Determination of driver dangerous driving behavior based on the consensus, stage and solution of deep learning research

Yan Fan[1]，Yilan Wu[*, 2, 3]

[1]Institute for advanced studies in humanities and social science, Chongqing University, Chongqing, China;

[2]Criminal Law School, East China University of Political Science and Law, Shanghai, China;

[3]Police Department, Anhui Vocational College of Police Officers, Hefei, China.

[*]Correspondence: wuyilan@ahjgxy.edu.cn

Abstract

With the development of road traffic and the popularity of private cars, the number of traffic accidents and deaths in the world is also increasing.Under the World Health Organization (WHO) Road Traffic Injury Act, about 1.35 million people worldwide die every year from traffic accidents, accounting for 3% of the GDP of most countries.As the world's largest autommaker, China's car ownership and drivers have grown rapidly in recent years. According to the Ministry of Public Security's Traffic Management Bureau, the number of motor vehicles in China hit 260 million at the end of 2019, rising 8.83 percent from the previous year. The number of private cars among them has risen to 207 million, up an average of 1.966 billion per year over the past five years. The country drove to 435 million people, including 397 million, accounting for 91.26% of all drivers.（Kuang，2012）The increase in the number of motor vehicles and drivers further complicates the road driving environment, partly increasing the likelihood of traffic accidents.According to the statistics of the People's Bureau of China under the National Bureau of Statistics, the total number of traffic accidents reached 244,900, an increase of 40,000 over 2017,40,000 traffic accidents and 1.384 billion yuan.How to ensure road safety and reduce the number of traffic accidents is an urgent problem to be solved in China's social development.So, based on that basis,In the face of this problem, this paper adopts deep learning technology, combined with target recognition and behavior recognition method, and video data as input to realize the comprehensive real-time monitoring of distracted driving behavior (mobile phone, drinking when driving) and fatigue driving behavior (dozing), which is of great significance to reduce the possibility of traffic accidents and improve the drivers' driving awareness and safe driving.

Keywords

Deep learning, Dangerous driving, and Behavior determination technology.

1. Introduction

A The causes of road traffic accidents include subjective and objective factors, which from the data research is indeed mainly related to human factors, but the objective factors can not be underestimated.Objective factors include vehicles, roads, traffic and other factors.A large number of analytical studies show that human factors are the main cause of traffic accidents, including fatigue driving (dozing) and distracted driving (phone calls, text messaging, drinking, etc.).Fatigue driving refers to the phenomenon of physical and mental dysfunction after long periods of driving.According to statistics, China's annual fatigue rate is reduced to 40% due to major traffic accidents.Distracting driving refers to the phenomenon that drivers pay attention to activities unrelated to the operation of the vehicle, resulting in a decline in driving ability.Studies show that drivers are 2.8 times more likely to have a phone accident while driving, and 23 times more likely to have a traffic accident while driving.

Among the multiple influencing factors, distracted driving can also easily lead to an increase in the number of accidents.（Tran，2016）UK statistics show that 70% of fatal traffic accidents are caused by driver negligence.From these data, it can see that preventing driver dangerous driving behaviour (fatigue driving and distracted driving) is a worldwide concern and attention.In response to a large number of traffic accidents caused by dangerous driving behaviors, countries around the world have formulated a series of laws and regulations that must be dealt with. However, the difficulty of law enforcement agencies, evidence collection and high regulatory labor costs make it difficult to eliminate dangerous driving behavior. If you can effectively monitor mobile phones, drinking, sleeping and other behaviors during driving, to ensure the timely detection and correction of the behavior, it can not only reduce the cost of dangerous driving monitoring, but also significantly reduce the possibility of traffic accidents.（Knapik，2019）Due to the risk and high incidence of dangerous driving behavior, scientists at home and abroad put forward various technical schemes to monitor driving behavior, according to the driving data, due to the high cost of sensors in the car, must install the driving behavior monitoring device, and the method is vulnerable to the road environment,（Girshick，2014）so it is difficult to put it into practical application: it also affects the driver's driving, not very practical.Vision-based driving behavior monitoring is increasingly mainstream due to its non-contact and economic efficiency. Traditional computer vision-based methods of monitoring dangerous driving behavior mainly include monitoring fatigue driving behavior based on the driver's facial characteristics, and monitoring distracted driving behavior according to the driver's body posture characteristics.However, most of the dangerous driving behavior detected by these methods depends on the manually extracted texture or shape properties, and the accuracy and the real-time requirements of driving behavior monitoring cannot be considered, which has certain limitations in practical applications.

Deep learning-based driving behavior monitoring techniques are divided into two categories. The first class is to realize the classification of the driver driving behavior by using the spatial properties of the driver images in the deep neural convergence network model.A comparative analysis of Alex, Google, VGG-16 and Rith Net performing distracted driving tasks, Tal showed that Google recognized them most. Irachi trained drivers to use various deep neural symbiosis networks to train drivers in raw images, cracked skin, face images, hand images and driver hand face combinations, and then assessed the weights of all network costs using a genetic algorithm to detect distracted driving attempts.Leung designed a 10-level sink neural network that utilized the parameters of the open surface network model and applied the PERCLOS guide to the training of sink neural networks.The extraction and identification function of the driver fatigue characteristics are realized.Liu Keke used the To F camera to obtain the image of the driver setting depth, designed a deep neural longitudinal network based on Le net to extract the hierarchical image function, realize the identification of the driver's joint points, and then evaluate the driver's driving behavior.The second type is to create deep learning-based behavioral recognition models (e. g.B TSN, C3D, LRCN et al.), the model combines spatial and temporal feature extraction from spatial feature extraction in the network to identify the driver's driving behavior.Ma Ying-jeou（2019) et al used Kinect to construct fatigue height video recordings and proposed a fatigue detection model based on dual longitudinal neural networks, which takes into account the spatial properties of the current deep frames and the temporal properties of the adjacent depth images represented by motion vectors.The driver's fatigue driving behavior was detected.

In order to improve the accuracy of driver fatigue detection, the study introduces visual attention mechanism into the spatiotemporal network model, and balances its features through learning methods of forward propagation and backward feedback, enabling the spatiotemporal network model to focus on important features and suppress unimportant features, thus improving the information processing ability of the model.（Kuang，2016）In deep learning algorithms, attention mechanisms can be regarded as a special average weighted attention mechanism compared to pooling operations, and the attention mechanism can be regarded as a general concentration method with input allocation preference.Following this idea, attention mechanisms are introduced into spatiotemporal network models through an attention-weighted pooling approach.

2. IMAGE—BASED DRIVING BEHAVIOR MONITORING

Image-based driving behavior monitoring technology mainly uses image processing and traditional machine learning methods to extract the driver's face or body posture characteristics collected by the on-board camera to monitor the driver's driving behavior.The approximate position of the eye is positioned according to the geometric distribution of the driver's facial features, and using DI value processing and secondary filter to calculate the pixels of the eyebrows and eye area, so as to judge the open and closed state of the eye, thus realizing fatigue detection. A thermal imaging-based yawn recognition technology is proposed, first completing face alignment by canthus recognition, then realizing yawning recognition by average temperature analysis, and then judging the driver's fatigue state, which can achieve good recognition effect both day and night. The AdaBoost algorithm was used to divide the human eye area according to the distribution characteristics of the face "three places and five eyes".The open and closed state of the eyes was determined by the fuzzy comprehensive evaluation algorithm, and the fatigue state of the driver was detected according to the Perclos principle.Zhao Minghui proposed a driving attitude recognition method based on local DPM fusion characteristics, which first detects the core area of driver driving behavior, then constructs a DPM scoring model and realized the extraction of local DPM fusion characteristics. Finally, a RBF nuclear support vector machine was adopted to realize driving attitude classification, so as to detect scattered driving behavior.Wenteng people and others found that white has good clustering on CB-Cr, established a Gaussian white division model based on YCB-Cr color space, and used the model to divide the eye white within the image of the face area, using the white area of the eye as an indicator of eye opening, and combined with Perclos to judge the driver fatigue status.

3. DRIVING BEHAVIOR MONITORING BASED ON PHYSIOLOGICAL SIGNALS

Driving behavior monitoring technology based on physiological signals mainly take over the driver's physiological signal collection suite to collect physiological signal data, such as EEG and heart, these data and distraction and fatigue state have a stronger correlation between electricity, eye current, muscle current and other information, using machine learning methods to analyze the data, and then realize driving for monitoring.（Redmon，2016）analysis by Waller compared the driver EMG, the skin electrical activity , respiratory rate and ECG during normal and fatigue driving (ECG), It was shown that Eda and ECG parameters had a good effect on the distinction between normal and fatigue states. Coins etc.with wave sheet decomposition and example entropy algorithm for the driver in awake and fatigue states in ethnographic signals (according to EEG analysis). After completion of the experiment, the results showed that fatigue is that the α wave has relatively high energy and β waves relative to low energy.Analysis of the driver's electronic signal while driving was distracted by monitoring (EEG changes, hair ABC Iaxial waves had a significant association with driver distraction.（Eraqi，2019）driver analyzed the electrocardiogram data between drivers for a long time, obtained the driver electrocardiogram time domain index, and proposed the fatigue driving monitoring algorithm of multi-finger base standard fusion theory to determine whether the driver is in a fatigue state. Thank the same person for extracting and combining the characteristics of EEG, ECG and dielectric signals during driver driving and proposing a base assessment method for multifunctional fusion and migration learning for driving fatigue in physiological signals.（Ma，2015）insisted on analyzing the driver's heart rate and ECG change rate during the driving distraction, and investigated the distracted driving according to the sample entropy.The results showed that the rate of ECG change in distracted driving is more likely to indicate an ECG change than heart rate.

4. METHOD OF SPATIOTEMPORAL FEATURE FEATURE EXTRACTION OF FATIGUE—DRIVEN DRIVING

For the task of fatigue-driven state recognition, we first analyze the advantages and disadvantages of the current mainstream behavior recognition algorithms, and study them based on convolutional neural networks and short-time memory networks.In order to improve the recognition accuracy and real-time of the behavioral recognition algorithm and reduce the computation of the algorithm,（Eraqi，2019）the spatial dimension information of video frames extracted by the mobile network V3 light convolutional neural network is studied as spatial features, using the variable structure gated cycle unit of short long memory network acts as a function of time to extract time-dimensional information of video image sequence, and establish space-time network model by combining spatial feature extraction network and time feature extraction network to realize the accurate identification of E; Regression-based detection algorithms do not require pre-extract candidate regions and can regression directly from the images.The category probability and the edge position of the labeled object can eventually be detected by one detection.Such a party the law includes Jolo (Singh,2013)series algorithms and SSD(Billah,2018)serial algorithm. Jolo It is the beginning of the one-step detection method It divides images intoS×S-Cell, and then predicted whether each cell contains a target. The body of the final filtered target object, the category and the location coordinates of the target object prediction field for the target object.The algorithm recognizes very quickly, but has a relatively low measurement accuracy, objects and adjacent objects are not well detected,SSD and Jolo the same holds for regression-based detection algorithmic, the difference is that,SSD the use of multi-scale feature maps to detect different sizes is to achieve a certain degree of reduction and precision.

After comprehensive evaluation, the dangerous driving behavior is graded for warning.In order to realize the comprehensive evaluation and graded early warning of dangerous driving behavior, the distracted driving behavior detection algorithm is combined with the fatigue driving condition detection algorithm to construct the dangerous driving behavior detection algorithm.According to the driving category determined by the model, the comprehensive evaluation indicators of dangerous driving behavior were summarized and weighted by hierarchical analysis.Then a comprehensive evaluation model of dangerous driving behavior was established with Mamdani to test the comprehensive evaluation results. Finally, literature algorithms such as dangerous driving behavior models were tested on dangerous driving behavior datasets, summarize the advantages and disadvantages of the model in identifying accuracy and speed, and put forward the general warning process of real-site dangerous driving behavior models.

5. DEEPIY LEARNS THE OBJECTIVE RECOGNITION ALGORITHM ANALYSIS

Target recognition is a key task in the field of computer vision, which has been studied for nearly two decades. Deep learning mode has progressively overtaken traditional machine vision methods and become the standard algorithm in the field of target recognition in recent years, thanks to the rapid growth of deep learning technology. （Girshick，2014） Deep learning-based target identification algorithms now consist mostly of two-step detection techniques based on candidate areas and single-step detection methods based on regression. The detection algorithm based on the candidate regions divides the detection problem into two stages, first generating the candidate regions, and then performing category prediction and boundary adjustment for the candidate regions.Such algorithms feature a A-CNN series of algorithms, such as R-CNN, SPP-Net.

For example, the R-CNN, the generating approximately 2000 candidate fields from the image using a selective search algorithm, and then unifying each candidate box with a uniform size.It enters the sink neural network and eventually achieves target detection by support vector machines and linear regression.

Although improving the accuracy of target detection is very good, there are also more serious problems of slow detection speed, large memory consumption and more information loss.For these questions, various researchers propose R-CNN-based enhanced target identification algorithms.SPP-Net aborts the process of image normalization, introduces spatial pyramid aggregation to accommodate feature maps of different dimensions, and completes the dimensional unification of functional candidate boxes when extracting functions from the original map.The idea of fast R-cnn leasing from SPP-Net is to fold only the original image and assign it to candidates when publishing the feature.At the same time, it utilizes the multitask loss function to achieve classification and regression of weight distribution, significantly improving recognition speed and accuracy compared to R-CNN. Fast R-CNN based on fast R-CNN uses the RPN(region suggested network) rather than the original selective search method to cone its own candidate boxes, thus significantly reducing the number of candidate boxes.The quality of the candidate boxes was also significantly improved.Thereafter, two-step detection algorithms such as mask R-CNN, cascade R-CNN further optimize speed or accuracy based on the existing models.In general, target region-based target recognition algorithms perform well in recognition accuracy, but processing candidate boxes takes more time and is difficult to meet real-time requirements. However, with recent developments, this type of algorithm also improves recognition accuracy.

6. JUDGING ON DISTRACTED DRIVING BEHAVIOR ACCORDING TO THE WEIGHT OF THE INDEX

Driver's distracted and tired driving behavior include phone calls, texting and drinking and yawn material while driving.Based on these four types of dangerous driving behaviors, a comprehensive evaluation of dangerous driving behavior is constructed

Indicators: United States.To ensure the accuracy of the U weight values for the comprehensive evaluation metrics, this section calculates the index weights using the analytical hierarchy process (AHP), by comparing two pairs of multiple metrics.In real cases, datasets of dangerous driving behavior were collected for model testing.The comprehensive behavior of dangerous driving behaviors is summarized based on the category of hazardous behaviors identified by the hazardous driving behavior detection model.

Metrics were evaluated, and a hierarchical analysis was used to give them different weights.Then, a comprehensive assessment model of dangerous driving behavior was created based on the Mamdami argument system.

7. CONCLUSIONS

Monitoring and warning of driver distraction and fatigue is important to reduce the likelihood of traffic accidents and reduce the safety risk of drivers.However, existing distraction and fatigue driving monitoring methods, including driving behavior monitoring based on driving data, physiological signals, image processing-based driving behavior monitoring, etc., have defects in practical application.Based on computer visual fatigue detection and dangerous driving behavior detection has a broad perspective and application space, based on the face feature point recognition, the paper introduces multiple fatigue detection and depth-based video dangerous driving behavior detection, has achieved a relatively ideal detection effect, but there are still some deficiencies, to be improved.

References

[1] Singh A, Kaur J. Driver fatigue detection using machine vision approach[C]. 2013 3rd IEEE International Advance Computing Conference (IACC). IEEE, 2013: 645-650.

[2] Knapik M, Cyganek B. Driver’s fatigue recognition based on yawn detection in thermal images[J]. Neurocomputing, 2019, 338: 274-292.

[3] Billah T, Rahman S M M, Ahmad M O, et al. Recognizing distractions for assistive driving by tracking body parts[J]. IEEE Transactions on Circuits and Systems for Video Technology, 2018, 29(4): 1048-1062.

[4] Pan Zhigeng, Liu Rongfei, Zhang Minmin.Study on the fatigue driving detection algorithm based on fuzzy comprehensive evaluation [J].Software Journal, 2019,30 (10): 2954-2963.

[5] Zhao Minhui.Image-based study on driver attitude recognition methods [D].Southeastern University, 2018.

[6] Kuang Wenteng, Mao Kuancheng, Huang Jiacai, etc.Fatigue driving detection based on a Gaussian eye white model [J].Chinese Image Graphic Graphics Journal, 2016,21 (11): 1515-1522.

[7] Tran D, Do H M, Sheng W, et al. Real-time detection of distracted driving based on deep learning[J]. IET Intelligent Transport Systems, 2018, 12(10): 1210-1219.

[8] Eraqi H M , Abouelnaga Y , Saad M H , et al. Driver Distraction Identification with an Ensemble of Convolutional Neural Networks[J]. Journal of Advanced Transportation, 2019,

[9] 2019:1-12.

[10] Liu Ke.Driver pose recognition based on depth images [D].Nanjing University of Technology, 2018.

[11] Ma X, Chau L P, Yap K H. Depth video-based two-stream convolutional neural networks for driver fatigue detection[C]. 2017 International Conference on Orange Technologies (ICOT). IEEE, 2017: 155-158.

[12] Ma X, Chau L P, Yap K H, et al. Convolutional three-stream network fusion for driver fatigue detection from infrared videos[C]. 2019 IEEE International Symposium on Circuits and Systems (ISCAS). IEEE, 2019: 1-5.

[13] Wu Hao.Deep Learning-Based Driving Behavior Recognition Methods [D].Southeastern University, 2018.

[14] Liang Xiaoyu.Study on the Driver Fatigue Detection Method Based on Deep Learning [D].Tianjin University of Technology, 2018.

[15] Zhao Yujuan.Distraction driving behavior analysis and management actions [N].People's Public Security Daily Traffic Safety Weekly, 2016-09-27 (003).

[16] Girshick R, Donahue J, Darrell T, et al. Rich feature hierarchies for accurate object detection and semantic segmentation[C]. Proceedings of the IEEE conference on computer vision and pattern recognition. 2014: 580-587.

[17] Redmon J, Divvala S, Girshick R, et al. You only look once: Unified, real-time object detection[C]. Proceedings of the IEEE conference on computer vision and pattern recognition. 2016: 779-788.