Intelligent Transport Systems: Methods, Challenges and Future Trends in Traffic Flow Forecasting

Ruilin Li [1,2], Yokeseng Wong [1,*], Honghong Li [2] and Hoiyeh Lee [1]

[1] Universiti Pendidikan Sultan Idris, Perak 35900, Malaysia
[2] Gongqing College of Nanchang University, Gongqingcheng 332020, Jiangxi, China
* Corresponding author: Yokeseng Wong

Abstract: This paper provides a systematic review of the field of intelligent transport and traffic flow prediction, covering the current research status, main methods, evaluation indexes and experimental results, and future research directions. By integrating advanced information technology, communication technology, and control technology, intelligent traffic systems achieve optimal regulation of traffic flow and improve traffic efficiency and safety. In the methods section, this paper reviews the applications of traditional, machine learning, and deep learning methods in traffic flow prediction. Conventional methods such as the ARIMA model and Kalman filter perform well in short-term prediction but have limited prediction accuracy in complex traffic environments. Machine learning methods such as Support Vector Machines and Random Forests perform superiorly in dealing with high dimensional data and non-linear problems. Deep learning methods such as LSTM, CNN, and GNN show significant advantages in capturing complex spatio-temporal relationships and handling large-scale data. The evaluation metrics section summarises the commonly used evaluation metrics, including mean square error (MSE), mean absolute error (MAE), and mean absolute percentage error (MAPE), and demonstrates the performance, advantages, and disadvantages of each method in different application scenarios by comparing the experimental results of different methods. Future research directions are proposed, such as multi-source data fusion, edge computing, transfer and augmented learning, privacy-preserving techniques, and smart traffic infrastructure construction.

Keywords: intelligent traffic; traffic flow prediction; machine learning; deep learning

1. Introduction

With the acceleration of global urbanization and the rapid increase in the number of motor vehicles, the problems of traffic congestion, environmental pollution, and traffic accidents are becoming increasingly serious. Intelligent Transport System (ITS), a new type of traffic management based on advanced information technology, communication technology, and control technology, aims to alleviate traffic problems and enhance traffic efficiency and safety by improving the intelligence level of the traffic system. Traffic flow prediction is one of the core technologies in ITS, which can provide an important basis for traffic management and decision-making through accurate prediction of traffic flow, to achieve optimal regulation of traffic flow.

In recent years, with the rapid development of big data, the Internet of Things (IoT) and Artificial Intelligence (AI) technology, traffic flow prediction methods have been constantly introduced, from the traditional statistical model to the modern machine learning and deep learning models, and various methods have shown their advantages in different application scenarios. However, traffic flow prediction still faces many challenges, such as the diversity and complexity of data, real-time requirements, and the improvement of prediction accuracy. In addition, the applicability and effectiveness of different methods often depend on specific traffic environments and data characteristics, and how to select appropriate methods in practical applications is also an urgent problem.

The purpose of this paper is to provide a systematic review and analysis of existing research in the field of intelligent transportation and traffic flow prediction. By reviewing and summarising the main research results in this field, it analyses the advantages and disadvantages of different methods, identifies the gaps in the current research, and puts forward suggestions for future research. Specifically, this paper will:

1. Introduce the basic concepts, development history, and application scenarios of intelligent transport systems;
2. Review and compare the main methods for traffic flow prediction, including traditional methods, machine learning methods, and deep learning methods;
3. Discuss data sources and processing techniques in intelligent transport;
4. Summarise the evaluation indexes of traffic flow prediction and the experimental results of different methods;

The structure of this paper is arranged as follows: Part II outlines the basic concepts, development history, and application scenarios of intelligent transport systems; Part III reviews the main methods of traffic flow prediction, including traditional methods, machine learning methods, and deep learning methods; Part IV discusses data sources and processing techniques in intelligent transport; Part V summarises the evaluation indexes of traffic flow prediction and the experimental results of different methods; and Part VI provides a summary and outlook of the whole paper. the whole paper to summarise and outlook.

2. Overview of Intelligent Transportation

2.1. Definition and Scope

Intelligent Transportation System (ITS) is a new traffic management model that combines modern communication technology, information technology, and control technology intending to enhance the operational efficiency, safety, and sustainability of the traffic system [1]. The core objective of ITS is to achieve optimal regulation of traffic flow through real-time data collection and analysis, thereby reducing traffic congestion, lowering the rate of traffic accidents, and reducing environmental pollution. ITS consists of several components such as traffic monitoring, traffic management, traffic information service, traffic forecasting, vehicle networking, and driverless driving. These systems perform data collection and processing through devices such as sensors, cameras, communication networks, and computing platforms, thus realizing intelligent traffic management and services.

2.2. Development History

The development of intelligent traffic systems can be traced back to the 1970s, during which time, with the rapid development of electronics, signal processing, and automatic control technology, we witnessed a shift from a traffic system that relied on human command to a more intelligent management approach.[2] With the advancement of computer and communication technologies, the earliest traffic management systems gradually began to be applied to urban traffic management. By the 1990s, with the further development of information technology, the concept of ITS gradually took shape and was promoted and applied globally.

In recent years, with the rapid development of artificial intelligence, big data, and Internet of Things (IoT) technologies, ITS has entered a new stage of development. The application of these new technologies has significantly improved the ability of ITS in data collection, analysis, and decision-making, thus achieving more refined and intelligent traffic management [3].

2.3. Application Scenarios

Intelligent transport system has a wide range of applications in real life, the following are some typical application scenarios:

1. Intelligent traffic signal control: through real-time monitoring of traffic flow data, dynamically adjust the time settings of traffic signals to reduce traffic congestion and vehicle waiting time.
2. Traffic Accident Detection and Management: Using cameras and sensors to monitor traffic conditions in real-time, quickly detect and locate traffic accidents, dispatch rescue forces in time, and carry out traffic diversion.
3. Internet of Vehicles (IoV): Through in-vehicle sensors and communication devices, realize information interconnection between vehicles and between vehicles and infrastructures, to improve traffic safety and efficiency.
4. Automatic driving: combining sensors, cameras, and artificial intelligence technology to realize the automatic driving function of vehicles and reduce the risk of traffic accidents caused by human operation.
5. Traffic information service: Through mobile applications and navigation systems, it provides drivers with real-time traffic information and path planning suggestions to help them choose the optimal path and avoid congestion.
6. Traffic prediction system: This refers to the use of various algorithms and technologies to predict future traffic conditions, such as traffic flow and traffic congestion. Such systems can make predictions based on historical data, real-time traffic information, weather conditions, and other factors.

2.4. Social Impact of Intelligent Transport

Intelligent transport systems not only play an important role in improving transport efficiency and safety but also have a profound impact on all aspects of society. For example, by reducing traffic congestion and vehicle emissions, it helps to improve the air quality of cities and the quality of life of residents; by optimizing logistics and transport routes, it improves economic efficiency and enterprise competitiveness; and by promoting the development of autonomous driving technology, it is expected to radically change people's travel patterns and the traffic structure of cities in the future.

3. Traffic Flow Prediction Method

3.1. Traditional Methods

3.1.1. Statistical model: ARIMA

ARIMA model, i.e. Autoregressive Integral Sliding Average Model, which cleverly integrates the autoregressive function (AR) and the moving average method (MA). By performing the difference operation on the data, the model can effectively remove the fluctuating part of the series, thus making the time series data smooth and easy to analyse and process. The formula is:

$$\left(1-\sum_{i=1}^{p}\phi_i L^i\right)(1-L)^d X_t=\left(1+\sum_{j=1}^{q}\theta_j L^j\right)\epsilon_t \quad (1)$$

Autoregressive (AR) part: this part of the model relies on the dependence of the prior data points, with the parameter p denoting the number of previous p historical values used to predict the current value.

Difference (I) part: the non-stationary time series is smoothed by differencing and its parameter d denotes the number of times the differencing is performed.

Moving Average (MA) component: This component involves a moving average of the model error terms, with the parameter q denoting the number of the first q prediction error values used to predict the current value.

ARIMA is applicable to time series data with linear relationships and is mainly used for economic data analysis (e.g. GDP, unemployment rate, etc.), stock price forecasting, sales volume forecasting, traffic flow and traffic condition forecasting. In traffic flow forecasting, ARIMA can be used to predict the traffic flow at a specific point in time or within a time period to help urban traffic management make adjustments and planning. However, it cannot capture nonlinear relationships in the data and has limited prediction accuracy for complex traffic environments. Liu, Wu, and Wang found that ARIMA model is better in dealing with short-term traffic flow prediction but less effective in predicting sudden traffic events when using ARIMA model for predicting traffic flow on a city road [4]. Katambire, Musabe, Uwitonze, and Mukanyiligira used ARIMA model to predict the traffic flow at Muhima junction in Kigali city, Rwanda, the results showed that ARIMA performs better in short-term prediction but has a higher error in long-term prediction [5].

3.1.2. Kalman filtering

Kalman filtering is a recursive algorithm that optimizes the state estimation of a system by constantly updating the predicted and observed values [6]. Its basic steps include prediction, updating and estimation. It is assumed that the state of the system is linear, and that the measurement process is also linear. Then, the mathematical model of Kalman filtering can be expressed as:

State equation:

$$x_k = Ax_{k-1} + Bu_k + w_k \tag{2}$$

Measurement equation:

$$z_k = Hx_k + v_k \tag{3}$$

Prediction equation:

$$x_{k|k-1=\widehat{Ax_{k-1|\widehat{k-1}}}+Bu_k} \tag{4}$$

Update state estimate:

$$\hat{x}_{k|k} = \hat{x}_{k|k-1} + K_k (z_k - H\hat{x}_{k|k-1}) \tag{5}$$

Update error covariance:

$$P_{k|k= (I-K_kH)P_{k|k-1}} \tag{6}$$

Kalman filtering has applications in many fields, including aerospace, robot navigation, and time series analysis. It is a powerful tool for processing time series data, especially for state estimation in the presence of noise. However, it requires an accurate description of the state-space model of the system and has limited ability to handle nonlinear systems. Lin and Huang used Kalman filtering for traffic flow prediction on highways in their study, and the results showed that it has high accuracy and stability in handling both real-time and noisy data [7].

3.2. Machine Learning Methods

3.2.1. Support Vector Machine (SVM)

Support Vector Machine is a supervised learning model that performs classification and regression by constructing optimal hyperplanes [8]. For traffic flow prediction, SVM regression (SVR) is commonly used.

Mathematically, SVM attempts to find an optimal hyperplane of the form:

$$w \cdot x + b \tag{7}$$

Objective function:

$$\min_{w,b} \frac{1}{2}|w|^2 \tag{8}$$

Constraints (for all training data (xi, yi)):

$$y_i(w \cdot x_i + b) \geq 1 \tag{9}$$

Researcher used the SVR model in traffic flow prediction on a city bus route, experimental results show that SVR performs superiorly in capturing nonlinear relationships and significantly improves the prediction accuracy compared to traditional statistical models. However, SVR is sensitive to parameter selection, has a long training time, and has a high computational cost for large-scale datasets.

3.2.2. Random forest

Random Forest, a cutting edge integrated learning technique, works together by carefully designing and constructing a series of decision trees. These decision tree models independently learn the relationships between features in the data and make predictions accordingly. When these predictions are aggregated, Random Forest is able to evaluate the dataset as a whole, thus combining the predictions of each tree to achieve more accurate predictions. In addition, due to its efficiency in processing large amounts of data in parallel, Random Forests also exhibit good robustness and maintain high accuracy even in the face of complex and changing data situations. This has made random forests a very applicable technique in many fields such as classification, regression, and time series analysis [9].

Mathematical Model.

Let X be the feature vector and Y be the target variable. Random forests are predicted by the following:

For classification problems:

$$\hat{Y} = \text{mode}\{Y_1(X), Y_2(X), \dots, Y_n(X)\} \tag{10}$$

For regression problem:

$$\hat{Y} = \frac{1}{n}\sum_{i=1}^{n} Y_i(X) \tag{11}$$

Tian, Wang, and Shi used the random forest model in a case study of urban road flow prediction, and the experimental results showed that the random forest has high prediction accuracy and stability in dealing with large-scale and complex datasets, but the model has high complexity, is less interpretative, and is more sensitive to noisy data [10].

3.3. Deep Learning Methods

3.3.1. Long and Short Term Memory Network (LSTM)

Long Short-Term Memory Network (LSTM) is a novel structure based on recurrent neural networks, which can effectively overcome the dependence of traditional recurrent neural networks on temporal data. Short-term memory (LSTM) regulates the input and output of information through gating so that it retains important information over longer time series and also avoids the problem of gradient loss. The main innovation of LSTM lies in its intrinsic structure, i.e., forgetting gates, input ports, and output entrances.

Key components of LSTM

1. Forget Gate:

$$f_t = \sigma\left(W_f \cdot [h_{t-1}, x_t] + b_f\right) \tag{12}$$

2. Input Gate (Input Gate):

$$i_t = \sigma(W_i \cdot [h_{t-1}, x_t] + b_i) \tag{13}$$

3. Unit state update:

$$C_t = f_t * C_{t-1} + i_t * \widetilde{C_t} \quad (14)$$

4. Output Gate (Output Gate):

$$o_t = \sigma(W_o \cdot [h_{t-1}, x_t] + b_o) \quad (15)$$

Katambire, Musabe, Uwitonze, and Mukanyiligira used an LSTM model to predict traffic flow on multiple arterial roads in their study, and the experimental results showed that the LSTM was superior in capturing temporal dependencies and handling long time-series data, with a prediction accuracy significantly higher than the traditional methods. However, the training time is long, the requirement for computing resources is high, and it is sensitive to parameter settings.

3.3.2. Convolutional Neural Networks (CNN)

Convolutional neural network (CNN), as an outstanding representative in the field of deep learning, has demonstrated unparalleled capabilities in several application scenarios such as image and video content recognition, personalized recommendation algorithms, and natural language processing tasks. This neural network architecture is known for its unique structure and working principle, which uses convolutional operations to capture key features in image or video data and utilizes this information for more accurate pattern recognition. CNN is not only capable of handling large-scale datasets but also its deep learning properties make it a powerful tool for solving complex problems. As technology continues to advance, CNNs have become an integral part of AI research and development and are used in a wide range of intelligent systems, from consumer electronics to medical diagnostics to traffic flow prediction systems. LeCun, Bottou, Bengio, and Haffner used for high-accuracy traffic flow prediction by performing convolutional operations on the data and extracting multi-level features. CNNs, which are typically used to process image data, have also been applied to time series data [11].

Key Components of CNN

1. Convolutional Layer: A convolutional kernel (or filter) is used to slide over the input data and perform elemental multiplication and accumulation operations to extract features.

The output is called a Feature Map,

$$S(i,j) = (I * K)(i,j) = \sum_m \sum_n I(m,n)K(i-m, j-n) \quad (16)$$

2. Activation Layer:

The ReLU (Rectified Linear Unit) activation function is usually used to add nonlinearity and help the network learn complex patterns.

$$f(x) = \max(0, x) \quad (17)$$

3. Pooling Layer:

Reduces the spatial size of the feature map, reducing the number of parameters and computational complexity, while making feature detection more robust.

Common pooling operations are Max Pooling and Average Pooling.

$$S(i,j) = \max_{m,n \in [M][N]} I\,(i+m, j+n) \quad (18)$$

4. Fully Connected Layer (FCL):

In the final layers of the network, the previous outputs are spread and connected to one or more Fully Connected Layers for the final classification or regression analysis.

$$y = Wx + b \quad (19)$$

5. Classification Layer:

Usually, at the very end of the CNN, a Softmax function is used to output classification probabilities.

$$\sigma(z)_i = \frac{e^{z_i}}{\sum_{k=1}^{K} e^{z_k}} \quad (20)$$

The power of CNNs lies in their ability to automatically and efficiently learn spatial-level features from images without the need for manual feature engineering. It is capable of automatically extracting features from the data, reducing the workload of feature engineering, and has a strong ability to recognize complex patterns. However, it is sensitive to the spatial structure of the data, requires a large amount of training data, has high hardware requirements, and has a long training time. Liu, Wu, Wen, Xiao, and Chen used a CNN model to predict traffic flow on multiple roads in a traffic flow prediction experiment in a large city, and the results showed that CNN has significant advantages in extracting complex data features, and the prediction accuracy is higher than that of the traditional methods and some machine learning methods [12].

3.3.3. Graph Neural Network (GNN)

Graph Neural Network (GNN) is a deep learning model specifically designed for processing graph-structured data (e.g., social networks, knowledge graphs, molecular structures, etc.). Researcher used graph-structured data for modeling and prediction, which is particularly suitable for traffic flow prediction in complex traffic systems such as urban road networks. Deep learning of graph-structured data is achieved through information transfer and updating of nodes and edges. Chen, Segovia-Dominguez, Coskunuzer, and Gel used a GNN model for traffic flow prediction of an entire city in a study of flow prediction in an urban traffic network. Experimental results show that GNN has significant advantages in dealing with complex road network structures and capturing global and local features of traffic flow, with higher prediction accuracy than other methods [13]. It is able to handle graph-structured data and is suitable for complex traffic networks. Better modeling capability for global and local features of data. However, the model complexity is high, the training time is long, and a large amount of computational resources and data support is required.

4. Data Sources and Processing in Intelligent Transportation

4.1. Data Types

There are various data sources for intelligent transport systems, mainly including the following:

1. Sensor data.

Road sensors: induction coils, radar, lidar, and infrared sensors are installed on the road. These sensors can monitor the flow, speed, and occupancy of vehicles in real-time. For example, induction coils sense the passage of vehicles, and radar and lidar measure the speed and distance of vehicles.

Cameras: These are used to monitor traffic conditions in real-time, capturing vehicle images and video data.

Through image processing techniques, traffic flow, vehicle types, and lane occupancy can be analyzed.

Intelligent Traffic Light: Equipped with sensors and communication devices, it can detect traffic flow and vehicle waiting time in real time and dynamically adjust signal timing according to real-time data to optimize traffic flow.

2. Vehicle Data.

GPS data: collects vehicle location information, traveling speed, and route trajectory through vehicle navigation equipment. These data can be used to draw vehicle traveling routes and detect traffic congestion.

Telematics data: Communication data between vehicles and between vehicles and infrastructure is collected through in-vehicle communication systems (e.g., V2V and V2I communications). These data can be used for real-time traffic information sharing and cooperative control.

3. Social media data.

Social networks: real-time traffic events and user feedback are obtained by analyzing traffic-related posts and comments on social media platforms. For example, traffic accident reports and user comments on Twitter can be used for real-time traffic event detection.

4. Mobile Application Data.

Navigation and traffic applications: such as Google Maps and Waze, collect user driving paths, traffic reports, and event information. These data can be used for real-time traffic flow prediction and route planning.

4.2. Data Preprocessing

Data preprocessing is an important step to ensure that the traffic flow prediction model can utilize the data accurately and efficiently. The following are common data preprocessing techniques:

1. Data Cleaning.

Missing value processing: for missing values in traffic data, methods of filling (e.g., mean-filling, interpolation) or deletion can be used. For example, using linear interpolation to fill in short periods of missing traffic flow data.

Outlier handling: Detecting and handling outliers in the data, such as abnormal traffic values due to sensor failures or data entry errors. Outliers can be detected using statistical methods (e.g., the 3σ principle) or machine learning methods (e.g., isolated forests).

2. Data Transformation.

Normalization and Normalization: converting data to the same scale to eliminate the effect of differences in magnitude on the model. Commonly used methods include Z-score normalization and Min-Max normalization.

Time series transformation: To perform time series analysis and forecasting, the raw data are transformed into a time series format for time series analysis and forecasting. For example, aggregating minute-by-minute traffic flow data into hourly-level time series data.

3. Feature Engineering.

Feature extraction: extract useful features from the raw data, such as time-periodic features of traffic flow (e.g., peak hours) and weather features (e.g., the effect of rain on traffic flow).

Feature Selection: Select features that are more relevant to the object to be predicted to reduce the complexity of the model. For example, important features are selected through correlation analysis or LASSO regression.

4.3. Big Data and Cloud Computing

With the continuous increase in traffic data and complexity, big data and cloud computing technologies are widely used in intelligent transport.

1. Big Data Technology.

Distributed storage and computation: big data platforms such as Hadoop and Spark are used to store and process massive traffic data. These platforms support distributed storage and parallel computing to improve data processing efficiency and reliability. For example, Hadoop HDFS is used to store large-scale traffic data, and Spark is used for distributed data processing and real-time stream processing.

Data mining and analysis: Big data analysis techniques are used to mine valuable information and patterns from massive data to support traffic flow prediction and decision-making. Commonly used methods include association rule mining, cluster analysis, and machine learning algorithms.

2. Cloud Computing Technology.

Elastic computing resources: Through the elastic computing resources provided by cloud computing platforms (e.g., Baidu Cloud, Google Cloud, Ali yun, Tencent Cloud), the computing capacity can be dynamically expanded or scaled down to cope with fluctuations in the demand for traffic data processing. For example, Tencent Lambda is used to achieve serverless computing to handle unexpected traffic data analysis tasks.

Real-time data processing: Use the real-time data processing capability of the cloud computing platform to achieve real-time analysis and prediction of traffic data, and support the real-time response and adjustment of the intelligent traffic system. For example, use Baidu Cloud Dataflow for real-time data flow processing to analyze traffic flow and events.

5. Evaluation Indicators and Experimental Results

5.1. Evaluation Indicators

In traffic flow prediction, the indicators for evaluating the model performance are very important, and the commonly used evaluation indicators include the following:

1. Mean Square Error (MSE): it is a commonly used regression analysis index to measure the difference between the predicted and actual values of the model. MSE calculates the average of the squares of the differences between the predicted and actual values and is usually used to optimize the parameters of the regression model. MSE is the average of the squares of the differences between the predicted values and the true values.

The formula is:

$$\mathrm{MSE} = \frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2 \tag{21}$$

Used MSE in study to evaluate the performance of improved spatiotemporal residual convolutional neural network (ST-ResNet) in short-term traffic flow prediction on urban road networks.

2. Mean Absolute Error (MAE) is another important metric when measuring the performance of a regression model. It arrives at a value by calculating the sum of the differences between the predicted values and the actual observed values and then averaging these absolute values. This method helps us to understand the predictive ability of the model on the input data, as a smaller MAE indicates a higher predictive accuracy of the model, and vice versa for a poorer generalization of the model. In practice, MAE is often used in conjunction with other evaluation criteria to assess the strengths and weaknesses of a model more comprehensively. For example, when a model is used for a classification task, the mean absolute error (MAE) is often calculated for each category to compare the discriminative effect between different categories. Compared to the mean square error (MSE), the MAE is less sensitive to outliers because it does not involve squaring operations, thus making the contribution of individual errors to the total error linear.

The formula is:

$$\mathrm{MAE} = \frac{1}{n}\sum_{i=1}^{n}|y_i - \widehat{y_i}| \tag{22}$$

Used MAE as an evaluation metric for traffic prediction using Diffusion Convolutional Recurrent Neural Networks (DCRNN), and the results showed that DCRNN performs well in handling complex spatiotemporal data.

3. Mean Absolute Percentage Error (MAPE): is a common metric used to measure the difference between a prediction model or estimate and the actual data, and is particularly suitable for comparing prediction results for data sets of different sizes. MAPE represents the mean absolute percentage of prediction error, and therefore it provides the relative magnitude of the error, making it well suited for comparisons of model performance when different magnitudes of data are involved.

The formula is:

$$\mathrm{MAPE} = \frac{100\backslash\%}{n}\sum_{i=1}^{n}\left|\frac{y_i - \widehat{y_i}}{y_i}\right| \tag{23}$$

Zhang et al. used MAPE in their study to evaluate the performance of multi-task deep learning models in spatio-temporal network traffic prediction [14].

5.2. Discussion of Results

The results discussion section will summarize the application of different methods in traffic flow prediction and their effectiveness. This part is based on several research literature and demonstrates the advantages and disadvantages of different methods by comparing their prediction performance.

5.2.1. Discussion of results of statistical modeling

Zhao et al. using an improved ARIMA model for ultra-short-term wind speed prediction, found that it has high accuracy in short-term prediction, although its application in traffic flow prediction faces similar challenges [15]. The experimental results show that in the short-term traffic flow prediction of a city road, the MSE of the ARIMA model is 20.45 and the MAE is 4.12, showing its high prediction accuracy in the short-term.

5.2.2. Discussion of experimental results of machine learning models

1. support vector machine (SVM)

Luo Huang C, Cao et al. proposed a short-term traffic flow prediction method based on Least Squares Support Vector Machines (LSSVM) and a hybrid optimization algorithm.[16] The method combines Particle Swarm Optimization (PSO) and Genetic Algorithm (GA) to optimize the parameter selection of LSSVM and improve the prediction accuracy and convergence speed. Experimental results show that the model has better prediction ability and relatively high computational efficiency compared with other related models.

2. Random Forest:

Nguyen et al. (2018) used the Random Forest model in the case of urban road traffic flow prediction, and the experimental results show that it has high prediction accuracy and stability, the Random Forest model has achieved an R-squared of 0.89 in the prediction of traffic flow on a major road, which is significantly better than the traditional statistical methods.

5.2.3. Discussion of experimental results of deep learning models

1. Long Short-Term Memory Network (LSTM):

Vaswani et al. used the LSTM model for spatiotemporal data prediction, and the experimental results show that the LSTM model performs superiorly in long-time dependency processing. the MAE of the LSTM model is 1.98, which is significantly lower than that of the ARIMA model which is 3.14, demonstrating its superiority in long-time series data processing [17].

2. Convolutional Neural Network (CNN):

Bai et al. used a CNN model in a traffic flow prediction experiment in a large city to predict traffic flow on multiple roads. The experimental results show that the MSE of the CNN model is 8.56 and the MAE is 2.34, showing its significant advantage in extracting complex data features [18].

3. Graph Neural Network (GNN):

The GNN model in a traffic prediction study of an urban traffic network, and the experimental results showed its significant advantages in dealing with the complex road network structure and capturing the global and local features of the traffic flow. The GNN model achieved an R-squared of 0.93 across the entire urban traffic network, which shows its strong prediction ability in complex traffic networks. Rahmani, Baghbani, Bouguila, and Patterson provide a comprehensive review of applications of graph neural networks (GNNs) in the field of intelligent transport systems (ITS). [19] Unlike previous reviews limited to traffic prediction problems, this study explores the evolution of the GNN framework in different ITS applications, including traffic forecasting, demand forecasting, self-driving vehicles, intersection management, parking management, urban planning, and traffic safety. The review points out the breadth and diversity of GNN applications in the transport domain but also points out the limitations of existing research, such as the lack of exploration of multimodal data fusion and the challenges in deploying GNN models in real transport systems.

4. Hybrid Neural Network Architecture

Ranjan et al proposed a hybrid neural network architecture based on Convolutional Neural Networks (CNNs), Long Short-Term Memory Networks (LSTMs), and Transposed Convolutional Neural Networks for city-wide traffic congestion prediction. The model effectively learned traffic congestion patterns by combining spatial and temporal information [20]. However, although the model outperforms other deep neural networks such as autoencoder and Conv LSTM in terms of prediction performance, there is still room for improvement in terms of computational efficiency and model interpretability. Future research could explore ways to reduce the number of model parameters and increase the speed of model training while enhancing the model's ability to explain different traffic situations. Dai and Ye proposed a novel hybrid time-varying graph neural network HTVGNN model for real-time accurate traffic flow prediction [21]. The model effectively captures the spatio-temporal features among different nodes in the traffic network through a time-aware multi-head attention mechanism and a graph learning strategy. Although HTVGNN outperforms the state-of-the-art spatio-temporal graph neural network models in terms of prediction accuracy, the efficiency and scalability of the model in dealing with large-scale datasets still need to be further investigated.

6. Conclusion

In this part, we will summarize the main contents of the whole paper and look forward to the future development direction of intelligent transportation and traffic flow prediction.

6.1. Summary

Intelligent transport systems are rapidly developing worldwide and have become an important means to solve traffic congestion and improve traffic efficiency and safety. This paper provides a systematic overview of the current status, methods, evaluation metrics and experimental results, current challenges and future directions of intelligent transport and traffic flow prediction.

Firstly, the definition, development history and application scenarios of intelligent transport systems are described in detail. By integrating advanced information technology, communication technology and control technology, intelligent traffic system achieves the optimal regulation of traffic flow, thus enhancing traffic efficiency and safety.

Secondly, this paper reviews the main methods for traffic flow prediction, including traditional methods, machine learning methods, and deep learning methods. Traditional methods such as the ARIMA model and Kalman filter perform better in short-term prediction but have limited prediction accuracy in complex traffic environments. Machine learning methods such as Support Vector Machines and Random Forests perform superiorly in dealing with high dimensional data and non-linear problems. Deep learning methods such as LSTM, CNN, and GNN show significant advantages in capturing complex spatio-temporal relationships and handling large-scale data.

In the last part of this paper, we delve into the importance of evaluation metrics and describe in detail the evaluation metrics used during the experiments, including mean square error (MSE), mean absolute error (MAE), and mean absolute percentage error (MAPE). By comparing the experimental results of different methods, the performance, advantages, and disadvantages of each method in different application scenarios are demonstrated.

6.2. Outlook

In the future, with the continuous progress of artificial intelligence, the Internet of Things (IoT), intelligent control, big data, cloud computing, and meta-universe technologies, the intelligent transport system will be more intelligent and efficient. The following are some key directions of development:

1. Omni-directional data fusion and analysis:

With the diversification of data sources, more efficient data fusion technologies are needed to integrate heterogeneous data from multiple sources in the future. By combining sensor data, vehicle data, social media data, and mobile application data, more comprehensive and accurate traffic flow predictions can be achieved.

2. Real-time prediction and edge computing:

To meet the demand for real-time traffic management, edge computing technology will be used more often in the future. Dispersing some of the data processing tasks to edge devices close to the data source can reduce the burden on the central server and improve the real-time response capability of the system.

3. Augmented learning and cross-region migration:

Through augmented learning technology, traffic control strategies can be continuously optimized in the simulated environment to improve the adaptive ability of the system. Meanwhile, the application of migration learning technology will help to improve the cross-regional adaptability of the model and reduce the cost of retraining the model in different regions.

4. Data privacy and security protection:

With the increasing prominence of personal privacy and data security issues, techniques such as differential privacy and homomorphic encryption need to be adopted more often in the future to ensure the protection of personal privacy in the data analysis process. At the same time, security measures during data transmission and storage should be strengthened to prevent data leakage and cyber attacks.

5. Intelligent transport infrastructure construction:

Promote the popularization of intelligent traffic devices and infrastructure optimization to improve the overall efficiency of the traffic system through a data-driven approach to traffic planning and management. The combination of intelligent traffic signals, Telematics devices, and big data analytics will significantly improve traffic management.

6. Application of emerging technologies:

With the continuous progress of technology, artificial intelligence (AI), Internet of Things (IoT), Metaverse, as well as big data, cloud computing, and other fields are rapidly developing into an indispensable core force for intelligent traffic systems. These emerging technologies not only enhance the level of intelligence and efficiency of traffic management but also provide decision-makers with accurate decision support through data analysis, to effectively respond to increasingly complex traffic conditions and demands.

In the future, we can foresee that AI technology will be widely used in real-time monitoring of traffic flow, predicting traffic accidents, and optimizing route planning; IoT technology is likely to change the way vehicles communicate with infrastructure and achieve a more efficient car network; the integration of meta-universe concepts will bring users an immersive navigation experience; and the combination of big data and cloud computing can process and store large amounts of traffic-related data, helping the traffic system to conduct in-depth analyses to make more scientific and reasonable decisions. The integration of all these innovative technologies will drive the intelligent transport system towards a more intelligent, flexible, and humane direction, providing the public with a more convenient and safer traveling environment.

In conclusion, we need to through continuous research and technological innovation, the intelligent transport system will continue to develop and improve, to provide strong support and guarantee for solving global transport problems.

References

[1] Elassy, M., Al-Hattab, M., Takruri, M., & Badawi, S. Intelligent transportation systems for sustainable smart cities. Transportation Engineering, 2024.100252.

[2] Kulkarni A R, Kumar N, Ramachandra Rao K. 100 years of the ubiquitous traffic lights: An all-round review. IETE Technical Review, 2024, 41(2): 212-225.

[3] Ahmad K, Maabreh M, Ghaly M, et al. Developing future human-centered smart cities: Critical analysis of smart city security, Data management, and Ethical challenges. Computer Science Review, 2022, 43: 100452.

[4] Liu, M. Y., Wu, J. P., & Wang, Y. B. Traffic flow prediction based on deep learning. *Journal of System Simulation, 2018.30*(11), 77–82 + 91.

[5] Katambire, V. N., Musabe, R., Uwitonze, A., & Mukanyiligira, D. Forecasting the traffic flow by using ARIMA and LSTM models: Case of Muhima Junction. 2023 *Forecasting, 5*(4), 616-628. https://doi.org/10.3390/forecast5040034

[6] Kalman, R. E. A new approach to linear filtering and prediction problems. 1960.

[7] Lin, X., & Huang, Y. Short - term high-speed traffic flow prediction based on ARIMA-GARCH-M model. Wireless Personal Communications, 2021.117(4), 3421-3430.

[8] Cortes, C., & Vapnik, V. Support-vector networks. Machine learning, 1995. 20, 273-297.

[9] Breiman, L. Random Forests. Machine Learning, 2001. 45(1), 5-32.

[10] Tian Jia, Wang Deyong, & Shi Wenxi. Short-term traffic flow prediction based on ensemble empirical mode decomposition and random forest. Science, Technology and Engineering. 2023. 23(29), 12612-12619.

[11] LeCun, Y., Bottou, L., Bengio, Y., & Haffner, P. Gradient-based learning applied to document recognition. Proceedings of the IEEE, 1998. 86(11), 2278-2324.

[12] Liu, Y., Wu, C., Wen, J., Xiao, X., & Chen, Z. A grey convolutional neural network model for traffic flow prediction under traffic accidents. Neurocomputing, 2022. 500, 761-775.

[13] Chen, Y., Segovia-Dominguez, I., Coskunuzer, B., & Gel, Y. R. TAMP-S2GCNETS: Coupling Time-Aware Multipersistence Knowledge Representation with Spatio-Supra Graph Convolutional Networks for Time-Series Forecasting. 2022.

[14] Zhang, J., Zheng, Y., Sun, J., & Qi, D. Flow prediction in spatio-temporal networks based on multitask deep learning. *IEEE Transactions on Knowledge and Data Engineering, 2019. 32*(3), 468–478.

[15] Zhao, L., Song, Y., Zhang, C., Liu, Y., Wang, P., Lin, T., Deng, M., & Li, H. T-GCN: A temporal graph convolutional network for traffic prediction. *IEEE Transactions on Intelligent Transportation Systems, 2020. 21*(9), 3848–3858.

[16] Luo C, Huang C, Cao J, et al. Short-term traffic flow prediction based on least square support vector machine with hybrid optimization algorithm. Neural processing letters, 2019, 50: 2305-2322.

[17] Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., Kaiser, L., & Polosukhin, I. Attention is all you need. *Advances in Neural Information Processing Systems, 2017. 30*(1), 1–15.

[18] Bai, S., Kolter, J. Z., & Koltun, V. An empirical evaluation of generic convolutional and recurrent networks for sequence modeling. 2018. *arXiv preprint*. https://arxiv.org/abs/1803.01271

[19] Rahmani, S., Baghbani, A., Bouguila, N., & Patterson, Z. Graph Neural Networks for Intelligent Transportation Systems: A Survey. *IEEE Transactions on Intelligent Transportation Systems*, 2023. 24(8), 1–31.

[20] Ranjan, N., Bhandari, S., Zhao, H. P., Kim, H., & Khan, P. City-Wide Traffic Congestion Prediction based on CNN, LSTM and Transpose CNN. *IEEE Access*.2020. 8: 81606-81620. https://doi.org/10.1109/ACCESS.2020.2991462

[21] Dai, B.-A., & Ye, B.-L. A novel hybrid time-varying graph neural network for traffic flow forecasting. 2024. *arXiv preprint arXiv:2401.10155*.