The Intelligent Traffic Safety System Based on 6G Technology and Random Forest Algorithm

Jianhua Dai, Xiaofan Tian, Lei Liu, Hao Zhang, Jian Fu, and Miao Yu

***Abstract*— The main objective of this study is to explore and evaluate the application potential of combining Sixth Generation (6G) technology and random forest (RF) algorithm in the intelligent traffic safety system. This study designs and implements an intelligent traffic safety system using the RF algorithm and 6G technology, to improve the traffic conditions' real-time monitoring and prediction ability. The advantages of 6G technology in real-time data transmission and efficient data processing. Moreover, the application of the RF algorithm in traffic congestion and accident prediction is discussed. The value of these techniques in improving prediction accuracy, system stability, and safety performance is analyzed. Extensive experimental tests are carried out in multiple traffic scenarios by constructing modules such as data collection and preprocessing, model training and optimization, real-time data processing, and system integration and display. In the experimental test, two main scenarios of traffic congestion warning and accident prediction are designed. The results reveal that in the traffic congestion warning scenario, under the condition that the traffic flow is 1800-2000 vehicles/hour and the average speed is 45-55 km/h, the prediction accuracy reaches 96%, the recall is 99%, and the F1 score is 97%. In the traffic accident prediction scenario, the system's prediction accuracy, recall, and F1 score are 92%, 95%, and 93% when the traffic flow is 1200-1400 vehicles/hour on rainy days. The results of this study provide practical technical solutions for smart city traffic management and explore the prospect of future intelligent transportation system development, thus offering a theoretical and empirical basis for research and practice in related fields.**

***Index Terms*— 6G technology, random forest algorithm, intelligent transportation, traffic safety system, real-time monitoring.**

I. Introduction

TRAFFIC safety is one of the major global concerns [1]. With the acceleration of urbanization and the rapid increase in the number of motor vehicles, the frequency and severity of traffic accidents are also constantly increasing [2], [3]. According to the World Health Organization (WHO), over 1.3 million people worldwide die from traffic accidents and millions are injured each year, causing enormous social and economic losses. Therefore, improving traffic safety has become a vital goal of governments and scientific research institutions [4].

With the process of urbanization and the increase of traffic flow, the importance of the intelligent traffic safety system is becoming increasingly prominent. These systems use advanced technology and data analytics to improve traffic safety, efficiency, and environmental sustainability. However, the current system still faces challenges in real-time data processing, prediction accuracy, and system stability. The research of the intelligent traffic safety system is not only a response to modern urban traffic management but also a vital support for the future development of smart cities. Through real-time data transmission and efficient data processing, traffic congestion and accidents can be predicted more quickly and accurately, thus offering critical decision support for urban traffic management departments. This helps reduce the impact of traffic accidents and traffic jams on urban life and the environment. In addition, it promotes the integration of technology and society and advances the development of smart cities.

In this context, the Intelligent Transportation System (ITS) has emerged [5]. ITS is a system that utilizes advanced technology and communication facilities to improve traffic safety, efficiency, and sustainability. ITS utilizes advanced Information and Communication Technology (ICT) to monitor and manage traffic flow, thereby improving traffic efficiency and safety [6]. The current ITS mainly relies on fourth-generation (4G) and fifth-generation (5G) mobile communication technologies and traditional data analysis methods. However, with the explosive growth of data volume and the increasing complexity of application scenarios, the existing technologies face great challenges regarding real-time performance and accuracy [7], [8].

The existing traffic system has significant deficiencies in processing real-time data, predicting traffic conditions, and preventing traffic accidents [9]. It includes, but is not limited to, bottlenecks in real-time data transmission, limitations in data processing efficiency, and insufficient prediction accuracy. Traditional communication technologies and prediction models are not good at handling large data volumes and complex traffic patterns, which are difficult to meet the growing needs of urban traffic management. In the face of huge datasets and complex traffic environments, traditional data analysis methods often show low efficiency and insufficient prediction accuracy [10]. In addition, the limitations of 4G and 5G technologies in terms of data transmission speed and network coverage also limit the further development of ITS [11], [12].

Received 30 May 2024; revised 18 July 2024; accepted 30 July 2024. This work was supported in part by the National Natural Science Foundation of China under Grant 72273152, in part by the Qian Duansheng Distinguished Scholars Program of China University of the Political Science and Law, and in part by the Social Science Project from China University of the Political Science and Law under Grant 24KYGH002. The Associate Editor for this article was I. Ashraf. *(Corresponding author: Miao Yu.)*

The authors are with the Business School, China University of Political Science and Law, Beijing 100000, China (e-mail: miaoy@cupl.edu.cn).

Digital Object Identifier 10.1109/TITS.2024.3451184

Hence, finding new technical means to improve the performance of the traffic safety system has become crucial. Sixth Generation (6G) technology, with its ultra-high speed and ultra-low latency characteristics, provides technical support for real-time data transmission, greatly enhancing system response speed and data processing efficiency. Meanwhile, the Random Forest (RF) algorithm constructs multiple decision trees within large datasets, employing feature selection and node splitting techniques, thereby significantly improving the accuracy and reliability of traffic congestion and accident prediction.

Therefore, this study aims to explore how to utilize 6G technology and RF algorithm to construct a more efficient intelligent traffic safety system. Features of 6G technology include ultra-low latency, ultra-high speed, and large-scale connectivity. This technology is expected to provide higher bandwidth, lower latency, and wider network coverage, thus significantly enhancing data transmission and processing capabilities [13]. As an ensemble learning method, the RF algorithm has powerful data analysis and prediction capabilities, especially suitable for dealing with complex traffic data [14]. It is hoped to achieve prominent improvements in traffic situation prediction, real-time data processing, and accident prevention by combining 6G technology and RF algorithm, thereby facilitating the overall level of traffic safety.

Specifically, the objectives of this study include the following. Firstly, an intelligent traffic safety system architecture based on 6G technology and RF algorithm is designed. Secondly, the data collection, pre-processing, and analysis methods are developed to realize efficient traffic data processing. Additionally, the system's performance in the actual traffic scenario is verified by experiments, and the traditional method is compared and analyzed. Finally, the research results are summarized, and the future improvement direction and application prospects are put forward.

II. Literature Review

ITS uses ICT to improve traffic management and operational efficiency [15]. ITS covers many aspects such as traffic monitoring, information dissemination, intelligent signal control, and accident management. For example, Reza et al. [16] studied ITS applications based on big data and artificial intelligence, showing that these technologies could markedly improve the efficiency and safety of traffic management. However, the existing ITS system mainly relies on 4G/5G communication technology, and its data transmission rate and network coverage are limited to cope with large-scale and real-time data processing [17].

6G is considered to be one of the key technologies for future communication networks. Mahmoud et al. [18] argued that 6G could provide peak data rates of up to 1 Terabit per second (Tbps), sub-millisecond latency, and nearly full network coverage. These characteristics make 6G suitable for real-time, large-scale data transmission and processing to meet ITS demand for high bandwidth and low latency [19].

With its ultra-high speed, low delay, and large capacity, 6G technology provides the ability of fast real-time data transmission, which can effectively support the real-time monitoring and response of intelligent traffic management systems to complex urban traffic environments. The RF algorithm is known for its efficiency and accuracy in processing large data sets and variable factors. Moreover, it can effectively predict key indicators such as traffic flow and accident risk, providing a scientific basis for traffic management decisions.

The RF algorithm is a powerful ensemble learning method, which improves the accuracy and stability of prediction by constructing multiple decision trees [20]. Breiman [21] proposed the RF algorithm for the first time and proved that it had significant advantages in processing high-dimensional data and nonlinear relations. In the traffic data analysis field, Capitaine et al. [22] adopted the RF algorithm for traffic accident prediction, and the results showed that its prediction performance was superior to traditional linear regression and decision tree methods. However, applying the RF algorithm in real-time data processing and large-scale data sets still had some challenges.

There have been some attempts in existing research to combine advanced communication techniques with machine learning (ML) algorithms to apply to ITS. Zhou et al. [23] explored the application of 5G and deep learning in traffic flow prediction and achieved remarkable results. However, due to the limitations of 5G technology, the system still had bottlenecks when processing high-density, real-time data. In addition, Li et al. [24] discussed the potential of applying 6G technology to smart cities, but the specific research on traffic safety systems was not delved into.

For example, Khan and Das [25] analyzed relevant studies classified according to the pillars of the safety systems approach: safe road users, safe vehicles, safe speeds, safe roads, and post-crash care. Ma et al. [26] utilized encryption systems such as N-th degree Truncated Polynomial Ring Unit (NTRU) and McEliece cipher system to explore the potential of blockchain-based human intelligence systems for urban security.

It can be found that the current ITS system mainly depends on 4G/5G technology, which cannot fully meet the requirements of large-scale and real-time data transmission. Therefore, the wide coverage, high bandwidth, and low latency characteristics of 6G technology are employed to improve ITS's data transmission and processing capabilities. For real-time, large-scale traffic data, the optimization of the RF algorithm can enhance its prediction accuracy and processing efficiency, and it is expected to make significant progress in improving traffic safety and efficiency.

By thoroughly reviewing and analyzing the latest research findings, the limitations of existing studies in real-time data processing efficiency and prediction accuracy have been highlighted. Specifically, current research shows limited capability in handling large volumes of data, complex traffic patterns, and variable weather conditions, lacking effective technical means to address real-time challenges in traffic management. Combining 6G technology with RF algorithms not only enhances the data processing speed and accuracy of the intelligent traffic safety system but also optimizes its adaptability and predictive capabilities in complex traffic scenarios. This integration leverages the advantage of rapid data transmission with 6G technology to provide real-time data support for

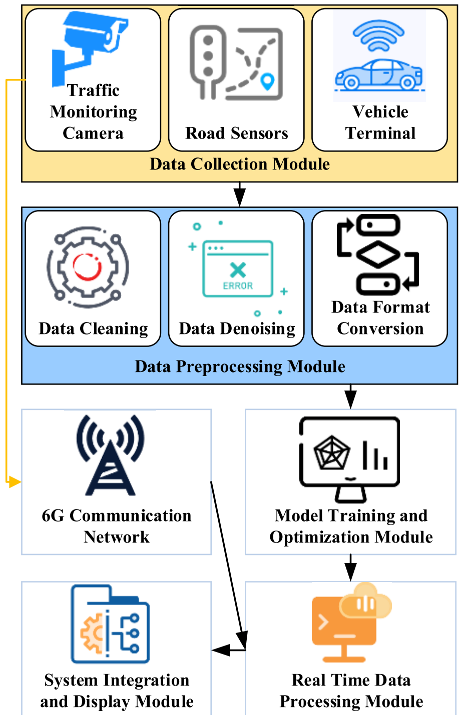


Fig. 1. The intelligent traffic safety system architecture.

RF algorithms, thereby synergistically improving intelligent traffic safety. It offers new technological means and solutions for urban traffic management and safety. Thus, this study achieves efficiency in real-time data transmission through the introduction of 6G technology and enhances the accuracy and reliability of traffic congestion and accident prediction by integrating RF algorithms, thereby filling research gaps. In particular, this study proposes a new method for data cleaning and feature extraction that effectively overcomes the shortcomings of traditional approaches in data processing efficiency and prediction accuracy. These innovations not only address current research gaps but also pave the way for further technological advancements and methods in the intelligent traffic safety system.

III. Design of an Intelligent Traffic System Combining 6G With RF

A. Overall Design of the System Architecture

The intelligent traffic safety system designed in this study is mainly composed of the following modules: data collection, data preprocessing, model training and optimization, real-time data processing, and system integration and display modules. The core of the system architecture is to integrate 6G technology and RF algorithm to realize real-time monitoring and prediction of traffic conditions through efficient data transmission and processing. The specific structure is presented in Fig. 1:

In Fig. 1, the data collection module is the system's front end and is responsible for obtaining traffic data from various data sources. These data sources include: 1) Traffic surveillance cameras: real-time video streams for monitoring road conditions and traffic flow; 2) Road sensors: For instance, pressure and infrared sensors can provide vehicle passing information, covering speed and distance. 3) On-board terminal: Through Vehicle-to-Everything technology, information such as vehicle location, speed, and direction are collected. Data is transmitted via a 6G network to ensure its high bandwidth and low latency to meet real-time requirements.

The main task of the data preprocessing module is to clean and format the original data, including the following steps: 1) Data cleaning is a key step in data preprocessing, where the quality of data is concerned with detecting and processing missing and outlier values. This process ensures the accuracy, consistency, and reliability of the data in preparation for subsequent analysis or modeling. When there are missing values in the dataset, a variety of methods can be used to fill in, one of the commonly used methods is mean fill. It can be assumed that there is a dataset X with missing values, whose mean is $\bar{X}$. Then, the missing value X_{missing} can be filled in with the mean $\bar{X}$:

$$X_{\text{missing}} = \bar{X} = \frac{1}{n}\sum\nolimits_{i=1}^{n} X_i \tag{1}$$

Linear interpolation can be used for time series data, assuming that a value X_t is missing between time point t and $t+1$, which can be estimated using Eq. (2):

$$X_t = X_{t-1} + \frac{X_{t+1} - X_{t-1}}{2} \tag{2}$$

Outlier detection can be achieved through statistical methods, such as using the z-score method, to determine whether a data point is an outlier:

$$z = \frac{X - \mu}{\sigma} \tag{3}$$

μ represents the mean of the data, and σ refers to the standard deviation. When $|z| > 3$, data point X is considered to be an outlier.

2) Data denoising is another vital step in data preprocessing aimed at removing noise components from data to enhance data quality, ensuring more accurate and effective subsequent data analysis and modeling. Noise in data can stem from various factors such as measurement errors, equipment malfunctions, transmission distortions, or external interference, all of which degrade the signal-to-noise ratio, affecting the authenticity and reliability of the data. Among numerous data denoising techniques, wavelet denoising stands out as a highly effective tool, particularly suitable for signal and image processing domains. The fundamental principle of wavelet denoising involves decomposing the original signal into different scales using wavelet transform, and analyzing and processing the signal at different scales. Then, the high-frequency noise is removed by threshold processing, and the signal is restored by wavelet inversion. Supposing the signal $S(t)$ contains noise $N(t)$, which can be expressed as:

$$S(t) = X(t) + N(t) \tag{4}$$

With the wavelet transform, the signal can be decomposed into components of different frequencies, and the high-frequency noise components can be removed by threshold

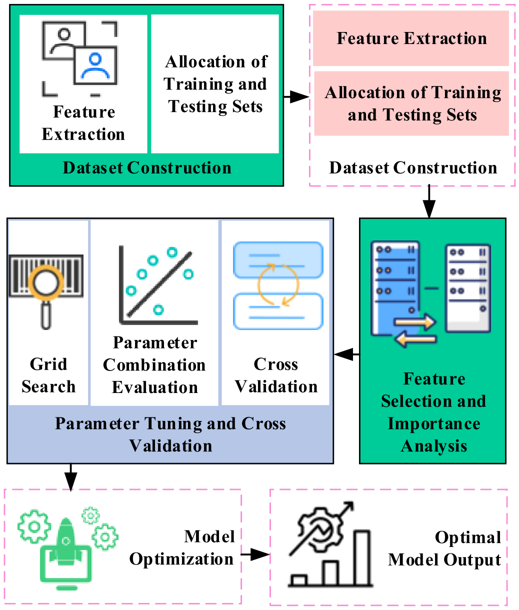


Fig. 2. The raining and optimization process of the RF model.

processing. Assuming a wavelet coefficient of W, the denoised signal $\hat{X}(t)$ can be written as:

$$\hat{X}(t) = \sum_i W_i \cdot \psi_i(t) \tag{5}$$

$\psi_i(t)$ denotes the wavelet basis function, and W_i means the wavelet coefficient after threshold processing.

3) Data format conversion is another critical step in data preprocessing, aimed at transforming data from diverse sources with varied formats into standardized formats, facilitating subsequent data integration, analysis, and processing. This process typically involves adjusting and optimizing the structure of raw data to ensure that all data can be understood and operated within the same framework, thereby enhancing the efficiency and effectiveness of data processing.

Traffic video data often exists in the form of a continuous sequence of frames, where each frame represents a static image. When analyzing video data, it is necessary to convert the video data into a frame sequence format to facilitate independent analysis of each frame image. This conversion process involves video decoding, which decodes compressed video files into a series of image frames. For example, sensor data requires extracting specific sensor readings from raw binary or text log formats and sorting them by timestamp. Similarly, Global Position System data, which includes latitude-longitude coordinates and timestamps, needs to be converted into a time series format for analyzing movement trajectories, speed, direction, and other related information.

Data preprocessing is followed by the model training and optimization module, where the RF model is trained and optimized using historical traffic data. The specific process is displayed in Fig. 2:

Fig. 2 demonstrates that features are extracted from historical traffic data first, and training sets and test sets are constructed, including feature extraction and dataset segmentation. The RF model is then trained using the training set,

TABLE I
PSEUDO-CODE OF THE RF ALGORITHM

Serial number	Step
1	**Input:** Training dataset D = {(x1, y1), (x2, y2), ..., (xn, yn)}, Number of features m, Number of trees T // The training dataset D consists of n samples, each with feature vector xi and corresponding label yi. m is the number of features, and T is the number of decision trees.
2	**Output:** The Random Forest (RF) model // The output is a Random Forest model composed of T decision trees.
3	**For t = 1 to T:** // Loop to create T decision trees.

4	1. Samples from training set D are randomly selected to create a new training set D', with the same number of samples as D, allowing duplicates.
	// A new training set D' is created by randomly selecting samples from D with replacement, so D' has the same size as D but may contain duplicates.
	2. A subset of feature F is randomly selected from m features.
	// A subset of features F is selected from the total m features.
	
	3. A decision tree is constructed using D' and F. For each node:
	- A feature is randomly selected for splitting to minimize node impurity.
	// One feature from F is chosen randomly to split the node, aiming to reduce impurity (e.g., Gini impurity or entropy).
	- The above steps are repeated until reaching maximum tree depth or the number of samples in the node is below a threshold.
	// This process is repeated until the tree reaches a maximum depth or the node has fewer samples than a defined threshold.
	4. The constructed decision tree is added to the RF.
	// The constructed tree is then added to the Random Forest.

5	**Return** the RF model containing T decision trees. // The final Random Forest model, containing all T decision trees, is returned.

encompassing initialization of model parameters and parallel computation acceleration. Next, the recursive feature elimination method is utilized to select the features that have the greatest influence on traffic condition prediction, and the importance analysis is carried out. Moreover, the mesh search method tunes the model parameters, involving parameter combination evaluation and cross-validation, to select the best parameter combination. Furthermore, based on model training, model optimization is performed, comprising model evaluation and improvement, overfitting detection, and processing. Lastly, the optimized model is saved and output by the real-time data processing module. The RF algorithm involved in this process is represented by pseudo-code, as exhibited in Table I:

Subsequently, the real-time data processing module is responsible for processing and forecasting real-time data in practical applications. First, real-time traffic data from the data collection module is received via the 6G network. Then, the received real-time data is quickly cleaned and formatted to ensure data quality. Next, the trained RF model is used to analyze and predict the real-time data and generate real-time feedback and early warning information about traffic conditions. Ultimately, prediction results and warning information are sent to the system integration and display module, to provide decision support.

Lastly, the system integration and display module integrates all functional modules into a unified platform to display traffic monitoring and prediction results through a visual interface. It includes the following functions: 1) Data visualization: Traffic data and prediction results are visually displayed through charts, maps, and dashboards. 2) Early warning notification: When the system detects a potential traffic accident or congestion, it sends an early warning notification to the relevant departments and users. 3) Decision support: Data analysis and prediction results are offered, thereby providing decision support for traffic management departments, and optimizing traffic management strategies. 4) System maintenance and upgrade: System performance is regularly checked, and necessary maintenance and upgrades are carried out to ensure stability and efficiency.

Through the design and function description of the above modules, this study builds an efficient intelligent traffic safety system, which combines 6G technology and RF algorithm to realize real-time monitoring and prediction of traffic conditions, thus furnishing technical support for improving traffic safety.

B. System Implementation Process

After determining the design and development scheme of the system, the actual implementation process of the intelligent traffic safety system is introduced in detail, including hardware deployment, software development, and real-time data processing processes.

The hardware deployment of the system includes the following key components. 1) The 6G communication module is applied for real-time data transmission, sending traffic data from sensors and other devices to backend (BE) servers. 2) High-performance computing server is used for data processing and model operation, with powerful computing and storage capabilities, to ensure that the system can efficiently process large-scale data and run complex algorithms. 3) A data storage device is employed to store historical and real-time data, furnish data storage and access functions, and support the system's rapid data retrieval and analysis.

The software part of the system encompasses the following key modules: 1) Data collection software is responsible for collecting real-time traffic data from traffic sensors, surveillance cameras, and other devices, and transmitting it to the BE server. 2) The data preprocessing software performs preprocessing operations such as cleaning and format conversion of the collected original data to ensure data quality and consistency. 3) Model training and prediction software uses ML models such as the RF algorithm to train and predict traffic data and provide real-time prediction traffic condition functions. 4) Visual display software presents the forecast results and traffic conditions to users in an intuitive way, providing an interactive interface and decision support functions. The software modules interact with each other through the API interface, which constitutes a complete workflow of the system. The processing flow of real-time data is indicated in Fig. 3:

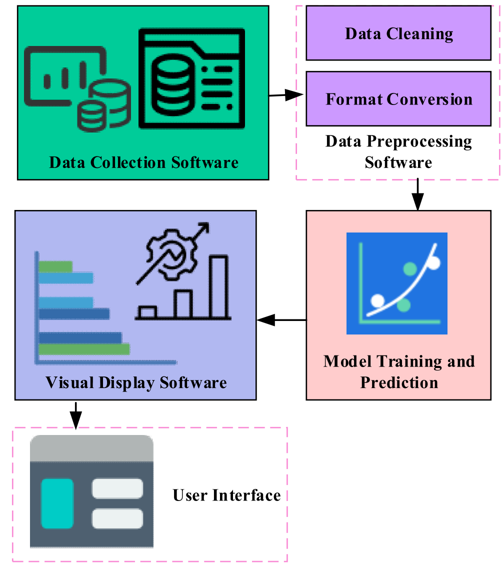


Fig. 3. Real-time data processing flow.

Fig. 3 underscores the system's real-time data processing flow in detail. First, the data collection software acquires traffic data from devices such as traffic sensors in real-time through a 6G communication module and transmits it to a BE server. The data preprocessing software then cleans and converts the raw data received to prepare the data for model training and prediction. Next, the prediction software and model training utilize ML models to predict real-time traffic data and generate real-time predictions of traffic conditions. Finally, the forecast result is displayed to the user intuitively through the result output software, which helps them to understand the traffic situation in real-time and make corresponding decisions. Eventually, users can view and interact with traffic conditions in real-time through the user interface.

The training process of the RF model is as follows:

1) Feature extraction: Firstly, features are extracted from real-time traffic data. These features include but are not limited to traffic flow, vehicle speed, weather conditions, etc. The process of feature extraction adopts time series analysis methods. 2) Construction of training and testing sets: The features extracted from the raw data are divided into training and testing sets. The training set is used to train the RF model, while the testing set is utilized to evaluate the model's performance. 3) Model parameter optimization: Before constructing the RF model, model parameters need to be optimized. Grid search cross-validation is employed to adjust the model parameters, aiming to improve the model's prediction accuracy and generalization ability. 4) Construction of the RF model: The RF model is constructed using the optimized parameters. During the model construction process, considerations include random sampling of samples and random selection of features to ensure the model operates robustly in complex traffic environments.

C. Experiment Settings

Real-time traffic data from an urban traffic monitoring system is selected as the experimental dataset, including traffic flow, speed, weather conditions, and other information, to simulate the real urban traffic scene. A city traffic monitoring system utilizing sensors and surveillance cameras is employed. These devices are strategically placed at critical locations on both major and minor roads throughout the city to ensure comprehensive coverage of the traffic situation. The data used in this study are obtained from the city's transportation department's open data portal, covering traffic sensor data collected over the past two years. Data is collected at a frequency of one-minute intervals to ensure real-time traffic information is adequately captured. This dataset provides a representative view of urban traffic conditions, offering a comprehensive perspective on city traffic patterns. However, several limitations need to be considered: the data span covers two years, potentially limiting the ability to explain long-term trends or seasonal variations beyond this period. Moreover, the data is specific to an urban area and may not fully reflect traffic conditions in rural or other urban environments. Due to sensor malfunctions, some data points may be missing or contain errors, potentially impacting model performance. Data-cleaning techniques are applied to mitigate these issues.

Data preprocessing involves outlier detection and filtering methods to remove noise and anomalies from the collected data. For missing data, methods such as mean imputation or interpolation are used to fill in the gaps, ensuring data integrity. To eliminate differences in scale among different datasets, Z-score standardization or Min-Max normalization is performed on the data, facilitating more effective data analysis and model training in subsequent stages.

The experiment is conducted on a high-performance server equipped with a powerful processor (Intel Xeon) featuring multiple cores and high clock speed to handle the demands of processing large-scale data and complex computations. Additionally, the server is equipped with 64GB of high-capacity memory to support data processing and storage requirements during the experiment. The experimental environment runs on the Ubuntu 20.04 operating system. Ubuntu 20.04 is chosen for its stability, open-source nature, and extensive support from the research community, ensuring the stability and consistency of the experimental environment. To ensure the comparability and accuracy of experimental results, all experiments are conducted under the same hardware environment.

The following experimental scenarios are mainly concerned. 1) Traffic congestion warning: In the experiment, different congestion scenarios are simulated on urban main roads and highways, including traffic flow variations during peak and off-peak hours. Different weather conditions (such as sunny, rainy, and overcast days) are analyzed to assess the response speed and effectiveness of the congestion alert system, investigating the impact of various alert strategies on traffic fluidity. 2) Traffic accident prediction: Potential accident-prone road sections and times are accurately predicted based on historical traffic data and detailed weather information (such as temperature, humidity, visibility, etc.). Considering the different characteristics of urban internal roads and highways, the accuracy and reliability of prediction models are evaluated in different environments. 3) Traffic route optimization: Real-time traffic data changes are analyzed to dynamically adjust driving routes, considering factors such as road construction, accident congestion, and sudden weather changes. Under different traffic modes (such as urban commuting and long-distance travel) and weather conditions, the performance and effectiveness of optimal route recommendation algorithms are explored. 4) Analysis of weather impacts on traffic: Detailed analysis of the effects of rain and snow weather on different types of roads (urban road networks and highways), including the impact of rainfall intensity on visibility and vehicle travel speed. Specific data support is offered, such as the influence of precipitation and snowfall on traffic flow and road friction, to provide a scientific basis for traffic management decisions.

TABLE II
THE PARAMETER SETTINGS OF THE RF ALGORITHM

Parameter	Setting
The number of trees	100
The maximum depth of the tree	10
The partition number for the minimum sample	2
The sample size of the minimum leaf node	1
Maximum number of features	sqrt(n_features)

During the experiment, real-time traffic data, including weather conditions, traffic flow, and speed are obtained from the urban traffic monitoring system and stored in the system database. The original data is converted and cleaned by the data preprocessing module, covering the processing of missing values and outliers, to ensure quality and consistency. Then, the RF algorithm is utilized to train the pre-processed data, and model optimization is conducted to adjust the model parameters to enhance the prediction accuracy. Finally, the trained model can forecast real-time traffic data and generate prediction results of traffic conditions.

To comprehensively evaluate the system's performance and effect, the following evaluation indicators are primarily concerned: 1) Model prediction accuracy: Compared with the actual traffic situation, the accuracy of model prediction is evaluated. 2) Processing time: The system's time required is recorded to process data and evaluate its processing efficiency and performance. 3) Real-time: The system's processing and response speed to real-time traffic data are assessed to ensure that the system can make timely predictions and feedback on traffic conditions. the RF algorithm's parameter settings are outlined in Table II:

IV. Results and Discussion

A. Overall System Performance

The prediction of different urban roads under the experimental scenario of traffic congestion warning is denoted in Fig. 4:

Fig. 4 indicates that in the urban main road scenario, the RF algorithm integrated with the 6G system's real-time data transmission and efficient data processing functions presents high prediction accuracy and stability. Whether in sunny, cloudy weather, or different traffic flow ranges, the system can maintain a high accuracy and recall level, with a relatively high ratio of correct congestion predictions. However, the proportion of wrong congestion prediction and false alarm rate can be controlled at a low level. This shows that the system can

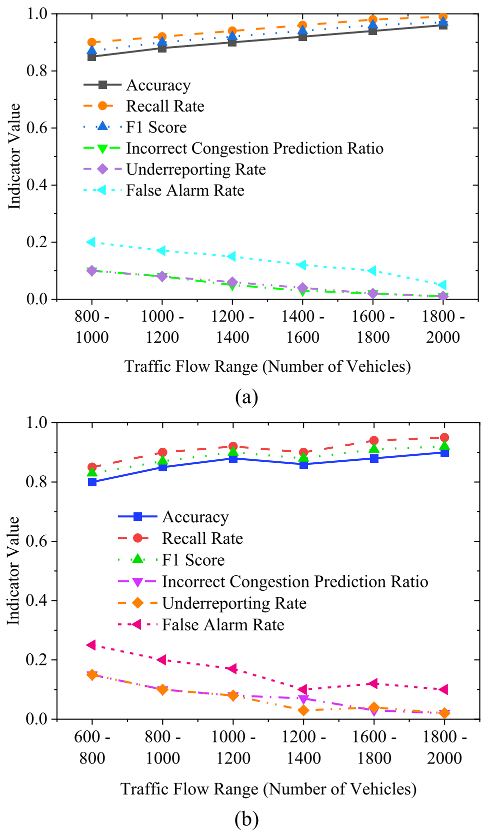


Fig. 4. Predictions under traffic congestion warning scenarios (a. urban main roads; b. Urban secondary roads).

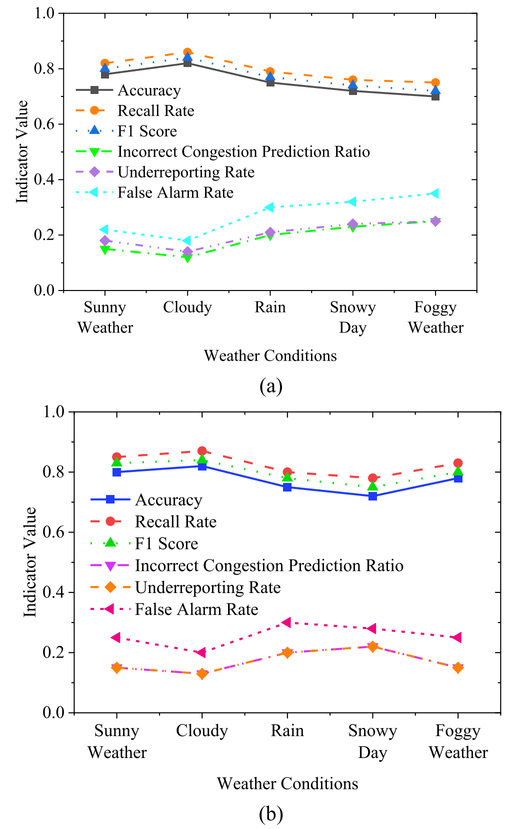


Fig. 5. Prediction of traffic accidents under diverse weather conditions (a. urban main roads; b. Urban secondary roads).

effectively predict traffic congestion under diverse conditions, and offers critical support for urban traffic management.

In the scenario of an urban secondary road, the system also shows high stability and prediction accuracy. Despite a slight decrease in accuracy and recall under rainy and cloudy weather conditions, as well as with recorded accidents, the overall level remains relatively high. The system's missing and false alarm rates with different weather conditions and traffic flow ranges are also kept at a relatively low level, illustrating that the system has a robust ability to recognize real congestion and non-congestion conditions.

These results highlight the advantages of an intelligent traffic safety system combining 6G systems with RF algorithms, especially regarding real-time data transmission and efficient data processing. This system can effectively predict traffic congestion in various scenarios, provide vital decision support for urban traffic management departments, help alleviate traffic congestion, and improve the efficiency of urban traffic operations. The prediction of traffic accidents is plotted in Fig. 5:

Fig. 5 denotes that different weather conditions impact the prediction performance of the system in both urban main roads and secondary roads scenarios. In clear and cloudy weather, the system exhibits high accuracy, recall, and F1 scores, while its prediction performance slightly decreases in adverse weather conditions like fog. For example, in the urban main road scenario, the average accuracy in clear weather is 0.78, whereas in rainy and snowy conditions, the average accuracy drops to 0.75 and 0.72, respectively. Across all weather conditions, the system demonstrates a high proportion of correct accident predictions, a relatively low ratio of congestion predictions, and low rates of missed and false alarms. For instance, in the urban secondary road scenario, under cloudy weather conditions, the system's correct accident prediction proportion is 0.70, with a relatively low false alarm rate of 0.20, indicating its high accuracy in identifying traffic accidents.

Hence, the system shows the advantages of real-time data transmission and efficient data processing in all weather conditions. This enables the system to obtain the latest traffic data in time, and quickly process and analyze the data, to achieve timely prediction and early warning of traffic accidents. For example, on sunny days, the average missing and false alarm rates are 0.18 and 0.22, thus highlighting the system's pivotal role in the traffic safety field.

B. Test of Real-Time Data Processing Capability

Figs. 6 and 7 reveal the system's processing performance on real-time data in different experimental scenarios.

In Fig. 6, the data collection speed test results indicate that the system's data collection times vary across different test scenarios. For example, during peak hours on Urban Main Road 1, the data collection time is approximately 45 seconds, while on Urban Main Road 2 during peak hours, its time is slightly longer, at about 50 seconds. This difference could be due to increased traffic flow or reduced visibility. In contrast, during off-peak hours, early morning, and evening, the data

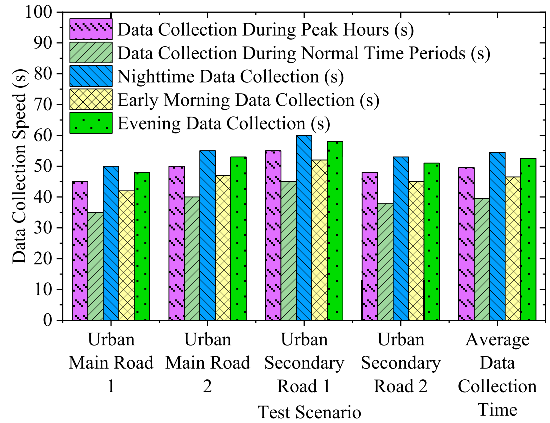


Fig. 6. Test results of data collection speed.

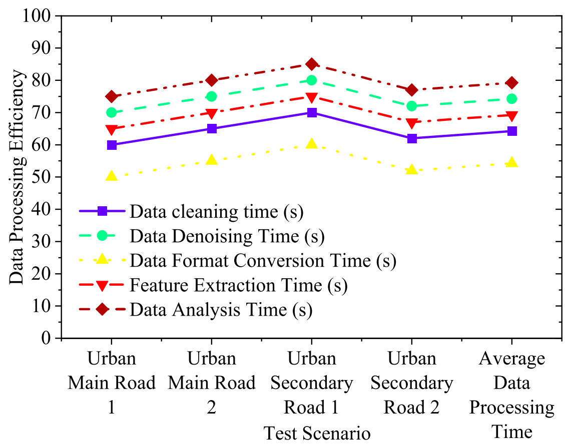


Fig. 7. Test results of data processing efficiency.

collection times are relatively shorter, being approximately 35, 42, and 48 seconds, respectively. Regarding the average data collection time, the system's performance varies across diverse scenarios, with an average time of about 49.5 seconds. This indicates that the system can complete data acquisition tasks relatively consistently across diverse scenarios.

Fig. 7 illustrates that the time required by the system for different data processing stages can be observed for test results of the data processing efficiency. For instance, the data cleaning and denoising stages take relatively longer, approximately 60 and 70 seconds, respectively, due to the need to clean the raw data and remove noise to improve data quality. The data format conversion stage is comparatively faster, around 50 seconds, likely because the data format conversion process is simpler. The feature extraction and data analysis stages require more time, around 65 and 75 seconds, respectively, possibly due to the demand for more complex data processing and analysis. In summary, although the system spends a certain amount of time on different processing stages, the overall data processing efficiency remains high, meeting the requirements for real-time data processing. The proposed system can maintain high processing efficiency under different data processing tasks, thanks to the optimized algorithm design and hardware support of this system. These advantages not only ensure the reliable performance of the system under high

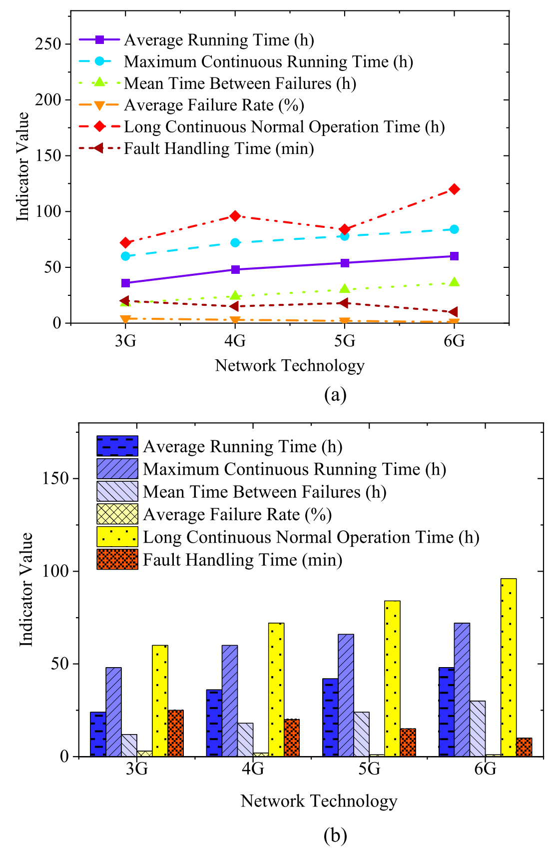


Fig. 8. Comparison of system stability and reliability (a peak period; b. Daily time).

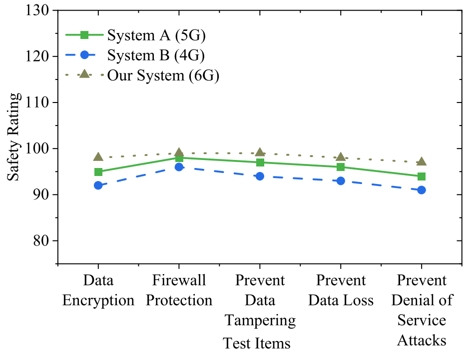
load and real-time processing requirements but also lay a solid foundation for improving the system performance and expanding the application field in the future.

C. System Stability and Reliability Testing

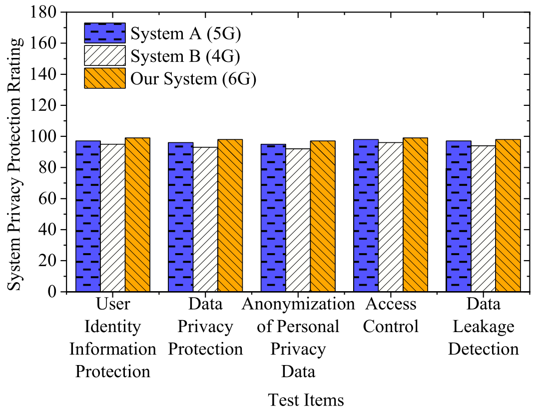
Fig. 8 compares the stability and reliability of the traffic safety system supported by different network technologies.

Fig. 8 displays that during peak hour tests, the system's average running time has increased from 36 hours with 3G to 60 hours with 6G, while the longest continuous running time has extended from 60 to 84 hours. The average time between failures has improved from 18 to 36 hours, and the average failure rate has decreased from 4% with 3G to 1% with 6G. Furthermore, the longest continuous normal running time has risen from 72 to 120 hours, and the fault handling time has been reduced from 20 minutes with 3G to 10 minutes with 6G. During tests under normal road conditions, the system's stability performance shows similar improvements. The upgrade to 6G technology results in increased average running time, longest continuous running time, and average time between failures, while the average failure rate decreases. Concurrently, the longest continuous normal running time increases, and fault handling time decreases with the advancement of network technology.

Therefore, during peak hours and under normal road conditions, the 6G network demonstrates higher stability and reliability compared to other network technologies. This is



(a)



(b)

Fig. 9. Comparison of system security (a: Data transmission security; b: System privacy protection).

due to significant improvements in data transmission speed, bandwidth, and capacity in the 6G network, as well as optimizations in system hardware and software. Specifically, the 6G network can improve data transmission speed and expand bandwidth to effectively address challenges during high-traffic periods, ensuring rapid data transmission and real-time processing. Simultaneously, optimizations in system hardware and advancements in software algorithms further enhance network responsiveness and stability, enabling outstanding performance in complicated traffic management environments.

D. System Security Testing

The comparison of traffic safety protection using various network technologies is suggested in Fig. 9.

Fig. 9(a) shows that regarding data transmission security, the proposed 6G system performs excellently in data encryption, achieving a score of 98, notably higher than other systems. This signifies that the 6G network can better ensure data security during transmission, effectively preventing data leaks and attacks. Additionally, the proposed system excels in firewall protection, preventing data tampering, data loss prevention, and protection against denial-of-service attacks, with scores all above 98. In comparison, other systems have lower ratings, particularly in firewall protection and data tampering prevention, where they are markedly inferior to the 6G network.

Fig. 9(b) expresses that in terms of system privacy protection, the proposed system performs outstandingly in protecting user identity information, data privacy, anonymizing personal private data, access control, and data leak detection, with scores all above 98. This illustrates that the 6G network offers higher reliability and credibility in user privacy protection. In contrast, other systems have relatively lower scores in privacy protection, especially in user identity information protection and data privacy protection, where they are remarkably lower than the 6G network. This discrepancy may be due to the 6G network's stricter privacy protection mechanisms and more advanced data encryption algorithms.

In summary, through detailed data and analysis, this study comprehensively demonstrates the application potential of 6G technology in ITS, especially its significant advantages in improving system security, data processing speed, and privacy protection. These results not only offer innovative solutions for smart cities' traffic management in the future but also lay a solid foundation for the application of 6G technology in other key areas.

E. Discussion

This study integrates 6G technology with RF algorithm for application in the intelligent traffic safety system and conducts in-depth experimental evaluation and discussion. The experimental results demonstrate outstanding predictive performance in various scenarios, particularly on major and minor urban roads. Specifically, under clear and cloudy weather conditions, the system exhibits high accuracy and stable predictive capabilities, accurately forecasting traffic congestion and accidents. However, under adverse weather conditions (such as rain or snow) and during special events (such as accidents), there is a slight decrease in predictive accuracy and recall, though overall performance remains at a high level.

On the one hand, previous studies, such as those by Alekseeva et al. [27] and Lohrasbinasab et al. [28], which mostly focus on traditional wireless networks or single machine learning algorithms for traffic prediction. This study for the first time combines high-speed, low-latency 6G technology with powerful RF algorithms to achieve real-time monitoring and precise prediction of traffic conditions. On the other hand, unlike other studies, such as those by Boukerche and Wang [29] and Elfar et al. [30], which may be limited to specific environments or weather conditions. However, the proposed system demonstrates stable and efficient predictive capabilities across different scenarios, particularly achieving unprecedented accuracy in predicting traffic congestion and accidents under clear and cloudy weather conditions. Even under adverse weather and special event conditions, the system's overall performance remains industry-leading.

Therefore, the experimental results have significant application potential and significance in practical traffic systems. Through real-time data analysis and prediction, these systems help in the early detection of traffic congestion and potential accident risks, enabling timely measures to reduce accident rates and traffic congestion time. This not only saves fuel and reduces vehicle emissions but also improves road utilization and commuting time, enhancing the travel experience for urban residents. Additionally, this study explores the role of real-time route optimization technology in improving driver travel efficiency. By dynamically adjusting driving routes based on real-time traffic data and road conditions, optimal driving paths can be identified to reduce travel time and vehicle congestion, thus enhancing the overall efficiency of the traffic system. These applications not only hold significant importance for urban traffic management and road safety but also play a positive role in reducing adverse environmental impacts such as exhaust emissions and noise pollution.

Furthermore, one of the key factors affecting prediction accuracy is data quality. Despite data cleaning and processing efforts, noise and outliers in sensor data may still affect model performance. Future research could explore more efficient data-cleaning techniques to minimize these negative effects. Moreover, the importance of feature selection cannot be overlooked; the current feature set may not fully capture the complex variations in traffic conditions. Introducing more relevant features or utilizing advanced feature engineering techniques, such as time-series feature analysis and composite feature construction, can further enhance the predictive capabilities of the model.

Parameter optimization of the RF algorithm is also crucial for improving system performance. Adjusting parameters such as the number of trees and maximum depth can optimize the model's generalization ability and stability, thereby facilitating the consistency and accuracy of prediction results.

To further enhance the practicality and application value of the system, future research could focus on several aspects. Firstly, data sources and volumes are expanded, especially integrating traffic data and historical data from multiple cities, to extend the model's applicability and enhance predictive generalization. Secondly, methods of multi-model fusion are explored, such as combining RF with deep learning models, to further improve prediction accuracy and robustness in complex traffic scenarios. Additionally, continuous optimization of real-time data processing capabilities and system stability is crucial for ensuring long-term operation and performance enhancement.

In conclusion, by thoroughly analyzing the significance of experimental results and identifying influencing factors, along with proposing specific improvement directions, this study provides robust support and guidance for the development of the intelligent traffic safety system, promoting the effectiveness and efficiency of technology in urban traffic management.

V. Conclusion

An intelligent traffic safety system is designed and implemented based on 6G technology and the RF algorithm. Efficient data transmission is achieved through the 6G communication module. Moreover, data processing and model execution are handled by high-performance computing servers. The traffic data undergoes preprocessing steps such as cleaning, format conversion, and feature selection, and the RF algorithm is used to train and optimize the traffic data. Suitable parameters are set and cross-validation is performed in experiments to ensure the model's high accuracy and reliability. Extensive experiments are conducted in various traffic scenarios, involving traffic congestion warnings, traffic accident predictions, and tests of the system's real-time data processing capabilities.

The proposed system has great potential in practical traffic applications. In the intelligent traffic management domain, leveraging the ultra-high speed and low latency characteristics of 6G networks can greatly enhance real-time communication between vehicles. In urban planning, utilizing urban operational data collected by big data analysis and Internet of Things devices, city managers can more accurately predict population flow trends and optimize the layout of public facilities. This not only improves the efficiency of city management and the quality of residents' lives but also enhances urban security and sustainability, which is crucial for building smart, green, and safe cities.

However, the experimental data mainly come from traffic data of a specific city, which may have regional limitations and may not fully represent the traffic conditions of different cities. Although the RF algorithm performs well in this study, its performance with other types of traffic data or in extreme situations (such as emergencies or extreme weather) has not been fully verified. Future plans include incorporating more types of data sources, such as social media data and in-vehicle data, to enhance the system's prediction accuracy and comprehensiveness.

Current research in the intelligent traffic safety system field demonstrates the potential and advantages of integrating 6G technology with RF algorithms. However, future research can further explore how to optimize data processing efficiency and enhance the complexity and accuracy of prediction models to address more complex traffic scenarios and data challenges. Additionally, exploration into multi-model fusion, deep learning technologies for applications in intelligent traffic systems, and broader applications for smart city development can also be pursued.

References

[1] A. Morimoto, A. Wang, and N. Kitano, "A conceptual framework for road traffic safety considering differences in traffic culture through international comparison," *IATSS Res.*, vol. 46, no. 1, pp. 3–13, Apr. 2022.

[2] Y. Amiour, E. O. D. Waygood, and P. E. W. van den Berg, "Objective and perceived traffic safety for children: A systematic literature review of traffic and built environment characteristics related to safe travel," *Int. J. Environ. Res. Public Health*, vol. 19, no. 5, p. 2641, Feb. 2022.

[3] S. Wu, "Spatiotemporal dynamic forecasting and analysis of regional traffic flow in urban road networks using deep learning convolutional neural network," *IEEE Trans. Intell. Transp. Syst.*, vol. 23, no. 2, pp. 1607–1615, Feb. 2022.

[4] K. Yu, L. Lin, M. Alazab, L. Tan, and B. Gu, "Deep learning-based traffic safety solution for a mixture of autonomous and manual vehicles in a 5G-enabled intelligent transportation system," *IEEE Trans. Intell. Transp. Syst.*, vol. 22, no. 7, pp. 4337–4347, Jul. 2021.

[5] J. N. Njoku, C. I. Nwakanma, G. C. Amaizu, and D. Kim, "Prospects and challenges of metaverse application in data-driven intelligent transportation systems," *IET Intell. Transp. Syst.*, vol. 17, no. 1, pp. 1–21, Jan. 2023.

[6] Z. Lv and W. Shang, "Impacts of intelligent transportation systems on energy conservation and emission reduction of transport systems: A comprehensive review," *Green Technol. Sustainability*, vol. 1, no. 1, Jan. 2023, Art. no. 100002.

[7] A. Gohar and G. Nencioni, "The role of 5G technologies in a smart city: The case for intelligent transportation system," *Sustainability*, vol. 13, no. 9, p. 5188, May 2021.

[8] M. Yu, "Construction of regional intelligent transportation system in smart city road network via 5G network," *IEEE Trans. Intell. Transp. Syst.*, vol. 24, no. 2, pp. 2208–2216, Feb. 2023.

[9] R. Jabbar et al., "Blockchain technology for intelligent transportation systems: A systematic literature review," *IEEE Access*, vol. 10, pp. 20995–21031, 2022.

[10] R. W. Liu, Y. Guo, Y. Lu, K. T. Chui, and B. B. Gupta, "Deep network-enabled haze visibility enhancement for visual IoT-driven intelligent transportation systems," *IEEE Trans. Ind. Informat.*, vol. 19, no. 2, pp. 1581–1591, Feb. 2023.

[11] N. Yuvaraj, K. Praghash, R. A. Raja, and T. Karthikeyan, "An investigation of garbage disposal electric vehicles (GDEVs) integrated with deep neural networking (DNN) and intelligent transportation system (ITS) in smart city management system (SCMS)," *Wireless Pers. Commun.*, vol. 123, no. 2, pp. 1733–1752, Mar. 2022.

[12] G. Yi and S. Wu, "Graphical visual analysis of consumer electronics public comment information mining under knowledge graph," *IEEE Trans. Consum. Electron.*, vol. 70, no. 1, pp. 2197–2924, Jan. 2024.

[13] C. Roy, R. Saha, S. Misra, and K. Dev, "Micro-safe: Microservices- and deep learning-based safety-as-a-service architecture for 6G-enabled intelligent transportation system," *IEEE Trans. Intell. Transp. Syst.*, vol. 23, no. 7, pp. 9765–9774, Jul. 2022.

[14] Q. Xue, X. Wang, Y. Li, and W. Guo, "Young novice drivers' cognitive distraction detection: Comparing support vector machines and random forest model of vehicle control behavior," *Sensors*, vol. 23, no. 3, p. 1345, Jan. 2023.

[15] Y. Regragui and N. Moussa, "A real-time path planning for reducing vehicles traveling time in cooperative-intelligent transportation systems," *Simul. Model. Pract. Theory*, vol. 123, Feb. 2023, Art. no. 102710.

[16] S. Reza, H. S. Oliveira, J. J. M. Machado, and J. M. R. S. Tavares, "Urban safety: An image-processing and deep-learning-based intelligent traffic management and control system," *Sensors*, vol. 21, no. 22, p. 7705, Nov. 2021.

[17] S. K. Jagatheesaperumal, M. Rahouti, K. Ahmad, A. Al-Fuqaha, and M. Guizani, "The duo of artificial intelligence and big data for industry 4.0: Applications, techniques, challenges, and future research directions," *IEEE Internet Things J.*, vol. 9, no. 15, pp. 12861–12885, Aug. 2022.

[18] H. H. H. Mahmoud, A. A. Amer, and T. Ismail, "6G: A comprehensive survey on technologies, applications, challenges, and research problems," *Trans. Emerg. Telecommun. Technol.*, vol. 32, no. 4, p. e4233, Apr. 2021.

[19] A. Salh et al., "A survey on deep learning for ultra-reliable and low-latency communications challenges on 6G wireless systems," *IEEE Access*, vol. 9, pp. 55098–55131, 2021.

[20] N. Jain and P. K. Jana, "LRF: A logically randomized forest algorithm for classification and regression problems," *Expert Syst. Appl.*, vol. 213, Mar. 2023, Art. no. 119225.

[21] L. Breiman, "Random forests," *Mach. Learn.*, vol. 45, pp. 5–32, Oct. 2001.

[22] L. Capitaine, R. Genuer, and R. Thiébaut, "Random forests for high-dimensional longitudinal data," *Stat. Methods Med. Res.*, vol. 30, no. 1, pp. 166–184, Jan. 2021.

[23] S. Zhou, C. Wei, C. Song, X. Pan, W. Chang, and L. Yang, "Short-term traffic flow prediction of the smart city using 5G Internet of Vehicles based on edge computing," *IEEE Trans. Intell. Transp. Syst.*, vol. 24, no. 2, pp. 2229–2238, Feb. 2023.

[24] C. Li, Y. Xu, H. Zheng, Z. Wang, H. Han, and L. Zeng, "Artificial intelligence, resource reallocation, and corporate innovation efficiency: Evidence from China's listed companies," *Resour. Policy*, vol. 81, Mar. 2023, Art. no. 103324.

[25] M. N. Khan and S. Das, "Advancing traffic safety through the safe system approach: A systematic review," *Accident Anal. Prevention*, vol. 199, May 2024, Art. no. 107518.

[26] Y. Ma, S. B. Goyal, A. S. Rajawat, P. Bedi, and S. Yasmeen, "Blockchain-based human intelligent systems for smart city safety," *Trans. Emerg. Telecommun. Technol.*, vol. 35, no. 2, Feb. 2024, Art. no. e4939.

[27] D. Alekseeva, N. Stepanov, A. Veprev, A. Sharapova, E. S. Lohan, and A. Ometov, "Comparison of machine learning techniques applied to traffic prediction of real wireless network," *IEEE Access*, vol. 9, pp. 159495–159514, 2021.

[28] I. Lohrasbinasab, A. Shahraki, A. Taherkordi, and A. D. Jurcut, "From statistical-to machine learning-based network traffic prediction," *Trans. Emerg. Telecommun. Technol.*, vol. 33, no. 4, 2022, Art. no. e4394.

[29] A. Boukerche and J. Wang, "Machine learning-based traffic prediction models for intelligent transportation systems," *Comput. Netw.*, vol. 181, Nov. 2020, Art. no. 107530.

[30] A. Elfar, A. Talebpour, and H. S. Mahmassani, "Machine learning approach to short-term traffic congestion prediction in a connected environment," *Transp. Res. Rec., J. Transp. Res. Board*, vol. 2672, no. 45, pp. 185–195, Dec. 2018.

Jianhua Dai is currently a Professor with the Business School, China University of Political Science and Law, China. His research interests include big data and machine learning.

Xiaofan Tian was born in Zhongwei, Ningxia, China, in 1994. She received the bachelor's degree in chemical engineering and technology from Ningxia University in 2016. She is currently pursuing the master's degree in business administration with China University of Political Science and Law.

Lei Liu is currently pursuing the master's degree in business administration with the Business School, China University of Political Science and Law. Her research interests include artificial intelligence algorithms and big data.

Hao Zhang was born in Beijing, China, in 1997. He received the master's degree in commerce and economic development from Northeastern University in 2021. He is currently pursuing the Ph.D. degree with China University of Political Science and Law. His research interests include innovation economics, the Internet of Things, and big data.

Jian Fu was born in Henan, China, in 1991. She received the master's degree in business administration from China University of Political Science and Law in 2024. She is currently pursuing the Ph.D. degree with China University of Political Science and Law. Her research interests include innovation economics, big data, and artificial intelligence.

Miao Yu was born in Harbin, Heilongjiang, China, in 1970. He received the Ph.D. degree in trade economics from the Renmin University of China in 1999. He was a Post-Doctoral Fellow with Beihang University in 2001. He is currently a Professor with the Business School, China University of Political Science and Law. His research interests include management science and engineering.