A Review of Computer Vision and Algorithmic Advancements in Autonomous Vehicle Technology

Saad Shalbi and Ali Aboria

June 23, 2023

A Review of Computer Vision and Algorithmic Advancements in Autonomous Vehicle Technology

Saad Shalbi[1] **Ali Aboria**[1]

[1]*Student at Faculty of Artifical Intelligence, Kafr elsheikh university, Kafr elsheikh, Egypt*

Abstract: Self-driving cars are emerging fast and will be a bright reality soon. For large companies and companies willing to search, years of integrated and comprehensive lists have been devoted. This is the ability to work in the future. While the utopian dreamer searches for the path to this imagined future. As driving has evolved dramatically, the development of computer vision algorithms and combined driving has led to the creation of comprehensive teams to train this information. In this review, we provide a brief look at research trends in the intersectional fields of autonomous vehicles and computer vision. In addition, advanced algorithms for solving common challenges faced by autonomous systems, including object detection and semantic landscape segmentation, are discussed.

Keywords: *objects tracking, scene segmentation, and human interaction.*

I. Introduction

In the current era, cars have become a widespread means of transportation because of their benefits in saving time and effort. The two ranks in cars, plus car fuel costs These problems can be addressed through the use of feedback in the self-driving car industry. Cars can be saved in the surrounding night hours by night security. Self-driving cars also have predictive driving capabilities, which make for more casual, smoother rides. The following events can be predicted: It helps with better traffic management, which helps with road congestion. Intelligence can identify the best signs and avoid traffic jams. In addition, ease of fuel consumption and reduced vehicle costs.

II. Self-Driving Cars Algorithms

Self-driving cars rely on a wide range of algorithms to achieve their efficiency.

The main algorithms it relies on:

1. Deep Neural Networks: used to recognise images and analyse sensory data to make driving decisions.
2. Reinforcement Learning: It is used to train models to make decisions in a specific environment by reinforcing correct behaviour and reducing unwanted behaviour.
3. Deep Learning: It is used to train models to make decisions based on the expected return value of different actions in the driving environment.
4. Transfer Learning: allows knowledge and experience gained in previous learning environments to be transferred to new tasks, accelerating the learning process in an autonomous context.
5. Classification and Optimization Algorithms: They are used to learn about various elements in the driving environment, such as identifying obstacles and other vehicles and optimizing the vehicle's trajectory.
6. Computer Vision Algorithms:. help in recognizing images and videos and extracting information necessary for driving, such as determining lanes and reading traffic signs.
7. Natural Language Processing (NLP) technologies: These are used to understand voice instructions and communicate with passengers or other systems.

Together, these algorithms enable self-driving cars to perceive their surroundings, make decisions, and control their movements in real time. In this paper, we will talk about computer vision algorithms.

III. Computer Vision in Self-driving Cars

An autonomous vehicle needs to navigate to its intended destination without relying on external guidance. This entails the vehicle's ability to independently navigate the lane while avoiding obstacles. To achieve this, autonomous vehicles use sensors such as cameras, radar, and lidar to perceive their surroundings and develop an understanding of the potential actions of various elements within the environment. Computer vision technologies play an important role in this process. In this section, we highlight research related to computer vision in the field of self-driving cars. A. Object detection in previous attempts to solve computer vision problems, manual feature engineering techniques were commonly used. These techniques included Haar Wavelets [1], Histogram of Gradient Descent [2], Scale Invariant Feature Transform [3], and others. The HOG features, introduced in [1], include extracting features from images and using an SVM classifier for pedestrian detection. Similarly, for traffic light detection, methods such as top-hat morphology and template matching processes have been explored [4–6]. OverFeat [7] modified the CNN model by replacing the fully connected layer with a regression layer, which is trained to output bounding box coordinates. However, the R-CNN method presented in [8] is superior to the OverFeat method. It involved creating region proposals, extracting feature vectors for each region using CNN, and then classifying the region as containing or not containing an object using SVM classifiers. Modifications have been developed on R-CNN algorithms, such as Fast R-CNN [9][2][19][18] and Faster R-CNN [10] and have shown significant improvements in performance and speed compared to the original R-CNN. Object detection is mainly focused on images captured by cameras. However, data from other sensors such as depth maps, optical flow, thermal images, etc. can also be used to produce results in different conditions [11][17][16][18].

A. *Object detection*

In earlier attempts to solve computer vision problems, manual feature engineering techniques were commonly used. These techniques included Haar Wavelets [1], Histogram of Gradient Descent [2], Scale Invariant Feature Transform [3], and others. HOG features, introduced in [1], involved extracting features from images and using an SVM classifier to detect pedestrians. Similarly, for traffic light detection, methods such as top-hat morphology operations and template matching [4-6] have been explored.

OverFeat [7] modified the CNN model by replacing the fully-connected layer with a regression layer, which was trained to output bounding box coordinates. However, the R-CNN method introduced in [8] outperformed the OverFeat method. It involved generating region proposals, extracting feature vectors for each region using a CNN, and then classifying the region as containing or not containing an object using SVM classifiers. Modifications to the R-CNN algorithms, such as Fast R-CNN [9] and Faster R-CNN [10], were developed and demonstrated significantly improved performance and speed compared to the original R-CNN.

Object detection has mainly focused on images captured by cameras. However, data from other sensors like depth maps, optical flow, thermal images, and more can also be utilized to achieve results in different conditions [11].

B. *Semantic landscape segmentation*

- Image segmentation is a computer vision task that assigns predefined classes to each pixel in an image. Semantic image segmentation specifically categorizes pixels based on the object classes they belong to. Various algorithms have been employed for image segmentation, such as K-Mean clustering and Expectation Maximization, treating it as a pixel clustering problem. Conditional Random Field (CRF) algorithms have also been widely utilized in this context.
- Convolutional Neural Networks (CNNs) have played a significant role in semantic scene segmentation. Some approaches combine superpixels with CNNs for feature extraction, while others employ fully connected CNN models to map image pixels to labels. The combination of CRF and CNN methods has proven beneficial, as seen in methods where a CRF is connected to the output of a CNN layer and both layers are trained separately to obtain segmentation results. Another approach involves integrating CRF into a Recurrent Neural Network (RNN), resulting in improved performance compared to previous methods [19][20].

- Mask R-CNN, an advancement over traditional R-CNN, has demonstrated higher accuracy in object detection and pixel-level semantic segmentation. It outperforms other algorithms not only in bounding box localization but also in semantic segmentation tasks [21][22].

C. *Contro*

In this module, we develop a control system to guide our car along a planned trajectory. While humans find driving along a predetermined path intuitive, our system must consider the future consequences of our current actions. For instance, failing to decelerate before a turn could result in a collision during the maneuver.

Our self-driving car utilizes front steering, gearless transmission, and acceleration (throttle) and steering angle as actuators. The effectiveness of the model predictive controller (MPC) for active steering control is confirmed in [15]. To implement MPC, we adopt the kinematic bicycle model described in [16]. The model is optimized by minimizing various costs, including cross track error (CTE), deviation from the reference velocity, and abrupt changes in acceleration and steering angle [23][24][25].

D. *Lane line Detection*

Lane line detection is a fundamental task in computer vision and autonomous driving systems aimed at identifying and tracking the lane boundaries on a road. Accurate detection of lane lines is crucial for tasks such as lane keeping, lane departure warning, and autonomous navigation.

Lane line detection algorithms typically operate on images or video frames captured by cameras mounted on vehicles. The process involves several steps:

1. Preprocessing: The input image is often preprocessed to enhance relevant features and reduce noise. Common techniques include color space conversion, gradient computation, and image filtering.

2. Feature Extraction: Lane line features are extracted from the preprocessed image. Common features include edges, color gradients, and texture information. These features help distinguish lane lines from the surrounding road environment.

3. Region of Interest (ROI) Selection: A specific region of the image, typically corresponding to the road area ahead, is selected as the ROI for lane line detection. This helps improve efficiency and reduces false detections in irrelevant areas.

4. Line Detection: Various techniques can be used to detect lane lines within the ROI. Popular approaches include the Hough Transform, which identifies line segments, and the use of probabilistic models like the Random Sample Consensus (RANSAC) algorithm to fit lines to the detected segments.

5. Line Tracking and Filtering: Detected lane lines are often tracked over time to provide smoother and more robust estimates. Filtering techniques such as Kalman filters or moving averages are commonly used to reduce noise and improve line stability.

6. Lane Modeling: Detected lane lines can be used to model the lane geometry, such as lane curvature and width. This information is valuable for advanced driver assistance systems and autonomous driving applications.

Lane line detection algorithms can vary in complexity and performance depending on factors such as lighting conditions, road markings, and environmental challenges like occlusions or shadows. Additionally, machine learning approaches, such as convolutional neural networks (CNNs), have been increasingly employed for lane line detection, offering improved accuracy and generalization capabilities.

Overall, accurate and reliable lane line detection is essential for ensuring the safety and effectiveness of autonomous vehicles and advanced driver assistance systems.

- Levels of automation and computer vision

Different levels of automation for self-driving cars have been determined based on the extent of human driver involvement with the systems. While achieving full Level 5 automation may still be a long way off, there have been significant advances in lower-level autonomous systems that require human attention and input during operation.

Advanced Driver Assistance Systems (ADAS), such as Advanced Cruise Control and Lane Keeping Assist, allow drivers to relinquish control of basic driving tasks while remaining behind the wheel. However, this can divert drivers' attention away from the road when the autonomous system is in control. To address this problem, the systems are designed to detect drivers' attention and alertness while driving, which is critical to ADAS and simpler systems that prevent accidents caused by driver inattention.

Human attitude estimation can be used in both semi-autonomous and ADAS systems, providing insight into the driver's current activity, hand position on the steering wheel, and head orientation towards the road. Situation recognition can be performed using 2D or 3D data, with the latter offering advantages in overcoming challenges such as overlap and variability in lighting conditions.

Autonomous systems could also benefit from analysing drivers' reactions to their performance. For example, a feature-based vector graph (HOG) method can be used to extract facial features such as eyes, nose, mouth, and eyebrows. By evaluating these features, it is possible to determine how satisfied or dissatisfied the driver is with the performance of the assistance system.

Overall, the integration of human drivers and computer vision technologies is critical to enhancing the safety and effectiveness of autonomous driving systems. By understanding and leveraging human capabilities, autonomous vehicles can be developed to work in unison with human drivers, ultimately advancing the future of transportation.

V. Conclusions

The Currently, there are already existing autonomous system solutions, and some are being prepared for deployment on roads in the near future. However, most of these systems are categorized as level 2 or level 3 automation systems. Achieving level 5 autonomy, which represents fully autonomous driving, will take a significant amount of time. It's important to recognize that human capabilities have evolved gradually over many of years to reach our current standards. Similarly, autonomous systems will also require extensive training and testing data to become as efficient as humans in driving and other tasks.

In the current trends, human attention is still required, and we have discussed how self-driving cars can benefit from interacting with human drivers. Through these interactions, both humans and self-driving cars can coexist and develop harmoniously, leveraging the strengths and capabilities of each.

References

[1] Oren, M., Papageorgiou, C., Sinha, P., Osuna, E., Poggio, T.: Pedestrian detection using wavelet templates. In: Proceedings of the 1997 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, IEEE, 17 Jun 1997, pp. 193–199.

[2] Hassan E, El-Rashidy N, Talaat FM (2022) Review: Mask R-CNN Models. https://doi.org/10.21608/njccs.2022.280047.

[3] E. Hassan, M. Y. Shams, N. A. Hikal and S. Elmougy, “A novel convolutional neural network model for malaria cell images classification,” Computers, Materials & Continua, vol. 72, no. 3, pp. 5889–5907, 2022.

[4] Jensen, M.B., Philipsen, M.P., Mogelmose, A., Moeslund, T.B., Trivedi, M.M.: Vision for looking at traffic lights: issues, survey, and perspectives. In: IEEE Transactions on Intelligent Transportation Systems (2015).

[5] Agarwal, N., Sharma, A., Chang, J.R.: Real-time traffic light signal recognition system for a self-driving car. In: International Symposium on Signal Processing and Intelligent Recognition Systems 13 Sep 2017, pp. 276–284. Springer, Cham.

[6] Talaat, Fatma M., and Esraa Hassan. "Artificial Intelligence in 3D Printing." Enabling Machine Learning Applications in Data Science: Proceedings of Arab Conference for Emerging Technologies 2020. Springer Singapore, 2021.

[7] Hassan, Esraa, et al. "Enhanced Deep Learning Model for Classification of Retinal Optical Coherence Tomography Images." Sensors 23.12 (2023): 5393.

[8] Sermanet, P., Eigen, D., Zhang, X., Mathieu, M., Fergus, R., LeCun, Y.: Overfeat: Integrated Recognition, Localization and Detection Using Convolutional Networks. arXiv preprint.

[9] Girshick, R., Donahue, J., Darrell, T., Malik, J.: Rich feature hierarchies for accurate object detection and semantic segmentation. In: Proceedings of the IEEE Conference on

Computer Vision and Pattern Recognition 2014, pp. 580–587.

[10] Gamel, S.A., Hassan, E., El-Rashidy, N. et al. Exploring the effects of pandemics on transportation through correlations and deep learning techniques. Multimed Tools Appl (2023). https://doi.org/10.1007/s11042-023-15803-1

[11] . Ren, S., He, K., Girshick, R., Sun, J.: Faster r-cnn: towards real-time object detection with region proposal networks. In: Advances in Neural Information Processing Systems, pp. 91– 99 (2015).

[12] Pop, D.O., Rogozan, A., Nashashibi, F., Bensrhair, A.: Pedestrian recognition through different cross-modality deep learning methods. In: IEEE International Conference on Vehicular Electronics and Safety, Jun 2017, Vienna, Austria.

[13] Hassan, Esraa, et al. "The effect of choosing optimizer algorithms to improve computer vision tasks: a comparative study." Multimedia Tools and Applications (2022): 1-43.

[14] Hassan, Esraa, et al. "COVID-19 diagnosis-based deep learning approaches for COVIDx dataset: A preliminary survey." Artificial Intelligence for Disease Diagnosis and Prognosis in Smart Healthcare (2023): 107.

[15] Dalal, N., Triggs, B.: Histograms of oriented gradients for human detection. In: IEEE Computer Society Conference on Computer Vision and Pattern Recognition. CVPR 2005, IEEE, 25 Jun 2005, vol. 1, pp. 886–893.

[16] Lowe, D.G.: Object recognition from local scale-invariant features. In: The Proceedings of the Seventh IEEE International Conference on Computer Vision, IEEE, 1999, vol. 2.

[17] Shi, Z., Zou, Z., Zhang, C.: Real-time traffic light detection with adaptive background suppression filter. IEEE Trans. Intell. Transp. Syst. 17(3), 690–700 (2016).

[18] Girshick, R.: Fast r-cnn (30 2015 Apr). arXiv preprint arXiv:1504.08083.

[19] Battina, Dhaya Sindhu. "Artificial intelligence in software test automation: a systematic literature review." International Journal of Emerging Technologies and Innovative Research (www. jetir. org| UGC and issn Approved), ISSN (2019): 2349-5162.

[20] Kumar, Boggula Varun, et al. "Artificial Intelligence Based Algorithms for Driver Distraction Detection: A Review." 2021 6th International Conference on Signal Processing, Computing and Control (ISPCC). IEEE, 2021.

[21] Chavarro, Javier Mauricio Morales. "Analyzing the Effect of Deceiving Agents in a System of Self-Driving Cars at an Intersection-a Computational Model."

[22] Chavarro, Javier Mauricio Morales. "Analyzing the Effect of Deceiving Agents in a System of Self-Driving Cars at an Intersection-a Computational Model."

[23] Mijwil, Maad M., Guma Ali, and Emre Sadıkoğlu. "The Evolving Role of Artificial Intelligence in the Future of Distance Learning: Exploring the Next Frontier." Mesopotamian Journal of Computer Science 2023 (2023): 98-105.

[24] Kowlgi, Gurukripa N., Fatima M. Ezzeddine, and Suraj Kapa. "Artificial intelligence applications to improve risk prediction tools in electrophysiology." Current Cardiovascular Risk Reports 14 (2020): 1-9.

[25] Kowlgi, Gurukripa N., Fatima M. Ezzeddine, and Suraj Kapa. "Artificial intelligence applications to improve risk prediction tools in electrophysiology." Current Cardiovascular Risk Reports 14 (2020): 1-9.