Short-Term Traffic Flow Prediction Using Optimized Least Squares Support Vector Machine

Xiaobo Yang

Abstract—**This study proposes an optimized Least Squares Support Vector Machine (LSSVM) model to enhance the accuracy of short-term traffic flow prediction. We first analyze the LSSVM algorithm's implementation and introduce an improved leave-one-out cross-validation (LOOCV) method for parameter selection, significantly enhancing the model's performance. Next, we tailor the LSSVM framework specifically for traffic flow prediction tasks. To validate the model's efficacy, we conduct comparative experiments against two state-of-the-art algorithms, DeepSORT and YOLOv5, using real-world traffic datasets. Experimental results demonstrate that our optimized LSSVM model achieves a prediction accuracy of 97.29% and a lower mean absolute percentage error (MAPE) than both benchmark methods. These findings underscore the model's robustness in capturing the continuous and complex dynamics of traffic flow, offering a reliable solution for advancing intelligent transportation systems.**

Index Terms—**Least squares support vector machine, Cross validation, Comparative experiment, Traffic flow prediction**

I. Introduction

SHORT-term road traffic flow analysis stands as a cornerstone of intelligent transportation systems (ITS), playing a pivotal role in traffic information services and traffic control management. As such, the accurate analysis and prediction of short-term traffic flow for road condition assessment have emerged as a critical challenge in modern ITS, demanding effective solutions.

In recent years, numerous advanced methodologies have been developed to tackle this challenge. Li Jia et al. [1] proposed an enhanced Kalman filter approach specifically designed for real-time traffic prediction under normal operating conditions. Pei Yulong et al. [2] constructed a short-term traffic flow prediction model for urban road network nodes by combining the BP learning algorithm with an error-corrected SPDS algorithm. Guo D. Y. et al. [3] developed an innovative object-oriented time-delayed recurrent neural network model, which showed high accuracy in short-term traffic flow forecasting.

For addressing complex traffic patterns, Xie Y. H. et al. [4] proposed a hybrid approach that integrates wavelet analysis with fuzzy Markov theory to effectively predict random fluctuations in short-term traffic flow time series. Meanwhile, Li J. Y. et al. [5] explored the fractal characteristics of traffic flow variations by combining phase space reconstruction techniques with chaos and fractal theory, yielding promising results.

In practical applications, Stephen Clark et al. [6] extended non-parametric regression methods to multivariate analysis and successfully implemented multivariate non-parametric regression algorithms for predicting actual traffic conditions in London. Through comparative experiments, Tian Jing et al. [7] demonstrated that both BP neural networks and chaotic time series prediction methods can achieve high accuracy in short-term traffic flow forecasting, with the latter showing superior real-time performance.

Although the aforementioned traffic flow prediction methods exhibit relatively low computational complexity, they still fail to meet the rigorous requirements for accuracy and real-time adaptability in traffic systems with complex road conditions. These algorithms may face challenges such as slow convergence and diminished predictive accuracy when tackling complex nonlinear problems. Furthermore, insufficient emphasis has been placed on the integration and forecasting of traffic parameters across multiple time horizons, as most existing studies focus primarily on single-step predictions rather than continuous multi-step forecasting. To address these limitations, this paper proposes an approach based on Least Squares Support Vector Machine (LSSVM) regression. Leveraging its advantages—including low structural risk, minimal parameter tuning needs, and strong capability in handling nonlinear regression — this method aims to enhance traffic flow prediction performance under complex road conditions.

II. Least squares support vector Machine (LSSVM) algorithm

Traditional support vector machines (SVMs) are effective in addressing practical challenges such as small sample sizes, non-linearity, high dimensionality, and the problem of local minima. However, in real-world applications, the computational complexity of SVM algorithms is directly proportional to the number of samples. As the sample size increases, the computational process becomes more intricate, resulting in reduced computational efficiency. To alleviate this limitation, the least squares support vector machine (LSSVM) has been proposed as an enhancement to traditional SVM algorithms. The formulation of the LSSVM algorithm can be summarized as follows.

Given a set of l training samples $\{x_i, y_i\}, i = 1,2,\cdots, l$, where each input training sample is denoted by $x_i \in R$ and the corresponding output training sample by $y_i \in R$, the objective optimization function $f(x)$ of the LSSVM algorithm can be formulated as follows.

$$f(x) = min\frac{1}{2}\|w\|^2 + \frac{1}{2}\gamma\sum_{i=1}^{l}\xi_i^2 \tag{1}$$

$$\gamma = w^T\emptyset(x_i) + b + \xi_i, i = 1,2,\ldots, l \tag{2}$$

Manuscript received April 17, 2025; revised August 6, 2025.

This work was supported by the China Natural Science Foundation. Grant number: Y1110023

Xiaobo Yang is a professor in the Information Technology Department at Zhejiang Shuren University, Hangzhou, ZIP Code 310015, China (email: yxb71520@163.com).

Among them, $\emptyset(\cdot)$ denotes the kernel space mapping function, w represents the weight vector, ξ_i refers to the relaxation variable, b is the bias term, and γ stands for the normalization parameter, where $\gamma > 0$.

To find the minimum of the objective function, the Lagrangian function $L(x)$ is formulated as follows.

$$L(x) = \frac{1}{2}\|w\|^2 + \gamma \sum_{i=1}^{l} \xi_i^2 - \sum_{i=1}^{l} \alpha_i \{w^T \emptyset(x_i) + b + \xi_i - y_i\} \quad (3)$$

Where α_i denotes the Lagrange multiplier.

The partial derivatives of Equation (3) are computed following the Karush-Kuhn-Tucker (KKT) optimality conditions outlined in [7], yielding the following results.

$$\frac{\partial L}{\partial w} = 0 \rightarrow \; w = \sum_{i=1}^{l} \alpha_i \emptyset(x_i) \quad (4)$$

$$\frac{\partial L}{\partial b} = 0 \rightarrow \; \sum_{i=1}^{l} \alpha_i = 0 \quad (5)$$

$$\frac{\partial L}{\partial \xi_i} = 0 \rightarrow \; \alpha_i = \gamma \xi_i, i = 1,2,\ldots,l \quad (6)$$

$$\frac{\partial L}{\partial \alpha_i} = 0 \; \rightarrow \; w^T \emptyset(x_i) + b + \xi_i - y_i = \gamma \xi_i, i = 1,2,\ldots,\mathrm{l} \quad (7)$$

By eliminating w, the optimization problem of the objective function $f(x)$ is transformed into solving a linear equation.

$$\begin{bmatrix} K + 1/\gamma & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} \alpha \\ b \end{bmatrix} = \begin{bmatrix} y \\ 0 \end{bmatrix} \quad (8)$$

Among them, $K(x_i, x_j) = exp(-\|x_i - x_j\|^2 / 2\sigma^2)$, and the kernel function used is the radial basis function (RBF kernel).

By solving Equation (8), parameters α and b can be obtained. Accordingly, the Least Squares Support Vector Machine (LSSVM) model for function estimation is established as follows.

$$y = \sum_{i=1}^{l} \alpha_i K + b \quad (9)$$

From the derivation process, it is evident that the LSSVM algorithm involves determining only two parameters, γ and σ. Furthermore, by adopting the least squares loss function combined with equality constraints, the LSSVM algorithm transforms the optimization process into a system of linear equations. This structural reformulation significantly reduces the computational complexity of the algorithm [8–9].

III. Parameter selection for LSSVM modeling

In applying the Least Squares Support Vector Machine (LSSVM) algorithm to traffic flow prediction, parameter selection plays a crucial role. The suitability of kernel function parameters and the regularization factor significantly affects the performance of the LSSVM model. Typically, parameter selection is performed via two main approaches: empirical selection and automated optimization. While empirical selection offers certain advantages in model development, it relies heavily on subjective judgment, which may undermine the consistency and reliability of modeling results. In contrast, automated parameter determination through programming enables the systematic identification of optimal parameters, thereby improving model performance. To ensure both accuracy and robustness, this study adopts the cross-validation method [10] for parameter selection.

The core principle of cross-validation can be summarized as follows: Given a dataset with independently and identically distributed (i.i.d.) data points in the input space, the dataset is partitioned into two distinct subsets. The first subset serves as the training set for model development, while the second acts as the test set to evaluate the trained model's performance. The optimal parameter set is then selected based on the configuration that yields the highest performance metric on the test set.

Cross-validation methods can be broadly categorized into two main types: k-fold cross-validation [11] and leave-one-out cross-validation (LOO-CV) [11]. While k-fold cross-validation is computationally efficient and thus preferred for large datasets, LOO-CV offers a more exhaustive evaluation and is typically applied to smaller datasets. In this study, we adopted LOO-CV to determine the optimal model parameters.

The Leave-One-Out Cross-Validation (LOO-CV) method sets the number of iterations (k) equal to the total sample size (n). The resulting error estimate is computed as follows.

$$Loo - CV_{error} = \frac{1}{n} \sum_{i=1}^{n} |f(x_i) - y_i| \quad (10)$$

Here, $f(x_i)$ denotes the predicted value, y_i represents the corresponding observed value, where the data pair (x_i, y_i) is held out as the test set while the remaining data points are used for model training.

The conventional LOO-CV method exhibits two key limitations: (1) it can only evaluate model errors for individual parameter pairs, yielding incomparable error metrics, and (2) its iterative parameter assignment process significantly increases computational complexity and resource demands. To address these issues, we propose augmenting the traditional LOO-CV framework with an integrated parameter optimization module. This enhancement facilitates systematic parameter tuning while improving computational efficiency, ultimately leading to superior model performance. Figure 1 illustrates the proposed parameter optimization methodology.

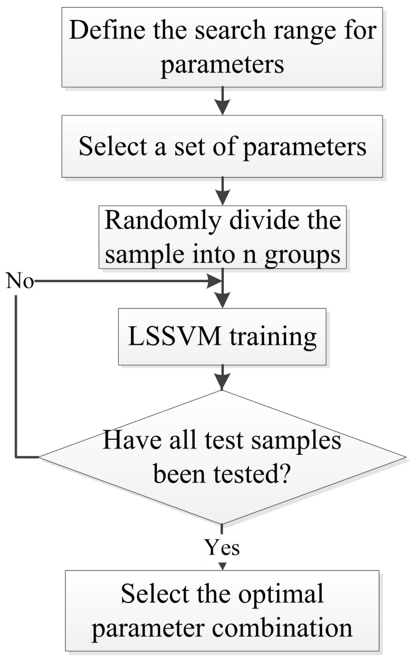

Fig. 1. Flowchart of optimal parameter selection

Figure 1 illustrates the parameter optimization workflow, which consists of four key steps: First, the parameter search space is defined and candidate values are selected. Second, the input samples are randomly divided into n mutually exclusive subsets, with one subset reserved for testing and the remaining n-1 subsets used for training. Third, the LSSVM model is trained and validated, with the prediction accuracy quantified by the root mean square error (RMSE). Finally, after iterating through all test subsets, the optimal parameter combination is identified based on minimal RMSE criteria.

IV. LSSVM ALGORITHM PREDICTION

The Least Squares Support Vector Machine (LSSVM) algorithm is employed for traffic flow prediction, with its complete implementation procedure depicted in Figure 2.

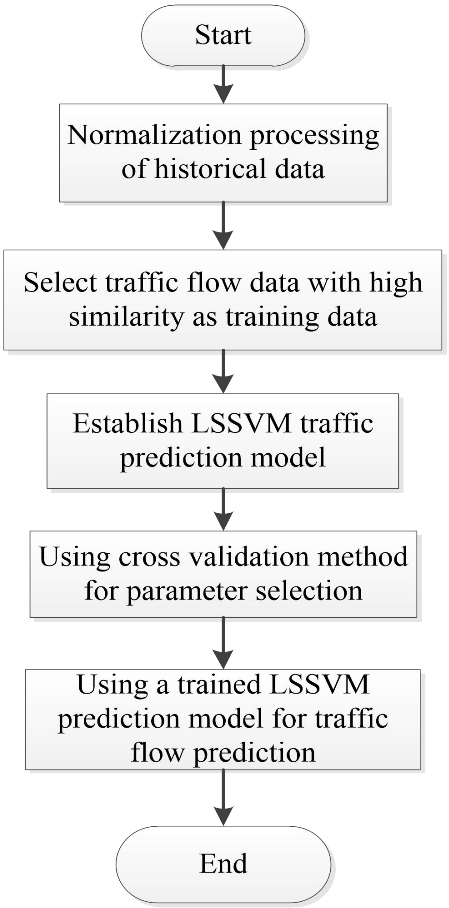

Fig. 2. Flowchart of LSSVM algorithm for traffic flow prediction

Figure 2 presents the LSSVM-based traffic flow prediction framework, which comprises four key phases: (1) normalization of historical traffic flow data, (2) computation of inter-sample similarity and selection of high-similarity training samples, (3) construction of the prediction model with parameters optimized via leave-one-out cross-validation, and (4) application of the trained model for traffic flow forecasting.

To assess the predictive performance of the LSSVM model, this study employs traffic flow data collected from the Shanghai North-South Elevated Road. The dataset comprises 15-minute interval records spanning from November 27 to December 15, 2023. For model development and evaluation, the dataset is divided into two distinct temporal subsets: (1) a training set (November 27 - December 10, 2023) for model calibration, and (2) a testing set (December 11 - 15, 2023) for predictive accuracy validation.

For the LSSVM prediction model implementation, we employ the radial basis function (RBF) as the kernel function, with optimal parameters γ= 5.1125 and σ=1.4871 determined through cross-validation. The model demonstrates robust performance in continuous five-day traffic flow forecasting, as evidenced by the close agreement between predicted and measured mean values shown in Figure 3.

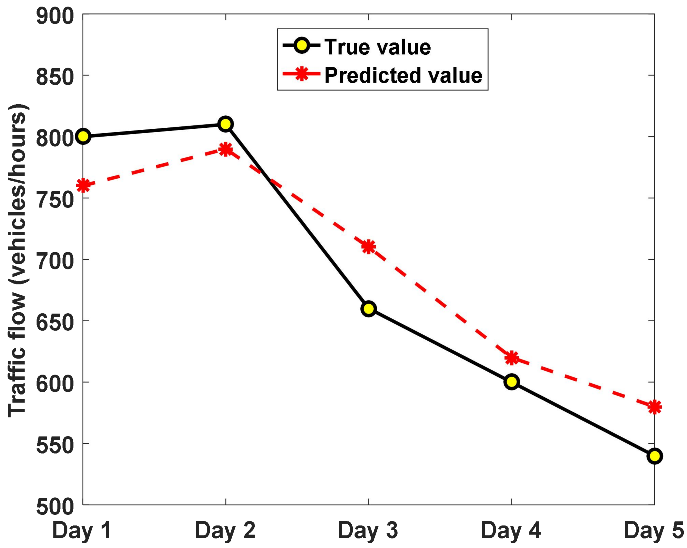

Fig. 3. Comparison of Real and Predicted Traffic Flow for 5 consecutive days

Figure 3 demonstrates strong agreement between the predicted and measured five-day average traffic flows, achieving a prediction accuracy of 97.29%. These results validate the LSSVM model's capability to accurately characterize real-world traffic flow dynamics.

V. COMPARATIVE EXPERIMENT

To evaluate the reliability of the proposed LSSVM algorithm, this study compares its performance with two state-of-the-art prediction methods: Deep SORT [12] and YOLOv5 [13]. The comparative analysis employs the benchmark HighD dataset [14], which contains 190,100 vehicle trajectories (totaling 64,000 km) recorded over 19.5 hours at six highway locations, including 6,500 documented lane changes. The Mean Absolute Percentage Error (MAPE) [15] serves as the primary evaluation metric for quantifying prediction accuracy. Figure 4 presents the comparative results, demonstrating the superior performance of the LSSVM approach.

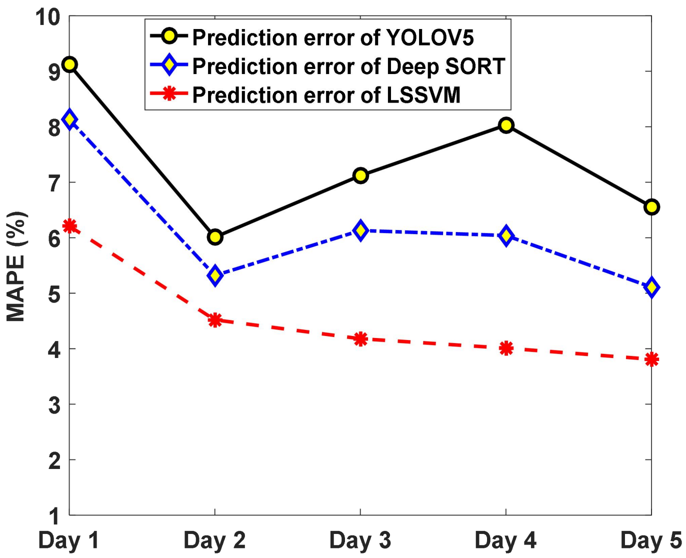

Fig. 4. Comparison of traffic flow prediction using three different algorithms

Figure 4 demonstrates that during the five-day prediction period, the proposed LSSVM algorithm achieves 62% and 35% reductions in MAPE values compared to the two benchmark algorithms, respectively. These significantly lower error rates indicate superior prediction accuracy and enhanced suitability for continuous traffic flow forecasting applications.

To provide additional validation of the algorithm's effectiveness, we evaluate prediction accuracy using mean square error (MSE) [16] as a complementary performance metric. Figure 5 presents the comparative MSE results for all three algorithms, further demonstrating the robustness of our proposed approach.

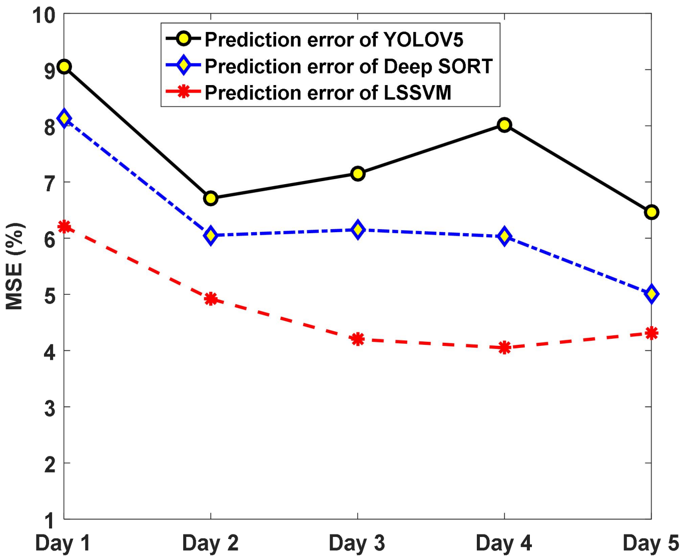

Fig. 5. Comparison of traffic flow prediction using three different algorithms

As evidenced in Figure 5, the LSSVM model achieves significantly lower MSE values than both comparative forecasting methods. This result further validates the superior predictive accuracy of our proposed approach for traffic flow forecasting, demonstrating its potential to enable more reliable traffic management decisions.

VI. Conclusion

This study develops a traffic flow prediction framework using the Least Squares Support Vector Machine (LSSVM) algorithm. Through comprehensive model implementation and comparative performance analysis, we draw the following key conclusions:

1) Parameter selection plays a pivotal role in LSSVM-based traffic flow prediction. Our study introduces an enhanced leave-one-out cross-validation methodology to optimize model parameters. This approach significantly improves model performance through precise parameter calibration, demonstrating particular effectiveness when working with limited dataset sizes.
2) To evaluate predictive performance, the LSSVM model trains on 14 days of traffic flow data and tests on 5 consecutive days of holdout data. The comparison between predicted and measured average values achieves 97.29% accuracy, demonstrating the model's effectiveness in capturing and forecasting traffic flow dynamics.
3) To validate the reliability of the LSSVM model, this study conducts a comparative analysis with two widely-used prediction algorithms: Deep SORT and YOLOv5. The experimental results show that the proposed LSSVM model achieves significantly lower Mean Absolute Percentage Error (MAPE) values compared to these benchmark methods, demonstrating superior prediction accuracy and enhanced applicability for continuous traffic flow forecasting tasks.

References

[1] J. Li, L. X. Liu. "A prediction model of traffic information and its application based on Kalman filtering theory," Central South Highway Engineering, 2005,30(4), pp. 108-109.
[2] Y. L. Pei, Y. Zhang. "Research on Short term Traffic Volume Prediction Model for Urban Road Network Nodes," China Civil Engineering Journal, 2003, 36(1), pp. 12-15.
[3] D. Y. Guo, Z. Zhang, H. Y. Zhou, and X. B. Chen, "Finite-Time Sliding Mode Control of Vehicle Formation with Specified Performance," Engineering Letters, 2025,33(3), pp. 620-629.
[4] Y. H. Xie, G. S. Mi, D. Tan, and C. N. Liu, "WOA-XGBoost Based Railway Accident Type Prediction and Cause Analysis," Engineering Letters, 2025,33(3), pp. 732-742.
[5] J. Y. Li, J. G. Xi, C. W. Xie, and J. P. Gao, "Research on Hybrid Control Methods for the Path Tracking of Driverless Vehicles," Engineering Letters, 2025,33(2), pp. 476-489.
[6] Stephen. "Traffic prediction using multivariate non-parametric regression," Journal of Transportation Engineering, 2003, 129(2), pp. 161-168.
[7] J. Tian, Y. Z. Yang, Y. Z. Chen. "Research on two prediction methods for short-term traffic flow," Highway Transportation Technology, 2006, 23(4), pp. 103-106.
[8] W. Kang, J. Li, G. Y. Cao, et al. "Dynamic temperature modeling of an SOFC using least squares support vector machines," Journal of Power Sources, 2008, 179(2), pp. 683-692.
[9] D. Hanbay. "An expert system based on least square support vector machines for diagnosis of the valvular heart disease," Expert Systems with Applications, 2009,36(4), pp. 8368-8374.
[10] Michael. "Cross-validation method," Journal of Mathematical Psychology, 2000, 44(1), pp. 108-132.
[11] H. M. Jae, C. L. Young. "Bankruptcy prediction using support vector machine with optimal choice of kernel function parameters," Expert Systems with Applications, 2005, 28, pp. 603-614.
[12] N. Wojke, A. Bewley, D. Paulus. "Simple online and realtime tracking with a deep association metric," 2017 IEEE International Conference on Image Processing (ICIP), Beijing: IEEE, 2016, pp. 3645–3649.
[13] Y. Jin, C. L. Zhang. "Extended YOLOv5 for Multi-level Target Classification Detection of Helmet," Computer Systems and Applications, 2023, 32(2), pp. 139-149.
[14] The highD dataset, by RWTH Aachen University's ika, has 110,500 vehicles' natural trajectories on German highways via drones, with <10cm positioning error. It includes 6 locations, traffic states, THW, TTC, maneuvers. Free for non-commercial use; commercial access via leveLXData. Cite its 2018 IEEE ITSC paper. Available: https://www.highd-dataset.com/
[15] Z. Y. Cui, K. Henrickson, et al. "Traffic graph convolutional recurrent neural network: A deep learning framework for network-scale traffic learning and forecasting," IEEE Transactions on Intelligent Transportation Systems, 2020, 21(11), pp. 4883-4894.
[16] B. P. Carlin, T. A. Louis. "Bayes and Empirical Bayes Methods for Data Analysis," Statistics and Computing, 1997, 7(2), pp. 153-154.

Xiaobo Yang: The individual, born on May 20, 1971, in Shanghai, China, holds Chinese nationality and possesses a Ph.D. in Computer Application from Donghua University. The doctorate was successfully obtained in May 2002. Key research areas encompass data mining and intelligent transportation.

He is currently employed as an associate professor at a prestigious college in Hangzhou, Zhejiang province, China. With over 30 research articles, he has established himself as a prolific researcher. His current areas of focus encompass intelligent transportation systems and computer applications.

Dr. Xiaobo Yang holds the esteemed position of being a third-level trainee for the New Century 151 Personnel Project and is also an active member of IEEE.