ADVANCES IN TRANSPORTATION STUDIES
An International Journal

Guest Editor: Y.C. Zhou and D. Yu

2025 Special Issue, Vol.1

Contents

aracne

Advances in Transportation Studies an international Journal 2025 Special Issue, Vol. 1

ISBN 979-12-218-2056-0 – ISSN 1824-5463 – DOI 10.53136/97912218205601 – pag. 3-16

A decision support system for predicting and relieving traffic congestion in urban road networks from the perspective of connected vehicles

Y.L. Zhang[1,2,*] X.M. Wang[1]

[1]*College of Civil Engineering, Southwest Jiaotong University Hope College, Chengdu 610400, Sichuan, China*
[2]*College of Education, Shinawatra University, Bangkok 10400, Thailand*
**corresponding author; email:zhangyanling1701@swjtuhc.cn*

Abstract

This paper proposed a decision support system for predicting and alleviating urban road network traffic congestion from the perspective of the Internet of Vehicles. Firstly, the Internet of Vehicles utilizes GPS and OBU to collect urban road network data, which is then fused through Kalman filtering to extract congestion features and predicted using a CNN model. Secondly, design a decision support system for traffic congestion management, covering multiple modules such as data collection and preprocessing. Finally, a decision model is constructed with the goal of minimizing congestion time and maximizing traffic capacity, and solved using particle swarm optimization algorithm to achieve decision support for traffic congestion mitigation. The experimental results show that the proposed method has low traffic congestion time and road capacity utilization, indicating that it can provide scientific decision support for traffic congestion alleviation.

Keywords - internet of vehicles, urban road network, traffic congestion prediction, decision support, road network data fusion

1. Introduction

The urban road network is complex, with large traffic flow and spatiotemporal dynamics, leading to increasingly frequent and complex traffic congestion [15, 9]. The traditional methods for predicting and alleviating traffic congestion mainly rely on fixed monitoring equipment, which makes it difficult to accurately grasp the dynamic changes in traffic conditions, resulting in lag and inefficiency in traffic congestion mitigation decisions [16, 4, 11]. Therefore, it is necessary to study the urban road network traffic congestion prediction and diversion decision support system, which can help improve the accuracy and timeliness of prediction, achieve more scientific and efficient traffic diversion decisions, and enhance the efficiency of urban traffic operation and the quality of residents' travel.

Currently, the research on urban road network traffic congestion prediction and diversion decision support system has received widespread attention, and scholars have conducted extensive research and achieved certain results. Liang and Peng [6] proposed a traffic congestion prediction method that considers the time-varying topology of the road network. This method achieves flexible learning of the preset static structure of the road network by constructing an adaptive auxiliary adjacency matrix, Meanwhile, utilizing adaptive embedding adjacency matrix to deeply explore potential information in the static structure of the road network. In addition, this method uses gate controlled loop units to accurately extract the temporal features of road network traffic flow. Research has confirmed that this method can efficiently identify and locate changes in road network topology, demonstrating strong adaptive performance for dynamic changes in road network topology. However, due to the inclusion of multiple adaptive matrices and gating loop units in the model, the tuning process of these parameters may be cumbersome, and inappropriate parameter settings may affect the predictive performance of the model, resulting in high road capacity utilization. Li and Zhao [7] proposed a deep learning based urban traffic congestion prediction model, which utilizes convolutional neural networks to extract spatial features, Subsequently, the representation capability of the model was further enhanced through the use of fully connected neural networks (FNN), On this basis, the model introduces a similarity location encoding mechanism to integrate location information into traffic data, Ultimately, the Transformer network is utilized to capture time-dependent features in traffic data. Research has confirmed that this method has high flexibility and scalability, and can adapt to the needs of different traffic scenarios. However, the predictive performance of this method is affected by the accuracy and completeness of the input data. If there are problems such as noise, missing or outlier values in the input data, it may affect the predictive performance of the model and cause prolonged traffic congestion. Li et al. [10] constructed a dual layer optimization model for evacuation and scheduling collaborative decision-making. In the upper layer model, evacuation routes were selected by considering travel distance. The lower level model aims to minimize travel time. To achieve these goals, the model sets multiple constraint conditions and optimizes the vehicle scheduling scheme in combination with time window requirements, At the same time, plan traffic evacuation strategies based on the maximum flow limit of the road section. In order to efficiently solve the lower level collaborative decision-making model, genetic algorithm is introduced, which incorporates gene expression regulation strategy coding. The calculation results show that this method adopts a two-layer optimization model, which can simultaneously consider the selection of paths and the optimization of vehicle scheduling, making the entire evacuation process more efficient and orderly. However, this two-layer optimization model involves multiple variables and constraints, which may result in a large amount of computation during the model establishment and solving process, leading to prolonged traffic congestion and delaying the optimal evacuation timing.

From the above analysis, it can be seen that although research on predicting and alleviating traffic congestion in urban road networks has achieved certain results, it still faces problems such as low accuracy in traffic congestion prediction and long evacuation time. Therefore, from the perspective of the Internet of Vehicles, the study of urban road network traffic congestion prediction and diversion decision support system aims to improve prediction accuracy and diversion efficiency, ensuring smooth and unobstructed traffic flow.

2. Prediction of traffic congestion in urban road network

2.1. Urban road network data collection

The Internet of Vehicles can obtain a large amount of real-time, extensive, and accurate traffic information. With the help of various sensors and communication devices installed on vehicles, it can directly collect micro level dynamic data such as real-time position, speed, and driving direction of vehicles. These data can reflect the actual operating status of vehicles on the road in detail [3, 13]. At the same time, the Internet of Vehicles has a wide coverage area and can obtain information on different areas and roads in the city, providing a rich data foundation for a comprehensive understanding of the traffic conditions of the urban road network.

This article mainly collects urban road network data through Global Positioning System (GPS) and Vehicle Mounted Unit (OBU). The specific collection methods are as follows:

(1) Global Positioning System (GPS)

GPS determines the position of a vehicle by receiving signals from multiple satellites, and the GPS receiver on the vehicle continuously monitors signals from at least four satellites to obtain the vehicle's three-dimensional coordinates.

During the data collection process, the collection frequency of GPS data can be flexibly set according to specific application requirements. For example, in high-precision traffic congestion monitoring scenarios, the collection frequency may be set to once per second or even higher to obtain real-time changes in vehicle location information. In some long-term traffic flow analysis applications that require relatively low accuracy, the collection frequency can be set to once per minute.

The GPS receiver records the collected location data in a preset format and transmits the data in real-time to the data center or cloud platform through the vehicle's built-in communication module (such as 4G, 5G, and other wireless communication networks).

(2) Onboard Unit (OBU)

The OBU communicates with the roadside units (RSUs) installed along the road. When the vehicle reaches the coverage area of the RSU, the OBU automatically establishes a connection with the RSU and exchanges data. The information it collects specifically includes:

1) Speed acquisition: The OBU obtains real-time speed information of the vehicle through its electronic control system. The speed sensor in the electronic control system measures the wheel speed and converts it into a speed signal. The OBU reads the signal and records the vehicle's speed data.

2) Driving direction collection: Using the vehicle's inertial measurement unit (IMU), the OBU can obtain the vehicle's driving direction information. IMU measures the acceleration and angular velocity of the vehicle, and combines algorithms to calculate the vehicle's attitude and direction of travel.

3) Other information collection: In addition to location, speed, and driving direction, OBU can also collect other relevant information of the vehicle, such as mileage, engine speed, braking status, etc.

OBU will organize and package various collected data, and then upload the data to the data center through wireless communication network.

2.2. Urban road network data fusion processing

On this basis, the Kalman filtering method is used to fuse GPS positioning data with vehicle driving status data collected by OBU. This multi-sensor data fusion method can effectively compensate for the limitations of a single sensor, improve the accuracy and stability of data collection, and help to more accurately analyze traffic flow changes [12, 5], thereby achieving effective prediction of traffic congestion.

Establish a motion state model of the vehicle and describe its motion state on a two-dimensional plane. Assuming that the motion of a vehicle can be approximated as a combination of uniform linear motion and uniform acceleration linear motion, its state vector X can be expressed as:

$$X = \begin{bmatrix} x \\ y \\ v_x \\ v_y \\ a_x \\ a_y \end{bmatrix} \quad (1)$$

where x and y respectively represent the horizontal and vertical coordinates of the vehicle, v_x and v_y represent the speed components of the vehicle, a_x and a_y represent the acceleration components of the vehicle.

Based on the kinematic principles of vehicles, establish a state transition equation to describe the relationship between the vehicle's state and time:

$$X_k = F_{k-1}X_{k-1} + B_{k-1}u_{k-1} + w_{k-1} \quad (2)$$

where X_k represents the vehicle state vector at time k, F_{k-1} is the state transition matrix, B_{k-1} is the control input matrix, u_{k-1} is the control input vector, w_{k-1} is the process noise vector.

Establish observation models for GPS data and OBU data separately, and describe the relationship between sensor measurements and the true state of the vehicle. GPS measures the position information of vehicles, and its observation vector z_{GPS} can be expressed as:

$$z_{GPS} = \begin{bmatrix} \lambda_{GPS} \\ \varphi_{GPS} \end{bmatrix} \quad (3)$$

The observation equation is:

$$z_{GPS,k} = H_{GPS}X_k + v_{GPS,k} \quad (4)$$

where H_{GPS} is the GPS observation matrix, which maps the vehicle state vector to the GPS observation space, $v_{GPS,k}$ is the GPS observation noise vector.

OBU measures information such as vehicle speed and direction of travel, and its observation vector z_{OBU} can be expressed as:

$$z_{OBU} = \begin{bmatrix} v_{OBU} \\ \theta_{OBU} \end{bmatrix} \quad (5)$$

The observation equation is:

$$z_{OBU,k} = H_{OBU}X_k + v_{OBU,k} \quad (6)$$

where H_{OBU} is the OBU observation matrix, which maps the vehicle state vector to the OBU observation space, $v_{OBU,k}$ is the OBU observation noise vector.

According to the state transition equation, predict the state of the vehicle:

$$\hat{X}_k = F_{k-1}\hat{X}_{k-1} + B_{k-1}u_{k-1} \quad (7)$$

Calculate the covariance matrix of the predicted state:

$$\hat{P}_k = F_{k-1}P_{k-1} + F_{k-1}^T + Q_{k-1} \quad (8)$$

Calculate Kalman gain:

$$K_k = P_k H_k^T (H_k P_k H_k^T + R_k)^{-1} \quad (9)$$

where K_k is the Kalman gain at time k, H_k is the fused observation matrix, which is composed of GPS and OBU observation matrices, R_k is the fused observation noise covariance matrix, which is combined based on the GPS and OBU observation noise covariance matrices.

Update the state prediction based on the observed values:

$$\hat{X}_k = X_k + K_k(z_k - H_k X_k) \quad (10)$$

where $\hat{X}_k$ represents the estimated state value at time k, z_k is the fused observation vector, composed of GPS observation vector and OBU observation vector.

Update the covariance matrix of state estimation:

$$P_k = I - K_k H_k \quad (11)$$

where P_k represents the covariance matrix of state estimation at time k, I is the identity matrix.

By iteratively calculating the Kalman filter, the estimated state value $\hat{X}_k$ of the vehicle is continuously updated. Output the estimated vehicle position, speed, direction of travel, and other information as fusion results to obtain more accurate vehicle trajectory and status information.

Through the above data fusion process, GPS positioning data can be effectively fused with vehicle driving status data collected by OBU, improving the accuracy and reliability of the data and providing better data support for predicting traffic congestion in urban road networks.

2.3. Traffic congestion feature extraction

The accuracy and reliability of transportation network data have been further improved through data fusion processing. However, the original data is relatively complex, and direct use for prediction is inefficient and ineffective. By extracting traffic congestion features, massive data can be transformed into representative and indicative key information, which can more clearly reflect the traffic congestion situation and provide strong support for accurate prediction in the future. Therefore, traffic congestion features are extracted from the fusion results of traffic network data [1, 8, 14]. The specific extracted features include the average speed of road sections, vehicle density, and speed standard deviation. These features depict the characteristics of traffic congestion from different perspectives, which helps to accurately grasp the traffic congestion situation and make effective predictions.

By utilizing the aforementioned data collection and fusion techniques, accurate vehicle travel paths and status details can be obtained, covering key information such as the current position, speed, and direction of the vehicle. Next, calculate the average speed of vehicles on the road section. Assuming there are n vehicles on the road section and the speed of the i-th vehicle is V_i, the expression for the average speed V_{avg} of the road section is:

$$V_{avg} = \frac{1}{n}\sum_{i=1}^{n} V_i \tag{12}$$

For vehicle density, assuming the length of the road segment is L, the expression for vehicle density ρ is:

$$\rho = \frac{n}{L} \tag{13}$$

The standard deviation of speed reflects the degree of dispersion of vehicle speed on a road section, and the expression for the standard deviation of speed σ_v is:

$$\sigma_v = \sqrt{\frac{1}{n}\sum_{i=1}^{n}\left(V_i - V_{avg}\right)^2} \tag{14}$$

2.4. Traffic congestion prediction

Based on the extracted traffic congestion features, a convolutional neural network (CNN) is used to predict urban road network traffic congestion. The extracted traffic congestion features, namely the average speed V_{avg} of the road section, vehicle density ρ, and speed standard deviation σ_v, will be used as inputs for the model. Assuming the urban road network is divided into M segments, at each time step t, the input data can be represented as a three-dimensional tensor $X_t = R^{M\times 3\times 1}$, where 3 represents the number of features (average speed, vehicle density, and speed standard deviation), and 1 represents the number of channels.

The output of the model is the predicted result of whether traffic congestion will occur on each road segment at a certain time step in the future, which can be represented by an N-dimensional vector $Y_{t+\Delta t} \in R^N$, where $Y_{t+\Delta t}[i]$ represents the congestion prediction result of the i-th road segment at a future time step $t + \Delta t$, and the value range can be [0,1].

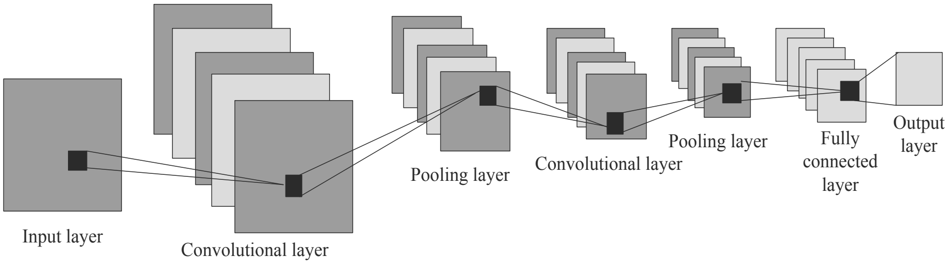

Fig. 1 - Schematic diagram of CNN model

Build a CNN model that includes convolutional layers, pooling layers, and fully connected layers, as shown in Figure 1.

In Figure 1, convolutional layers are used to extract local patterns and features of input features. Assuming that the size of the convolution kernel in the convolutional layer is K, the stride is s, and the number of convolution kernels is C. For input data X_t, the convolution operation can be expressed as:

$$Z_t^l = f\left(\sum_{i=0}^{K-1}\sum_{j=0}^{K-1} X_t[i + m \times s, j + n \times s, c] \times W^l[i,j,c] + b^l\right) \quad (15)$$

where Z_t^l represents the output feature map of the l-th convolution kernel at time step t, f is the activation function, m is the position index of the convolution kernel on the feature map, c is the input channel index, W^l is the weight matrix of l convolutional kernels, b^l is the bias term.

The pooling layer is used to reduce the dimensionality of feature maps, decrease computational complexity, and extract more representative features. Assuming the pooling window size of the pooling layer is p and the stride is q, for the output feature map Z_t^l of the convolutional layer, the maximum pooling operation can be expressed as:

$$P_t^l[m,n] = max_{i=0,\dots,p-1} max_{j=0,\dots,p-1} Z_t^l[m \times q\_i, n \times q + j] \quad (16)$$

After the convolution and pooling process, the feature map is transformed into a one-dimensional vector form, which is then passed to the fully connected layer for performing classification and prediction tasks. Assuming the weight matrix of the fully connected layer is W_{fc} and the bias term is b_{fc}, the output of the fully connected layer can be expressed as:

$$O_t = g\left(W_{fc} \times F(P_t) + b_{fc}\right) \quad (17)$$

where $F(P_t)$ represents flattening the output feature map P_t of the pooling layer into a one-dimensional vector, g is the activation function, O_t is the output of the fully connected layer.

To optimize the performance of the CNN model, a loss function is set to quantify the error between the model's predicted values and the true labels. Assuming the true label is $Y_{t+\Delta t}$ and the model prediction result is O_t, the cross entropy loss function can be expressed as:

$$\mathrm{L} = -\frac{1}{\mathrm{M}}\sum_{\mathrm{i}=1}^{\mathrm{M}}\left(Y_{t+\Delta t}[i]log(O_t[i]) + (1 - Y_{t+\Delta t}[i])log(1 - O_t[i])\right) \quad (18)$$

By using the backpropagation mechanism, the gradient of the loss function relative to the model parameters is calculated, and the parameters are updated accordingly, aiming to minimize the loss and gradually approach the real situation of the model's traffic congestion prediction, thereby enhancing the accuracy of the prediction.

3. Design of decision support system for traffic congestion relief

By extracting the characteristics of traffic congestion in urban road networks (average speed of road sections, vehicle density, and speed standard deviation) and using CNN models for analysis and prediction, it is possible to know in advance the likelihood of traffic congestion occurring in each road section at future time steps. Based on this, a decision support system for traffic congestion mitigation is constructed, which analyzes and simulates different congestion scenarios to effectively alleviate traffic congestion and improve the overall operational efficiency of urban transportation.

3.1. Overall system architecture and hardware design

The decision support system for traffic congestion mitigation mainly consists of a data collection and preprocessing module, a traffic congestion prediction module, a congestion scenario analysis and simulation module, a mitigation strategy generation module, and a decision support and display module. Each module works together to achieve real-time monitoring, prediction, and decision support for traffic congestion. The overall system architecture is shown in Figure 2.

(1) Data collection and preprocessing module: Responsible for collecting raw traffic data from various traffic monitoring devices for each section of the urban road network, including average speed, vehicle density, and speed standard deviation. Subsequently, these data are preprocessed to ensure data quality and provide accurate and consistent data for subsequent modules.

The hardware of the data acquisition and preprocessing module mainly consists of a geomagnetic sensor (GM-100), a video surveillance camera (Hikvision DS-2CD3T46WD-L), and a data acquisition terminal (DTU-8000). The geomagnetic sensor is installed under the road surface, which can accurately detect the passage of vehicles to obtain flow information and adapt to harsh environments, The video surveillance camera is installed above the road and has high-definition resolution, wide-angle view, and intelligent image analysis function, which can recognize various vehicle information, The data collection terminal is responsible for collecting sensor and camera data, and transmitting it to the data processing center through a wireless communication network to ensure real-time and complete data.

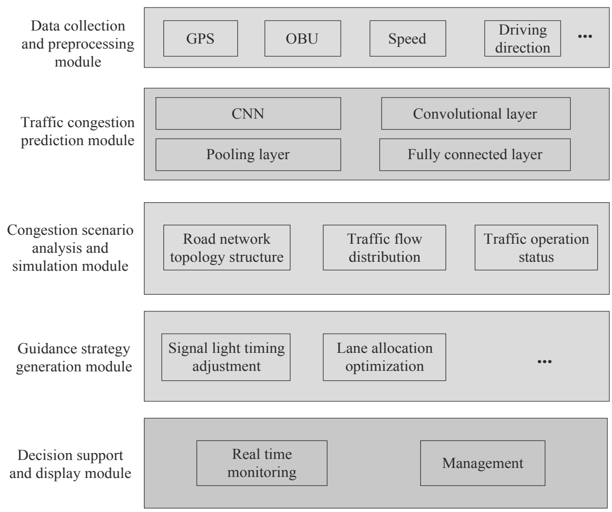

Fig. 2 - Overall system architecture diagram

(2) Traffic congestion prediction module: Using a convolutional neural network (CNN) model, preprocessed traffic congestion feature data is used as input. Through the operation of convolutional layers, pooling layers, and fully connected layers, the probability of traffic congestion occurring on each road segment at a specific time step in the future is output, providing key prediction basis for congestion scenario analysis and mitigation strategy formulation.

The hardware core of the traffic congestion prediction module is a server (Dell PowerEdge R740) paired with a graphics processing unit (GPU, NVIDIA Tesla V100). The server is equipped with high-performance CPUs (such as Intel Xeon Gold 6248R) and large capacity memory (128GB DDR4), as well as high-speed storage devices for storing data and model parameters, meeting the complex computing and storage requirements of CNN models, GPUs have powerful parallel computing capabilities that can significantly accelerate model training and inference processes, enabling real-time prediction of traffic congestion.

(3) Congestion scenario analysis and simulation module: Based on the predicted results of traffic congestion, combined with information such as urban road network topology and traffic flow distribution, a traffic flow model is constructed to analyze and simulate the traffic operation under different congestion scenarios, clarify the propagation rules and impact range of congestion, and provide support for generating reasonable diversion strategies.

The congestion scenario analysis and simulation module uses a high-performance workstation (HP Z8 G4) as hardware support. The workstation is equipped with multi-core CPUs (such as Intel Core i9-10980XE), large capacity memory (64GB DDR4), and professional graphics cards (NVIDIA Quadro RTX 8000), as well as high-speed hard drives (NVMe SSD), which can efficiently run congestion scenario analysis and simulation software, meet the calculation and visualization needs of complex traffic flow models, and provide strong support for analyzing congestion propagation patterns and ranges.

(4) Traffic diversion strategy generation module: Based on the analysis and simulation results of congestion scenarios, intelligent algorithms and optimization models are used to comprehensively consider factors such as traffic congestion, road resource utilization efficiency, and social costs, and generate traffic congestion diversion strategies that include signal timing adjustment, lane allocation optimization, traffic control, and guidance information release.

The diversion strategy generation module relies on a cluster of computing servers (consisting of multiple Lenovo SR650 servers) to provide powerful computing power. Each server is equipped with a high-performance CPU (AMD EPYC 7742), large capacity memory (256GB DDR4), and high-speed network interfaces. Through distributed computing, complex intelligent algorithms and optimization model computing tasks are allocated to multiple nodes for parallel processing, thus comprehensively considering multiple factors and efficiently generating traffic congestion diversion strategies that include signal timing adjustments.

(5) Decision support and display module: Provides decision support and visual display interface for traffic management departments, presenting traffic congestion prediction results, congestion scenario analysis reports, and diversion strategy solutions in an intuitive form. At the same time, it supports decision-makers to interactively adjust and optimize diversion strategies, achieving real-time monitoring and dynamic management of traffic congestion.

The hardware of the decision support and display module mainly includes a display terminal (Dell U4919DW) and interactive devices (such as touch all-in-one machines, Hisense H65E3G). The display terminal adopts an ultra wide screen design, with high resolution (5120 × 1440) and excellent color performance, and can display multiple traffic information windows simultaneously, The touch all-in-one machine supports multi touch operation, equipped with high-performance processors and rich interfaces, and can communicate in real-time with the system server, making it convenient for decision-makers to adjust and optimize traffic diversion strategies through touch screen interaction, achieving real-time monitoring and dynamic management of traffic congestion.

3.2. Decision making for traffic congestion relief

The above system hardware modules provide real-time and accurate collection capabilities for urban road network traffic data. The data preprocessing module standardizes the collected data, and the congestion prediction module provides the future congestion probability of each road section [2]. On this basis, a decision model for traffic congestion mitigation is constructed to achieve the goals of minimizing congestion time and maximizing traffic capacity, while considering constraints such as road capacity and traffic rules. By establishing and solving specific models such as traffic flow allocation models and signal timing optimization models, scientific and effective decision support is provided for traffic congestion mitigation.

In traffic congestion management, establish objective functions to minimize congestion time and maximize traffic capacity, and optimize the operational efficiency of the transportation system, as follows:

If T_{ij} represents the congestion time of road segment i during time period j, the objective function for minimizing congestion time can be expressed as:

$$maxZ_1 = \sum_{i=1}^{n}\sum_{j=1}^{m} T_{ij} \tag{19}$$

In the process of traffic congestion alleviation, assuming C_{ij} represents the actual traffic capacity of road section i during time period j, and C_i^{max} represents the maximum traffic capacity of road section i, the objective function of maximizing traffic capacity can be expressed as:

$$maxZ_2 = \sum_{i=1}^{n}\sum_{j=1}^{m} \frac{C_{ij}}{C_i^{max}} \tag{20}$$

By weighting and combining the above two objectives, a comprehensive objective function is obtained:

$$minZ = w_1 Z_1 \text{-} w_2 Z_2 \tag{21}$$

where w_1 and w_2 are the weight coefficients for minimizing congestion time and maximizing traffic capacity, respectively, and $w_1 + w_2 = 1$.

Set the following constraints for the established objective function:

The number of vehicles on each road section cannot exceed its road capacity. Let q_{ij} represent the vehicle flow of road section i during time period j, and Q_i^{max} represent the maximum road capacity of road section i. Therefore, the road capacity constraint can be expressed as:

$$q_{ij} \leq Q_i^{max} \quad \forall i = 1,2,\ldots,n; j = 1,2,\ldots,m \tag{22}$$

If x_{ij} represents the traffic rule variables related to road segment i during time period j (such as signal duration, lane allocation, etc.), and x_{ij}^{max} and x_{ij}^{min} represent their lower and upper limits, respectively, then the traffic rule constraints can be expressed as:

$$x_{ij}^{min} \leq x_{ij} \leq x_{ij}^{max} \tag{23}$$

In the urban road network, the traffic flow entering and leaving each node (intersection) should be conserved. If q_{ij}^{in} and q_{ij}^{out} respectively represent the traffic flow of vehicles entering and leaving node i during time period j, then the traffic conservation constraint can be expressed as:

$$\sum_{k\in\Omega_i^{in}} q_{kj}^{out} = \sum_{l\in\Omega_i^{out}} q_{lj}^{in} \tag{24}$$

where Ω_i^{in} and Ω_i^{out} respectively represent the sets of input and output edges connected to node i.

Based on the above analysis, a traffic congestion mitigation decision model is constructed, including a traffic flow allocation model and a signal timing optimization model.

The purpose of constructing a traffic flow allocation model is to reasonably allocate vehicles to different road sections based on the traffic conditions and demand, in order to achieve the goals of minimizing congestion time and maximizing traffic capacity. Let f_{ij} represent the traffic flow from source node i to destination node j, e_{ij} represent the traffic cost of road segment ij, and b_{ij} represent the flow allocation variables on road segment ij. The traffic flow allocation model can be expressed as:

$$\min Z = \sum_{\mathrm{i}} \sum_{\mathrm{j}} e_{ij} \, b_{ij} \tag{25}$$

The signal timing optimization model is mainly aimed at optimizing the signal timing of intersections. Let t_i represent the signal cycle duration of intersection i, g_{ij} represent the green light duration of phase j at intersection i, q_{ij} represent the traffic flow of phase j at intersection i, and h_{ij} represent the saturation flow of phase j at intersection i (i.e. the maximum number of vehicles passing through per unit time). The signal timing optimization model can be expressed as:

$$\min Z = \sum_{\mathrm{i}} \sum_{\mathrm{j}} \frac{q_{ij}(t_i - g_{ij})}{2h_{ij}} \tag{26}$$

In practical applications, particle swarm optimization algorithm is used to solve the above evacuation decision model. Through continuous iterative optimization, the optimal evacuation strategy that meets the constraints is found to achieve effective evacuation of traffic congestion. The model solving process is shown in Figure 3.

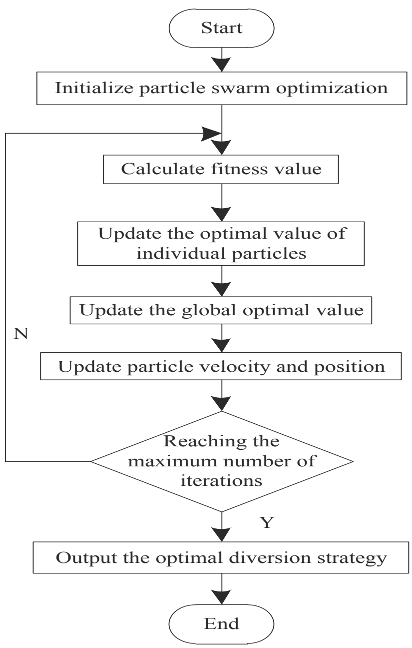

Fig. 3 - Flow chart for solving the decision model of traffic congestion relief

4. Experiments and results analysis

4.1. Experimental plan design

(1) Experimental area information

Taking a certain city as the experimental area, the total area of the area is about 150 km2, including 30 main roads and 50 branch roads, with a total road length of about 200 km. There are 5 large commercial centers in the area, with an average daily passenger flow of 100000 people, 10 residential communities with a permanent population of approximately 80000 people, There are three schools with high traffic flow for students and parents during daily commuting hours.

(2) Experimental sample data source

There are a total of 100 traffic monitoring points set up in the region, including 60 road sensors, which are used to collect real-time data such as average speed and vehicle density of road sections, 30 cameras that can assist in analyzing vehicle travel trajectories and traffic flow conditions, Vehicle mounted GPS devices cover approximately 50000 motor vehicles, providing real-time location and speed information of the vehicles. The data collected by traffic monitoring points is recorded every 5 minutes, 24 hours a day without interruption, forming a rich sample of basic data such as traffic flow, speed, density, etc. The camera uses image recognition technology to analyze vehicle types, lane occupancy, and other factors, and adds them to the sample data. The in car GPS device uploads real-time dynamic information of the vehicle, including driving path, speed changes, etc., further enriching the dimensions of the sample. At the same time, combined with weather data provided by the meteorological department, such as temperature, humidity, whether there is rainfall, and special event records from the traffic management department, such as traffic accidents, road construction, etc., a complete experimental sample dataset is formed. These data cover traffic conditions under different time periods, weather conditions, and traffic events, providing a comprehensive and authentic sample basis for the experimental verification of traffic congestion prediction and evacuation decision support systems.

(3) Experimental indicators

Select congestion time and road capacity utilization rate as experimental indicators to verify the application effects of the proposed method, the method of Liang and Peng [6] and Li and Zhao [7]. Among them, the lower the congestion time value, the more effective the diversion strategy is, and the lower the road capacity utilization rate, the more effective the diversion strategy is in avoiding road overload situations.

4.2. Analysis of experimental results

The comparison results of congestion time between the proposed method, the method of Liang and Peng [6] and Li and Zhao [7] are shown in Table 1.

Tab. 1 - Comparison results of congestion time/min

Experiment number	Proposed method	Liang and Peng [6] method	Li and Zhao [7] method
1	35.2	42.8	48.5
2	33.9	45.1	50.2
3	36.7	43.9	49.8
4	34.5	44.6	47.9
5	37.1	46.3	51.4

From the overall situation of multiple sets of data, the proposed method performs the best in reducing congestion time. Taking Experiment No. 1 as an example, the congestion time of the proposed method is 35.2 minutes, the method in Liang and Peng [6] is 42.8 minutes, and the method in Li and Zhao [7] is 48.5 minutes. The congestion time of Liang and Peng [6] is 7.6 minutes longer than the proposed method, while the method in Li and Zhao [7] is 13.3 minutes longer. This difference is also evident in other experimental sequences, indicating that the proposed method can more effectively optimize traffic diversion strategies in dealing with traffic congestion problems, thereby shortening the overall congestion time. Meanwhile, observing the congestion time data of the proposed method, the fluctuation is relatively small, indicating that the method has good stability. In contrast, the congestion time fluctuation range of Liang and Peng [6] and Li and Zhao [7] is relatively large, which means that these two methods have poor adaptability in different traffic scenarios and are easily affected by various factors, resulting in unstable effects. The proposed method has better application effects in controlling congestion time compared to the methods in Liang and Peng [6] and Li and Zhao [7], not only effectively reducing congestion time, but also having stronger stability.

The comparison results of road capacity utilization rates between the proposed method, the method of Liang and Peng [6] and Li and Zhao [7] are shown in Figure 4.

From the overall situation of multiple sets of data, the proposed method has a relatively low utilization rate of road capacity, which means that this method can allocate road resources more reasonably in the process of traffic diversion, avoid excessive road congestion, and make the use of road capacity more balanced and efficient. The higher road capacity utilization rates of Liang and Peng [6] and Li and Zhao [7] indicate that these two methods may have certain irrationality in resource allocation, leading to roads approaching or exceeding their carrying capacity. From this, it can be seen in actual traffic scenarios, the proposed method can fully consider various factors such as changes in traffic flow on different road sections, differences in road carrying capacity, and dynamic characteristics of traffic demand, scientifically allocating traffic flow to each road section, avoiding a large number of vehicles from rushing into certain specific road sections, and effectively preventing excessive congestion on the road. This balanced allocation of road resources enables each road to operate efficiently within its reasonable carrying capacity, resulting in a more balanced and efficient use of road capacity, greatly improving the stability and smoothness of the entire transportation system.

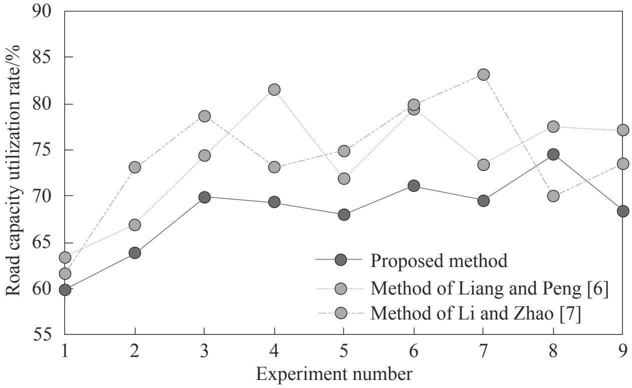

Fig. 4 - Comparison results of road capacity utilization rate

5. Conclusion

The research results of the urban road network traffic congestion prediction and diversion decision support system proposed in this article from the perspective of the Internet of Vehicles are as follows:

(1) By combining the Internet of Vehicles with GPS and OBU to collect data, Kalman filtering is used for fusion processing, and key congestion features are extracted. The CNN model is used for traffic congestion prediction, which improves the accuracy and reliability of the prediction.

(2) We have constructed a traffic congestion mitigation decision support system that includes multi module collaborative work, with the goal of minimizing congestion time and maximizing traffic capacity. We have built a decision model and used particle swarm optimization algorithm to solve it, achieving comprehensive and intelligent traffic congestion mitigation decision support.

(3) Through experimental verification, this method can effectively reduce traffic congestion time and road capacity utilization, providing scientific decision support for traffic congestion diversion.

References

1. Abdoos, M., Bazzan, A.L.C. (2021) Hierarchical traffic signal optimization using reinforcement learning and traffic prediction with long-short term memory. Expert Systems with Applications, 171(6), PP. 1145-1153.
2. Charalambos, M., Stelios, T., Panayiotis, K., Christos, G. (2022) Joint Route Guidance and Demand Management for Real-Time Control of Multi-Regional Traffic Networks. IEEE Transactions on Intelligent Transportation Systems, 23(7), PP. 8302-8315.
3. Choi, C., Wang, T., Esposito, C., Gupta, B.B., Lee, K. (2021) Sensored Semantic Annotation for Traffic Control Based on Knowledge Inference in Video. IEEE Sensors Journal, 21(10), PP. 11758-11768.
4. Karnati, Y., Sengupta, R., Rangarajan, A., Ranka, S. (2022) Subcycle Waveform Modeling of Traffic Intersections Using Recurrent Attention Networks. IEEE transactions on intelligent transportation systems, 23(3), PP. 2538-2548.
5. Karthik, S., Loui, T.R. (2021) Development of traffic congestion model to evaluate the level of congestion and its associative cost for an urban arterial road: a case study. Journal of The Institution of Engineers (India): Series A, 102(2), PP. 623-632.
6. LIANG, J., PENG, J.H. (2022) Research on an Adaptive Traffic Congestion Prediction Method Considering a Time-varying Network Topology. China Journal of Highway and Transport, 35(09), PP. 157-170.
7. Li, S.W., Zhao, X.H. (2023) Prediction Algorithm of Short-term Traffic Congestion in Urban Area Based on Deep Learning. Science Technology and Engineering, 23(25), PP. 10866-10878.
8. Liu, H., Claudel, C., Machemehl, R. (2021) Robust traffic control using a first order macroscopic traffic flow model. IEEE Transactions on Intelligent Transportation Systems, 23(7), PP. 8048-8062.
9. LIU, X., LI, G.M., WU, X.C., CAI, X.Y., HUANG, J.H., ZHOU, S.Y. (2023) Short-term Traffic Flow Prediction Based on Spatial-temporal Feature Fusion. Journal of Natural Science of Hunan Normal University, 46(02), PP. 137-142.
10. LI, Y.Q., CHEN, X.J., BAN, Y., HAN, X.Q., YANG, J.L. (2024) Collaborative Decision Making for Urban Traffic Evacuation and Emergency Vehicle Dispatch Based on Bi-level Programming. Journal of Wuhan University of Technology: Information & Management Engineering, 46(02), PP. 188-194.
11. Maddineni, V.K., Ansarinejad, M., Ahmed, M.M. (2024). Safety and operational impacts of different Autonomous Vehicle operations on freeway work zones, Advances in Transportation Studies, Special Vol4, PP. 117-126.
12. Mohanty, A., Mohanty, S.K., Jena, B., Mohapatra, A.G., Rashid, A.N., Khanna, A. Gupta, D. (2022) Identification and evaluation of the effective criteria for detection of congestion in a smart city. IET Communications, 16(5), PP. 560-570.

13. Oreg, Z., Shin, H., Tsourdos, A. (2022) Traffic conflict reduction based on distributed stochastic task allocation. The Aeronautical Journal, 126(6), PP. 993-1025.
14. Perumal, S.S.G., Lusby, R.M., Larsen, J. (2022) Electric bus planning & scheduling: A review of related problems and methodologies. European Journal of Operational Research, 301(2), PP. 395-413.
15. XIONG, W.F., DU, S.S., LIU, A., YANG, L. (2023) Prediction Model of Expressway Congestion Spread and Dissipation based on Multi-dimensional Comprehensive Traffic Patterns. Highway, 68(02), PP. 371-378.
16. XU, X.F., YANG, F., LIU, Z.Z., LI, L.J., XIA, Z. (2023) Combination model of short-term traffic flow prediction based on two-level screening mechanism. Control and Decision, 38(01), PP. 84-92.

Advances in Transportation Studies an international Journal 2025 Special Issue, Vol. 1

ISBN 979-12-218-2056-0 – ISSN 1824-5463 – DOI 10.53136/97912218205602 – pag. 17-30

The recognition method of highway vehicle driving dangerous behavior based on CNN-LSTM

Y.S. Yang[1*] F.M. Shang[2]

[1]*School of Electronics and Information, Zhengzhou University of Light Industry, Zhengzhou, 450002, China*
[2]*International Education College, Zhengzhou University of Light Industry, Zhengzhou, 450002, China*
**corresponding author; email:150024766@qq.com*

Abstract

To solve the problem of low accuracy in identifying dangerous driving behaviors of highway vehicles, a method for identifying dangerous driving behaviors of highway vehicles based on CNN-LSTM is proposed. Obtaining driving behavior data of highway vehicles through drones equipped with cameras, and extracting keyframe data of dangerous driving behaviors of highway vehicles, Obtain multidimensional features such as spectral characteristics of dangerous driving behavior and temporal spatial characteristics of driving behavior. Using multi-dimensional feature results as input, utilize CNN to enhance feature information fusion, Use the enhanced features as inputs to the LSTM network for temporal analysis, Using the Softmax function, the output of LSTM is converted into probability distributions of different categories of dangerous behaviors to complete the recognition of dangerous driving behaviors on highways. The results indicate that the proposed method has strong ability to identify driving risks of highway vehicles, with MAP index results reaching up to 98% and FPS index results reaching up to 63%, demonstrating high recognition accuracy and efficiency.

Keywords - expressway, car driving, CNN feature enhancement, LSTM network, identification of dangerous behavior

1. Introduction

As the cornerstone of modern transportation system, highways have significantly improved the efficiency and convenience of traffic circulation [2]. However, with the continuous extension of the total mileage of highways and the sharp increase in vehicle flow, traffic safety challenges are becoming increasingly prominent. In recent years, frequent highway traffic accidents have dealt a heavy blow to people's lives and property safety [9]. In these unfortunate accidents, the driver's dangerous driving behavior plays a crucial role. Therefore, accurately identifying dangerous behaviors during highway driving has extremely important practical significance. From the perspective of drivers, driving behavior on highways is complex and varied. Some drivers may feel fatigued due to prolonged driving, resulting in slower reaction times, impaired judgment, and inability to quickly respond to unexpected situations [6], Some drivers may exceed the speed limit, resulting in poor braking performance when facing emergency situations, Some drivers also illegally change lanes and drive distracted during the driving process, which significantly increases the likelihood of accidents [14, 10]. Due to the differences in driving habits and behavior patterns among different drivers, how to accurately and comprehensively identify these dangerous behaviors has become an urgent problem to be solved.

Li et al. [3] proposed a method for identifying dangerous driving behavior during peak hours based on the Apriori algorithm. Comprehensively collect driving behavior images of drivers through homogeneous matrix transformation, using histogram equalization algorithm to enhance images, By identifying the facial and hand variation features of dangerous driving behavior to locate it, and using the Apriori algorithm to determine the frequent itemsets of dangerous driving phenomena, the principal components of dangerous driving behavior images are extracted to achieve effective recognition. However, the Apriori algorithm cannot meet the requirements for rapid recognition of dangerous driving behaviors, resulting in delayed warnings and ineffective prevention of accidents. Zhang and Wang [16] proposed a deep learning based method for monitoring driver's dangerous driving behavior. Constructing a driving behavior dataset by collecting images and preprocessing the driving behavior data, Adopting the improved filtering redundant box module algorithm in deep learning to process the original information features of images, and introducing ResNet-50 with residual structure in deep learning to extract feature information, Using CRITIC weighting method to determine the weight allocation of various dangerous driving behaviors, and integrating machine vision technology to implement monitoring of drivers' dangerous driving behaviors. The weights calculated by CRITIC weighting method are relatively fixed and difficult to adapt to these dynamic changes in real time, resulting in inaccurate and untimely judgments of dangerous driving behavior in some situations. Chu et al. [1] proposed a road traffic dangerous driving behavior risk recognition method based on sliding window feature fusion. Extract driving behavior features using the sliding window method, Then, use feature multiplication to calculate the temporal features and spatial features for each channel. Using feature fusion methods to achieve spatiotemporal feature fusion, Construct a risk identification function based on the ConvLSTM cascade method and obtain the risk identification results. Fixed window settings are difficult to adapt to the dynamic changes in driving behavior on highways, resulting in poor detection performance. Zhang and Liao [17] collected driver images and implemented low light image enhancement processing using KinD network. Then, capture the facial feature points of the driver to extract the characteristics of dangerous driving behavior. Finally, the YOLOv5s model was optimized to enhance its recognition accuracy and feature capture capability, achieving effective recognition of dangerous driving behaviors. This method has high complexity and may cause recognition delays, making it difficult to issue timely danger warnings.

Therefore, based on the above research, in order to improve the recognition effect, a method for identifying dangerous driving behaviors of highway vehicles based on CNN-LSTM is proposed.

2. Extraction of keyframe data for dangerous driving behaviors of highway vehicles

With the continuous advancement and leap of sensor technology, data collection methods, and AI algorithms, new opportunities and technological means have been provided for identifying dangerous driving behaviors of highway vehicles. Various cameras, millimeter wave radars, etc. can obtain real-time multi-dimensional data such as the vehicle's operating status, surrounding environment information, and driver's operational behavior. Meanwhile, advanced data collection systems can efficiently collect and store this data, providing abundant data resources for subsequent analysis and identification. Based on this background, this article uses drones equipped with cameras to collect data on highway car driving behavior, laying the foundation for obtaining keyframes of dangerous driving behavior in the future [17, 5], thus achieving accurate recognition of dangerous driving behavior.

Based on the adaptive keyframe screening criteria, obtain keyframe data for dangerous driving behaviors of highway vehicles. The main method is to identify the difference between the original reference image and the distorted image by calculating the mean square error of the feature maps of two adjacent frames, in order to evaluate the image quality and more appropriately judge the driver's action changes in the driving behavior sequence.

Firstly, calculate the mean square error between two adjacent frame images. Assuming that the video frame at time t is $f_t(x,y)$ and the video at time t-1 is $f_{t-1}(x,y)$, then the mean square error S_{MSE} is:

$$S_{MSE} = \frac{1}{mn}\sum_{i=1}^{m}\sum_{j=1}^{n}\left(f_t(x,y) - f_{t-1}(x,y)\right)^2 \tag{1}$$

where mn is the size of the feature map.

Then, based on the calculated S_{MSE} between two adjacent points, construct a S_{MSE} curve to more intuitively explain the changes in the target's motion state. Next, select all extreme points from the S_{MSE} curve and store them as candidate keyframes in the combination Z, and construct two adaptive thresholds α_{avg} and β_{avg} based on the candidate keyframe $i\{i\epsilon Z|i = 1,2,\ldots,s\}$:

$$\alpha_{avg} = \sum_{i=1}^{s}\frac{a_i}{s} \tag{2}$$

$$\beta_{avg} == \sum_{i=1}^{s}\frac{b_i}{s} \tag{3}$$

where a_i represents the difference between the previous frame in the pole value domain, b_i represents the difference between the next frame in the pole value domain, and s represents the number of candidate keyframes. When the difference between the current video and the previous video frame is greater than $M\alpha_{avg}$, and the difference between the current video and the next video frame is greater than $M\beta_{avg}$, extract the frame as a driving behavior keyframe.

3. Multi dimensional feature acquisition of keyframe data on dangerous driving behaviors of highway vehicles

Based on the keyframe data of dangerous driving behaviors of highway vehicles obtained above, in order to provide a rich data foundation for comprehensive and accurate identification of dangerous driving behaviors of highway vehicles, multi-dimensional features such as highway vehicle driving danger degree characteristics, dangerous driving behavior spectrum characteristic values, and driving behavior spatial temporal characteristics are obtained.

3.1. Spectral characteristic values of dangerous driving behavior

By classifying conflicts between vehicles based on the driving conditions on highways, and using the obtained keyframe data of dangerous driving behaviors on highways, the traffic conflict ratio is calculated using the following formula to objectively measure the degree of danger of various driving behaviors at highway roundabouts [7, 11, 12]. Then, considering the impact of different types of conflicts on the degree of danger, a weighted method is used to obtain the traffic conflict ratio, as follows:

$$TCR_w = \sum_{i=1}^{n} w_i \times \frac{N_{ci}}{N_t} \tag{4}$$

where TCR_w is the weighted traffic conflict ratio, w_i is the weight of the i-th conflict type, N_{ci} is the number of occurrences of the i-th conflict type, n is the total number of conflict types, and N_t is the total number of vehicles passing through the same range.

The total score of the characteristic values of the dangerous driving behavior spectrum was selected to characterize the overall level of danger that drivers face when driving at the entrance and exit sections of the highway roundabout [13]. The larger the total score of the behavior spectrum, the higher the degree of danger of the driver's driving behavior at the entrance and exit sections, Correspondingly, the lower the overall score of the behavior spectrum, the more likely it is that the driver is in a low-risk driving state and has a higher level of safety at the entrance and exit sections.

Formula (5) is the calculation formula for the proportion of the characteristic parameter value of the driver's dangerous driving behavior exceeding the threshold value $S_i(t)$. It calculates the score of a certain dangerous driving behavior of the driver at a certain time, and uses this score to determine the severity of the driver's dangerous driving behavior. The higher the value, the higher the risk level of the driver's dangerous driving behavior, which is more likely to cause safety accidents. When $N_i < N_i^*$ is present, it indicates that the driver's current driving behavior is in a normal state and does not fall within the scope of dangerous driving behavior research. The calculation formula for $S_i(t)$ is as follows:

$$S_i(t) = \begin{cases} \frac{N_i(t)-N_i^*}{N_i^*} & N_i(t) > N_i^* \\ 0 & N_i(t) < N_i^* \end{cases} \tag{5}$$

The time cumulative average of $S_i(t)$ is calculated to obtain the time average score of the i-th dangerous driving behavior of the driver during the observation period, which is X_i. The calculation formula is as follows:

$$X_i = \frac{1}{T}\sum_{t=0}^{T} S_i(t) \tag{6}$$

where T represents the observation duration.

Due to the different calculation methods of the characteristic parameter values corresponding to each driving behavior, there are differences in the dimensions of the parameter values obtained. To calculate the total score of the driving behavior spectrum characteristic parameters, the problem of different dimensions needs to be solved. Therefore, the time averaged scores of the four dangerous driving behavior feature parameters obtained from the calculation are normalized to ensure that the processed feature parameter values are distributed within the (0,1) interval, thereby calculating the probability distribution of the total score of the dangerous driving behavior spectrum for all drivers, and analyzing the distribution of driving risks for drivers. The formula is as follows:

$$X_{Ni} = \frac{X_i - min(X_i)}{max(X_i) - min(X_i)} \tag{7}$$

By combining the standardized scores of driver behavior characteristic parameters with the weight ratios of each dangerous driving behavior, a weighted calculation method can be used to obtain the total score Q of the driving behavior spectrum characteristic values. The calculation formula is as follows:

$$Q = \sum_{i=1}^{m} \omega_i X_{Ni} \tag{8}$$

where ω_i is the weight of the i-th dangerous driving behavior, the sum of the weights of all dangerous driving behaviors is 1, and m is the number of types of dangerous driving behaviors.

3.2. Extraction of temporal spatial features of driving behavior

When extracting spatial features of driving behavior, the obtained keyframes of dangerous driving behavior of highway vehicles are divided into spatial sequence V and time sequence F, with a total frame count of T frames for each sequence. The spatial sequence V at this time is:

$$\mathrm{V} = [\mathrm{v}_1, \dots, \mathrm{v}_t, \dots, \mathrm{v}_T] \tag{9}$$

where $\mathrm{v}_T \in \mathrm{R}^{w \times h \times c}$ is the t-th frame image in the spatial sequence V, w and h are the width and height of the image, and c is the number of channels in the image color space.

Divide each frame of image sequence V into $\mathrm{K} \times \mathrm{K}$ sub regions, and use the 1-17 bottleneck of MobileNet-2V network to extract the spatial features $F_S(t,k)$ of each frame of image in V [15].