Dean&Francis
ISSN 2959-6157

# Traffic Flow Prediction Using Deep Learning: Advances, Challenges, and Future Directions

**Kai Liu**

Macau University of Technology and Science, Macau, China
Corresponding author:
1210031664@student.must.edu.mo

**Abstract:**

As urbanization accelerates globally and intelligent transportation technologies rapidly progress, the accurate prediction of road traffic has emerged as a critical issue in the field of smart transportation. This paper introduces a comprehensive review of current research on traffic flow prediction utilizing deep learning and related technologies. It analyzes the limitations of current methods, such as poor generalization to spatiotemporal heterogeneity, reliance on external influencing factors, data quality and quantity issues, and insufficient model explainability and computational scalability. Furthermore, the paper outlines the development trends of the field from the perspective of multimodal data fusion, proposing a three-layer fusion framework of data, models, and systems. Emphasis is placed on ensuring data security and privacy through multimodal data fusion and federated learning. The study also discusses future directions, including dynamic feature modeling, system deployment at the edge, and real-time prediction. By analyzing the architecture and challenges of current predictive models, this article offers theoretical direction and technological insights for the advancement of intelligent transportation technology.

**Keywords:** Deep learning; Traffic flow prediction; Spatiotemporal dependencies; Multimodal data fusion; Edge computing.

## 1. Introduction

In recent decades, with the ongoing global metropolitan growth and the surge in vehicle ownership, congested roadways have become a prominent obstacle to sustainable urban development. By early 2025, the number of registered vehicles worldwide is expected to exceed 1.5 billion. The resulting increase in traffic demand puts immense pressure on road infrastructure, leading to reduced travel efficiency, elevated accident rates, and a decline in overall quality of life. Traffic flow forecasting, as a fundamental element of intelligent transportation technologies, is essential for optimizing road usage and improving transportation

ISSN 2959-6157

management.

Traffic flow prediction leverages historical data and machine learning techniques to forecast future traffic conditions. It enables proactive traffic control and intelligent traffic management by providing data-driven decision support. Accurate prediction of traffic flow can enhance route planning [1], reduce congestion, improve road safety, and support the development of smart cities. Traditional traffic prediction models are primarily grounded in statistical analysis and shallow machine learning techniques. Representative examples include Autoregressive Integrated Moving Average (ARIMA) models [2], Support Vector Regression (SVR) [3], and other linear regression-based approaches. However, such methods struggle to effectively model the complex spatial and temporal correlations, as well as nonlinear and dynamic traffic patterns. Consequently, their prediction accuracy and generalization performance remain unsatisfactory.

Recent advancements in deep learning and big data technologies have led to the increased application of different deep neural network architectures—such as Recurrent Neural Networks (RNNs), Convolutional Neural Networks (CNNs) [4], Long Short-Term Memory (LSTM), and Gated Recurrent Units (GRUs). These methods offer strong feature representation capabilities and can model complex spatiotemporal dependencies, demonstrating significant advantages in improving prediction performance in traffic prediction tasks.

Nevertheless, current research still faces a range of challenges. These include external influencing factors (e.g., weather, social events), multimodal data integration, model interpretability, and privacy protection in distributed environments. Therefore, the effective integration of heterogeneous data and the practical implementation of prediction models remain active research topics. This study conducts a thorough evaluation of current research on deep learning-based traffic forecasting and summarizes prevailing challenges as well as future directions. It intends to function as a significant reference for both theoretical research and practical implementations in intelligent transportation technologies.

# 2. Problem Definition, Datasets, and Evaluation Metrics

## 2.1 Problem Definition

Prediction of traffic flows involves the task of analyzing and modeling traffic networks and their historical flow data in order to forecast the traffic volume of specific road segments over future time intervals. In this task, the traffic is commonly shown using a structure G = (V, E, A), where V refers traffic monitoring nodes, E denotes edges connecting these nodes, A is the adjacency matrix encoding connectivity and weights among nodes in the network.

Based on the prediction range, prediction of traffic flow is generally categorized into two types: (1) Short-term prediction, denoted as $\hat{T}_{short}$, typically with a time granularity of hours; (2) Long-term prediction, denoted as $\hat{T}_{long}$, which may span days, weeks, or even longer periods.

Input of prediction task is formulated as a time series sequence $X_T = (x_1, \ldots x_T) \in R^{N \times C \times T}$, where $x_t \in R^{N \times C \times T}$ indicates the traffic status of all monitoring nodes at timestamp t; $N$ denotes the nodes count within the traffic monitoring network; $C$ represents the number of traffic-related features collected at each node, such as speed, weather conditions, or traffic events.

The core challenge of this problem lies in learning from both current and historical traffic states while incorporating the spatial topology of traffic network to capture the complex spatiotemporal dependencies among road segments. The modeling aim is to develop a mapping function $f$ to predict the status at the subsequent T' timestamp, formulated as:

$$\left(X_{t-\tau}, \ldots, X_t; G\right) \overset{f}{\rightarrow} \left(Y_t, \ldots, Y_{t+T}\right). \quad (1)$$

## 2.2 Key features

Traffic flow prediction is a complex task due to the following key characteristics: (1) Non-linearity: The evolution of traffic flow is inherently non-linear, as it involves complex interactions among temporal and external factors. For instance, road congestion often exhibits threshold effects, where even small perturbations can lead to significant, abrupt changes. Additionally, external factors such as weather and accidents contribute further non-linear dynamics. As a result, traditional linear prediction models (e.g., autoregressive models) are often inadequate for capturing the complex non-linear behaviors in traffic systems; (2) Dynamic nature: Traffic systems are highly dynamic, with traffic flow continuously fluctuating and being influenced by sudden events, resulting in abrupt changes and intense volatility. For example, rush hours, road closures, or extreme weather events can lead to unpredictable and transient disruptions. Accurate forecasting requires models that can encapsulate both transient variations and enduring trends. However, traditional models that rely on stable or smoothed historical patterns often struggle to adapt to such dynamic environments; (3) Heterogeneity: Traffic data exhibits significant spatial and temporal heteroge-

neity. Different regions and time periods often display distinct traffic patterns and variations. For example, expressways and urban roads, weekdays and weekends, may follow drastically different flow patterns. This heterogeneity makes it difficult for models trained on one scenario to generalize well to others. Thus, models must be capable of adapting to diverse spatial and temporal distributions; (4) Multi-scale characteristics: Traffic flow exhibits dependencies across multiple spatial and temporal scales. Temporally, traffic may show periodicity on hourly, daily, or weekly scales, alongside random fluctuations. Spatially, traffic patterns can vary significantly across regions, from individual road segments to the entire urban road network. Therefore, effective models must integrate features at various spatial and temporal scales to better capture the hierarchical and the inherent multi-level nature of traffic flow.

## 2.3 Traffic Flow Prediction Datasets

To evaluate and compare different traffic flow prediction methods, the research community has established a series of publicly available benchmark datasets. Among them, several representative datasets include:

METR-LA Dataset: The METR-LA dataset is a traffic speed dataset for the Los Angeles County highway network, collected and released by the University of Southern California and other collaborators. It includes traffic speed readings from 207 sensor stations across the Los Angeles area, covering the period from March 2012 to June 2012, with a sampling interval of 5 minutes. The dataset features traffic conditions on complex urban highways and is widely adopted as a benchmark for spatiotemporal prediction models due to its data quality, spatial coverage, and open accessibility. It is particularly useful for evaluating model performance under urban traffic dynamics.

TaxiBJ Dataset: TaxiBJ is a large-scale open dataset for urban traffic volume prediction. It contains GPS trajectory and related contextual information from Beijing taxi trips. A typical subset used for research transforms the raw data into a 32×32 spatial grid format, with aggregated traffic volume recorded at 30-minute intervals. The dataset spans four distinct time periods from 2013 to 2016. Due to its inclusion of large-scale, high-resolution urban mobility data, TaxiBJ has been widely employed for modeling and predicting city-level traffic demand, human mobility, and flow patterns. It is one of the most frequently used open-source datasets for spatiotemporal deep learning in traffic prediction tasks.

PeMS Dataset: The PeMS (California Performance Measurement System) is a traffic flow data collection platform maintained by the California Department of Transportation. It provides live traffic information, including vehicle count and speed metrics, gathered from more than 45,000 sensors installed across major California freeways. PeMS offers abundant historical records. For instance, the PeMS-Bay dataset includes traffic flow readings from 325 sensor stations across the Bay Area, recorded every 5 minutes from January to May 2017. Commonly used subsets in research include PeMS-D4, D7, and D8, each corresponding to data from different regions (e.g., PeMS-D4 covers 307 sensors in the San Francisco Bay Area over a 2-month period). Owing to its large scale, fine granularity, and public accessibility, PeMS has become a widely used benchmark for validating traffic prediction models.

## 2.4 Model Evaluation Metrics

In traffic flow prediction research, model performance is commonly evaluated by quantifying the error between forecasted values and ground truth observations. The most often adopted assessment metrics comprise the following:

Mean Absolute Error (MAE): MAE quantifies the average extent of the absolute difference among forecasted and true results. It is defined as:

$$MAE = \frac{1}{n}\sum_{i=1}^{n}\left|y_i - \widehat{y_i}\right|, \tag{2}$$

where $y_i$ represents the true values, whereas $\widehat{y_i}$ signifies the predicted values, and $N$ represents the total number of predictions. $MAE$ reflects average level of prediction bias. A smaller $MAE$ indicates better total predictive accuracy of the model.

Root Mean Square Error (MSE): RMSE is the square root of the average of squared prediction errors, defined as:

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i - \widehat{y_i})^2}. \tag{3}$$

Compared to MAE, RMSE places greater emphasis on larger deviations because it involves squaring the prediction errors, which increases its sensitivity to outliers. A lower $RMSE$ suggests that the model has higher precision and better stability in its predictions.

Mean Absolute Percentage Error (MAPE): MAPE evaluates average relative forecasting error expressed as percentage, defined as:

$$MAPE = \frac{100\%}{n}\sum_{i=1}^{n}\left|\frac{y_i - \widehat{y_i}}{y_i}\right|. \tag{4}$$

It indicates the mean percentage divergence of the predicted values from the actual values. As it is formulated in percentage terms, MAPE allows comparison of model performance across datasets of different scales. However, care should be taken when $y_i$ is close to zero, as this can result in disproportionately large errors or instability in

ISSN 2959-6157

MAPE.

In general, the lower the values of the above metrics, the more accurate the model's predictions. These metrics are frequently used in the literature to assess model effectiveness and facilitate fair comparisons across different approaches.

## 3. Deep Learning-Based Models for Traffic Flow Forecasting

Deep learning has emerged as a key methodology for traffic flow forecasting. Due to the rapid expansion of data and increased computing power, traditional prediction methods have shown limitations in feature extraction. Deep neural networks, by contrast, can automatically learn intricate spatiotemporal patterns from massive traffic data, yielding far superior predictive power. This section provides a summary of traffic flow forecasting models based on deep learning. As shown in Table 1, it compares their methodological advances and performance.

**Table 1. Comparative Analysis of Different Models in Traffic Flow Prediction**

| Model Name | Main Strategy | Applicable Scenario | Characteristics |
|---|---|---|---|
| MF-CNN | Multi-source feature fusion | Complex traffic environments influenced by multiple factors | Strong adaptability; capable of handling weather, holidays, and other disturbances |
| DGCNN | Temporal feature-aware graph convolution | Scenarios requiring high temporal dependencies | Effectively captures complex spatial structures in road networks |
| 3D-CNN | Spatiotemporal convolution (TF-3DNet) | Long-term data missing or sparse scenarios | High spatiotemporal representation capability; improved prediction performance on incomplete data |
| DST-3D-CNN | Spatiotemporal coordination + region-aware sensitivity | Scenarios with diverse traffic patterns across different regions | Captures regional behavioral differences and conducts fine-grained dynamic modeling |
| MGDAM | Multi-graph attention mechanism + spatial saliency | Mid-to-long-term traffic trend forecasting | Captures dependencies between key nodes; applicable to both traffic volume and speed transmission |
| DCRNN | Diffusion convolutional recurrent structure | Regions with clear traffic flow propagation patterns | Leverages realistic traffic diffusion characteristics; suitable for modeling directional traffic signals |
| Multivariate LSTM | Multivariate time series modeling | Early-morning traffic with strong inter-variable correlations | Supports multivariate input; effective for capturing correlations among variables |
| LSTM+ | Attention mechanism integrated into LSTM | Complex and fine-grained long-term forecasting tasks | Enhances model sensitivity to key features and improves memory capability |
| GRU-RNN | GRU-based recurrent neural network | Urban traffic prediction with textual or time-series signals | Lightweight architecture; suitable for real-time and nonlinear traffic prediction |
| TrafficTransformer | Transformer-based sequence modeling | Urban-level traffic incident prediction | Supports parallelism, multi-head attention, and long-term sequence learning |
| TrafficGAN | Adversarial training mechanism | Forecasting under highly dynamic or rare traffic events | Captures complex distributions and improves adaptability to anomalous scenarios |

### 3.1 MCNN-Based Traffic Flow Prediction Models

CNNs have exhibited exceptional efficacy in computer vision and have been applied to traffic flow forecasting as well. In this context, traffic flow data can be structured analogously to images: for example, a city map is divided into grids so that each pixel represents the traffic volume or speed in a specific region over time. By leveraging techniques from image recognition, CNN models can effectively capture the spatial relationships and spatiotemporal features in traffic data. This allows CNNs to learn patterns of traffic flow distribution across both space and

time, which is crucial for accurate prediction.

However, vanilla CNN approaches still face challenges in modeling complex temporal dynamics and incorporating external influences on traffic. To overcome these limitations, researchers have suggested many advanced CNN-based architecture in traffic flow prediction. Key CNN-based models and their contributions include:

Data Grouping CNN (DGCNN) – *Yu et al.* introduced a data grouping strategy that partitions the input traffic data along dimensions such as time intervals and spatial regions before feeding it into a CNN [5]. By grouping data based on temporal and spatial characteristics, their DGCNN model enables the network to better capture complex localized traffic patterns and improves adaptability and prediction accuracy. Empirical studies showed that this method significantly enhanced short-term traffic flow forecasting accuracy compared to standard approaches.

Multi-Feature Fusion CNN (MF-CNN) – *Yang et al.* proposed an MF-CNN model to integrate external factors (e.g. weather conditions, holidays, temperature, wind speed) into CNN-based traffic prediction [6]. The MF-CNN architecture employs an early-fusion strategy with time alignment: external variables are concatenated with main traffic features – such as short-term temporal continuity (denoted C), daily periodicity (D), weekly periodicity (W) – along feature dimension, forming a unified multi-dimensional input tensor. This design ensures that external factors are temporally synchronized with the traffic sequence, allowing the CNN to jointly model their influence on traffic flow. In experiments on two datasets (JPEA and PeMS), the inclusion of these external features led to continuous performance improvements; notably, after adding external factors (denoted E), the Mean Absolute Error (MAE) dropped from 0.0098 to 0.0096 on JPEA and from 0.0257 to 0.0254 on PeMS, effectively improving prediction accuracy. These results empirically validate the positive contribution of incorporating external factors in multi-step traffic flow prediction.

3D-CNN – To further capture dynamic temporal features and spatial dependencies in traffic data, *Yu et al.* developed a 3D Convolutional Neural Network for large-scale traffic flow prediction [7]. Unlike a conventional 2D CNN that models only spatial patterns on a static grid, a 3D-CNN uses three-dimensional convolutional kernels to simultaneously capture and integrate across both space and time. This joint spatiotemporal modeling approach enables the network to learn the temporal evolution of traffic directly within the convolution layers. The 3D-CNN method was shown to improve prediction accuracy for large traffic networks, and it also demonstrated greater robustness in the face of missing data by learning temporal continuity – in other words, the model can better interpolate or withstand gaps in sensor data due to its understanding of how traffic flows evolve over time.

Deep Spatial-Temporal 3D-CNN (DST-3D-CNN) – Building on the 3D-CNN, *Guo et al.* proposed an enhanced model called DST-3D-CNN (also referred to as ST-3DNet) tailored for traffic data forecasting [8]. The DST-3D-CNN inherits the joint spatiotemporal convolutional modeling of 3D-CNN, but further introduces a novel recalibration (Rc) block to dynamically adjust for regional differences in traffic patterns. This Rc block enables the model to explicitly account for spatial heterogeneity – i.e. the fact that different road segments or regions may exhibit different traffic dynamics – which earlier CNN models with uniform convolution often ignored. By recalibrating feature responses based on location-specific characteristics, DST-3D-CNN better captures complex, non-uniform spatial influences. Empirical results indicate that this approach achieves superior performance in complex urban traffic scenarios, outperforming prior CNN models by effectively addressing the issue of spatial heterogeneity. In summary, DST-3D-CNN demonstrates more robust predictive capability in city-scale traffic networks with diverse regional traffic behaviors.

### 3.2 Traffic Flow Forecasting Using Graph Convolutional Networks (GCN)

Compared with traditional CNNs, GCNs are better suited for modeling the complex spatial structures of non-Euclidean domains like road networks [9]. The core idea behind GCNs is to enable each node in the traffic graph to aggregate its own features along with those of its neighbors, thereby learning spatial dependencies across the network. The standard layer-wise propagation rule for GCNs, as proposed by Kipf et al. [10], is formulated as:

$$H^{l} = f\left(H^{l-1}, A\right) = \sigma\left(D^{-\frac{1}{2}} A D^{-\frac{1}{2}} H^{l-1} W^{l-1}\right) \tag{5}$$

In Equation (2), $X \in R^{N \times F_0}$ is the input feature matrix, with N indicating the number of nodes and $F_0$ denoting the quantity of starting features per node. $A \in R^{N \times N}$ is an adjacency matrix encoding spatial connections between nodes. However, multiplying $A$ with the hidden feature matrix only allows information from neighboring nodes to propagate, and excludes the node's own features. To resolve this, each node is connected to itself via a self-loop, modifying the adjacency matrix to $A = A + I$, where $I$ denote the identity matrix. The diagonal degree matrix $D$ is calculated by $D_{ii} = \sum_j A_{ij}$. The initial hidden represen-

ISSN 2959-6157

tation is set as $H^{(0)} = X$, $W^{l-1} \in R^{F^{l-1} \times F^l}$ represents a trainable weight matrix and $\sigma(\cdot)$ denotes an activation function, typically sigmoid or ReLU.

While this propagation rule has been widely applied in GCN-based models, it has limitations in traffic flow prediction. Specifically, it assumes a static graph structure and cannot model time-varying upstream and downstream traffic relationships. To address this limitation, *Abu-El-Haija et al.* proposed incorporating graph attention mechanisms [11], which assign learnable attention weights to neighbors, enabling the model to adaptively modulate each node's impact over time. This approach helps the model better represent the real-world, time-sensitive variations in traffic networks.

In order to better model spatial–temporal correlations, *Li et al.* proposed the Diffusion Convolutional Recurrent Neural Network (DCRNN) for traffic prediction [12]. This model employs diffusion convolution to represent the directed nature of traffic propagation, integrated with recurrent components to capture temporal dynamics.

Extending this line of work, *Wang et al.* developed the Multi-Graph Diffusion Attention Network (MGDAN) [13]. Unlike DCRNN, which uses a static graph, MGDAN constructs multiple adaptive graphs and employs attention mechanisms to model spatial correlations over long range and nonlinear temporal relationships. This enhances the model's ability to comprehend intricate spatiotemporal patterns and improves its performance in long-term traffic prediction tasks. This greatly enhances the modeling capability for long-range spatial dependencies and nonlinear temporal patterns. This enables the effective capture of long-term temporal dynamics, thereby enhancing the performance of medium and long-term traffic flow prediction.

### 3.3 RNNs for Traffic Flow Forecasting

Traffic flow data are a typical form of sequential data, and extracting its temporal characteristics is critically important for accurate traffic flow prediction. RNNs and their variants, such as LSTM networks and GRU-RNNs, have therefore been widely utilized in this domain. *Baskar et al.* developed a multivariate LSTM-based model for short-term traffic flow prediction in intelligent-driven transportation technologies [14], and their results indicated that compared to univariate LSTM models, the proposed approach achieves higher prediction accuracy.

However, as the length of the input time series grows (e.g. beyond two hours of data), even LSTM models struggle to capture such long-range temporal relationships in traffic flow. *Yang et al.* addressed this limitation by incorporating an attention mechanism into the LSTM's temporal modeling in order to identify the long-sequence traffic flow features that have a significant impact on prediction results [15]. On this basis, they proposed a feature-enhanced LSTM model (denoted as LSTM+), which strengthens the LSTM's capability to retain information over ultra-long time dependencies.

Many traffic flow prediction models also suffer from the vanishing gradient problem, resulting in information loss throughout the training process. To address the issue, *Yu et al.* introduced a RNN model using a GRU-RNN approach to elucidate the interdependencies within traffic flow data [16]. This model leverages specialized activation and gating functions to mitigate the issues of gradient explosion and vanishing in RNN training, thereby achieving high-accuracy predictions even for long-range traffic flow forecasting.

### 3.4 Transformer-Based Traffic Flow Forecasting

Compared with traditional CNN and RNN architectures, the Transformer model demonstrates stronger capabilities in handling sequential tasks. Unlike RNNs, Transformers offer superior parallel computation and can capture global contextual information more efficiently through self-attention mechanisms, enhancing the modeling of complex dependencies in traffic data. The Transformer builds a complete dependency structure between input and output solely via attention mechanisms.

*Al-Thani et al.* introduced a Transformer-based traffic forecasting model [17], Traffic Transformer, to perform multivariate and multi-step prediction. The model utilizes an attention mechanism of the transformer to capture long-term temporal relationships by turning irregular traffic data into a regular time-series structure. Compared with RNN-based architectures such as LSTM, this model shows enhanced parallelism and modeling ability, resulting in improved performance in multi-step forecasting tasks.

### 3.5 Traffic Flow Prediction Based on GANs

GANs constitute a cutting-edge advancement in deep learning and have been utilized for traffic flow forecasting. Neural network models can generally be categorized as discriminative or generative. A discriminative model learns the conditional probability distribution between inputs and outputs, thereby allowing it to predict outputs given certain inputs. In contrast, a generative model learns the joint distribution of inputs and outputs and can generate new sample data that resemble the distribution of the training data. By leveraging an adversarial training process between a discriminator and a generator, GANs effectively fuse the spatial and temporal dependencies inherent in traffic flow data, which improves the modeling of

temporal correlations and enhances the accuracy of traffic state estimation [18]. In other words, the GAN framework enables the model to capture complex time-dependent patterns in traffic flow more effectively than conventional predictive models.

However, traditional traffic flow prediction models have limited capability to learn such complex and dynamic data distributions. They often struggle to handle the highly variable and non-stationary traffic patterns observed in real-world road networks. To address these challenges, *Zhang et al.* introduced a GAN-based traffic flow forecasting model called TrafficGAN [19], aimed at short-term traffic prediction at the urban road-network scale. TrafficGAN integrates a CNN and a LSTM network to jointly capture temporal and spatial correlations in traffic data. Within this architecture, the CNN component is responsible for extracting and modeling spatial characteristics of the roadway system (such as connectivity or proximity of road segments), and the LSTM portion captures sequential temporal dynamics of traffic flow over time. Moreover, TrafficGAN employs deformable convolutional kernels to more efficiently process the spatial heterogeneity in the road network, allowing model to adapt to irregular traffic network structures. Through its adversarial training mechanism (in which a generator produces predicted traffic flow sequences and a discriminator evaluates their realism), TrafficGAN learns the underlying distribution of historical traffic flow patterns. Consequently, it can generate predictions that more closely align with true future traffic conditions, achieving higher accuracy in forecasting future traffic volumes compared to conventional methods.

# 4. Challenges in Road Traffic Flow Forecasting

Precise forecasting of traffic flow is essential for efficient transportation management. Despite the advancements brought by deep learning models—such as better utilization of large-scale data and improved accuracy—several real-world challenges remain:

Dynamic and Uncertain Influences: Traffic flow is affected by diverse external variables, including road conditions, accidents, weather, special events. These influences interact and contribute to the highly uncertain and dynamic nature of traffic states, making stable prediction difficult.

Spatio-temporal Dependency Modeling: Traffic data exhibits complex spatio-temporal patterns. Variations arise from periodic trends and sudden changes, while the structural complexity of road networks introduces significant spatial heterogeneity. Even geographically similar areas may show different flow patterns due to factors like population and infrastructure differences.

Data Quality Issues: Deep learning relies on high-quality data, yet traffic datasets often suffer from missing values, noise, or sensor failures. These issues degrade model performance and complicate consistent evaluation across methods, especially when data is collected from heterogeneous sources.

Deployment Constraints: Deploying deep learning models in real-world environments is challenging due to limited computing resources on edge devices like in-vehicle or roadside units. The complexity of many models hampers real-time application, highlighting the need for lightweight architectures and model compression techniques.

# 5. Challenges in Road Traffic Flow Prediction

## 5.1 Data Layer

In the data layer, future research should emphasize multi-source data fusion to integrate diverse traffic data and extract richer information. This involves incorporating heterogeneous data from various sources (for example, combining loop detector data with GPS probe data, camera feeds, and even mobile phone signaling data) to provide a more comprehensive basis for prediction.

It is important to include relevant external factors like traffic incidents or weather conditions—so that models can account for these influences on traffic flow. Additionally, improving data quality through cleaning and preprocessing, as well as establishing real-time data updating mechanisms, will enhance the reliability of predictions.

By enriching data diversity and quality in this way, we can better capture underlying traffic patterns and spatiotemporal features needed for accurate forecasting.

## 5.2 Model Layer

In the model layer, the focus is on developing advanced models and algorithms that can more effectively capture the intricate spatial and temporal dynamics of traffic flow. Future models should leverage sophisticated neural network architectures to exploit these patterns—for instance, applying graph convolutional networks (GCN) to model spatial relationships among road segments, and recurrent networks or attention-based Transformers to represent temporal dependencies and evolving trends.

There is value in combining the strengths of different approaches (such as CNN, GCN, RNN, GRU, and Transformer models) to form hybrid models that can learn multi-scale and multi-dimensional features of traffic data. Incorporating attention mechanisms can help the model

ISSN 2959-6157

concentrate on key temporal–spatial information, and integrating domain knowledge (e.g. known traffic flow laws or constraints) may improve interpretability and robustness.

Moreover, researchers are beginning to explore the use of large-scale pre-trained models and transfer learning in this field—foundation models like GPT and advanced patio-temporal networks such as DeepSTN+ could be adapted to traffic prediction to further boost performance. These directions aim to improve prediction accuracy, stability, and interpretability by leveraging comprehensive data attributes and cutting-edge AI methodologies.

### 5.3 Edge-End Collaboration

Another important future direction is the edge-end collaboration in intelligent traffic prediction systems. With the proliferation of IoT devices and advances in edge computing, parts of the predictive computation can be deployed on distributed edge devices (such as sensors, connected vehicles, or roadside units) in coordination with centralized cloud servers. This collaborative computing paradigm reduces latency and bandwidth usage by processing data locally at the edge and only transmitting necessary information to the cloud, enabling more timely and responsive traffic flow predictions.

For example, an edge device at an intersection could perform initial data filtering or local prediction updates, which are then integrated with a global model at the cloud, combining real-time local insights with a broader network-wide perspective. Such an approach can also enhance reliability and privacy – local processing means critical data can be analyzed on-site without constant cloud communication, and systems remain functional even if connectivity is temporarily lost.

Moving forward, research will need to address challenges like model optimization for resource-constrained edge hardware, efficient coordination protocols between edge and cloud, and possibly federated learning techniques to train models across multiple edge nodes. Overall, cloud-edge-end collaborative frameworks are expected to form a key part of future traffic prediction architectures, allowing systems to deliver faster, more scalable, and context-aware forecasting of road traffic conditions.

## Conclusion

This paper presents a comprehensive review of deep learning-based approaches for traffic flow prediction in the context of smart transportation. It highlights key limitations of existing methods, including poor spatiotemporal generalization, dependence on external factors, data quality and availability issues, and limited model interpretability and scalability.

## References

[1] Liebig T, Piatkowski N, Bockermann C, Morik K. Dynamic route planning with real-time traffic predictions. Information Systems, 2017, 64: 258–265.

[2] Shi, Q. X., & Zheng, W. Z. (2004). A comparison of short-term traffic flow prediction methods for road networks. Journal of Transportation Engineering, (04), 68–71, 83.

[3] Lin, H., Li, L. X., & Wang, H. (2020). A review of research and applications of support vector machines in intelligent transportation systems. *Journal of Computer Science and Exploration*, 14(06), 901–917.

[4] Ma X, Dai Z, He Z, Ma J, Wang Y, Wang Y. Learning traffic as images: a deep convolutional neural network for large-scale transportation network speed prediction. Sensors, 2017, 17(4): 818.

[5] Yu D, Liu Y, Yu X. A data grouping CNN algorithm for short-term traffic flow forecasting. Proceedings of the 2016 International Conference on Neural Information Processing (ICONIP). Berlin: Springer, 2016: 92–103.

[6] Yang D, Li S, Peng Z, et al. MF-CNN: Traffic flow prediction using convolutional neural network and multi-features fusion. IEICE Transactions on Information and Systems, 2019, E102. D(8): 1526-1536.

[7] Yu F, Wei D, Zhang S, Li J, Chen M. 3D CNN-based accurate prediction for large-scale traffic flow. Proceedings of the 2019 4th International Conference on Intelligent Transportation Engineering (ICITE). Singapore: IEEE, 2019: 99–103.

[8] Guo S, Lin Y, Li S, Chen Z, Wan H. Deep spatial-temporal 3D convolutional neural networks for traffic data forecasting. IEEE Transactions on Intelligent Transportation Systems, 2019, PP(early access): 1–14.

[9] Wu B, Liang X, Zhang S, Xu R. Advances and applications of graph neural networks: A survey. Journal of Computer Research and Development, 2022, 45(01): 35–68.

[10] Kipf TN, Welling M. Semi-supervised classification with graph convolutional networks. Proceedings of the 5th International Conference on Learning Representations (ICLR). Toulon, France: ICLR, 2017: 1–10.

[11] Abu-El-Haija S, Perozzi B, Al-Rfou R, et al. Watch your step: Learning node embeddings via graph attention. Proceedings of the 32nd International Conference on Neural Information Processing Systems (NIPS). Montréal, Canada: NIPS, 2018: 9198–9208.

[12] Li Y, Yu R, Shahabi C, Liu Y. Graph Convolutional Recurrent Neural Network: Data-Driven Traffic Forecasting. 2017-07-06 [2024-12-26].

[13] Wang Q, Lu Q X, Shi P. Multi-graph diffusion attention network for traffic flow prediction. Computer Applications, 2024, (0): 1–10 [2024-12-23].

[14] Praveen Kumar B, Hariharan K. Multivariate Time Series Traffic Forecast with Long Short-Term Memory based Deep Learning Model. IEEE. 2020 International Conference on Power, Instrumentation, Control and Computing (PICC). Thrissur, India: IEEE, 2020: 1–5.
[15] Yang B, Sun S, Li J, Lin X, Tian Y. Traffic flow prediction using LSTM with feature enhancement. Neurocomputing, 2019, 332(C): 320–327.
[16] Yu D X, Qiu S, Zhou H X, Wang Z R. Short-term traffic flow prediction at intersections based on the GRU-RNN model. Highway Engineering, 2020, 45(04): 109–114.
[17] Al-Thani MG, Sheng Z, Cao Y, et al. Traffic Transformer: Transformer-based framework for temporal traffic accident prediction. AIMS Mathematics, 2024, 9(5): 12610-12629.
[18] Liang Y, Cui Z, Tian Y, Chen H, Wang Y.ADeep Generative Adversarial Architecture for Network-Wide Spatial-Temporal Traffic-State Estimation. Transportation Research Record, 2018, 2672(45): 87–105.
[19] Zhang Y, Wang S, Chen B, Cao J, Huang Z. TrafficGAN: Network-Scale Deep Traffic Prediction with Generative Adversarial Nets. IEEE Transactions on Intelligent Transportation Systems, 2021, 22(1): 219–230.