related to self-driving cars facilitates the application of data-driven learning methods to autonomous vehicles. Data emphasizes the role of ML and DL as it is infeasible to craft all possible if-then-else rules that learn all possible situations a self-driving vehicle might encounter in the drive-terrain. Training on data allows self-driving cars to learn by driving, develop efficient inference algorithms, identify patterns in data models and relate complex dependencies in real-time.

A. Role of big data in self-driving cars

Defining a self-driving vehicle problem, formalizing it, collecting sufficient related data on it, and to devise solutions through general purpose AI such as reinforcement and unsupervised learning usually requires raw sensor information and low-level data. Deep learning, however, involves training and testing on labelled data, which can be labelled in case of self-driving cars and annotated by means of ground-truth bounding boxes. By training self-driving cars on these datasets, they are expected to respond to new input data they have never seen before. The self-driving cars need information on their surroundings such as one-way streets, navigation routes, no-entry status, and speech recognition.

B. Collecting big data for self-driving cars

Data collection on public roads is valuable to aid autonomy in self- driving vehicles. The vehicles used for data discovery and collection are equipped with a vast array of sensors and cameras, specifically LiDAR. The data from the sensors are logged on a disk or transmitted to the nearby cloud, which are used to train and test various algorithms for self-driving cars such as vehicle detection, pedestrian detection, or motion estimation. Sensors collect data from external environment, the software analyses the data, and recreates road conditions in three dimensions. Data collection is a long and costly process, and redundancy is avoided by directly exploiting existing datasets as well as collaborating with data collected by other researchers. To facilitate the analysis of ML or DL controlled driving, these datasets vary in terms of traffic conditions, application focus, detector setup, format, size, tool support, and performance aspects.

In self-driving cars, high accuracy is usually defined as 100% accuracy and less than 100% could lead to fatality. Human visual system is the benchmark for self-driving cars to classify objects, perform edge detection, track lanes and expand visibility range. The AI system in self-driving cars must be able to reveal when it does not see some aspects of a scene. The main sources of raw data in self-driving cars are the automotive sensors. Whereas LiDAR is the most powerful camera, it is expensive and researchers argue that images captured using RGB cameras are sufficient for self-driving applications in certain conditions. Research is underway to manufacture low-cost LiDAR. A brief comparison of images generated by the three cameras is presented below –

1) Camera: Cameras are image sensors that operate on RGB values. Cameras capture infrared visual data and offer high resolution information. Cameras can be used as readily available and cheap sensors to capture information that can be learned and inferred to interpret external scene. Human brain uses similar sensor technology, with eyes acting as sensors that work under illumination, and operate in RGB space to detect and segment lane markings, traffic lights, pedestrians, etc. Cameras work well in visible light but their performance degrades in darkness and extreme weather, and are bad at depth estimation
2) RADAR: Highly reliable and provide higher resolution and accuracy. They are ultrasonic, cheap, and work extremely well in extreme weather. However, under low resolution, they are mostly used as automotive sensors for object detection.
3) LiDAR: Although expensive, LiDAR provides extremely accurate depth information, has resolution much higher than RADAR, and provides 360 degrees visibility. LiDAR has been a successful source of 3D ground truth data in driving environment.

In summary, the need for annotated data increases the demands on camera technology. Although cameras operate in clear, well lit conditions, good visibility requirements over a long range in dark conditions, heavy rain, snow, and fog makes CV and image processing a critical and open research problem. A novel technique suggests cameras, RADAR, and LiDAR be integrated together through sensor fusion, where DL methods can be used to interpret the spatial-visual characteristics to understand, interpret and track the dynamics of the environment.

C. Multimodal sensor fusion in autonomous vehicles

Multimodal sensor fusion entails the combination of information from a set of different types of sensors. Exploiting complementary information from different sensors, shows that target detection and classification problems can greatly benefit from this fusion approach and result in a performance increase. Sensor modalities allow reconstruction of images for regularization and feature-based reconstruction on data from multiple sources and sensors, where each modality provides significant knowledge and valuable information pertaining to the object of interest.

This enables driving under uncertainty through dynamic sampling, characterization and image denoising, deblurring and segmentation. Multimodal sensor fusion is widely used to ensure robustness of data acquired from different sensors in different driving scenarios and cross modality generalization, and to approximate missing data through correlation between different available modalities. Once the data is ready, the next step is to choose a DL model and architecture.

IV. DEEP LEARNING ARCHITECTURES FOR PERCEPTION IN AUTONOMOUS VEHICLES

To achieve object-detection, cognition and scene perception, self-driving cars are expected to perceive surroundings in a way at least similar to the way human eye processes information leading to cognitive AI systems that can learn, relearn, and take action. To achieve human level driving from CV perspective, self-driving cars need to be able to recognize environment, interpret 3D representation of world, to discern the movement of objects, pedestrians, and other cars, and deal with human emotions. The feasibility of DL is being established as some state-of-the-art results have been achieved by Google cars and Uber cars in maps-based localization, that were trained to drive with little prior knowledge of the roads. These vehicles use DL for path planning, obstacle avoidance, and try to process camera based information to solve complex CV problems. While DL algorithms learn effective perception control from data, LiDAR costs and the expenses involved in manually annotating the maps restrict the application of DL to autonomous driving. DL and ML techniques can be classified into three categories; supervised, unsupervised, and reinforcement learning. The mathematical framework to implement these DL techniques involves the following steps:

1) Backpropagation: the primary method of learning. Calculates error, computes error function and tries to minimize error function in sub sequent forward passes.
2) Use gradient descent to back propagate the error function.
3) Subtract a fraction of the gradient from the weight. · Recalculate the weights responsible for making a correct or incorrect decision.
4) The objective is to minimize the error function, by updating the weights using minibatch or stochastic gradient descent.
5) More data and very large networks lead to too many parameters, and increased training times.

A. Deep learning architectures for object detection and computer vision in autonomous vehicles

1) Convolutional Neural Networks

Convolutional neural networks (CNN) have been extensively applied to image classification and computer vision, and have returned 100% classification rates on datasets such as ImageNet. In CNN architecture, successive layers of neurons learn progressively complex features in a supervised way by back-propagating classification errors, with the last layer representing output image categories. CNNs do not use a distinct feature extraction module or a classification module, i.e. CNNs do not have an unsupervised pre-training and the input representation is implicitly through supervised training but eliminate the need for manual feature description and feature extraction. It extracts features from raw data based on pixel values leading to final object categories. The CNNs explore the nature of AI and the role of artificially intelligent systems in the society, as full autonomy in self-driving cars involves creating intelligence. The inputs to CNN can be images, video, text, audio depending on different data with one-to-one, one-to-many, or many-to-many relation between input data and output classes. Depth in CNNs is provided by the number of layers, and is analogous to features, taking all the features generated by filters of different sizes and using backpropagation to arrive at the best features. Novel CNN architectures are designed using transfer learning where a series of predefined convolutions are followed by series of fully connected layers, without having to train CNNs from scratch. In a CNN, a driving image is convolved with activating functions to obtain feature maps, which can be further scaled down to identify patterns in an image or signal.

Each layer in the CNN finds successively complex features where the first layer finds a small, simple feature anywhere on the image, the second layer finds more complex features and so on. At the last layer, these feature maps are processed using fully connected neural networks (FCNN). In addition to reducing the driver's responsibilities and assist them in critical tasks, the end result envisioned is to eliminate the active need for driver engagement, which extends much beyond currently available semi-autonomous models with ADAS (Advanced Driver Assistance System).

2) Recurrent Neural Networks

Recurrent neural networks (RNNs) are a powerful and robust type of neural network, and belong to the most promising algorithms in use because it is the only one with an internal memory. Because of their internal memory, RNN's can remember important things about the input they received, which allows them to be very precise in predicting what's coming next. Recurrent neural networks recognize sequences and patterns in structures consisting recurrent computations to sequentially process the input data.

Long short-term memory (LSTM) is an RNN based method which uses feedback connections for sequences and patterns recognition using input, output, and forget gates. LSTM remembers the output computed from the previous time step, and provides output based on the current input. The connection between units' forms a directed cycle and the RNN input and output are related as the edges of RNN feed output from one time step into the next time step. RNNs have been applied for robust and accurate visual tracking in autonomous vehicles in constrained scenarios. Temporal correlation of RNNs predicts the object at next time frame based on region of interest (ROI) and uses that image as the input of the next frame, resembling a prediction model for object tracking.

3) *Deep Belief Networks*

A deep belief network (DBN) is a hybrid multi-layered, generative graphical model used for learning robust features from high dimensional data and consists multiple layers of stochastic, latent (hidden) variables connected between the neurons of different layers, but not between the units of each layer. The undirected Restricted Boltzmann Machines (RBM) where each layer trains separately to produce an expected input are the building blocks of DBN architecture. Each RBM layer communicates with the previous and next layers for accuracy and computational efficiency. DBN handles non convex objective functions and local minima through multiple layers of latent variables where a hidden layer acts as the visible input layer for the adjacent layer.

4) *Stacked Autoencoders*

Autoencoders are a class of unsupervised feature extractors that find a generalized transformation of the input and assist a classifier in a supervised task. Stacked autoencoders (SAE) are used in autonomous vehicles vision systems to visualise high-dimensional data to find clusters and to create similarity metrics between samples. SAE are sued to reduce the dimensions of the input image data captured using LiDAR sensors in self-driving vehicles. Dimension reduction avoids learning the identity function without any explicit changes to the driving accuracy and gives smaller reconstruction errors. SAE restrict the driving scene output to be sparse, imposing a sparsity constraint. SAE add random noise to the input, requiring a reconstruction of the original input. This forces the driving vision system to learn the structure of the input distribution to undo the effects of the added noise, which makes the system more robust to small changes of the input.

The learned features of the auto encoder are tolerant to the changes in the input space. DBN and SAE assist in building a non-linear, distributed representation of the input, where DBN captures the representation in stochastic distribution form, while the SAE learns a direct mapping of the input to another space.

5) *Singleshot Multibox Detection*

A previous technique known as sliding window utilizes CNNs for better image detection. In realistic self-driving scenarios, some images might have zero objects, and some might have up to 50 objects. The output of a CNN is still going to be a fixed set of numbers that the processing infrastructure or the vehicle AI system needs to interpret. Single shot multibox detection (SSD) is largely viewed as a milestone in digital image processing (DIP) and CV research to enhance real-time performance requirements. Self-driving cars have to recognize objects as soon as they see them, which is a CV problem as well as a security requirement in self-driving cars. Singleshot multibox detection improves both speed and accuracy by looking for the presence of an object, and its location. If an image has multiple class instances such as people, cars, trucks, bicycles, people, traffic lights, traffic signs, landmarks, lane markers in a single image, in such settings, image classification has limited applications. This leads to research into object localization, which tells a vehicle whether an object is present in an image and its location. This is done using bounding boxes or rectangles, and depending on applications these can be ellipses, facial key points known as landmarks, or fingerprints and retina structures. Comparable techniques are you only look once (YOLO) and region-CNN (R–CNN) and have been out performed by SSD.

B. *Deep learning libraries*

Modern deep learning libraries such as Theano, PyTorch, TensorFlow, and Keras make designing neural networks easier. TensorFlow allows for plug-and-play script. As training neural networks takes long time, ranging from days to weeks and months, these DL libraries make use of GPUs that speed up matrix multiplications and other mathematical operations by orders of magnitude. Utilizing these DL libraries in conjunction with GPU and vehicular cloud for real-time CV in self-driving cars and carrying out the task is an active area of research.

V. CONCLUSION

In this paper, we reviewed and studied the recent trends and developments in deep learning for computer vision, specifically vision, object detection, and scene perception for self-driving cars. The analysis of prevailing deep learning architectures, frameworks and models revealed that CNN and a combination of RNN and CNN is currently the most applied technique for object detection due to remarkable ability of CNNs to function as feature extractors. The CNNs can learn subtle patterns in an image, and are robust to translational and rotational variations. We outlined the ongoing initiatives taken by researchers to test self-driving cars and emphasized the role of DL in real-time object detection. With GPU and cloud based fast computation, DL could process captured information in real-time and communicate it to nearby cloud and other vehicles in the meaningful vicinity. The study also revealed that in order to improve performance metrics such as accuracy, precision, recall, and F1 scores, and transfer learning is used to enhance accuracy of object detection. In this survey, we focused on the recent advancements in CNNs that are principally used for images. In self-driving cars, CNN dependent strategies still need to be fine-tuned so as to achieve the precision level of human eye. The findings reported that although DL is a key catalyst to realize object detection and scene understanding in self-driving cars, there is a huge scope for additional advancements. It is yet to be investigated that when and under what conditions CNNs cease to perform well and can pose a threat to human life in self-driving scenarios.

The artificial driving intelligence is still incapable to annotate and categorize driving environment on its own, without need for human assistance. Also, much of the earlier tests conducted on autonomous driving were predominantly on open roads and good weather, but more recent tests include weather conditions such as driving in fog, adverse weather events, or snow. Limited exposure of the self-driving LiDAR cameras has been enhanced using multimodal sensor fusion and point cloud analysis for object classification. The findings of the survey summarize that self-driving cars are no longer a question of if but more of when and how. The penetration rate of these autonomous robots into human society depends on their ability to drive safely. This puts forth a critical need for reliable object detection techniques, mathematical models and simulations to mimic reality and arrive at best parameters and configurations that can adapt with changes in surroundings. Nevertheless, with big data, DL and CNNs, we have tools at our disposal that can achieve high levels of arbitrary accuracy to solve perception problems in self-driving cars. These tools have provided researchers with the ability to break complex problems into easier ones and previously impossible problems into solvable but slightly expensive ones such as capturing and annotating data to create ground truth.

REFERENCES

[1] "Deep learning for object detection and scene perception in self-driving cars: Survey, challenges, and open issues" Abhishek Gupta, Alagan Anpalagan, Ling Guan, Ahmed Shaharyar Khwaja Ryerson University, 350 Victoria Street, Toronto, M5B2K3, Ontario, Canada
[2] "Self-Driving Car A Deep-Learning Approach", Rutvik Shah, Mit Patel, Mayank Budhwani,
[3] "Sensor and object recognition technologies for self-driving cars" Mario Hirz [0000-0002-4502-4255]1, Bernhard Walzel [0000-0003- 4818-3726]2
[4] "A Review Of Deep Learning For Self-Driving Cars: Case Study", Joel Habib, Afrim
[5] "Deep Learning Techniques for Obstacle Detectionand Avoidance in Driverless Cars", Nischal Sanil, Pasumarthy Ankith Naga venkat, Rakesh V, Rishab Mallapur, Mohammed Riyaz AhmedSchool of Electronics and Communication EngineeringREVA University Bangalore India-64
[6] Pande, S. P., & Khandelwal, S. (2023). Scene Detection Classification and Tracking for Self-Driven Vehicle. International Journal on Recent and Innovation Trends in Computing and Communication, 11(7), 681–690. https://doi.org/10.17762/ijritcc.v11i7s.7529
[7] Deore, H., Agrawal, A., Jaglan, V., Nagpal, P., & Sharma, M. M. (2020). A new approach for navigation and traffic signs indication using map integrated augmented reality for self-driving cars. Scalable Computing, 21(3), 441–450. https://doi.org/10.12694:/scpe.v21i3.1742
[8] Zheng, X., & Su, W. (2023). Algorithm Design of Self-driving Vehicle Based on Visual Sensing Technology. In Proceedings - 2023 International Conference on Mechatronics, IoT and Industrial Informatics, ICMIII 2023 (pp. 432–436). Institute of Electrical and Electronics Engineers Inc. https://doi.org/10.1109/ICMIII58949.2023.00091
[9] Basit, A., Ejaz, M. U., Ayaz, Q., & Malik, F. M. (2023). Real-time object detection and 3D scene perception in self-driving cars. In 3rd IEEE International Conference on Artificial Intelligence, ICAI 2023 (pp. 109–115). Institute of Electrical and Electronics Engineers Inc.