Review of Vehicle Type Recognition Methods

Luchuan Yue

The High School Affiliated to Beijing Normal Univerdity, China

Abstract. With the booming social and economic development, the number of domestic and foreign motor vehicles has increased dramatically, resulting in increasing traffic flow, traffic congestion, and traffic safety problems. To better relieve traffic pressure, intelligent transportation has become the mainstream development direction and an important means of modern transportation, of which vehicle type recognition is the core component of intelligent transportation. Therefore, the research on vehicle identification is of great significance in both scientific research and real-life applications, and it is also a challenging task. Among the various detection and recognition methods, deep learning has been proved to be a more effective one, so this paper discusses the current research on deep learning and vehicle model recognition.

Keywords: Vehicle type recognition, deep learning.

1. Introduction

The automobile is one of the most important means of transportation in our lives. The current state of development of a country's automobile industry often represents the current state of the country's economic development as well. With the increasing development power of society and people's requirements for the quality of life, cars are becoming more and more popular in our lives. The main information feature of a car is its model number. The model information of a car can not only help in a monitoring security system to identify the target vehicle intelligently and help to find and identify the target vehicle. In an intelligent transportation system [1], the model identification also provides helps to establish the real traffic situation and solve the traffic problems.

Vehicle type recognition technology is one of the key technologies for intelligent transportation systems, public security monitoring systems, and unmanned systems. It can help intelligent traffic systems to judge the current traffic situation, help public security monitoring system to identify the target vehicle, and is an important element of unmanned technology.

In recent years, deep learning has evolved rapidly [1] to improve the accuracy of image classification by learning image features from low to high levels [2]. The methods of deep learning can be used to help the computer to implement the task of car model detection and car model recognition. The deep learning approach to implementing model detection and model recognition can greatly improve the efficiency of detection and recognition and reduce the cost of manual processing of the associated images.

At present, the main methods of vehicle type recognition are electromagnetic induction coil detection, radar, infrared, license plate recognition and video image recognition.

The electromagnetic induction coil detection is the most commonly used detection method. Lay a coil with high frequency current under the road in advance, and an electromagnetic field will be generated around the coil. When the vehicle passes by, different induction curves will be generated in the electromagnetic field due to its own shape and size, so as to judge different types of vehicles. The advantages of this method are high accuracy, not affected by the natural environment and climate, and low cost. However, in actual use, the induction coil will be affected by various conditions of the road surface, resulting in a short service life. At the same time, the detection and maintenance of the coil needs to damage the road surface, and the maintenance workload is large and the cost is high.

The principle of radar detection method is that when the vehicle and the radar are moving relative, the signal received by the radar and the signal sent by the radar are different in frequency. The signal received by the radar will vary when different vehicles pass by. The parameters of the passing vehicles can be calculated through the change of the received signal, so as to distinguish different vehicle

types. However, the disadvantage of this method is that the radar will reduce the detection accuracy when the vehicle speed is slow or there are too many irrelevant targets.

The infrared recognition method uses the infrared detectors on both sides of the road. When a vehicle passes through the infrared detection area, the vehicle will reflect the infrared light, thus obtaining the approximate outline of the vehicle. The obtained contour information can be compared with the data in the database to obtain the specific vehicle model results. The advantage of infrared detection method is that it can quickly and accurately identify the outline of the vehicle, which is convenient for installation and maintenance, but its disadvantage is that it is easy to be affected by dust and other environment, resulting in some large errors.

License plate recognition is an indirect vehicle recognition method. It processes the original image through image preprocessing, and then extracts the number and color of the license plate to search the vehicle corresponding to the license plate number in the database to obtain the vehicle model. This method has low hardware requirements and is convenient to use, but the premise is that there is a relatively complete database and the result of license plate recognition is required, otherwise this method is not applicable to reality.

Image-based vehicle type recognition technology is a non-contact passive recognition technology. By extracting images, using digital image processing and pattern recognition technology, vehicle information is obtained, and then the vehicle type is determined according to these features. The advantage of this method is that the digital image has a lot of vehicle information, the success rate of vehicle type recognition is relatively high, and the installation is convenient. Only one camera and computer are needed, and the maintenance is very convenient. After a problem occurs, only the camera needs to be checked, without checking other equipment, and the cost is low.

2. Importance of Identification Model

2.1. Model-identification is an important component of ITS

In the current application in intelligent transportation systems, car model identification methods include a variety of identification methods based on traditional hardware facilities, including radio frequency identification, magnetic field sensor measurement[3], ring induction coils[4], and other methods. With the continuous development of computer computing, the application of computer computing for model recognition is becoming more and more common. In particular, the application of computer vision-based model identification methods is becoming more and more widespread. Applying computer vision for car model recognition can improve the efficiency of car model detection and reduce costs. The development of deep learning also provides new methods for car model recognition, and car model recognition techniques based on deep learning for image processing are receiving more and more attention from researchers[5]. In the intelligent transportation system, the model recognition method based on deep learning can improve the efficiency of model detection, realize real-time information collection and judgment, and send data to the intelligent transportation management system promptly to help the system make a quick judgment and decision to help solve the actual traffic problems. The continuous development of intelligent transportation systems also provides a wider application space and development prospect for car model recognition.

2.2. Model-identification is an important technology for public security monitoring systems

In modern life, cars are frequently used as one of the most important means of transportation. In many criminal cases, criminal vehicle information has been an important clue to solve the case. Vehicle type recognition technology can help the public security police quickly identify the information of illegal vehicles and determine whether they are set vehicles or modified vehicles. In the traditional crime-solving method, the public security police often need to conduct a target vehicle search according to the vehicle information described by the victim. Or the use of manual search methods, from the huge surveillance video to find the target vehicle information, and then screening. And in the lengthy surveillance video, containing millions of images, the workload of using manual

methods for vehicle detection and identification of target vehicles is enormous. The application of computer technology, especially deep learning-based image processing technology, can greatly improve the search efficiency, reduce the cost of hot get, shorten the search time, for the public security police to solve the case to win valuable time.

2.3. Model recognition is an important component of intelligent driving systems

One of the most important tasks of intelligent driving systems is the detection of fast-moving vehicles. Especially in the field of autonomous driving, the results of real-time vehicle detection directly affect the life safety of the vehicle occupants. Fast real-time vehicle detection is also a hot content of research in the field of target detection. Vehicle type recognition, which can help intelligent driving systems to make intelligent judgments and prompt drivers to pay attention to the information of the vehicles in front, can effectively avoid some traffic accidents. Image processing technology based on deep learning also has important applications in the field of intelligent driving and has gradually become the main technical means. Model recognition based on deep learning can improve the accuracy of model recognition and can achieve faster model recognition in complex states according to different scenes.

3. Datasets

The development of vehicle datasets has evolved from coarse-grained to fine-grained vehicle identification. Earlier published vehicle datasets were coarse-grained datasets.

A typical dataset for coarse-grained vehicle identification is the BIT-VehicleDataset dataset from the Beijing Institute of Technology [6]. This includes six vehicle categories, including bus-buses, minibusses, minivans, cars, SUVs, and trucks, with a total of 9850 vehicle images, captured by cameras on the highway, in both 1920*1080 and 1600*900 resolutions.

CSU also released its vehicle dataset, VRID, which was taken at urban chokepoints and contains 10,000 vehicle images taken by 326 cameras over 14 days for 10 medium categories of vehicles. The distribution of image resolutions in the VRID dataset ranges from 400*424 to 990*1134. The vehicle images in the VRID dataset have complex real-world contexts that include different lighting conditions, different geographical locations, and different weather conditions. The dataset is richly annotated with vehicle attributes and can also be used in vehicle color recognition.

The University of Oxford collected video images over a total of 1,000 hours from May 2014 to December 2015, with hundreds of vehicle data captures of a continuous section of Oxford city center. The OxfordRobotCar dataset has a total of 20 million images, captured by six more cars in real driving conditions, totaling over 23 terabytes. The OxfordRobotCar dataset captures many different combinations of the same vehicle in the same area in different weather conditions, traffic conditions, and pedestrians, as well as changes over time including construction and roadworks.

Vehicle Color Recognitiononan Urban Roadby Feature Context [7] vehicle color recognition dataset is published by the Department of Electronic and Information Engineering, Huazhong University of Science and Technology. The vehicle color recognition dataset focuses on the problem of classifying the color of vehicles. The vehicle color recognition dataset contains 15601 vehicle images including eight colors, namely, black, blue, cyan, gray, green, red, white, and yellow. The images are captured in frontal view with a high-resolution camera at 1920 × 1080 resolution on an urban road. The collected dataset is very challenging due to illumination variations, haze, and noise caused by overexposure. The CompCar dataset [8] is one of the classic and most commonly used vehicle fine-grained recognition datasets, published by the Multimedia Laboratory of the Chinese University of Hong Kong, China. There are two images in the CompCar dataset, images of vehicles in the network and images of vehicles in surveillance videos. There are 163 vehicle types and 1716 vehicle models captured with the network, with a total of 136,726 vehicle images and 27,168 vehicle part images. Each car model is labeled with five attributes: maximum speed, displacement, number of doors, number of seats, and car type. A total of 50,000 car images were captured and monitored.

The PKUVehicleID dataset is publicly released by the National Engineering Laboratory of Video Technology (NELVT) of Peking University. VehicleID is a real image of a car captured by multiple cameras. There are 26267 target vehicles in the VehicleID dataset, with a total of 221,763 vehicle images. Among them, the authors manually tagged a total of 90,196 images of 10319 vehicles with vehicle model information, i.e., "MINI-cooper", "AudiA6L" and "BWM1 series" in the dataset.

4. Deep Learning based Method for Vehicle Type Recognition

4.1. Deep Learning Models for Vehicle Type Recognition

An artificial neural network is a deep neural network [9], and the deeper the depth, the more features are obtained [10]. Commonly used deep learning neural network models include feedforward neural networks [27], Boltzmann machines [11], convolutional neural networks, and recurrent neural networks [12] such as LSTM long short-term memory networks [13].

With the rapid development of deep learning, deep learning methods for extracting high-dimensional features from data have been widely used in areas such as image recognition, object detection, and speech recognition [14]. The deep learning network AlexNet [15] was invented by Prof. Geof Hiton and Dr. Alex and Dr. Iley. They first applied the AlexNet network algorithm to the classification task of the prestigious ILSVRC Visual Recognition Challenge, and the final result had a top-5 error rate of only 15.3%, making them the winner of the classification task of that year's competition. Compared to traditional manual feature extraction for classification tasks, AlexNet reduced the error rate by 11%. This created a good start for deep learning applied to image classification tasks [16]. Researchers, both at home and abroad, have started to apply deep learning network methods to a wide range of image classification tasks such as face recognition [17]. Deep learning-based model recognition algorithms are also increasingly studied and applied [18]. Its classification results far surpass traditional manual feature extraction [19].

The limitations of traditional single-branch CNN networks were presented in the paper by Rachmadi et al. The authors proposed a new deep learning-based algorithm for model recognition [20]. The algorithm features a single-branch CNN into two sets of CNNs, and then extracts the deep features learned by the image in these two CNN networks, and combines these two sets of features to achieve simultaneous learning in parallel networks. On the car color-based dataset, the new algorithm proposed in the paper improves the accuracy of car model recognition on the same dataset compared to the bag-of-words model of the baseline data algorithm combining contextual features by 2%. However, the application of the two-branch network makes it necessary to learn more network parameters as well, which greatly increases the training process, resulting in serious time consumption.

Chuanping Hu et al. proposed the idea of combining traditional models with deep network models in their paper [21]. They combined the SPP model SPM with the deep learning network AlexNet to introduce spatial features in the deep neural network. The advantages of traditional manual feature models and deep learning models are combined in the color recognition task of vehicles. The steps of this algorithm are first the original car image is segmented with SPP model for image segmentation, then trained with deep network model AlexNet to extract the convolutional feature image of the vehicle image, the traditional SPP features, and the depth features, which are obtained after learning by the deep convolutional neural network, are cascaded to form a new feature vector of the vehicle image, and finally, the SVM classifier is used for vehicle color classification and recognition. The algorithm is applied to the training and testing of the vehicle color dataset, and the classification accuracy of the final vehicle color is improved compared to the classification accuracy of the vehicle color alone.

Yaming Wang et al. human proposed a sequential-based triad mining algorithm based on order and shape condition constraints [22] to express the deep image features in the intermediate layer and then use an SVM classifier for fine-grained recognition of vehicle model recognition. The algorithm mainly optimizes the deep feature representation for deep learning networks.

Liu M et al. proposed a cascade-based CNN network model [23] for fine-grained recognition of vehicles in their paper. The main cascade of this algorithm is a Faster R-CNN and an end-to-end CNN network model, which first uses the Faster RCNN to suggest candidate frames for multiple regions of interest of the input vehicle image, and then uses the CNN network to extract deep learning features for the images within the regions. The cascade-based network model, which is not restricted to one network model trained to complete the vehicle model classification task, takes full advantage of the different benefits of the two networks in the cascade.

4.2. Network in Network (NIN) Model for Vehicle Type Recognition

Zhu et. al [24] proposed a novel deep network model for vehicle type recognition based on NIN [25] deep model (see in Fig. 1). They address the low identification problem of vehicle type recognition to feature representation. In order to solve the low identification accuracy problem of vehicle type recognition, they introduced the NIN network to extract the nonlinear features which are more powerful in feature representation. Their model of vehicle type recognition fully exploited the NIN network and integrated the features by 1x1 convolution kernal. For the complexity of vehicle images of different types, they improved the original NIN network by changing convolution kernal into small scale, which has increased the depth of the network and improved the accuracy of the recognition model.

The structure used global average pooling layer (GAP) to combine the features and avoid overfitting problem. Moreover, they increased the depth and arranged residual measures to prevent the network fall into gradient loss problem. The improved NIN network achieved a better performance than GoogleNet and VGG in classification of vehicle type recognition.

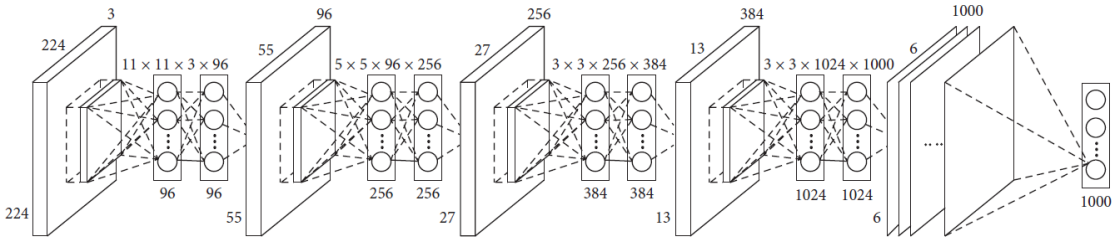


Fig 1. NIN based deep network for vehicle type recognition

In the improved NIN network, they use 4 convolution kernels of 3x3 size instead of 11x11 convolution kernel. They adopted the small scaled convolution kernels to formulate the structure of deep network. Moreover, the improved NIN network introduced the residual blocks to constraint the training errors of stacking identity maps, which improved the performance of NIN network. The NIN network construct its structure by 3x3 convolution kernel to achieve rapid convergence. The structure is as shown in Fig. 2.

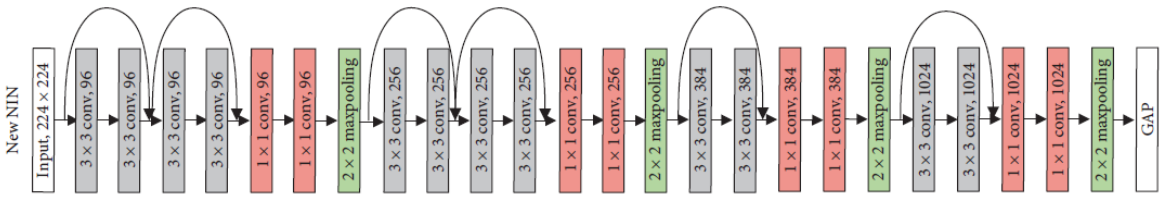


Fig 2. Structure of improved NIN network

Reference

[1] Li Y J, Zhang T, Hu HH. “Deep kernel mapping support vector machine based on multilayer perceptron.” Beijing University of Technology Journal, 2016, 42(11): 1652-1661.

[2] Li W, Qiao JF. “Structural design of fuzzy neural network based on recursive clustering and similarity.” Journal of Beijing University of Technology, 2017, 43(2):210-216.

[3] Vytautas M, Dangirutis N, Mindaugas Z, et al. “Dynamic Vehicle Detection via the Use of Magnetic Field Sensors.” Sensors, 2016, 16(1).

[4] Xu Zewei, Chen Xianqiao. “Vehicle recognition and classification method based on laser scanning point cloud data.” Transportation information and Safety (ICTIS),2015,10.

[5] Li Kaiming, Zhang Qun, Luo Ying, Liang Bishuai, Yang Xiaoyu. “A review of ground vehicle target recognition research.” Journal of Electronics, 2014, Vol. 42 (No. 3): 538-546.

[6] LIU H, TIAN Y, YANG Y, et al. “Deep relative distance learning: tell the difference between similar vehicles.” Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition New York: IEEE, 2016:2167-2175.

[7] CHEN P, BAI X, LIU W. “Vehicle color recognition on the urban road by feature context.” IEEE Transactions on Intelligent Transportation Systems, 2014, 15(5) : 2340-2346.

[8] YANG L, LUO P, CHANGE LOY C, et al. “A large-scale car dataset for fine-grained categorization and verification.” Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. new York: IEEE, 2015: 3973-3981.

[9] HAYKIN S. “Neural networks.” New York: Prentice hall New York, 1994.

[10] Hisham El-Amir, Mahmoud Hamdy. “deep learning pipeline.” Apress. 2019.

[11] Goodfellow I, BENGIO Y, COURVILLEA. “deep leaming.” Cambridge: MIT Press, 2016.

[12] Ackley D H, Hinton G E, Sejnowski T J. “A learning algorithm for Boltzmann machines.” Cognitive science, 1985, 9(1): 147-169.

[13] U. Shah and A. Harpole. “a Review of Deep Learning Models for Computer Vision.” 2018 IEEE Punecon, Pune, India, 2018, pp. 1-6.

[14] Jiang, X., Hadid, A, et al. “Deep Learning in Object Detection and Recognition.” Springer. 2019.

[15] M. Baccouche, F. Mamalet, C. Wolf. “Sequential deep learning for human action recognition.” Human Behavior Understanding. Springer Berlin Heidelberg, 2011:29-39.

[16] KRIZHEVSKY A, SUTSKEVERI, HINTON G E. “Imagenet classification with deep convolutional neural networks.” Communications of the ACM, 2012, 60(2): 2012.

[17] SIMONYAN K, ZISSERMAN A. “Very deep convolutional networks for large-scale image recognition.” arXiv Preprint arXiv, 2014: 1409-1556.

[18] K.He,X.Zhang,S Ren,et al. “Spatial Pyramid Pooling in Deep Convolutional Networks for Visual Recognition.” IEEE Transaction on Pattern Analysis & Machine Intelligence,2015,37(9):1904-1916.

[19] Sun Yanfeng, Zhang Kun, Hu Yongli. “Deep video-based human behavior feature representation and recognition.” Journal of Beijing University of Technology, 2016, 42 (7):1001-1008.

[20] RACHMADI RF, PURNAMA I. “Vehicle color recognition using convolutional neural network.” arXiv Preprint arXiv, 2015: 1510. 07391.

[21] HU C, BAI X, QI L. “Vehicle color recognition with spatial pyramid deep learning.” IEEE Transactions on Intelligent Transportation Systems, 2015, 16(5):2925-2934.

[22] WANG Y, CHOI J, MORARIU V I, et al. “Mining discriminative triplets of patches for fine-grained classification.” Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. New York: IEEE, 2016: 1163-1172.

[23] LIU M, YU C, LING H, et al. “Hierarchical joint CNN-based models for fine-grained cars recognition.” International Conference on Cloud Computing and Security. Berlin: Springer, 2016: 337-347.

[24] Zhu E, Xu M, Pi D C. “Vehicle Type Recognition Algorithm Based on Improved Network in Network. Complexity,” 2021,1-10.

[25] M. Lin, Q. Chen, and S. Yan, “Network in network,” Computer Science, arXiv: 1312.4400, 2013.