Fatigue Driving Warning Internet System of Vehicles Based on Trajectory and Facial Feature Fusion

Zihan Li[1,*,a], Siqi Yang[2,b] and Teng Zhou[3,c]
[1]School of Software Engineering, Tongji University, Shanghai, China
[2]School of Computing and Artificial Intelligence, Southwest Jiaotong University, Sichuan, China
[3]New Engineering Industry College, Putian University, Fujian, China
*

Keywords: Image Recognition, Internet of Vehicles, Fatigue Driving.

Abstract: Fatigue driving is one of the common factors leading to traffic accidents. To mitigate its risks, this study proposes a vehicle network-based warning system for detecting fatigue driving through information fusion and develops a corresponding simulation prototype. First, the drivers' facial images and vehicle trajectory data are pre-processed to enhance model training accuracy. Then, a detection framework based on ResNet50 is constructed to integrate and analyze the facial features of drivers and vehicle trajectories for identifying fatigue driving behavior. Finally, the architecture of the vehicle network early warning system is designed to utilize the identification results. Once signs of fatigued driving are detected, the system will automatically issue an alert to help drivers prevent accidents caused by drowsy driving, ensuring road and driver safety. The code for the system part can be found at https://github.com/Theo-teng/Internet-of-Vehicles-Simulation.git.

1 INTRODUCTION

With the continuous increase in the number of vehicles, the road traffic safety problem is becoming increasingly serious. As one of the main causes of traffic accidents, fatigue driving is significantly harmful. Therefore, the research and development of fatigue driving detection and the early warning system are significant in reducing the incidence of traffic accidents.

Fatigue driving recognition technology mainly relies on the monitoring of drivers' physiological and behavioral characteristics, such as electrooculogram, facial key point, and motion capture (Tian & Cao, 2021; Chen et al., 2021; Liu et al., 2020). Although such detection techniques are mature enough, their effectiveness is still limited by a single type of data source. For example, facial recognition may be affected by factors such as light changes and facial occlusion. In existing studies of fused data sources, although multiple data sources such as blink frequency and driving route offset are considered, these methods rely on manually setting the positive sample threshold (Tang, 2024). Therefore, this study fuses the changes in the driver's facial state and vehicle trajectory through data fusion to make up for the shortcomings of a single detection method to a certain extent and improve the overall robustness of the system.

The rise of the Internet of Vehicles (IoV) provides new solutions and application scenarios for fatigue driving detection technology (Abbas & Alsheddy, 2020). Real-time monitoring of fatigue driving behavior can be achieved by collecting and analyzing vehicle driving data through the IoV platform. Based on the fatigue driving recognition method in this study, this paper proposes a real-time recognition and warning scheme for the vehicle networking system. When the system detects the driver's fatigued driving behavior, the warning will be carried out through both in-vehicle and inter-vehicle channels, to avoid the safety hazards brought by fatigued driving behavior.

The rest of the paper is organized as follows. Section 2 describes the techniques used and their advantages; Section 3 describes the methodology used in the study and the implementation details,

[a] https://orcid.org/0009-0004-1526-5148
[b] https://orcid.org/0009-0005-3714-5937
[c] https://orcid.org/0009-0008-2296-5416

Li, Z., Yang, S., Zhou and T.
Fatigue Driving Warning Internet System of Vehicles Based on Trajectory and Facial Feature Fusion.
DOI: 10.5220/0013525100004619
In *Proceedings of the 2nd International Conference on Data Analysis and Machine Learning (DAML 2024)*, pages 412-417
ISBN: 978-989-758-754-2

where the architecture is divided into data pre-processing, information fusion recognition, and the framework of the IoV system, and concludes with an evaluation of the model; Section 4 presents the experimental results and discusses the advantages and disadvantages in comparison with other models, as well as the performance changes before and after the model improvement; Section 5 summarises the research, explains the advantages and significance of this research, and the direction of future work.

2 METHODOLOGYS

ResNet50 is selected as the target detection model framework (He et al., 2015). ResNet50 was proposed by Kaiming He of Microsoft Research in 2015 and won the championship of the ILSVRC (ImageNet Large Scale Visual Recognition Challenge) image classification competition in that year, which proved ResNet50's effectiveness and superiority in processing large-scale image data. Meanwhile, due to its excellent generalization ability and availability of pre-trained models, ResNet50 is widely used in a variety of computer vision tasks.

The residual connections in the structure allow information to be passed in direct jumps between network layers, and batch normalization and activation functions are also used in each residual block to improve the performance of the model. A 7×7 convolutional kernel with a step size of 2 is used at the top of the model to reduce the resolution of the image to speed up the model training, after that each layer of ResNet50 is based on a 3×3 convolutional kernel with a step size of 1 for extracting the features, and at the end of the model there is a global average pooling layer and a full connectivity layer for the final classification.

3 DATASETS

3.1 Dataset Description

The datasets KITTI and Driver Drowsiness Dataset (D3S) are selected as model training samples (Geiger et al., 2012; Gupta et al., 2018). The KITTI dataset provides a large amount of sensor data from real driving, which contains the vehicle trajectory data needed for this study. The D3S dataset contains 22348 fatigue driving images and 19445 normal driving images, which have been all segmented with face regions for easy use in model training.

3.2 Facial Data Preprocessing

The pre-processing of the portrait is to convert the color portrait data in the D3S dataset into grey-scale images. The grey scale histogram equalization method is used to make the brightness of the image uniform and to highlight the features of the portrait. The image comparison before and after processing is shown in Figure 1.

(a) (b)

Figure 1: Original image and image after grey scale processing of the D3S dataset. (a) is the original image, (b) is the processed image (Photo/Picture credit: Original).

3.3 Trajectory Data Preprocessing

Firstly, canny edge detection is carried out, that is, after binarizing the image, a suitable threshold is set to prepare for obtaining the edge of the lane line, after detection, ROI extraction is carried out to obtain the accurate edge of the lane line, the edge detection result of this region is cropped out, and then the Hough transform is carried out for straight line detection in the ROI region, the central idea of this operation is to represent a straight line with a coordinate point and to find out the lane line and then plot it on the original map, as shown in Figure 2.

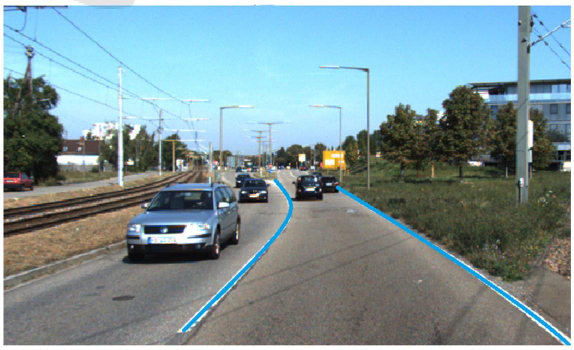

Figure 2: Lane line detection and mapping on the KITTI dataset (Photo/Picture credit: Original).

Write a function to calculate the slope of the lane line segment and calculate the radius of curvature of the lane to facilitate the subsequent accurate determination of driver fatigue by this offset. The formula is shown in Equation (1).

$$K = \frac{|f(\ddot{x}_0)|}{(1 + f(\dot{x}_0)^2)^{\frac{3}{2}}} \quad (1)$$

where $f(\mathrm{x}_0)$ is a close-circle expression, $f(\dot{x}_0)$ is the first-order derivative of this expression, and $f(\ddot{x}_0)$ is the second-order derivative.

4 INFORMATION FUSION IDENTIFICATION

In this paper, a ResNet50 model is trained based on the tracking dataset KITTI and the facial image dataset D3S. It can independently recognize whether the lane segments in the tracking dataset are offset or not, and judge fatigue driving according to the degree of offset, and the human face in the facial image dataset can be recognized, and judgement can be made according to the expression of the face. The recognition results are fused in the fully connected layer of ResNet50, and the final model can detect fatigue driving based on these two types of datasets.

This paper uses ResNet50 to recognize the features of the driver's face, compared with the traditional neural network, ResNet50 has a deeper network structure, and the problem of gradient disappearance in the training of the deep network is mitigated by the introduction of the residual connection, which effectively improves the performance of the model.

5 INTERNET OF VEHICLES SYSTEM

5.1 System Design

An IoV system has been designed that applies the fatigue driving recognition mentioned above in this paper. Figure 3 shown the flowchart of feature processing and fusion. The system architecture diagram shown in Figure 4 briefly describes the components of the system, while Figure 5 briefly describes the workflow of the system's detection and warning modules.

As shown in Figure 4, the system consists of a Hardware Layer, Interface Layer, Service Layer, Data Layer, and External Systems.

Hardware Layer mainly consists of in-vehicle cameras, sensors, and warning devices. The in-vehicle camera is mainly used to capture the facial image of the driver. Sensors are mainly used to collect vehicle status and position information, including but not limited to GPS positioning, vehicle speed sensors, and attitude information collection cameras. Out-of-vehicle warning devices mainly include lights and car horns. The onboard display is used to be able to provide warning information and possible operation options, including sending warnings to the co-driver for the vehicle and warnings for other vehicles.

Interface Layer consists of a data acquisition interface, a command execution interface, and a communication interface. Data acquisition interface is responsible for collecting data from the hardware layer and transmitting it to the functional layer. Command execution interface is responsible for receiving commands from the functional layer and sending them to the hardware layer. Communication interface is responsible for the external communication interface of the vehicle, which is used for data transmission with other vehicles.

The service Layer is mainly composed of a local data processing module, fatigue detection module, and decision engine. The data processing module processes the images captured by the camera and the sensor signals according to the data preprocessing process and provides grey-scale images and other data after processing. The fatigue detection module performs fusion recognition detection and outputs driver fatigue detection results. The decision engine decides whether to trigger a warning based on the fatigue detection result.

The Data Layer consists of a local database and a cloud database. The local database is used to store vehicle status, driver behavior data, and fatigue detection results. The cloud database is used for distributed permanent storage of data for subsequent data analysis and model training.

External Systems, i.e. other vehicles, can receive warning messages and take appropriate measures.

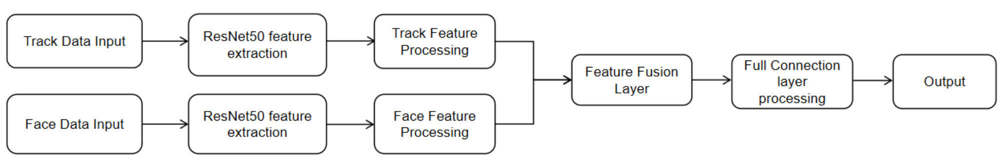

Figure 3: Flowchart of feature processing and fusion (Photo/Picture credit: Original).

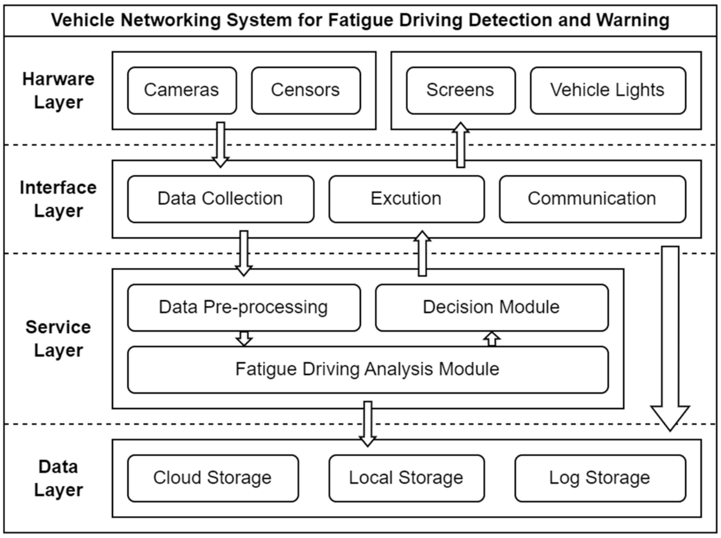

Figure 4: Architecture diagram of the IoV system (Photo/Picture credit: Original).

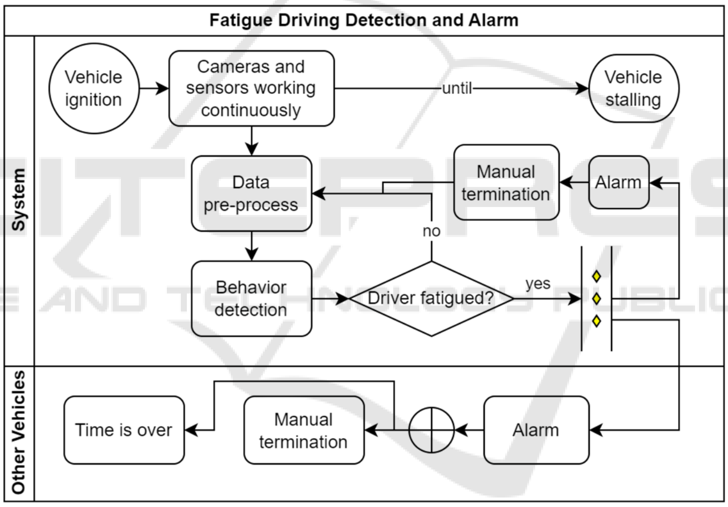

Figure 5: Flowchart of detection and warning module (Photo/Picture credit: Original).

As shown in Figure 5, the system works continuously after the vehicle ignition. Once a fatigued driving behavior is detected, the system will simultaneously send warnings both in-vehicle and inter-vehicle. The in-vehicle alarm continues to operate until manually canceled, indicating that the driver is sober enough. The alarms in other vehicles can also be terminated by time limits.

5.2 System Simulation

The two simulators, the OMNeT++ network simulation platform, and the SUMO road traffic simulation platform, are connected via TCP sockets, allowing for bi-directional coupled simulation of road traffic and network traffic (Sommer et al., 2011; Behrisch et al., 2011; Varga & Horning, 2008). In this case, the movement of vehicles in SUMO is mapped to the movement of nodes in the OMNeT++ simulation. The goal is to build a communication simulation framework for the Internet of Vehicles based on SUMO and OMNeT++ and to build a simulation environment in Oracle VM VirtualBox VM environment through Veins' mirrors for signal warning simulation. The environment for this experiment is Oracle VM VirtualBox 7.0.20, Windows 11. The Veins version framework is instant-veins-5.2-i1.

5.3 Simulation Content

In OMNET++, the INET standard host is used to simulate wireless communication between mobile hosts, and two static wireless communication hosts are created. The network contains two hosts, one of which sends a UDP data stream to the other host wirelessly, aiming to simplify the physical layer and low-level modules so that the network can be expanded in the future. To achieve wireless communication, the INET module in OMNET++ needs to create a radio medium, which represents the shared physical medium on which all communications in the simulation rely and is responsible for signal propagation, loss, interference, etc. Through code implementation, the communication range of host A and host B can be visualized, and the propagation process of UDP signals can be displayed at the same time. The transmission process is presented in the form of colored rings. The movement of the host is managed by the mobile submodule, which extends the mobility of the standard host and displays its speed and movement trajectory through visualization tools.

6 EXPERIMENTAL

In this experiment, the dataset is first pre-processed and then divided into a validation set, training set, and test set, the validation set test set is trained by ResNet50, the weights are adjusted and then tested in the test set, and finally, the validation and testing are performed on another dataset. The experimental flow of this paper is shown in Figure 6.

6.1 Experimental Settings

D3S and KITTI are used as the datasets, which are divided into training and testing sets in a ratio of 9:1, and the important hyper-parameter settings during the training process are shown in Table 1.

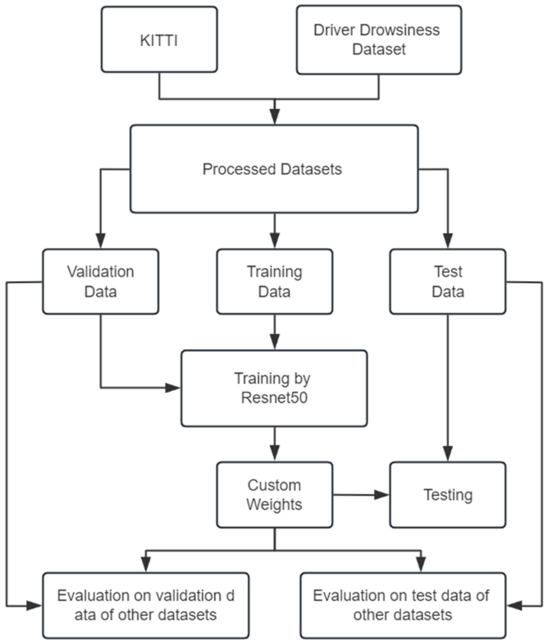

Figure 6: Training and assessment process (Photo/Picture credit: Original).

6.2 Experimental Results

The performance performance of the model is shown in Table 2 below, and by analyzing the model's accuracy, precision, and recall, it can be seen that the model performs better in detecting driver fatigue.

The model effectively distinguishes between fatigued and non-fatigued driving behaviors in most cases. The system produces fewer false alarms when detecting fatigued driving, reducing the risk of incorrectly detecting normal driving as fatigued driving. For a fatigue driving warning system, too many false alarms may affect the driver's trust. Strong sensitivity in detecting fatigued driving behaviors, effective in capturing most fatigued driving behaviors, and the model performs well in balancing precision and recall. Balancing forecast reduction and capturing more fatigued driving behaviors.

Table 1: Hyperparameter settings for model training.

Hyperparamete-rization	Setting	Instructions
learning rate	0.005	The step size of the model weight adjustment allows for fast convergence of the training
optimizer	Adam	Lower loss values to reduce the gap between model output values and real labels
weight decay	0.0001	Preventing overfitting
batch size	32	Number of samples entered into the model in each iteration of training
loss function	Cross-EntropyLoss	Calculate the gap between the forward computed result and the true value for each iteration of the neural network

Table 2: Model performance.

Evaluation	Value
Accuracy	87.63%
Precision	89.69%
Recall	86.14%
F1-score	87.88%

7 CONCLUSIONS

This work focuses on improving the fatigue driving recognition model and designing an alert system for fatigue driving in the IoV system. There are three main advantages of this work. The first is combining image and trajectory information for fatigue driving. The facial expression of the driver is captured by image information, while the driving pattern of the vehicle is analyzed by trajectory information, which improves the accuracy of identifying the fatigued driving state. Second, the fusion of the two kinds of data at the model level is achieved, which does not rely on manually setting thresholds for judgment. Compared with the existing multimodal data decision-making algorithms, the recognition process is more flexible and stable. Finally, by combining the fatigue driving recognition system with IoV, the design of a fatigue driving warning system in-vehicle and inter-vehicle is proposed, which provides a new solution for the application of fatigue driving recognition algorithms. Based on this study, the future research directions are as follows. Further training of the model, the existing model version is limited by time and resources, and the accuracy can be improved. Experiment with other model frameworks and data sources and train on more datasets to find a better solution for fatigued driving recognition. Continue to complete the development of the simulation application for the IoV system in the short-term plan and further apply the design to a real IoV system in the long-term plan.

AUTHORS CONTRIBUTION

All the authors contributed equally and their names were listed in alphabetical order.

REFERENCES

Tian, Y., & Cao, J., 2021. *Fatigue driving detection based on electrooculography: A review. EURASIP Journal on Image and Video Processing*, 2021(1), 33.

Chen, L., Xin, G., Liu, Y., Huang, J., & Chen, B., 2021. *Driver Fatigue Detection Based on Facial Key Points and LSTM. Sec. and Commun. Netw*., 2021.

Liu, L., Wang, Z., & Qiu, S., 2020. *Driving Behavior Tracking and Recognition Based on Multisensor Data Fusion. IEEE Sensors Journal*, 20(18), 10811–10823.

Tang, X., 2024. *Research on Dangerous Driving Behavior Recognition based on the fusion of localization and vision technology.* [Unpublished master's thesis]. Nanjing University of Science and Technology Awarding the Degree.

Abbas, Q., & Alsheddy, A., 2020. *Driver Fatigue Detection Systems Using Multi-Sensors, Smartphone, and Cloud-Based Computing Platforms: A Comparative Analysis. Sensors (Basel, Switzerland)*, 21.

He, K., Zhang, X., Ren, S., & Sun, J., 2015. *Deep Residual Learning for Image Recognition. 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 770-778.

Geiger, A., Lenz, P., & Urtasun, R., 2012. *Are we ready for autonomous driving? The KITTI vision benchmark suite. 2012 IEEE Conference on Computer Vision and Pattern Recognition*, 3354-3361.

Gupta, I., Garg, N., Aggarwal, A., Nepalia, N., & Verma, B., 2018. *Real-Time Driver's Drowsiness Monitoring Based on Dynamically Varying Threshold. 2018 Eleventh International Conference on Contemporary Computing (IC3)*, 1-6.

Sommer, C., German, R., & Dressler, F., 2011. *Bidirectionally Coupled Network and Road Traffic Simulation for Improved IVC Analysis. IEEE Transactions on Mobile Computing*, 10, 3-15.

Behrisch, M., Bieker, L., Erdmann, J., & Krajzewicz, D., 2011. *SUMO - Simulation of Urban MObility. The Third International Conference on Advances in System Simulation*, pp. 63-68

Varga, A., & Hornig, R., 2008. *An overview of the OMNeT++ simulation environment. International ICST Conference on Simulation Tools and Techniques.*