Traffic & Transportation

Fusing Visual Quantified Features for Heterogeneous Traffic Flow Prediction

Qinyang WANG[1], Jing CHEN[2], Ying SONG[3], Xiaodong LI[4], Wenqiang XU[5]

Original Scientific Paper
Submitted: 28 Mar 2024
Accepted: 26 June 2024

[1] mauldianhu@gmail.com, Hangzhou Dianzi University, School of Computer Science
[2] Corresponding author, cj@hdu.edu.cn, Hangzhou Dianzi University, School of Computer Science
[3] songying@xidian.edu.cn, Xidian University, Hangzhou Institute of Technology
[4] hzxiaodong22@163.com, Hangzhou Dianzi University, School of Computer Science
[5] wqx.zjl@gmail.com, China Jiliang University, College of Economics and Management

Publisher:
Faculty of Transport and Traffic Sciences, University of Zagreb

ABSTRACT

This paper presents a novel traffic flow prediction method emphasising heterogeneous vehicle characteristics and visual density features. Traditional models often overlook the variety of vehicles, resulting in inaccuracies. The proposed method utilises visual techniques to quantify traffic features, such as mixed flow and vehicle accumulation, enhancing dynamic density estimation and flow fluidity. We introduce a spatio-temporal prediction model that integrates various data types, capturing complex dependencies and improving accuracy. This research advances traffic flow prediction by considering the diverse nature of vehicles and leveraging visual data, offering valuable insights for intelligent transportation systems. Experimental results demonstrate the superiority of this approach over conventional methods, especially in capturing traffic flow fluctuations.

KEYWORDS

heterogeneous traffic flow; spatio-temporal modelling; traffic flow prediction; visual traffic quantification.

1. INTRODUCTION

Accurate short-term traffic flow prediction is crucial for urban managers, as it can help them make informed decisions and alleviate issues such as traffic congestion [1]. Unlike long-term forecasting, which addresses broader planning and infrastructure development, short-term prediction focuses on managing rapidly changing traffic conditions influenced by diverse social and economic demands. The presence of pedestrians and vehicles with different transportation functions significantly affects traffic flow dynamics. In particular, the role of heterogeneous vehicles in traffic prediction is becoming an increasingly indispensable key factor that cannot be ignored.

Current short-term traffic flow prediction models often overlook these distinctions, typically assuming a homogenised vehicle system. Large vehicles, for instance, can significantly influence the driving behaviour of smaller vehicles, leading to mobility bottlenecks, uneven traffic flow distribution and reduced road traffic circulation capacity [2, 3]. Such conditions diminish road user satisfaction and escalate the risk of accidents due to the mixed vehicle types, discrete traffic flows and uneven speed distributions. Traditional heterogeneous traffic flow models typically use the personal capacity unit (PCU) to standardise vehicles of different types and sizes into a single measurement. It simplifies the analysis by treating all vehicles as equivalent to a standard passenger car. However, this method overlooks the unique impacts of vehicles of varying sizes and functionalities on traffic dynamics.

Although advancements in deep learning have helped to capture temporal and spatial dependencies in traffic flow prediction, current research exhibits two main shortcomings. (1) There is insufficient emphasis on visual methods, which offer the advantages of low installation and maintenance costs for visual sensors and can provide multidimensional quantitative data compared to traditional ground induction coils [4–6]. (2) Many models focus on integrating external factors at both macro and micro scales, overlooking the specific impacts of heterogeneous vehicle types on short-term traffic flow dynamics [7, 8]. To address these issues, a heterogeneous traffic flow prediction method based on visual density features is proposed. The contributions of this paper are as follows:

— We propose a vision-based method for quantifying heterogeneous traffic features. In real traffic scenarios, heterogeneous traffic features such as mixed flow features, free flow and queueing states have a significant impact on traffic flow. The mixed flow features are quantified by using a multi-scale target detection method. By quantifying vehicle accumulation over different periods, we achieve dynamic density estimation of road traffic. By analysing the distribution and changes in vehicle speeds, as well as the movement patterns of vehicles in the road network, we assess the fluidity of traffic flow. This realises a multidimensional quantification of traffic flow features using visual methods.
— We construct a spatio-temporal prediction model for multidimensional heterogeneous traffic features. This model effectively integrates various types of data, including spatial (location, distribution) and temporal (time series) data, as well as the heterogeneous features of different vehicle types. By combining macroscopic flow and microscopic heterogeneous features, it captures the complex dependencies and patterns within traffic flow. This approach allows for more comprehensive learning of historical traffic data, and more accurately reflects the real impact of heterogeneous features such as vehicle model differences on the dynamic changes in traffic flow.
— The experimental results show that our heterogeneous traffic flow prediction method, based on visual density features, significantly improves prediction accuracy. The proposed method can effectively capture significant fluctuations in traffic flow and provide more precise forecasts than previous methods.

The rest of this paper is organised as follows. Related works about deep traffic flow prediction are given in Section 2. Then, the proposed heterogeneous traffic flow prediction method based on visual density features is introduced in Section 3. Afterwards, experimental results are presented in Section 4. Finally, Section 5 concludes this paper.

2. RELATED WORKS

Research in short-term traffic flow prediction has traditionally leaned on statistical models like historical averaging (HA) [9], support vector regression (SVR) [10] and regression analysis. While these methods are celebrated for their straightforward interpretations and effectiveness in limited data scenarios, they are often complemented by newer technologies. For instance, Haghani et al. [11] have enhanced traffic management through the use of Bluetooth sensors for accurate freeway travel time data, while Adu-Gyamfi et al. [12] applied empirical mode decomposition (EMD) in analysing probe-sourced traffic speed data to assess reliability. These advancements underscore a broader trend of integrating sophisticated data collection and analysis techniques into traditional traffic prediction frameworks.

Building on these foundations, the focus has increasingly shifted towards advanced deep learning methods that utilise spatio-temporal features. He et al. [13] developed an end-to-end spatio-temporal 3D densenet multiscale convLSTM-resnet network (ST-3DDMCRN) for predicting future traffic flow accurately, capturing traffic data slices' local regional spatio-temporal information and breaking through the limitation of traditional residual neural network (ResNet) [14] networks in capturing long-range spatial correlations. Guo et al. [15] proposed a hierarchical graph convolutional network, considering the natural hierarchical structure of traffic systems which is composed of the micro layers of road networks and the macro layers of region networks. Zhang et al. [16] introduced a multi-scale self-attention network for investigating multi-grained temporal dynamics across various time resolutions, with an aggregation layer to model the underlying dependencies across multi-level temporal dynamics. Moreover, this hierarchical graph neural network via attentive graph diffusion paradigm, enables spatial semantics from local-level to global-level traffic pattern representations.

To more realistically simulate actual road conditions, some researchers have attempted to integrate more external features to better predict traffic. Yang et al. [17] proposed a multi-feature traffic prediction model based on convolutional neural networks, which integrates external factors like weather and holidays to predict traffic flow. To address the issue that traditional data-driven traffic flow prediction methods tend to ignore traffic self-features, and are usually influenced by various complex factors, this model categorises traffic flow and further aligns and integrates the learned features with external factors through a logistic regression layer to generate the final prediction results, achieving notable accuracy and efficiency. Similarly, Yang et al. [18] proposed a hybrid deep learning structure for short-term traffic speed prediction, combining convolutional neural networks and long short-term memory networks. External factors like weather conditions and air quality can influence travellers' driving behaviour and cause fluctuations in traffic speed. Based on traffic engineering theory, this model uses a data-fusion method to measure the impact of environmental factors. To enhance model performance, an attention mechanism is introduced. Through the convolutional block attention module, the model network can emphasise important channels and pixels of input features and suppress unnecessary ones, thereby improving the accuracy of traffic prediction.

Although these models have enhanced prediction accuracy to varying degrees, they have all neglected the distinct features of heterogeneous types in influencing traffic flow. In actual road situations, the impact of different heterogeneous types on traffic flow varies. Larger vehicles, due to their greater size and poorer dynamic performance compared to smaller vehicles, are prone to creating numerous moving bottlenecks, thus significantly impacting traffic flow. With an increasing proportion of large vehicles, the standard deviation of vehicle following distances in the traffic flow gradually reduces, increasingly affecting the stability of the traffic flow. This leads to significant increases in traffic delays and, in severe cases, can result in traffic congestion and accidents [19]. In view of this, this paper proposes a heterogeneous traffic flow prediction model based on visual features to explore the impact of heterogeneous features on traffic flow.

3. METHODOLOGY

This paper further explores the impact of integrating heterogeneous visual features on traffic flow prediction, building upon the consideration of flow states [20]. To better simulate actual road traffic conditions, we initially merge the acquired heterogeneous traffic feature data. Subsequently, we delve into uncovering the hidden spatial and temporal dependencies within these data features to predict traffic flow.

3.1 Heterogeneous flow mixing rate

Typically, larger vehicles have greater length and width than smaller vehicles, and they have a more significant impact on road traffic flow. The mixing rate of large vehicles is one of the crucial parameters for assessing road safety and traffic congestion [21]. The uneven distribution of traffic flow states across different areas of the road network, reflecting the variation in traffic flow composition, is an important feature that cannot be overlooked.

Quantifying heterogeneous traffic flow differs from homogeneous traffic flow, as it requires identifying the specific type of each vehicle to determine the impact of different vehicle types on traffic flow. In this paper, buses and trucks are considered large vehicles, while others are categorised as small vehicles. We use the machine vision detection algorithm called You Only Look Once v8 (YOLO v8) to count the number of vehicles in each category. We use a loss function L_{det} as defined in *Equation 2* to train the detection model:

$$y_c = softmax(x_c) = \frac{e^{x_c}}{\sum_{c=1}^{C} e^{x_c}} \tag{1}$$

$$L_{det} = -\sum_{c=1}^{C} g_c \, log(y_c) \tag{2}$$

where C denotes the total number of classes, x_c denotes the predicted value of class c, y_c denotes the predicted probability of class c, and g_c denotes the true label of class c. The feature of the mixing rate of heterogeneous flow ρ is defined as follows:

$$q_{n,t} = \sum_{c=1}^{C} q_{n,t}^{c} \tag{3}$$

$$\rho_{n,t} = \sum_{c=1}^{C} \mu_c \frac{q_{n,t}^{c}}{q_{n,t}} \tag{4}$$

$$\rho_n = [\rho_{n,1}, \rho_{n,2}, \cdots, \rho_{n,T}] \tag{5}$$

$$\rho = [\rho_1, \rho_2, \cdots, \rho_N]^T \tag{6}$$

where $q_{n,t}^c$ denotes the total number of heterogeneous vehicles in class c at time t at node n, and $q_{n,t}$ denotes the sum of the quantities of all heterogeneous types at time t at node n, μ_c denotes the scaling factor for class c, $\rho_{n,t}$ denotes the feature of heterogeneous mixing rate at time t at node n, ρ_n denotes the feature of heterogeneous mixing rate across T moments at node n, and N denotes the total number of nodes. It is related to the current traffic flow state of the road. Generally, in non-free flow states, large vehicles have a greater impact on traffic flow.

3.2 Heterogeneous flow density state

Obtaining traffic density states is crucial for traffic flow prediction, as well as reducing congestion and aiding in planning the routes of relevant vehicles. Traffic density is primarily measured as the number of vehicles per unit length. Additionally, microscopic analyses of traffic flow can also estimate density indirectly by measuring the headway, or the space between vehicles, which provides insights into traffic patterns and congestion. A problem with these methods is that they do not account for the specific impact of different types of vehicles on traffic density states. In reality, in real road traffic conditions, large vehicles, due to their longer size and weaker acceleration and deceleration capabilities, tend to seek stable driving distances rather than higher speeds, compared to smaller vehicles. Also, large vehicles psychologically impact most surrounding drivers, causing them to consciously adjust the distance between their vehicles and the large vehicle, thereby affecting traffic flow. This indicates that the impact of large vehicles on road traffic density is different from that of smaller vehicles.

In heterogeneous traffic flow density prediction models considering vehicle types, the goal of the loss function is to minimise the error between the predicted values and the actual traffic flow state. The loss function L_{tfp} is presented as follows.

$$L_{tfp} = -\,log\left(\frac{e^{prob(gt)}}{\sum_{i=0}^{1} e^{prob(i)}}\right) \tag{7}$$

This loss focuses on the prediction accuracy of traffic density states. Based on the original cross-entropy loss, it compares the difference between the model's predicted traffic state (such as free or queued) and the actual state. $prob(i)$ denotes the model's predicted probability of road traffic density state i, and gt denotes the actual road traffic density state. In this paper, road traffic density state is divided into two categories: free (0) and queue (1). Therefore, the feature of traffic density state D considering heterogeneous factors can be represented as follows:

$$d_{n,t} = \begin{cases} 1, & if\ prob_{n,t}(1) \geq prob_{n,t}(0) \\ 0, & else \end{cases} \tag{8}$$

$$d_n = [d_{n,1}, d_{n,2}, \cdots, d_{n,T}] \tag{9}$$

$$D = [d_1, d_2, \cdots, d_N]^T \tag{10}$$

where, $prob_{n,t}(i)$ represents the model's predicted probability of road traffic density state i at time t at node n, $d_{n,t}$ represents the traffic density state score at time t at node n, and d_n denotes the feature of traffic density state across T moments at node n.

3.3 Prediction for heterogeneous traffic flow

As shown in *Figure 1*, this paper presents a heterogeneous traffic flow prediction model based on visual methods. The model mainly consists of two modules: the heterogeneous quantification module and the dependencies fusion module.

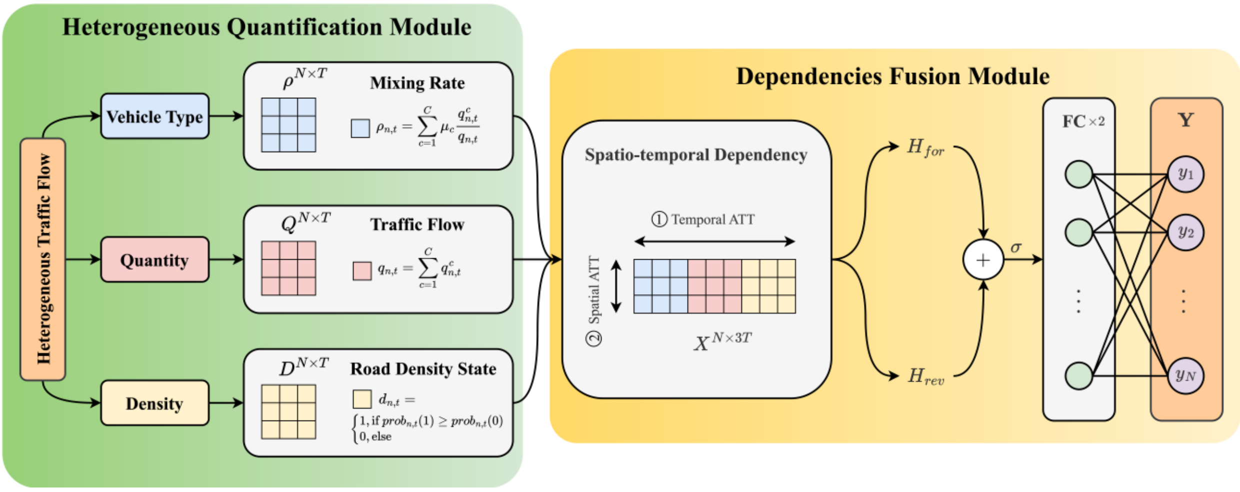

Figure 1 – The overview of our proposed model for heterogeneous flow prediction

1) The heterogeneous quantification module is utilised for the acquisition of heterogeneous traffic data features. By quantifying traffic flow, mixing rates and density states using visual methods, we obtain the mixing rate feature matrix $\rho^{N\times T}$, traffic flow feature matrix $Q^{N\times T}$ (as shown in *Equation 12*), and traffic density state feature matrix $D^{N\times T}$. We concatenate these heterogeneous features to form the input matrix $X^{N\times 3T}$, as shown in *Equation 13*. Unlike most models that suffer from the drawback of having a single type of input feature, our data features, after fusion, can more accurately simulate real road traffic conditions.

$$q_n = [q_{n,1}, q_{n,2}, \cdots, q_{n,T}] \tag{11}$$

$$Q = [q_1,\ q_2,\ \cdots,\ q_N]^T \tag{12}$$

$$X = [\rho, Q, D] \tag{13}$$

2) The dependencies fusion module searches for the spatio-temporal dependency of the fused features. We employ the widely used gated recurrent unit (GRU) [22] model, combined with a multi-head attention mechanism, for capturing temporal dependency. The features, the temporal dependency of which has been captured, can be represented as follows.

$$H_T = GRU\big(MultiHead(X)\big) \tag{14}$$

To capture spatial dependency, we combine the features processed by the GRU model with the adjacency matrix, inputting these features into the k-hop graph and $(k-1)$-hop graph of the central node respectively for graph convolution, which aggregates the spatial features. Then, we fuse these features after pooling. These procedures can be represented as follows:

$$H^{(1)}\hat{A}_2 = \left(\hat{A}_2 H_T W^{(0)}\right) W^{(1)} \tag{15}$$

$$H_{for} = \hat{A}_1 \sigma\left(P_{max}\left(H^{(1)}\right) + P_{mean}\left(H^{(1)}\right)\right) W^{(2)} \tag{16}$$

$$H_{rev} = \hat{A}_2^{rev} H_T W^{(3)} \tag{17}$$

$$H_{TS} = \sigma\left(H_{for} + H_{rev}\right) \tag{18}$$

where $W^{(i)}$ denotes the weight parameters of the i-th layer, $H^{(i)}$ denotes the hidden representation of the i-th hidden layer, $P_{max}(\cdot)$ means the max pooling function, $P_{mean}(\cdot)$ means the mean pooling function, $\sigma(\cdot)$ means the rectified linear unit (ReLU) activation function, $\hat{A}_k$ denotes the k-hop normalised adjacency matrix of the central node, $\hat{A}_k^{rev}$ denotes the flow-reverse k-hop normalised adjacency matrix of the central node, H_{for} means the flow-forward features, H_{rev} means the flow-reverse features and H_{TS} means the features the spatio-temporal dependency of which has been captured.

Finally, two fully connected layers are applied for traffic flow prediction, and the result Y can be represented as follows.

$$Y = FC(FC(H_{TS})) \tag{19}$$

4. EXPERIMENTS

4.1 Dataset and metrics

We utilise the STREETS [23] dataset for evaluation. STREETS is a novel traffic flow dataset from publicly available web cameras in the suburbs of Chicago, IL, providing over 4 million still images across 2.5 months and one hundred web cameras in suburban Lake County, IL. These cameras are divided into two distinct communities described by directed graphs, count vehicles to track traffic statistics and capture road images every 5 minutes.

In experiments, we use the Gurnee community road network in the STREETS dataset. The traffic flow data from 5 June 2019 to 13 June 2019 are used for training, and those on 14 June 2019 are used for evaluation. Moreover, we set the node 28-IL 21 at Washington East-inbound as a central node.

To comprehensively evaluate traffic flow prediction, we apply 3 metrics: mean absolute error (MAE), mean absolute percentage error (MAPE) and root mean squared error (RMSE). These metrics can be represented as follows:

$$MAE = \frac{1}{S}\sum_{s=1}^{S}\left|Y_s - \hat{Y}_s\right| \tag{20}$$

$$MAPE = \frac{1}{S}\sum_{s=1}^{S}\left|\frac{Y_s - \hat{Y}_s}{Y_s}\right| \times 100\% \tag{21}$$

$$RMSE = \sqrt{\frac{1}{S}\sum_{s=1}^{S}\left(Y_s - \hat{Y}_s\right)^2} \tag{22}$$

where Y_s denotes the true label of sample s, $\hat{Y}_s$ denotes the prediction of sample s, and S denotes the total number of samples in the test set.

It is worth mentioning that we only consider heterogeneous vehicles in this paper. Although other factors such as pedestrians can also influence the traffic flow, they are not taken into account to simplify the complexity of the model.

4.2 Experiments settings

By default, we set the batch size to 64, learning rate to 0.001 and the epoch to 10,000, using a stochastic gradient descent optimiser with momentum of 0.9. The length of historical traffic flow used for prediction is 12. We use only a single-layer GRU structure for capturing temporal dependency in traffic data and set the number of heads in the multi-head attention mechanism to 3. The scaling factor for large vehicles is set to 70, and that for small vehicles is set to 30.

All experiments are run on an Ubuntu 18.04.5 LTS system, equipped with an Intel i7-9700K CPU and an NVIDIA GeForce RTX 2080 GPU. We implement our method with Python 3.6 and Pytorch 1.7.

4.3 Traffic flow prediction

Table 1 shows the results of traffic flow prediction for the next 15/30/60 minutes on the STREETS dataset. We compare our method with 7 baselines, including HA [9], SVR [10], feedforward neural network (FNN) [24], GRU [22], a combined model of graph convolution network and long short-term memory (GCN-LSTM) [25], spatio-temporal residual graph attention network (ST-RGAN) [26] and vision based spatial-temporal forecasting (V-STF) [20]. Compared to other baselines, our method incorporates the feature of a heterogeneous traffic flow mixing rate, achieving the best results in both MAE and RMSE metrics. These results demonstrate that the feature of heterogeneous traffic flow mixing rate plays an important role in traffic flow prediction. The incorporation of this feature improves the accuracy of traffic flow prediction, benefitting traffic flow prediction.

Table 1 – Comparison of the proposed method with the state-of-the-art methods in traffic flow prediction

Method	**MAE**	**MAPE (%)**	**RMSE**
HA [9]	5.49/5.48/5.41	23.57/23.10/23.76	7.65/7.70/7.55
SVR [10]	5.04/5.01/5.26	24.98/24.52/24.94	6.62/6.66/7.02
FNN [24]	5.09/4.98/5.27	24.54/24.39/25.61	6.76/6.62/6.74
GRU [22]	5.05/5.05/5.14	25.27/23.92/25.03	6.53/6.52/6.77
GCN-LSTM [25]	5.04/5.03/5.13	24.82/23.76/24.90	6.61/6.75/6.74
ST-RGAN [26]	4.72/4.88/4.76	22.13/23.74/23.19	6.37/6.59/6.37
V-STF [20]	4.81/4.71/5.01	23.37/23.06/23.24	6.22/6.30/6.68
Ours	4.51/4.40/4.52	23.43/22.34/23.31	5.97/5.91/5.84

Note: The bold values are the best results for each metric.

Figures 2-4 illustrate the detailed traffic flow prediction from 5:00 to 17:00. Compared to HA [9] and SVR [10], our method can capture the sudden changes in traffic flow. For example, the ground truth of traffic flow is 5 at 9:45 and is above 13 from 9:50 to 10:30 as shown in *Figure 2*. It can be observed that the prediction of our proposed method is close to the ground truth, while the prediction of SVR does not exceed 9 and the prediction of HA does not exceed 6.

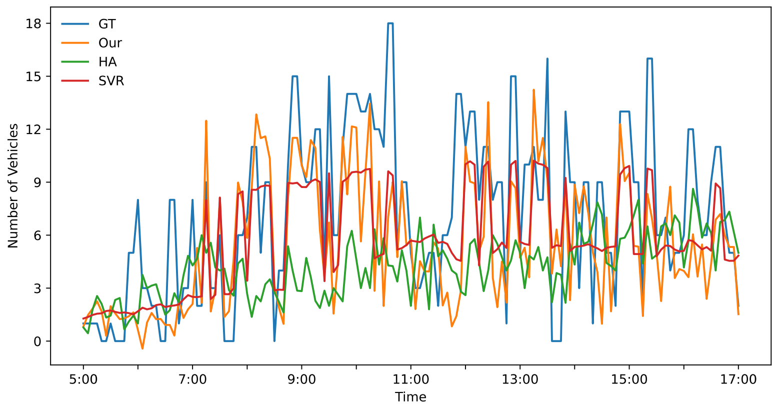

Figure 2 – Detailed comparison of our method with HA and SVR

As shown in *Figure 3*, the FNN [24] model has abnormal cases where the prediction is less than 0 and much larger than the ground truth. Compared to GRU [22], our method can make a more accurate prediction when the traffic flow abruptly changes, such as at 9:00, 10:00 and 12:30.

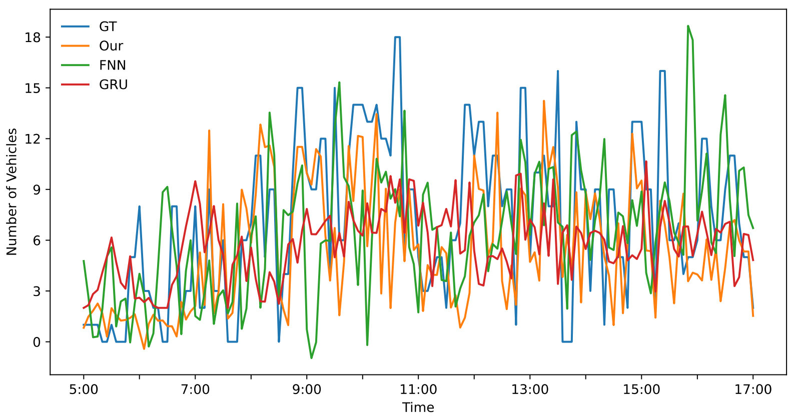

Figure 3 – Detailed comparison of our method with FNN and GRU

As can be seen in *Figure 4*, the performance of our method is close to GCN-LSTM [25] and ST-RGAN [26] in most cases. However, these two models tend to make a larger prediction, and our method outperforms them overall as depicted in *Table 1*.

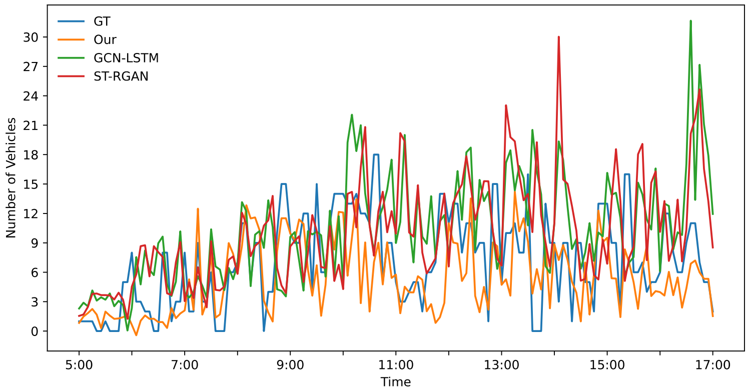

Figure 4 – Detailed comparison of our method with GCN-LSTM and ST-RGAN

4.4 Ablation study

Table 2 shows the results of ablation experiments on the STREETS dataset, which are acquired from traffic flow prediction for 15/30/60 minutes. Method (a) is the baseline GRU+GCN. Method (b) is based on method (a) and incorporates k-hop features. Method (c) is the integration of method (b) and traffic flow density features. Method (d) combines method (c) and traffic flow mixing rate features. With the addition of modules or features including GCN, k-hop, density state and mixing rate from method (a) to (d), the performance of the method increases, which proves the effectiveness of these modules or features on traffic flow prediction.

Table 2 – Ablation results on the STREETS dataset

Method	MAE	MAPE (%)	RMSE
(a) GRU+GCN	5.91/5.84/6.15	25.91/29.99/30.69	8.02/7.96/8.30
(b) GRU+GCN+k-hop	4.91/4.90/5.03	24.69/24.08/24.48	6.64/6.54/6.69
(c) GRU+GCN+k-hop+density state	4.81/4.71/5.02	23.37/23.06/23.24	6.22/6.30/6.68
(d) GRU+GCN+k-hop+density state+ mixing rate	4.51/4.40/4.52	23.43/22.34/23.31	5.97/5.91/5.84

Note: The bold values are the best results for each metric.

5. CONCLUSION

In this paper, we proposed a method for predicting traffic flow by leveraging visual heterogeneous traffic flow features. Initially, a visual object detection framework was employed to classify road objects and identify distinct traffic features such as large vehicles, small vehicles, non-motor vehicles and pedestrians. Following this, we quantified and fused features such as flow, heterogeneous density and road mixing rate to enhance the predictive accuracy of our model. This comprehensive approach enabled our model to not only consider flow features but also integrate heterogeneous density and road mixing rate features, thereby facilitating accurate predictions in varying traffic states, including free flow and queuing.

However, our method is subject to certain limitations. The reliance on visual sampling for the entire road network dataset restricts our capability to finely quantify the interactions between large and small vehicles. Additionally, factors such as lane count, speed limits, weather conditions and road anomalies also significantly impact traffic flow predictions but are not fully integrated into the current model.

Moving forward, future research should focus on refining the quantitative relationships between different types of heterogeneous traffic and extending the model to incorporate a broader array of influencing factors. This will not only improve the model's robustness but also enhance its practical usability in diverse traffic conditions and environments. By addressing these limitations and expanding the scope of variables included in the analysis, we aim to set a new standard in traffic flow prediction accuracy, particularly during non-periodic peak flow periods.

ACKNOWLEDGEMENTS

Partial financial support was received from the National Natural Science Foundation of China (No. U21A20484, No. 61976074), Fundamental Research Funds for the Provincial Universities of Zhejiang (No. GK239909299001-011), IOT Technology Application Transportation Industry R & D Centre Open Project Funding (No. 2023-05) and Zhejiang Provincial Natural Science Foundation of China (No. LY19F020042).

REFERENCES

[1] Kong X, et al. Urban traffic congestion estimation and prediction based on floating car trajectory data. *Future Generation Computer Systems*. 2016;61:97–107. DOI: 10.1016/j.future.2015.11.013.

[2] Montanino M, Monteil J, Punzo V. From homogeneous to heterogeneous traffic flows: Lp string stability under uncertain model parameters. *Transportation Research Part B: Methodological*. 2021;146:136–154. DOI: 10.1016/j.trb.2021.01.009.

[3] Chen J, et al. A review of vision-based traffic semantic understanding in ITSs. *IEEE Transactions on Intelligent Transportation Systems*. 2022;23(11):19954–19979. DOI: 10.1109/TITS.2022.3182410.

[4] Chen J, et al. Disparity-based multiscale fusion network for transportation detection. *IEEE Transactions on Intelligent Transportation Systems*. 2022;23(10):18855–18863. DOI: 10.1109/TITS.2022.3161977.

[5] Liu C, et al. A vision-based pipeline for vehicle counting, speed estimation, and classification. *IEEE Transactions on Intelligent Transportation Systems.* 2020;22(12):7547–7560. DOI: 10.1109/TITS.2020.3004066.

[6] Aboah A. A vision-based system for traffic anomaly detection using deep learning and decision trees. *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 19-25 June 2021, Nashville, TN, USA*. 2021. p. 4207–4212. DOI: 10.1109/CVPRW53098.2021.00475.

[7] Siregar ML, Tjahjono T. Speed characteristics of heterogeneous traffic on inter-urban roads in Indonesia. *Transportation research*. 2021;52:404-411. DOI: 10.1016/j.trpro.2021.01.047.

[8] Han J, Wang X, Wang G. Modeling the car-following behavior with consideration of driver, vehicle, and environment factors: A historical review. *Sustainability*. 2022;14(13):8179. DOI: 10.3390/su14138179.

[9] Hobeika AG, Kim CK. Traffic-flow-prediction systems based on upstream traffic. *Proceedings of VNIS'94-1994 Vehicle Navigation and Information Systems Conference, 31 Aug. 1994 - 02 Sept. 1994, Yokohama, Japan*. 1994:345–350. DOI: 10.1109/VNIS.1994.396815.

[10] Ganji A, Zhang M, Hatzopoulou M. Traffic volume prediction using aerial imagery and sparse data from road counts. *Transportation research part C: emerging technologies*. 2022;141:103739. DOI: 10.1016/j.trc.2022.103739.

[11] Haghani A, et al. Data collection of freeway travel time ground truth with bluetooth sensors. *Transportation Research Record*. 2010;2160(1):60–68. DOI: 10.3141/2160-07.

[12] Adu-Gyamfi Y, et al. Reliability of probe speed data for detecting congestion trends. *2015 IEEE 18th International Conference on Intelligent Transportation Systems, 15-18 Sept. 2015, Gran Canaria, Spain*. 2015. p. 2243-2249. DOI: 10.1109/ITSC.2015.362.

[13] He R, et al. Deep spatio-temporal 3D densenet with multiscale ConvLSTM-Resnet network for citywide traffic flow forecasting. *Knowledge-Based Systems*. 2022;250:109054. DOI: 10.1016/j.knosys.2022.109054.

[14] He K, Zhang X, Ren S, Sun J. Deep residual learning for image recognition. *Proceedings of the IEEE conference on computer vision and pattern recognition, 27-30 June 2016, Las Vegas, NV, USA*. 2016. p. 770–778. DOI: 10.1109/CVPR.2016.90.

[15] Guo K, et al. Hierarchical graph convolution network for traffic forecasting. *Proceedings of the AAAI conference on artificial intelligence*. 2021;35(1):151–159. DOI: 10.1609/aaai.v35i1.16088.

[16] Zhang X, et al. Traffic flow forecasting with spatial-temporal graph diffusion network. *Proceedings of the AAAI conference on artificial intelligence*. 2021;35(17):15008–15015. DOI: 10.1609/aaai.v35i17.17761.

[17] Yang D, et al. MF-CNN: traffic flow prediction using convolutional neural network and multi-features fusion. *IEICE TRANSACTIONS on Information and Systems*. 2019;102(8):1526–1536. DOI: 10.1587/transinf.2018EDP7330.

[18] Yang X, Yuan Y, Liu Z. Short-term traffic speed prediction of urban road with multi-source data. *IEEE Access.* 2020;8:87541–87551. DOI: 10.1109/ACCESS.2020.2992507.

[19] Qian ZS, et al. Modeling heterogeneous traffic flow: A pragmatic approach. *Transportation Research Part B: Methodological*. 2017;99:183–204. DOI: 10.1016/j.trb.2017.01.011.

[20] Chen J, et al. A flow feedback traffic prediction based on visual quantified features. *IEEE Transactions on Intelligent Transportation Systems*. 2023;24(9):10067–10075. DOI: 10.1109/TITS.2023.3269794.

[21] Wang P, Hao W, Jin Y. Fine-grained traffic flow prediction of various vehicle types via fusion of multisource data and deep learning approaches. *IEEE Transactions on Intelligent Transportation Systems*. 2020;22(11):6921–6930. DOI: 10.1109/TITS.2020.2997412.

[22] Cho K, et al. Learning phrase representations using RNN encoder-decoder for statistical machine translation. *Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP), 25-29 Oct. 2014, Doha, Qatar*. 2014. p. 1724–1734. DOI: 10.3115/v1/D14-1179.

[23] Snyder C, Do M. Streets: A novel camera network dataset for traffic flow. *Advances in Neural Information Processing Systems.* 2019;32. https://proceedings.neurips.cc/paper_files/paper/2019/file/ee389847678a3a9d1ce9e4ca69200d06-Paper.pdf

[24] Park D, Rilett LR. Forecasting freeway link travel times with a multilayer feedforward neural network. *Computer-Aided Civil and Infrastructure Engineering*. 1999;14(5):357–367. DOI: 10.1111/0885-9507.00154.

[25] Abdelraouf A, Abdel-Aty M, Mahmoud N. Sequence-to-sequence recurrent graph convolutional networks for traffic estimation and prediction using connected probe vehicle data. *IEEE Transactions on Intelligent Transportation Systems*. 2022;24(1):1395–1405. DOI: 10.1109/TITS.2022.3168865.

[26] Fang M, et al. FTPG: A fine-grained traffic prediction method with graph attention network using big trace data. *IEEE Transactions on Intelligent Transportation Systems*. 2021;23(6):5163–5175. DOI: 10.1109/TITS.2021.3049264.

王钦阳，陈婧，宋颖，李小冬，许文强

一种融合视觉量化特征的异质交通流预测方法

摘要：

本文提出了一种强调异质车辆特征和视觉密度特征的新的交通流预测方法。传统交通流预测模型往往会忽视车辆的多样性，导致预测不准确。提出的方法利用视觉技术量化混合流量和车辆堆积等交通特征，加强了动态密度估计和流量流动性。我们引入了一个时空预测模型，其集成了各种数据类型，能捕获复杂的依赖关系并提高准确性。本研究通过考虑车辆的多样性并利用视觉数据来提升交通流量预测性能，为智能交通系统提供有价值的见解。实验结果表明，提出的方法优于传统方法，特别是在捕获交通流波动方面。

关键词：

异质交通流，时空建模，交通流量预测，视觉交通量化