Proceedings of the 2023 International Conference on Machine Learning and Automation
DOI: 10.54254/2755-2721/33/20230264

Application of deep learning in image processing of unmanned vehicles

Tao Qu

School of Physics and Electronic Science, Chongqing University of Posts and Telecommunications, Chongqing, China

2019210825@stu.cqupt.edu.cn

Abstract. Artificial Intelligence has accelerated research on autonomous vehicles in the past few years. However, it is not easy to achieve full autonomy on account of the essence of a complex and dynamic driving environment. Fortunately, the rapid development of deep learning has made great progress, which can be used to solve problems about image processing in the autonomous vehicles field. This paper offers a comprehensive review of the recent deep-learning-based image processing methods that leverage data detected. The paper gives a brief overview of deep learning, discussing basic concepts and principles. We also discuss the common models and architectures of deep learning and introduce typical deep learning techniques: convolutional neural networks (CNN). Furthermore, we divided the application of deep learning in the autonomous vehicles into three parts and discuss them respectively. Finally, we review the disadvantages of the application of deep learning in image processing for autonomous vehicles. On this basis, we put forward our insights and point out promising research directions.

Keywords: autonomous vehicles, deep learning, convolutional neural networks, image processing.

1. Introduction

The number of casualties brought on by traffic congestion and accidents is rising along with the number of vehicles over time. When compared to 2018, the number of fatalities brought on by traffic accidents increased by 6.4% in 2019. Among them, the death toll caused by improper operation of the driver accounted for 60% and the main causes of traffic accidents are driver's fatigue driving, speeding, drunk driving, etc. [1]. Therefore, the concept of autonomous vehicles was proposed.

In order to execute the driving duty without the assistance of a person, autonomous driving refers to a vehicle that uses sensors to detect its surroundings and instantly alters its driving style. Autonomous vehicles have a broader environmental perception ability than humans and can quickly make a safe response to potential crises, which can fundamentally preventing traffic accidents brought on by human causes such as dangerous driving, distracted driving, and fatigue driving, and greatly reduce the traffic accident rate [2].

In recent years, the rapid development of Graphics Processing Unit (GPU) and the proposal of convolutional neural networks (CNN) have made the field of deep learning show explosive development [3]. Deep learning shows a strong ability to extract complex features from high-

© 2023 The Authors. This is an open access article distributed under the terms of the Creative Commons Attribution License 4.0 (https://creativecommons.org/licenses/by/4.0/).

dimensional data and expand data, and has achieved good results in many fields, including image processing, which plays a vital role in the process of realizing autonomous driving [4]. In order to accurately perceive the surrounding environment and identify roads, traffic signs, obstacles, etc., the deep learning method is used to make intelligent decisions to ensure driving safety and efficiency in the autonomous vehicles field. In terms of image processing, the application of computer vision to unmanned visual perception systems can be divided into three parts: target detection, image classification, and image segmentation.

This paper will focus on two components of autonomous vehicles: deep learning and image processing. The main aim is to provide a comprehensive review of the application of deep learning in the autonomous vehicles field. The paper is organized as follows.

The paper first introduces deep learning, including basic concepts, structure, principles of deep learning, and details of common deep learning models and architectures, and shows that it played a vital role in autonomous vehicles. Then, the paper discusses the application of deep learning in autonomous vehicles. Considering different types of the application, detection and obstacles avoidance or the detection and semantic segmentation of traffic signs, it can be divided into three parts to explain: objection identification, image segmentation, and target classification.

2. Application of deep learning in image processing of autonomous vehicles

2.1. Deep learning

2.1.1. Basic concepts and principles of deep learning.

Deep learning is a branch of machine learning, a neural network with three or more layers. To learn and extract high-level data features, deep learning works by simulating a brain nervous system with rich hierarchy and establishing a hierarchical model structure similar to the human brain. It processes multi-layer nonlinear information to accomplish supervised or unsupervised feature extraction, transformation, analysis and classification, commonly used for images, sounds and texts in the autonomous vehicle field.

There are three components in deep neural networks: the input layer, the hidden layer and the output layer, as shown in Figure 1.

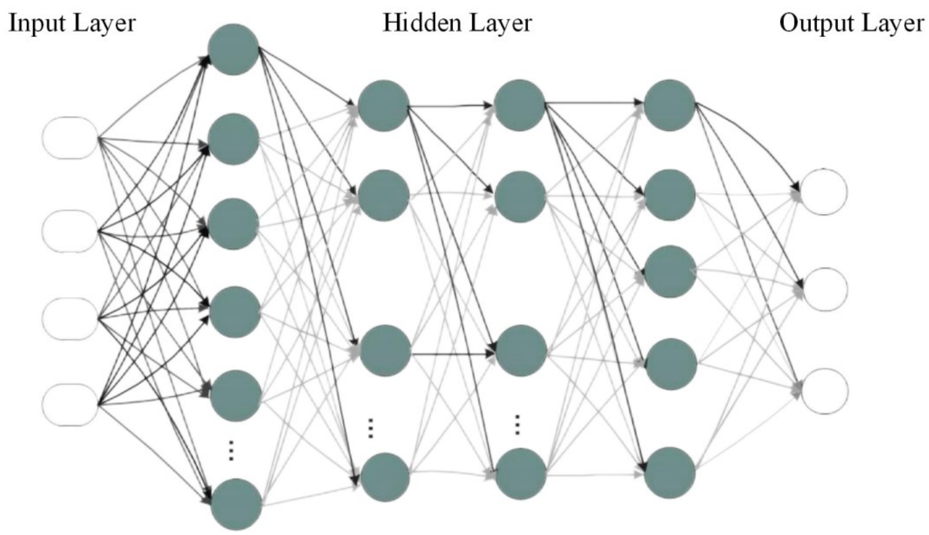

Figure 1. Deep learning diagram [5].

The input layer, which consists of several nodes that receive external data, is the first layer of the neural network. The hidden layer, between the input and output layers is responsible for feature extraction and transformation. Because there usually are multiple hidden layers in the neural network, that's where the conception of "deep" comes from. The output layer is the last part of the neural network, which takes charge of the output prediction and the input classification and regression.

Due to simple nonlinear objective function limitations, shallow abstraction can not identify some complex problems effectively. The essence of deep learning is to build up a variety of hidden layers and intricate network structure models, add a lot of training samples, and then alter the weights in the deep network to produce a high-dimensional feature representation.

Deep learning promotes the development of applications and services of Artificial Intelligence(AI), which can perform analytical and physical tasks without expert intervention. Deep learning is hidden behind our common products and services (such as speech recognition, digital assistants, and image processing) and emerging technologies (such as self-driving cars).

With continuous research and innovation in deep learning by researchers, deep learning has been widely displayed in autonomous vehicles, and its high accuracy has promoted the development of many core fields, such as target detection, decision making and sensor application. In terms of image processing, computer vision application to unmanned visual perception systems can be divided into three parts: target detection, image classification, and image segmentation.

2.1.2. Common deep learning models and architectures.

Deep learning techniques, typical of which are convolutional neural networks (CNN), are widely used in all kinds of image processing and are very suitable for autonomous driving. Here we will introduce CCN in brief.

According to Figure 2, the typical CCN model comprises a convolutional layer, pooling layer and fully connected layer, enabling it to utilize the two-dimensional structure of input data. CCN have shown excellent results in image and speech applications and can also be trained using standard backpropagation algorithms. CCN are easier to train than other deep structures because of fewer parameter estimates.

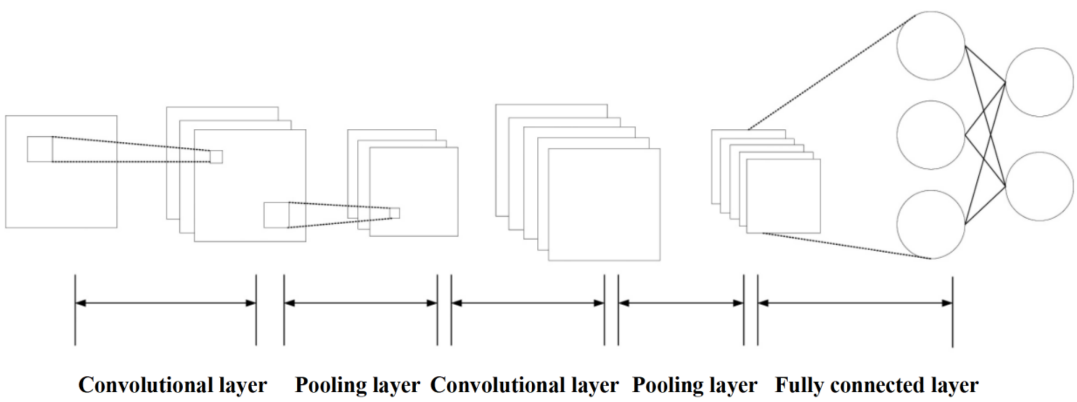

Figure 2. Typical convolutional neural network model [6].

2.2. Application of deep learning in autonomous vehicle

With the excellent performance of deep learning in image recognition, CNN have been widely used in image processing related to autonomous vehicles, such as the detection and obstacle avoidance of display obstacles or the detection and semantic segmentation of traffic signs. Convolutional neural network algorithm has become the most accurate algorithm, with strong robustness and generalization.

DOI: 10.54254/2755-2721/33/20230264

The image processing methods can be summarized into target recognition, image segmentation and image classification. The discussion of the three in autonomous vehicle field is follows.

2.2.1. Objection identification.

It is very important to realize obstacle detection and avoidance in reality. Nischal Sanil et al. applied CNN to IOT devices for video or image analysis and had proposed an autonomous vehicle based on monocular vision, in which the camera collects input images and the CNN model is used to make corresponding decisions [7]. They built a car framework by using a stripped-down RC car, which the Raspberry Pi, a useful and portable computer on a single board, can control [7]. During the training phase, the device's camera that is mounted at the front of the automobile is used to gather image data. A track, around 12 to 14 meters long, was built to train the car, which can enter commands for the neural network model and receive image data from the camera [7]. After training, a miniature version of a autonomous car has been implemented, which can make decisions with the help of CNN. It achieves an accuracy of 88.6% and is capable of improving accuracy by hyperparameter tuning and more data collection for adjusting to new scenarios [7].

In the field of computer vision, challenges with object detection and pedestrian recognition are becoming more and more complex. For the field of autonomous driving, it is required to provide accurate and dependable detection and recognition of nearby objects. Yakup Demir et al. have proposed a novel hybird Local Multiple system constructed using CNN and Support Vector Machines (LMCNN-SVM) due to their outstanding feature extraction functions and robust classification abilities [8]. Due to variations in illumination, shades, partial occlusion, and background clutters in a real environment, targets seem variously. This could result in inaccurate object recognition and object detection in applications involving autonomous vehicles, which could be dangerous. LMCNN-SVM system was presented to handle these challenges. Changing environments in driving, such as the variations in light conditions, may lead to wrong decisions by misleading object recognition and detection. The LMCNN-SVM system proposed can figure out these problems. Using a pre-trained AlexNet architecture and a new CNN architecture, images were divided into patches, and the CNN extracted their features [8]. The method proposed in object recognition can be well proved by the results. The best accuracy value was attained by the LM-CNN-SVM system, coming in at 89.80±0.50 and 92.80±0.5 for 15 and 30 images per class, respectively [8]. And the system excelled in terms of higher recognition performance, outperforming cutting-edge techniques (the HOG).

2.2.2. Image segmentation.

Image Segmentation is essential for the autonomous driving field, which can make intelligent and safe driving decisions by precisely perceiving the surrounding environment. For instance, detecting road surface and lane alignment can help unmanned vehicles to recognize geometric construction and the lane line. Therefore, the car can divide images collected into two parts: road and non-road areas, to the realization of navigation, path planning and lane keeping.

A method-based CNN and an image-processing approach, which can be used for advanced driving assistance systems, were proposed by Vedat and Burhan [9]. The method only used a camera view and could draw conclusions regarding the separations, intersections, and crosswalks. The suggested system could develop decision-making mechanisms for driver assistance systems and identify crucial highway information. The approach was originally created during the pre-treatment phase, and more precise and faster outcomes were found at the identification and classification step. The created CNN model showed a higher success rate, especially in detecting road junction and crosswalks, which were significant route types on highways. The crosswalk road was detected as 94.73% and the road junction as 100%. In the classification studies, the right road separation accuracy was 95.2%, however the accurate recognition for the left road separation is only 81.8% [9]. This study had successfully identified important regions of the highways and produced successful results of high level rates for autonomous vehicles field.

Semantic segmentation is a great method for processing the images from the cameras in order to process them in order to understand the scenario. It's noteworthy how deep learning for semantic segmentation has improved our ability to recognize items in images and their locations more precisely. In order to semantically segment images for autonomous driving, Ahmadi et al. introduced an architecture with four components that combines deep learning-based segmentation with traditional segmentation [10]. With the aid of traditional segmentation, this combined approach can separate the image into its component sections, resulting in the acquisition of useful information.

With the substantial increase in the number of vehicles,the requirements for traffic control and safety management are rising. The direction of intelligent transportation has been a developing tendency. Character segmentation, license plate location, and character recognition are the three main steps in the license plate recognition process, which are shown in Figure 3.

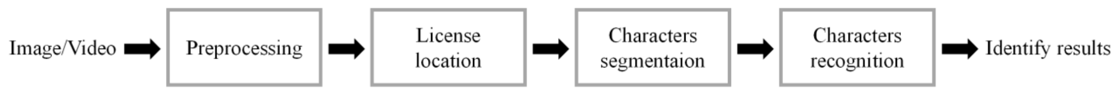

Figure 3. Deep learning diagram [11].

License plate location is used to determine the candidate area of license plate quickly and accurately. Character segmentation mainly divides the characters in the license plate area into a single character. Because the license plate obtained at the beginning may be inclined, it also needs to correct the rotation of the license plate. Character recognition is to solve the problem of single character recognition after segmentation and finally recognize continuous license plate characters. Kailong Wang proposed an license plate recognition algorithm based on Convolutional Recurrent Neural Network (CRNN). Compared to classical license plate recognition technique, his algorithm is an end-to-end recognition method without license plate correction, character segmentation and other processing [11].

Jun Li designed the autonomous driving area segmentation and target detection model in mining scene and deployed an embedded terminal to perform performance tests in real scenes [12]. The model can complete the map semantic segmentation task and target detection in mining scene. The structure was applied to the mining scene using the algorithm structure combining CNN and Recurrent Neural Network (RNN) [12]. CNN extracts spatial context information and extracts image feature mapping through convolutional downsampling. Then upsampled convolution is performed to restore the semantic information of the image. RNN extracts the context information in the time dimension and fuses the image information. The fusion stage is to fuse the feature maps after the convolution of subsamples. RNN can select the features after the image mapping and then apply the information useful at other times to the current frame. In the end, The segmentation of driving area on the map under mining environment is designed and the accuracy is more than 90% loU [12].

2.2.3. Target classification .

For advanced driver assistance systems and autonomous vehicles, traffic sign recognition is essential. It involves utilizing computer vision techniques to automatically identify and categorize traffic signs. Based on the data available, Jency et al. created a classification system for Indian traffic signs and proposed a useful technique for doing so. [13]. Gathering the data from Indian traffic signs and being instructed to construct the classifier, the system combined features with the CNN classifier was generalized and consolidated to deal with adjustments in size, shape, rotation, illumination and translation [13]. Detecting, segmenting, and classifying are the three operations in image analysis. After identifying the margins of the sign, segmenting separates the backdrop from the sign. This model provides a straightforward feature selection approach based on traffic sign recognition and identification data sets. These data offer a variety of complicated traffic signs, such as distracted, tilted, scenic maps, asymmetrical illumination and tilted images [13]. To verify that the algorithm is functioning, numerous intricate and unique traffic signs are used. Convolution Neural Network is used to test the system's functionality against a variety of complex and sophisticated traffic indicators.

3. Conclusion

In this study, we introduce a comprehensive overview of the application of deep learning in image processing for autonomous vehicles. By simulating the hierarchical structure of the human brain's nervous system, deep learning can achieve high-level feature extraction, transformation, analysis, and classification, making it widely used in image, audio, and text processing in the autonomous vehicles domain. The study discusses three applications of deep learning in image processing for autonomous vehicles: object identification, image segmentation, and target classification. In the field of object recognition and classification, researchers have successfully implemented obstacle detection and avoidance of autonomous vehicles by using CNNS. This not only improves vehicle safety, but also creates a safer road environment for drivers and pedestrians. By training the neural network model, the vehicle could quickly and accurately identify obstacles on the road and make timely decisions, thereby avoiding potential hazards. In terms of image segmentation, deep learning technology plays an important role in the intelligent perception of the road environment. The collected live images are divided into road areas and non-road areas, so that vehicles can more accurately identify important elements such as roads and lane lines, accurately perceive the surrounding environment, and provide strong support for intelligent navigation, path planning, and lane.

Overall, deep learning provides strong support and impetus for the development of autonomous driving technology. Through continuous innovation and research, we are expected to achieve a smarter and safer autonomous transportation system, bringing more convenience and comfort to human travel. Despite prevailing challenges, the successful application of deep learning marks a new era in the field of autonomous driving and shows us a better prospect.

References

[1] Meng X Hai, Ma Y Xin, Sun J Hao 2020 *J. Transportation Engineering* 20(3): 6-13
[2] Han X Hui 2018 *D. Sun Yat-Sen University* Learning Based Scene Recognition for Autonomous Driving
[3] LeCun Y, Bottou L, Bengio Y, et al. 1998 *J. Proceedings of the IEEE* 86(11): 2278-2324
[4] LeCun Y, Bengio Y and Hinton G 2015 *J. Nature* 521(7553): 436-444
[5] Xia W 2018 *D. University of Chinese Academy of Sciences* Simulation of Automatic Driving Strategy based on Deep Reinforcement Learning
[6] Hu Y and Cao Z 2021 *J. Environmental Technology* 39(03):100-105+110
[7] Sanil N, Venkat P A N, Rakesh V , Mallapur R and Ahmed M R 2020 *J. International International Conference on Artificial Intelligence and Signal Processing (AISP)* pp. 1-4, doi: 10.1109/AISP48273.2020.9073155
[8] Uçar A, Demir Y and Güzeliş C 2017 *J. SIMULATION* 93(9):759-769
[9] Tümen V and Ergen B 2020 *J. Physica A: Statistical Mechanics and its Applications* Volume 543 123510, ISSN 0378-4371
[10] Hamian M Hosein , Beikmohammadi A, Ahmadi A and Nasersharif B 2021 *J. Computer Society of Iran (CSICC)* pp. 1-6, doi: 10.1109/CSICC52343.2021.9420573
[11] Zhou K Long 2017 *D. Beijing University of Technology* Deep Learning based Image Recognition Application
[12] Li J 2020 *D. Yanshan University* Implementation of Diggings Drivable Area Segmentation and Vehicle Detection Technology Based on Deep Learning
[13] Jency S, Karthika S, Ajaykumar J, Selvaraj R and Aarthi A P 2023 *J. 2023 7th International Conference on Trends in Electronics and Informatics (ICOEI)* pp. 1116-1121