Frontiers in Computing and Intelligent Systems
ISSN: 2832-6024 | Vol. 15, No. 3, 2026

Research on Urban Traffic Flow Detection and Forecasting Based on YOLO Model and LSTM

Boyu Wu, Xinke Wang

Artificial Intelligence College, North China University of Science and Technology, Tangshan Hebei, 063210, China

Abstract: Urban traffic congestion is becoming increasingly prominent, and accurate traffic flow detection and short-term forecasting are crucial for achieving intelligent traffic scheduling and alleviating congestion. To enhance the real-time performance and accuracy of traffic flow detection in transportation scenarios while enabling effective traffic flow prediction, this paper employs an improved lightweight YOLO object detection model for real-time road vehicle recognition and traffic flow statistics. It further integrates a Long Short-Term Memory (LSTM) network to construct a traffic flow time-series prediction model. Experimental results demonstrate that the YOLO model efficiently performs vehicle detection on edge devices. When integrated with the LSTM model, it achieves precise traffic flow predictions for the next two hours. This provides data support for intelligent scheduling by traffic management authorities and travel planning for users, offering practical application value for smart transportation development.

Keywords: YOLO Model; Traffic Flow Detection; LSTM; Edge Computing; Smart Transportation.

1. Introduction

With the rapid advancement of China's urbanization and the steady improvement of residents' living standards, the ownership of private cars has witnessed an explosive growth. The construction speed of urban road infrastructure fails to keep pace with the surging traffic demand, leading to an increasingly conspicuous contradiction between road capacity and traffic demand. Problems such as traffic congestion, frequent traffic accidents, and low utilization efficiency of road resources have become common challenges for large and medium-sized cities. These issues not only result in enormous waste of time costs and reduce the overall operational efficiency of cities but also exacerbate exhaust emissions due to low-speed driving and frequent starts and stops of vehicles, exerting a negative impact on the urban ecological environment. According to relevant traffic reports, the average speed of road networks during peak hours in Chinese cities with over one million motor vehicles has been on a steady decline, and the duration of congestion on core road sections has increased year by year. Traffic congestion governance has thus become a crucial task in the refined management of cities.

Against this backdrop, the Outline for Building a Transportation Powerhouse clearly stipulates that a national smart transportation system will be fully established by 2035, realizing the digitalization, networking and intellectualization of transportation. Real-time traffic flow detection and accurate short-term prediction constitute the core foundational components of the smart transportation system, serving as a prerequisite for realizing dynamic traffic scheduling, early congestion warning and optimal resource allocation. Traditional traffic flow detection methods are mainly divided into two categories. One category is detection technologies based on geomagnetic, infrared, microwave and other sensors. Although such methods exhibit certain stability in specific scenarios, they suffer from high costs of equipment installation and maintenance, poor environmental adaptability, limited detection range, and are unable to cope with vehicle occlusion, mixed driving of multiple vehicle types and other situations in complex traffic scenarios. The other category is detection technologies based on basic computer vision such as background modeling and optical flow methods, which are susceptible to interference from sudden changes in illumination, weather variations, complex road backgrounds and other factors, leading to insufficient detection accuracy and robustness [1].

In terms of traffic flow prediction, traditional methods are mostly based on statistical models such as ARIMA and Kalman filtering.[2] These models rely on the assumption of linear data distribution and are unable to capture the inherent nonlinear, temporal, abrupt and highly correlated characteristics of urban traffic flow data. When facing scenarios such as morning and evening rush hours, holidays and sudden traffic incidents, they generate large prediction errors and thus fail to meet the demands of actual traffic scheduling [3]. In recent years, the rapid development of deep learning technology has provided new solutions for the technological upgrading of the transportation field. The YOLO series of object detection algorithms, by virtue of their end-to-end detection mode, ultra-fast inference speed and high detection accuracy, have become the mainstream algorithms for real-time vehicle detection in traffic scenarios. As an improved recurrent neural network, the Long Short-Term Memory (LSTM) network effectively solves the problems of gradient vanishing and gradient explosion in traditional RNNs when processing long-sequence data. It can accurately capture the long-term temporal dependencies in data and exhibits excellent performance in the field of time series prediction [4].

Meanwhile, the rise of edge computing technology has provided hardware support for the implementation of deep learning models in traffic scenarios. Deploying lightweight deep learning models on edge computing devices enables real-time processing of video streams and data analysis at the source of data collection, effectively addressing the problems of high transmission latency, heavy bandwidth pressure and high risk of privacy leakage associated with the traditional cloud-based processing mode. Based on this, this paper deploys an improved lightweight YOLO object detection model on the Jetson Xavier NX edge computing device, and combines it with the SORT tracking algorithm to realize real-time recognition, tracking and accurate statistics of road vehicles and traffic flow. On the basis of the time-series traffic flow data obtained from detection, an LSTM time series prediction model is constructed to achieve accurate prediction of traffic flow trends for the next 2 hours. The detection and prediction scheme proposed in this paper builds a complete technical chain of "edge real-time detection - data time-series analysis - short-term trend prediction", providing a feasible and implementable technical solution for the construction of urban smart transportation systems. It is of great significance for alleviating urban traffic congestion and improving the intellectualization level of traffic management.

2. Technical Foundations

2.1. YOLO Object Detection Algorithm

The YOLO series employs an end-to-end detection approach, transforming object detection into a regression problem. It completes object localization and classification in a single forward pass, achieving detection speeds far superior to two-stage detection algorithms and meeting real-time detection demands in traffic scenarios. The lightweight YOLO model balances detection accuracy with hardware adaptability by optimizing network architecture, reducing model parameters, and lowering computational complexity[5]. It can be deployed on edge computing devices like Jetson Xavier NX to enable local real-time processing of video stream data, effectively addressing transmission delays and bandwidth constraints inherent in traditional cloud-based processing.

In traffic flow detection scenarios, the YOLO model identifies vehicle objects in each frame of surveillance video, outputting detection bounding boxes and category information. Combined with the SORT tracking algorithm and line-crossing detection, it enables continuous tracking of road vehicles and precise traffic flow statistics, providing high-quality time-series data for subsequent traffic forecasting[6].

2.2. LSTM Time Series Forecasting Model

Long Short-Term Memory (LSTM) is an enhanced recurrent neural network (RNN)[7]. By introducing gating mechanisms—forgetting gates, input gates, and output gates—it effectively resolves the vanishing gradient and exploding gradient issues plaguing traditional RNNs when processing long sequences. This enables better capture of long-term temporal dependencies in data.

Traffic flow data exhibits distinct temporal, nonlinear, and burst characteristics. The LSTM model automatically learns traffic flow patterns from historical time-series data without manual feature extraction, making it suitable for short-term traffic volume forecasting. This paper preprocesses vehicle flow data detected by the YOLO model before feeding it into the LSTM model to predict road traffic trends for the next two hours, providing a basis for traffic scheduling and travel planning.

2.3. Edge Computing Technology

Edge computing is a computing model that sinks computing, storage and network resources from the cloud to edge devices close to the source of data collection, whose core goal is to complete data processing and analysis on the spot where data is generated, so as to reduce the distance and volume of data transmission. In the scenario of smart transportation, the combination of edge computing technology and deep learning models has become the key to solving the problem of real-time processing of traffic video streams.

The traditional cloud-based computing model requires uploading all the massive video stream data collected by traffic surveillance cameras to cloud servers for processing, which not only occupies a large amount of network bandwidth resources but also causes transmission delays of hundreds of milliseconds, making it difficult to meet the real-time requirements of traffic detection and scheduling. Meanwhile, the remote transmission of massive traffic video data also entails certain risks of privacy leakage. By contrast, deploying lightweight deep learning models on edge computing devices enables target detection, data analysis and result output of video frames at the source of data collection. Only a small amount of detection results and statistical data need to be uploaded to the cloud, which greatly reduces network bandwidth consumption and transmission delay and achieves millisecond-level real-time response.

In this paper, the Jetson Xavier NX is selected as the edge computing device. This device is an embedded development board specially designed by NVIDIA for edge intelligent computing, featuring high performance, low power consumption and miniaturization. It is equipped with a six-core ARM64 processor, a Volta-architecture GPU and a deep learning acceleration engine, which can efficiently support the inference and operation of deep learning models. In the meantime, the device is provided with a wealth of I/O interfaces and can be directly connected to traffic surveillance cameras to realize the collection and real-time processing of video stream data. The combination of edge computing devices and deep learning models provides reliable hardware and technical support for the real-time detection of traffic flow, and serves as an important guarantee for the realization of real-time monitoring in smart transportation.

3. Experiments and Results Analysis

3.1. Experimental Dataset and Environment

The VisDrone2019 dataset, specialized for traffic scenarios, was used for model training and testing. This dataset contains a large volume of vehicle images from complex traffic environments, covering scenarios such as vehicle occlusion, sudden lighting changes, and diverse vehicle types, ensuring high representativeness. The experimental environment comprises a model training environment and an edge deployment environment: The model training environment utilizes the PyTorch framework with an NVIDIA RTX 3060 GPU to train and optimize the YOLO model and LSTM model. The edge deployment environment employs a Jetson Xavier NX edge computing device running Ubuntu 18.04 OS, configured with deep learning frameworks such as CUDA and cuDNN, to achieve lightweight deployment and real-time detection of the YOLO model.

3.2. YOLO Model Vehicle Traffic Detection Results

The lightweight YOLO model was trained and tested on the VisDrone2019 dataset. Detection accuracy (P) and mean average precision at 0.5 (mAP0.5) served as performance evaluation metrics, while inference speed was used to assess real-time capability. Experimental results demonstrate that this YOLO model maintains lightweight characteristics while delivering outstanding detection performance: vehicle detection accuracy (P) exceeds 95%, mAP0.5 surpasses 92%, enabling effective recognition of vehicle targets in complex traffic scenarios and addressing detection accuracy degradation caused by vehicle occlusion and sudden lighting changes. Inference speed on the Jetson Xavier NX edge device exceeds 30 frames per second, meeting real-time video stream detection requirements in traffic scenarios.

When applied to actual road traffic flow detection, combined with the SORT tracking algorithm and lane crossing detection method, traffic flow statistics maintain an error rate below 5%. This enables precise counting of road traffic volume per unit time, providing high-quality time-series data for subsequent LSTM prediction models.

3.3. LSTM Model Traffic Volume Prediction Results

A 30-day continuous traffic volume dataset from a key urban road section was selected as the experimental dataset. 80% of the data served as the training set, while 20% constituted the test set for training and evaluating the LSTM model. The Mean Absolute Error (MAE) and Root Mean Square Error (RMSE) were used as performance metrics. Experimental results demonstrate that the LSTM model effectively captures temporal patterns in traffic flow, delivering excellent predictions for the next 2 hours: MAE is controlled within 8 vehicles/10 minutes, and RMSE within 10 vehicles/10 minutes. It accurately forecasts traffic flow trends during peak periods like morning and evening rush hours, providing reliable predictive support for traffic scheduling and travel planning.

4. Conclusion

To address the current predicament of urban traffic congestion and the shortcomings of traditional traffic flow detection and prediction methods, this paper proposes a traffic flow detection and prediction scheme based on a lightweight YOLO object detection model and an LSTM time series prediction model, and constructs an integrated smart transportation technology system featuring edge real-time detection - data time-series analysis - short-term trend prediction - dual-terminal service application. The main research conclusions are as follows:

The improved lightweight YOLO model integrates the characteristics of light weight and high precision. Its inference speed on the Jetson Xavier NX edge device exceeds 30 frames per second, with a vehicle detection accuracy of over 95%. Combined with the SORT tracking algorithm and the line-crossing detection method, the error rate of traffic flow statistics is less than 5%. It can meet the requirements for real-time and accurate traffic flow detection in traffic scenarios, and effectively solve the problems of poor real-time performance, insufficient accuracy and poor environmental adaptability of traditional detection methods.

The LSTM model can accurately capture the nonlinear and temporal characteristics of traffic flow data and achieves excellent performance in short-term traffic flow prediction for the next 2 hours, with the mean absolute error (MAE) and root mean square error (RMSE) controlled within 8 vehicles/10 minutes and 10 vehicles/10 minutes respectively. It is far superior to traditional statistical prediction models and provides reliable predictive support for traffic scheduling and travel planning.

The combination of edge computing technology and deep learning models effectively solves the problems of high transmission latency and heavy bandwidth pressure in the traditional cloud-based processing mode, realizing the local real-time processing of traffic video stream data and providing a feasible technical solution for the real-time monitoring of smart transportation. The constructed integrated detection-prediction technical chain enables the efficient utilization of traffic data, provides two-way information services for traffic management departments and residents, and promotes the transformation of traffic management from passive response to active regulation.

The detection and prediction scheme proposed in this paper realizes information collaboration between traffic management departments and travelers, and provides strong technical support for intelligent traffic scheduling, congestion early warning and personalized travel planning. It has significant practical application value and broad promotion prospects, and is of great significance for the construction and implementation of smart transportation systems and the governance of urban traffic congestion.

References

[1] Zheng Rongcai, Tan Dingwen, Xu Qing, et al. Research on Lightweighting of Salmon Detection Model Based on Improved YOLO v7 [J]. Transactions of the Chinese Society for Agricultural Machinery, 2024, 55(11): 132-139.

[2] Yang Sen, Zhang Pengchao, Wang Lei, et al. Lightweight Tomato Leaf Pest and Disease Identification Method Integrating Improved YOLOv8n and Channel Pruning [J]. Transactions of the Chinese Society for Agricultural Engineering, 2025, 41(02): 206-214.

[3] Huang Lingfeng, Yang Shilong, Xie Yaochin. YOLO-PointMap: Real-Time Acupoint Recognition on Human Back Based on Lightweight Dynamic Feature Fusion [J].Integration Technology, 2025, 14(02): 58-70.

[4] Pan Wei, Wei Chao, Qian Chunyu, et al. An Improved YOLOv8s Model for Small Object Detection from UAV Perspectives [J/OL]. Computer Engineering and Applications: 1-10 [2024-04-16].

[5] Tang Weibo, Fang Qiang, Li Peigen, et al. Small Object Detection in UAV Aerial Images Based on RSD-YOLO[J/OL]. Computer Engineering,1-15[2025-02-15].

[6] Yang Haitao, Zhao Junyu, Wang Rui, et al. Small Object Detection for UAVs Based on DBB-YOLOv10s [J/OL]. Infrared Technology, 1-8 [2025-02-13].

[7] Vaddi S, Kumar C, Jannesari A. Efficient Object Detection Model for Real-Time UAV Applications[J]. arXiv preprint arXiv:1906.00786, 2019.