



# Research on Support Vector Machine in Traffic Detection Algorithm

Renjie Zhang<sup>1</sup>, JinYang Huang<sup>2</sup>, YaXin Yan<sup>2</sup>, and Ying Gao<sup>2</sup>(✉)

<sup>1</sup> Hunan Post and Telecommunication College Network Center, Changsha 410015, China

<sup>2</sup> Hunan University of Finance and Economics, Changsha 410205, China

**Abstract.** In order to reduce the impact of traffic incidents on traffic operation, an Automated Traffic Incidents Detection (SVM-AID) algorithm based on Support Vector Machine (SVM) is proposed. This algorithm is of great significance for improving the efficiency of traffic management and improving the effect of traffic management. This article first introduces the background of the topic selection of the Traffic Incidents Detection algorithm, the research status at home and abroad. Then it focuses on the Optimal Separating Hyperplane, linear separable SVM, linear inseparable SVM, nonlinear separable SVM, and commonly used kernel functions. Then, the design flow chart based on the SVM-AID algorithm is given, and the principle component analysis method, Normalization Method and the selection method of Support Vector Machine parameters are introduced. Finally, using the processed data, 4 experiments were designed to test the classification performance of the SVM-AID algorithm, and the influence of each parameter in the SVM on the classification effect was analyzed. The results of the final experiment also showed us the design The effectiveness of the SVM-AID algorithm.

**Keywords:** Traffic Incidents Detection · Support Vector Machine · Kernel Function

## 1 Introduction

Since the world's first highway was completed and opened to traffic in the 1930s, countries around the world have vigorously implemented highway planning and construction. By the end of 2006, the total journey of highways in my country had reached 45,300 km [1]. Expressways can not only promote economic growth and social development, but also bring huge economic and social benefits [2]. However, with the rapid development of expressways, some problems have emerged in traffic operation and handling. With the continuous increase of traffic flow, the frequency of congestion and traffic accidents is also increasing, and the impact of the composition is also increasing [3]. In order to reduce the negative impact of traffic accidents on the operation of expressways, research on Traffic Incidents Detection algorithms, rapid detection of traffic incidents, recognition and recognition of the scene, and the use of traffic flow guidance, can effectively reduce the impact of traffic incidents on traffic incidents [4].

## 2 Related Works

Expressway incident detection is the key and center of the expressway traffic processing system, so the research on incident detection has impressive results [3, 5, 6]. First, we compare the traffic parameter data between neighboring stations to identify possible sudden traffic accidents [6].

In 1990, based on the catastrophe theory Prasuet et al. developed the McMeST algorithm. The algorithm uses multiple non-crowded and congested traffic data to establish a template of the distribution relationship between traffic and occupancy [6].

In 1995, CHEU and others developed an algorithm based on a multi-layer feedforward network (MLF), which is composed of three layers of input layer, center layer and output layer [6].

In recent years, with the increase in scientific research funds invested in this area, some experts and scholars in the transportation field have proposed Traffic Incidents Detection algorithms. The domestic research on the AID algorithm mainly converges on the application of new theories and new technologies, including wavelet transformation, BP network, SVM, etc. [7, 8].

Li Wenjiang and Jing Bian proposed an incident detection algorithm based on wavelet analysis: logic judgment to determine whether there is a traffic incident is the basic idea of this method [9].

Lv Qi and Wang Hui proposed a traffic detection algorithm based on dynamic BP network. The algorithm learns the training algorithm of the static BP network, and improves the shortcomings of the static BP network training algorithm that the convergence speed is slow [10].

Liang Xinrong and Liu Zhiyong proposed a SVM algorithm based on the principle of structural risk minimization. According to the different reasons for the traffic flow parameters of traffic operation and non-traffic work, the input samples of the whole algorithm added to the algorithm [11].

## 3 First Section

In 1993, Vapnik et al. developed a new trainable machine learning method-Support Vector Machine (SVM) theory. It is a model based on statistical learning theory, maximum classification principle and structural risk minimization principle. Support Vector Machine algorithm has good generalization ability and is very suitable for dealing with limited small sample problems in pattern recognition and regression analysis [1-5].

### 3.1 Description of Classification Problem

Consider such a classification problem: The set of sample sets contain  $l$  sample points is:

$$T = \{(x_1, y_1), \dots, (x_l, y_l)\} \in (X \times Y)^l \tag{1}$$

Where the input index vector is  $x_i \in X = R^n$ . The output is  $y_i \in Y = \{+1, -1\}$ ,  $i = 1, 2, \dots, l$ . The training sample is a collection of these  $l$  sample points. So for a given new input  $x$ , how to infer whether its output is  $+1$  or  $-1$  according to the training set.

The classification problem is described in mathematical language as follows:

For a given training sample, where  $T = \{(x_1, y_1), \dots, (x_l, y_l)\} \in (X \times Y)^l$ , among them  $x_i \in X = R^n$ ,  $y_i \in Y = \{+1, -1\}$ ,  $i = 1, 2, \dots, l$ . Find a real-valued function  $g(x)$  on  $X = R^n$ . It can be seen from the above that finding a rule that can divide the point of  $R^n$  into two parts is the essence of solving the classification problem.

The above two types of classification problems are called classification learning machines. Because when  $g(x)$  is different, the decision function  $f(x) = \text{Sgn}(g(x))$  will also be different.

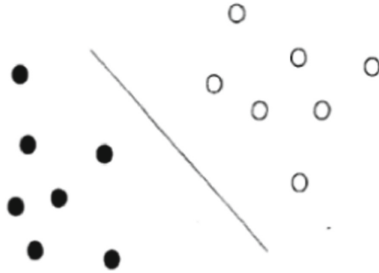


Fig. 1. Linearly separable problem.

As shown in Fig. 1, the linearly separable problem is a classification problem that can easily separate two types of training samples with a straight line.

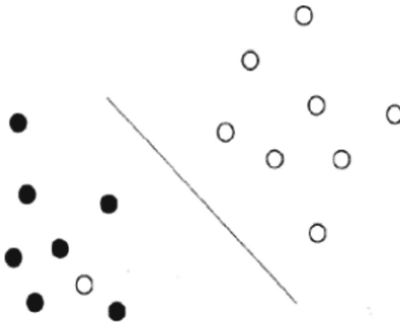


Fig. 2. Linear inseparable problem

As shown in Fig. 2, the linear inseparability problem is a classification problem that roughly separates the training samples with a straight line.

As shown in Fig. 3, the non-linear separable problem is a classification problem that will produce large errors when divided by a straight line. The Support Vector Machines corresponding to the above three classification problems are linear separable Support Vector Machines, linear inseparable Support Vector Machines and nonlinear separable Support Vector Machines.

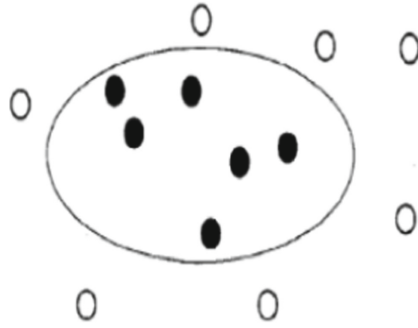


Fig. 3. Nonlinear separable problem

### 3.2 Linear Separable Support Vector Machine

#### Optimal Separating Hyperplane

In the two-dimensional linear separable situation shown in Fig. 4, the solid point and the hollow point are used to represent the two types of training samples. The classification line that completely separates the two types of samples is  $H_0$ , the normal vector of the classification line for  $w$ . First, suppose that the normal direction  $w$  of the optimal classification line  $H_0$  has been determined. At this time, the optimal classification line  $H_0$  is moved up and down in parallel until it touches a certain type of sample. The line between  $H_1$  and  $H_{-1}$  is the best classification line among these candidate classification lines.  $(W \cdot X) + b = -1$  is the normalized representation of the line  $H_1$ ;  $(W \cdot X) + b = 1$  is the normalized representation of the line  $H_{-1}$ . The distance between the two classification boundary lines is  $2/\|w\|$ . as mentioned above, the normal direction  $w$  with the largest distance between  $H_1$  and  $H_{-1}$  is obtained.

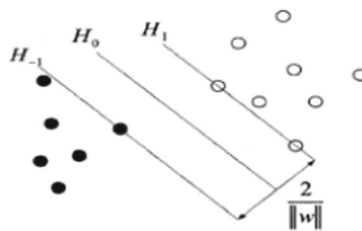


Fig. 4. The best classification surface

It can be seen from Fig. 4 that individual samples determine the maximum classification interval. The samples that happen to fall on the two classification boundary lines  $H_1$  and  $H_{-1}$  are called support vectors. Next, analyze why Support Vector Machines use the principle of maximizing classification interval:

For any training sample  $(x, y)$ , the form of the test sample is  $(x + \Delta_{x,y})$ , where the norm of perturbation  $\Delta_x \in H$  is upper bounded by a positive number. If we use an interval

as To divide the training samples with the hyperplane of  $p < r$ , we should separate all the test samples without error.

Since all training samples are at least classified hyperplane P, and the length of any input sample  $x_i$  ( $i = 1, \dots, l$ ) is bounded, the division of training samples will not be changed.

**Linear Separable Support Vector Machine**

$2/\|w\|$  is the maximum classification interval in the case of Linear Separability. According to the principle of maximizing the classification interval, the problem of finding the Optimal Separating Hyperplane becomes the problem of finding the minimum normal direction  $w.1/2 \|w\|^2$ , that is:

$$\begin{cases} (w \cdot X_i) + b \leq -1y_i = -1 \\ (w \cdot X_i) + b \geq 1y_i = 1 \end{cases} \tag{2}$$

Combining the (2), we can get  $y_i(w \cdot x) + b \geq 1$ . The Optimal Separating Hyperplane constructed in the case of Linear Separability is transformed into the following quadratic programming problem:

$$\begin{cases} s.t. y_i((w \cdot x) + b) \geq -1 \\ \min_{w,b} \frac{1}{2}(w \cdot w) \end{cases} \tag{3}$$

The original problem of solving the Optimal Separating Hyperplane. Duality theory converts the original problem of solving the Optimal Separating Hyperplane into a dual space into a dual problem solution. Define the Lagrange function as:

$$L(w, b, a) = \frac{1}{2}(w \cdot w) - \sum_{i=1}^l a_i [y_i(w \cdot x_i) + b] - 1, a_i \geq 0 \tag{4}$$

$$\frac{\partial(w, b, a)}{\partial w} = w - \sum_{i=1}^l a_i y_i x_i = 0 \tag{5}$$

$$\frac{\partial(w, b, a)}{\partial b} = \sum_{i=1}^l a_i y_i = 0 \tag{6}$$

According to the Karush-Kuhn-Tucker (KKT) theorem, the optimal solution should also satisfy:

$$a_i(y_i(w \cdot x + b) - 1) \geq 0 \forall i \tag{7}$$

From Eq. (7) can see that the coefficient  $a_i$  is required to be non-zero, so  $w$  can be expressed as:

$$w = \sum_{SV} a_i y_i x_i \tag{8}$$

Substituting Eqs. (6) and (5) into Eq. (4), we get:

$$L(w^*, b^*, a) = \sum_{i=1}^l a_i - \frac{1}{2} \|w^*\|^2 = \sum_{i=1}^l a_i - \frac{1}{2} \sum_{i=1}^l \sum_{j=1}^l a_i a_j y_i y_j (x_i \cdot x_j) \tag{9}$$

The optimization problem of Eq. (3) is transformed into a dual space, becomes a dual problem.

$$\begin{aligned} \max L(a) &= \sum_{i=1}^l a_i - \frac{1}{2} \sum_{i=1}^l \sum_{j=1}^l a_i a_j y_i y_j (x_i \cdot x_j) \\ \text{s.t. } a_i &\geq 0, i = 1, 2, \dots, l \\ \sum_{i=1}^l a_i y_i &= 0 \end{aligned} \tag{10}$$

If the coefficient  $a_i^*$  of the support vector is the optimal solution, then

$$w^* = \sum_{i=1}^l a_i^* y_i x_i \tag{11}$$

For any given test sample  $x$ , select the non-zero support vector coefficient  $a_i^*$ , substitute it into Eq. (7) to obtain  $b^*$ , and then use the obtained  $a_i^*$ ,  $b^*$ , training sample  $x_i$ , training result  $y_i$  and The test sample  $x$  is put into the Eq. (12).

$$f(x) = \text{sgn}[\sum a_i^* y_i (x_i \cdot x) + b^*] \tag{12}$$

### 3.3 Linear Inseparable Support Vector Machine

The basic idea of the Linear Non-Separability Support Vector Machine problem is to introduce a non-linear transformation  $\phi(x)$ , so that the input space  $R^n$  can be mapped to a high-dimensional feature space:  $\tilde{z}: X \subseteq R^n \xrightarrow{\phi(x)} Z \subseteq \tilde{z}$ . Then find the high-dimensional feature space as shown in Fig. 5.

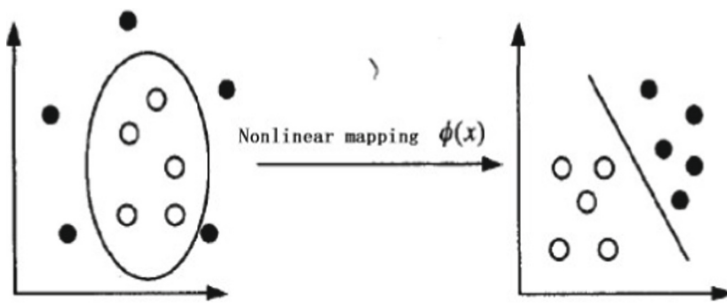


Fig. 5. Shows the nonlinear mapping

The method of obtaining the Optimal Separating Hyperplane of nonlinear separable Support Vector Machine. The Kernel functions mainly used for classification problem processing are:

- (1) Polynomial function Kernel function

$$K(x, x_i) = [\langle x \cdot x_i \rangle + 1]^q \quad (13)$$

- (2) Gaussian Radial Basis Kernel function

$$K(x, x_i) = \exp\left(-\frac{\|x - x_i\|^2}{6^2}\right) \quad (14)$$

- (3) Tangent Hyperbolic Kernell function

$$K(x, x_i) = \tanh[v \langle x \cdot x_i \rangle + c] \quad (15)$$

Because the final result is determined by a small number of support vectors, we can “reject” a large number of redundant samples by grabbing these key samples. It has better “robustness” [6].

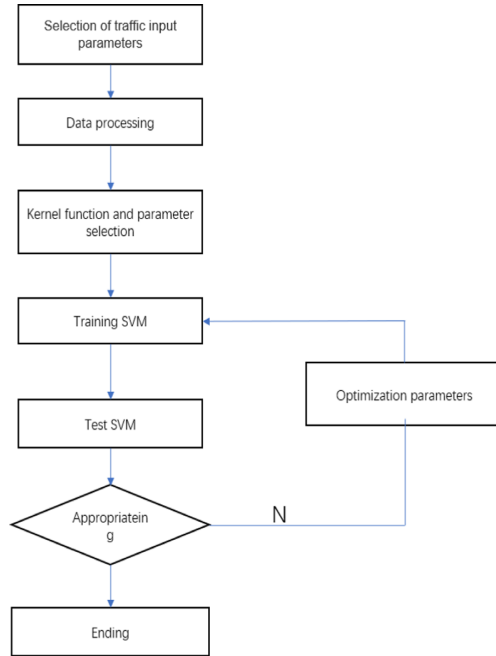
## 4 Design of Event Detection Algorithm Based on SVM

Based on the principle of traffic accident detection and support vector machine theory introduced in the previous chapter, the event detection algorithm based on SVM is designed to detect the occurrence of traffic events.

### 4.1 The Principle and Design Process of SVM-AID Algorithm

- (1) The principle of the SVM-AID algorithm is: According to the principle of maximum classification interval, the acquired data set is optimized and classified.
- (2) The design process of SVM-AID algorithm is as follows:

As shown in Fig. 6, the realization of the SVM-AID algorithm generally needs to go through the following steps: First, determine the traffic parameter indicators used in this design. Then the data is normalized, and all the data are transformed to between  $[-1, 1]$ . After that, selected train and test the SVM. Finally verify the suitability of the selected parameters.



**Fig. 6.** Design steps of SVM-based event detection algorithm

### 4.2 Traffic Parameter Selection

Because if you choose all the multiple related indicators of each traffic sample, it will greatly increase the complexity of the analysis problem. Consider whether you can use a few new indicators to replace the original indicators based on this relationship between the variable indicators.

#### Principal Component Analysis

- (1) Basic concepts: The Principal Component Analysis method replaces the original indicators with a few independent indicators to reflect the information to be presented.
- (2) Basic principles: Suppose there are p indicators and n samples to construct an n × p order matrix.

$$X = \begin{bmatrix} x_{11} & x_{12} & \cdots & x_{1p} \\ x_{21} & x_{22} & \cdots & x_{2p} \\ \vdots & \vdots & \vdots & \vdots \\ x_{n1} & x_{n2} & \cdots & x_{np} \end{bmatrix} \tag{16}$$



Definition: denote  $x_1, x_2, \dots, x_p$  as the original index, and  $z_1, z_2, \dots, z_m$  ( $m \leq p$ ) as the new index.

$$\begin{cases} z_1 = l_{11}x_1 + l_{12}x_2 + \dots + l_{1p}x_p \\ z_2 = l_{21}x_1 + l_{22}x_2 + \dots + l_{2p}x_p \\ \dots\dots\dots \\ z_m = l_{m1}x_1 + l_{m2}x_2 + \dots + l_{mp}x_p \end{cases} \quad (17)$$

How to determine the coefficient  $l_{ij}$ :

- 1)  $z_i$  and  $z_j$  ( $i \neq j; i, j = 1, 2, \dots, m$ ) are linearly independent of each other; Then the new indexes  $z_1, z_2, \dots, z_m$  are respectively called the first, second, ..., principal components of the original indexes  $x_1, x_2, \dots, x_p$ .  
From the above analysis, the following conclusions can be drawn: the essence of Principal Component Analysis is to determine the original index  $x_j$  ( $j = 1, 2, \dots, p$ ) on each principal component  $z_i$  ( $i = 1, 2, \dots, m$ ) The load  $l_{ij}$  ( $i = 1, 2, \dots, m; j = 1, 2, \dots, p$ ).
- 2) Calculation steps.  
Assuming that there are  $p$  indicators and  $n$  samples, the original data becomes a matrix of  $n \times p$  order.

$$X = \begin{bmatrix} x_{11} & x_{12} & \dots & x_{1p} \\ x_{21} & x_{22} & \dots & x_{2p} \\ \vdots & \vdots & \vdots & \vdots \\ x_{n1} & x_{n2} & \dots & x_{np} \end{bmatrix} \quad (18)$$

- 3) Standardize raw data. Use the standardized matrix to find the correlation coefficient matrix.

$$r_{ij} = \frac{\sum_{k=1}^n (X_{ki} - \bar{X}_i)(X_{kj} - \bar{X}_j)}{\sqrt{\sum_{k=1}^n (X_{ki} - \bar{X}_i)^2} \sqrt{\sum_{k=1}^n (X_{kj} - \bar{X}_j)^2}} \quad (19)$$

$$R = (r_{ij})_{p \times p} \quad (20)$$

- 4) Use the correlation coefficient matrix to obtain the characteristic root and the corresponding characteristic vector.

$$\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_p > 0 \quad (21)$$

$$a_1 = \begin{bmatrix} a_{11} \\ \vdots \\ a_{p1} \end{bmatrix}, a_2 = \begin{bmatrix} a_{12} \\ \vdots \\ a_{p2} \end{bmatrix}, \dots, a_p = \begin{bmatrix} a_{1p} \\ \vdots \\ a_{pp} \end{bmatrix} \quad (22)$$

- 5) Calculate the contribution rate and cumulative contribution rate of the principal components.

Contribution rate.

$$\frac{\lambda_i}{\sum_{k=1}^p \lambda_k} \quad i = 1, 2, \dots, p \tag{23}$$

Cumulative contribution rate

$$\frac{\sum_{k=1}^i \lambda_k}{\sum_{k=1}^p \lambda_k} \quad i = 1, 2, \dots, p \tag{24}$$

The index corresponding to the characteristic value whose cumulative contribution rate exceeds 85% is the 1..., mth ( $m \leq p$ ) principal components.

**Acquisition and Discrimination of Traffic Input Data**

This paper is determined by analyzing the experimental results. After that, the samples are divided based on the critical crowding density km.

**Selection of Training Samples and Test Samples**

The downstream the flow and density measured at the detection station  $t, t-1, t-2$  are compared with the flow and density measured at the upstream detection station  $t-2, t-3, t-4$ . When selecting training samples, several sets of data can be selected from the data as the design test samples [7].

**4.3 Choice of Model and Kernel Function**

**Selection of Kernel Function.** The most commonly used Kernel functions for SVM are Polynomial kernel functions and The Radial Basis Kernel function is selected in the SVM-AID experiment, and its form is as follows:

$$K(x, x_i) = \exp\left(-\frac{\|x - x_i\|}{\sigma^2}\right) \tag{25}$$

**Selection of Penalty Coefficient C and  $\delta$  core width.** In the SVM algorithm, We use one of the fixed Penalty Coefficient C or the kernel  $\delta$  width and change the other to detect the same set of data.

**4.4 Algorithm Simulation and Result Analysis**

**Simulation experiment 1-Use Principal Component Analysis to Select Input Traffic Flow Parameters.** In the traffic database, 350 sets of samples whose attributes are density, flow, and speed are selected. The experimental results are as follows:

**Table 1.** Results of simulation experiment 1.

Main ingredient	Contribution rate	Cumulative contribution rate
1	0.74828	0.74828
2	0.24927	0.99755
3	0.0024474	0.9999974

He principal components 1, 2, and 3 in Table 1 above correspond to the three variable indicators of density, flow, and speed respectively. In the following simulation experiments, two indicators of density and flow are used to detect traffic incidents.

**Simulation Experiment 2-With or Without Normalization Control Experiment.**

The Normalization Method is used to process them to reduce the data span. Two sets of data are selected, one set of data is normalized, and the other set of data is not processed. Results are as follows:

**Table 2.** Results of simulation experiment 2.

Category	Event detection rate	False picking rate	Classification accuracy	Detection time (s)
Normalized	100%	7.2%	92.8%	1.031
No normalization	5.6%	12.7%	87.3%	1.141

By analyzing Table 2, the following conclusions can be drawn:(1) Normalization is faster than no normalization, processing data faster and saving time. (2) By adjusting the non-uniform data scale to  $[-1, 1]$ . (3) Normalization speeds up the convergence when the program is running.

**Simulation Experiment 3-Choice of Penalty Coefficient C and  $\delta$  Core Width.**

Because of the selection of different Penalty Coefficients or core widths, the performance indicators of the algorithm will all be affected. Therefore, the following two experiments are designed to select the Penalty Coefficient and the kernel width. The experimental results are as follows:

It can be analyzed from Table 3 and Fig. 6, so  $C = 300$  is selected in SVM-AID. The experimental results are as follows:

**Table 3.** Experimental results of fixed  $\delta$  transformation C.

C value	Classification accuracy	Detection time (s)
0	100%	0.7298
100	95%	0.329
150	94.4%	0.345
200	93.5%	0.356
300	93.2%	0.361
400	93.2%	0.367
500	93.2%	0.372

**Table 4.** Experimental results of fixed C transformation  $\delta$ . $\delta$  value

$\delta$ value	Classification accuracy	Detection time (s)
0.5	15%	0.312
1	38%	0.313
2	59%	0.328
3	71%	0.339
4	85%	0.348
5	93.2%	0.355
6	95%	0.336
7	95.2	0.36

It can be analyzed from Table 4, as the kernel  $\delta$  width increases, the average detection time will continue to grow. When the  $\delta$  value reaches 6, the classification accuracy rate does not increase significantly, and the detection time is still increasing. Therefore, select  $\delta = 6$  in SVM-AID.

**Simulation Experiment 4-4 Road Sections Are Tested with Selected Parameters.**

Choose 4 sets of data from different road sections, each set of data consists of 30 sets of training data and 361 sets of test data. The experimental results are as follows:

**Table 5.** Results of simulation experiment 4.

Road section	Event detection rate	False detection rate	Classification accuracy
N <sub>1</sub>	100%	6.8%	93.18%
N <sub>2</sub>	98.68%	2.5%	97.5%
S <sub>1</sub>	100%	15.2%	84.8%
S <sub>2</sub>	100%	0	100%

It can be seen from Table 5 that the event detection rate is as high as 98%. By comparing the results of 4 road sections, it can be seen that the classification accuracy rate is slightly less than the event detection rate.

## 5 Summary

This article first analyzes the digs deep into the characteristics of expressway traffic flow under traffic incident conditions. Afterwards, using Support Vector Machine theory, a Traffic Incidents Detection algorithm based on Support Vector Machine was proposed. Then the traffic flow parameters, Penalty Coefficients and kernel width of the SVM-AID algorithm are determined through multiple experiments, and the influence of each parameter in the SVM algorithm on the identification of traffic incidents is analyzed, and the suitability of the selected parameters is verified.

**Acknowledgements.** This work was supported in part by the National Natural Science Foundation of China, grant number 72073041. Open Foundation for the University Innovation Platform in the Hunan Province, grant number 18K103.2011 Collaborative Innovation Center for Development and Utilization of Finance and Economics Big Data Property.

Hunan Provincial Key Laboratory of Finance & Economics Big Data Science and Technology. 2020 Hunan Provincial Higher Education Teaching Reform Research Project under Grant HNJC-2020-1130, HNJC-2020-1124.2020 General Project of Hunan Social Science Fund under Grant 20B16.

Scientific Research Project of Education Department of Hunan Province (Grand No. 20K021), Social Science Foundation of Hunan Province (Grant No. 17YBA049).

## References

1. Jiang, D.Y.: Freeway incident automatic detection system and algorithm design. *J. Transp. Eng.* **1**(1), 77–81 (2001)
2. Liu, W.M.: *Highway System Control Method*. China Communications Press, Beijing (1998)
3. Zhang, L.C., Xia, L.M., Shi, H.W.: Highway event detection based on fuzzy clustering support vector machine. *Comput. Eng. Appl.* **243**(17), 208–229 (2006)
4. Dai, L.L., Han, G.H., Jiang, J.Y.: An Overview of road detection event algorithms. *Int. Intell. Transpo.* (2005)
5. Tan, G.L., Jiang, Z.F.: Discussion on automatic detection algorithm for expressway accidents. *J. Xi'an Jiao Tong Univ.* **19**(3), 55–57.70 (1990)

6. Jiang, G.Y.: Road Traffic State Discrimination technology and Application. China Communications Press, Beijing (2004)
7. Jiang, Z.F., Jiang, B.S., Han, X.L.: Simulation study on on-ramp control of expressway. China Journal of Highway and Transport, Beijing, pp.83–89 (1997)
8. Teng, Z.S: Intelligent detection System and data Fusion. China Machine Press, Beijing (2000)
9. Li, W.J., Jing, F.S.: Event detection algorithm based on wavelet analysis. J. Xi'an Highway Univ. **17**(28), 134–138 (1997)
10. Lu, Q., Wang, H.: Traffic event detection algorithm based on dynamic neural network model. J. Xi'an Highway Univ. **20**(6), 105–108 (2003)
11. Zhou, L.Y.: Expressway event Detection Algorithm based on Support Vector Machine. Master's Thesis of Chang' a University (2009)
12. Wen, H.M., Yang, Z.S.: Research on the progress of traffic incident detection technology. Traffic Transp. Syst. Eng. Inf. **5**(1), 25–28 (2006)
13. Zhao, X.M., Jin, Y.L., Zhang, Y.: Theory and Application of Highway Monitoring System. Publishing House of Electronics Industry Press, Beijing (2003)
14. Shi, Z.K., Huang, H.X., Qu, S.R., Chen, X.F.: Introduction to Traffic Control System. Science Press, Beijing (2003)
15. Tian, Q.F.: Research on intelligent Algorithm for Traffic Incident Detection on Urban Expressway. Master's Thesis of Beijing University of Technology (2010)