Research on Traffic Accident Prediction Based on KG-CWT-RGCNN-BiLSTM

Changfeng Zhu, Yu Wang, Qingrong Wang, Jinhao Fang, Jie Wang, and Linna Cheng

Abstract—**Traffic accident is one of the rarest causes of death in the world. In order to reduce the frequency of traffic accidents, it is of great significance to predict traffic accidents. Due to the dependence of traditional methods, this paper uses deep learning methods to predict traffic accidents, which reduces manual errors. Therefore, this paper first uses Knowledge Graph (KG) to analyze the correlation between various factors of traffic accidents, and then proposes a Continuous Wavelet Transform (CWT). A model combining three methods of Relation Graph Convolution Neural Networks (RGCNN) and Bi-directional Long Short-Term Memory (BiLSTM) is used to predict traffic accidents. Through comparative analysis of the model, the accuracy rate, recall rate and F1-score value of the model are higher than other models, and the prediction accuracy rate reaches 99.32 %. Through the parameter analysis of the model, the optimal parameter values are obtained.**

Index Terms—**Accident prediction, Knowledge graph, Continuous wavelet transform, Relation graph Convolutional neural network, Bidirectional long short-term memory network**

I. Introduction

According to the Global Road Safety Report released by the World Health Organization in 2022, about 22,000 people worldwide die from traffic accidents every day, causing traffic accidents to be one of the rarest causes of death. However, there are many factors affecting traffic accidents. The frequency and environment of accidents in different regions will lead to traffic accidents. Due to the uncertainty of traffic accidents, such as the different number of accidents in prosperous cities and remote rural areas, many accidents in some areas on a monthly basis, and accidents in some areas on an annual basis, these have a serious impact on the prediction of traffic accidents. Therefore, how to establish an efficient traffic accident prediction model is of great significance for human normal travel.

In order to predict traffic accidents in advance, so as to avoid casualties. In 2006, García-Ferrer et al.[1] used a single model or a combination of multiple single models to predict the probability of traffic accidents in Spain each month. In 2013, Mehdi Hosseinpour et al. [2] took the traffic accident of Iran 's Qazvin-Loshan intercity road within two years as an example, and constructed the adaptive neuro-fuzzy inference system (ANFIS) for the first time to predict the probability of traffic accidents under uncertainty and complexity. In 2014, Markus Deublein et al. [3] took the traffic accident data of Austrian rural expressways as an example to summarize and compare the empirical Bayesian (EB) method and Bayesian probability networks (BPNs) method. Both methods can predict how far traffic accidents can be predicted, but BPNs have a higher correlation with traffic accident data. After that, Reference [4] used grey correlation analysis to evaluate the weight of each variable that causes traffic accidents on the impact of accidents, so as to analyze the correlation between various factors, and analyze the main factors from the correlation analysis, so as to construct the MGM (1,N) model and analyze the collinearity between variables.In the literature [5-6], the grey Markov model is used. Among them, the literature [5] compares and analyzes the grey system theory (GSTM) and the grey Markov model (GM), and mainly predicts the trend of road traffic accidents. Experiments show that the prediction results of the GM model are more accurate than the GSTM prediction results. The literature [6] mainly uses the Markov model to predict the traffic accidents in the plateau mountainous area, and combines the logistic regression method to analyze the correlation between the various factors. It shows that the road type has a significant impact on the traffic accident rate. In literature[7], a traffic safety state deep clustering network (TSDCN) method is proposed, including an end-to-end deep hybrid network combining feature extraction and cluster analysis, so as to quantify the risk level that will lead to accidents. Ospina Mateus Holman et al. In literature[8] constructed a model based on SPF-Bayes '. The prediction is mainly divided into two stages. The first stage uses the negative binomial regression method to construct SPF, and the second stage uses Bayes ' empirical method to predict the road section where motorcycles are prone to accidents, and completes the prediction of motorcycle accidents in Cartagena city. Muhammad Babar Ali Rabbani et al. In literature[9] compared the seasonal autoregressive moving

Manuscript received July 18, 2023; revised August 13, 2023.

This work was supported in part by the National Natural Science Foundation of China (No.72161024), "Double-First Class" Major Research Programs, Educational Department of Gansu Province (No.GSSYLXM-04) and Gansu Provincial Science and Technology Plan Project (22ZY1QA005).

Changfeng Zhu is a Professor at School of Traffic and Transportation, Lanzhou Jiaotong University, Lanzhou 730070, China. (e-mail: cfzhu003@163. com).

Yu Wang is a postgraduate student at School of Traffic and Transportation, Lanzhou Jiaotong University, Lanzhou 730070, China. (Corresponding author, phone: +86 15379994351, e-mail: 962078842@qq. com).

Qingrong Wang is a Professor at School of Electronic and Information Engineering,Lanzhou Jiaotong University, Lanzhou 730070, China. (e-mail:wangqr003@163.com)

Jinhao Fang is a doctoral candidate at School of Traffic and Transportation, Lanzhou Jiaotong University, Lanzhou 730070, China. (e-mail: fangjin_hao@163. com).

Jie Wang is a Ph.D. at School of Traffic and Transportation, Lanzhou Jiaotong University, Lanzhou 730070, China. (e-mail: 009696615@qq.com).

Linna Cheng is a Ph.D. at School of Traffic and Transportation, Lanzhou Jiaotong University, Lanzhou 730070, China. (e-mail: chenglinnajiao@126.com).

average (SARIMA) model with the exponential smoothing (ES) model, and analyzed the time series of traffic accidents. The performance of the ES model is better than the SARIMA model, thereby predicting the traffic accident rate in Pakistan. Deretić Nemanja et al. In literature[10] used time series to extract data features and constructed a SARIMA-based model to predict traffic accidents in Belgrade.

However, the factors leading to traffic accidents are complex and changeable in many cases. The traditional methods have low prediction stability for this situation, and the deep learning method is superior to the traditional method. Based on this, Marcillo Pablo et al. In literature[11] constructed a learning-based prediction model. The performance of the model mainly depends on the quality of data, the distribution of data, the exploration of data and the use of data, which lays a foundation for future deep learning prediction.

After that, Mamoudou Sangare et al. [12] constructed a traffic accident prediction model based on GMM-SVC combination. GMM is used to describe the data characteristics of traffic accidents and is used as the input of SVC.SVC is used as the classifier of the model. The model has higher performance than the baseline statistical method. Alqatawna Ali et al. [13] compared the multiple regression model (MRm) with the artificial neural network (ANNs) method to predict road traffic accidents in Spain. The ANNs mainly used the Levenberg-Marquardt algorithm and the sigmoid activation function. The results show that the performance of the ANNs model is higher than that of the MRm, and the estimated value predicted by the ANNs is more similar to the future traffic accidents in Spain. Wakatsuki Yuki et al. [14] used convolutional neural network (CNN) method to predict the probability of accidents in a specific road section in the next 2 hours by using chi-square test with traffic data, time data and weather data as factors. In literature[15] constructed an attention mechanism (AM) model that simultaneously extracts features of time and space to predict traffic accidents on urban roads and reduce losses.Lida Barba et al. [16], Gyanendra Singh et al. [17] and Li et al. [18] all used the theory of neural network. Among them, the literature[16] also combined the autoregressive moving average method to ensure the extraction and prediction of time series. the literature [17] used the deep neural network (DNN) method to quantify the influence of various variables on the accident frequency. the literature [18] used the combination of SNNs and CNNs to predict the traffic accident model. The model can efficiently capture the spatio-temporal characteristics of accidents and the advantages of CNNs in accurately characterizing the traffic environment.

Because there are many factors leading to traffic accidents, in order to characterize the correlation, some scholars use the knowledge graph method to analyze its correlation. Among them, the literature [19] takes open source data as an example to construct traffic knowledge graph and affair graph, which can help traffic management departments find traffic problems in time and give early warning, and give decisions as soon as possible. The literature [20,21] uses the structure of the knowledge graph to store the data. Among them, the literature [20] constructs the knowledge graph of urban rail transit travel and urban road travel, and uses the structure of the knowledge graph to store the data. The literature [21] mainly uses the knowledge graph to describe the correlation of multi-dimensional dynamic and static factors, and visualizes the results to facilitate the understanding of massive and complex data.

However, after the knowledge graph analyzes its correlation, it is only a structured storage of the data and cannot analyze the data features it contains. However, CNN is not very high for image feature extraction performance. Some scholars have improved the CNN methodology [22], which laid the foundation for the later. Some scholars combine knowledge graph with graph convolutional neural networks (GCNNs) [23].GCNN has better feature extraction performance than CNN.In literature[24] uses knowledge graph to construct urban rail transit emergencies. As input, the constructed knowledge graph is input into relational graph convolutional neural networks (RGCNNs) to predict the evolution results of urban rail transit emergencies. In literature[25], a deep spatio-temporal graph convolutional neural network (DST-GCN) is constructed, which is mainly composed of three parts. The spatial information is extracted, the dynamic characteristics of the spatio-temporal information are extracted and the external useful information is extracted. Finally, the prediction of traffic accidents is completed.

The spatial feature extraction of data features is completed, but the traffic accident data contains time features. Some scholars use long short-term memory network (LSTM) to extract the time features it contains. Among them, the literature [25,26] uses LSTM to extract time features. Among them, the literature [25] compares LSTM with traditional regression models and traditional neural networks to predict the safety level of traffic accidents. The literature [26] combines LSTM with gradient-enhanced regression tree (GBRT) to predict traffic accidents and compares it with various MR models and neural networks. The LSTM-GBRT model has better fitting and robustness. However, although LSTM has improved RNN [27], it can only extract data individually. After that, some scholars use the bidirectional LSTM (Bi-LSTM) [28] model, which can simultaneously forward and reverse propagation, with high prediction stability.

However, since the manually acquired data contains certain errors, in order to reduce the noise of the acquired data, the Fourier transform (FT)[29] proposed earlier laid the foundation for obtaining the frequency at the same time. On this basis, Morlet [30] proposed wavelet transform (WT), which further improved FT and overcame the problem that the window size does not change with the frequency transformation. On the basis of wavelet theory, the literature [31] established the grey GM(1,1) model. Through wavelet analysis, the road traffic accidents in a province from 2003 to 2009 were decomposed into multi-level stationary data sequences. GM(1,1) model was established for low-frequency reconstruction sequences to predict road traffic accidents. Continuous wavelet transform (CWT) [32] is another improvement of wavelet transform.

However, in the prediction of traffic accidents, some scholars seldom consider the four aspects of data correlation, data noise reduction, graph feature extraction and time feature extraction at the same time, which will lead to certain errors. Based on this, this paper comprehensively considers four aspects, introduces KG theory, CWT theory, RGCNN theory and BiLSTM theory, and solves the above four problems respectively. Finally, a model based on KG-CWT-RGCNN-BiLSTM is constructed to predict traffic accidents and complete classification.

The rest of this article is introduced as follows : In the second chapter, a model based on KG-CWT-RGCNN-BiLSTM is constructed. In the third chapter, according to the constructed KG-CWT-RGCNN-BiLSTM model, taking the traffic accidents of 49 states in the United States from 2016 to 2019 as an example, the traffic accidents of the state are predicted, and comparative analysis and parameter analysis are carried out. The fourth chapter, according to the case analysis conclusion.

II. Construction of Traffic Accident Prediction Model Based on CWT-KG-RGCNN-BiLSTM

The number of motor vehicles is increasing year by year, and the number of traffic accidents is also increasing year by year. There are many factors that affect traffic accidents. Weather conditions in fog and snow will affect road visibility and road congestion. The distribution of urban areas will also affect traffic accidents, and traffic flow will be larger in a certain period of time. Urban traffic accidents are much higher than remote rural areas, and the frequency of traffic accidents in the two areas is quite different. Based on this, there are many factors affecting traffic accidents, which is a great challenge to predict the occurrence of traffic accidents.

In order to predict traffic accidents, it is mainly divided into two parts : space and time. Firstly, the collected traffic accident data is denoised by CWT method, and then KG is used to describe the correlation between its factors. Complete the preliminary processing of the data, input the knowledge graph into the RGCNN layer to extract the spatial features of the data, and then input the BiLSTM layer to extract the temporal correlation features of the data, thereby completing the prediction of traffic accidents. The traffic accident prediction structure based on CWT-KG-RGCNN-BiLSTM model is shown in Fig.1.

A. Data Correlation Based on KG

After the CWT layer denoises the data, in order to predict traffic accidents more effectively, it is necessary to characterize the relationship between the data. KG describes the relationship between data in the way of edge and point, and can express massive and complex data in a simple form. Therefore, after denoising the data, KG is used to describe the relationship between traffic accidents. KG describes the relationship between two entities in a triple (head, relationship, tail) manner, where the head and tail represent entities in traffic accidents and are described by points in the knowledge graph. The relationship mainly connects two entities in traffic accidents and is depicted by edges in the knowledge graph. K = { R, S, A } is used to represent the triple of entity-relation-attribute, and the relationship between traffic accidents is described by points and edges in the corresponding knowledge graph, where R represents the relationship between two entities, S represents the attribute corresponding to the entity, and A represents the co-occurrence relationship between different attributes.

$$R=\left\{\left(s_i,con,s_j\right),i,j\in\left\{1,2,\cdots,n\right\}\right\} \tag{1}$$

$$S=\left\{\left(s_i,q_l,q_l^{s_i}\right),i\in\left\{1,2,\cdots,n\right\},l\in\left\{1,2,\cdots,N\right\}\right\} \tag{2}$$

$$A=\left\{\left(q_{l1},q_{l2},p\right),l_1,l_2\in\left\{1,2,\cdots,N\right\}\right\} \tag{3}$$

In the formula, s_i and s_j represent the road sections i and j in the city, *con* represents the connection relationship between their road sections, n represents the number of road sections contained in the city, q_l represents the attributes of class l, $q_l^{s_i}$ represents the corresponding attributes of the si road section. N represents the number of total attributes, and p represents the co-occurrence relationship between attributes. The flow chart of using knowledge graph to describe the relationship between traffic accident data is shown in Fig.2.

As shown in Fig.2, the construction of knowledge graph mainly includes five steps : data acquisition, data extraction, data fusion, data modeling and retrieval.

Step1 : Extract the factors that cause the accident from the accident case and denoise the data.

Step2 :Extract the entity-attribute-relationship information from the processed data in an automated form. Entity extraction is mainly to automatically identify entities from traffic accident data sets. Attribute extraction is mainly to identify the specific attribute information of the entity. Relation extraction is mainly to automatically find out the one-to-many or many-to-one relationship between entities, so as to connect entities and eventually form a traffic accident topology.

Step3 : After completing the entity-attribute-relationship extraction, it is spliced to complete the information fusion.

Step4 : The entities in the extracted and completed traffic accident data are automatically juxtaposed and formed the relationship between the upper and lower levels, and finally the ontology is generated to complete the triple construction of R-S-A.

Step5: After that, the quality of the built the entity-relationship-entity is evaluated, and the low-quality knowledge base is discarded. Finally, the relationship between traffic accidents is described by knowledge graph to complete the construction.

B. Data Denoising Based on CWT

KG analyzes the factors of traffic accident data. In order to further predict the occurrence of traffic accidents, the traffic accident data is input into the CWT layer for data noise reduction. Traffic accident data is obtained according to traffic accident cases, which is easy to cause artificial errors. Since CWT can obtain both time domain and frequency domain, the data is input into the CWT layer to remove the noise of traffic accident data. The noise reduction process of traffic accident data based on CWT is shown in Fig.3.

As shown in Fig.3, the factors leading to the accident are analyzed from the accident case, and the factors are integrated into the CWT layer. Firstly, the function in the CWT layer is determined. The waveform of the Morlet function is similar to that of the traffic accident image, so the Morlet function is used to denoise the traffic accident. After that, the time domain $\phi_{a,b}(t)$ and frequency domain $wt(a,b)$ of

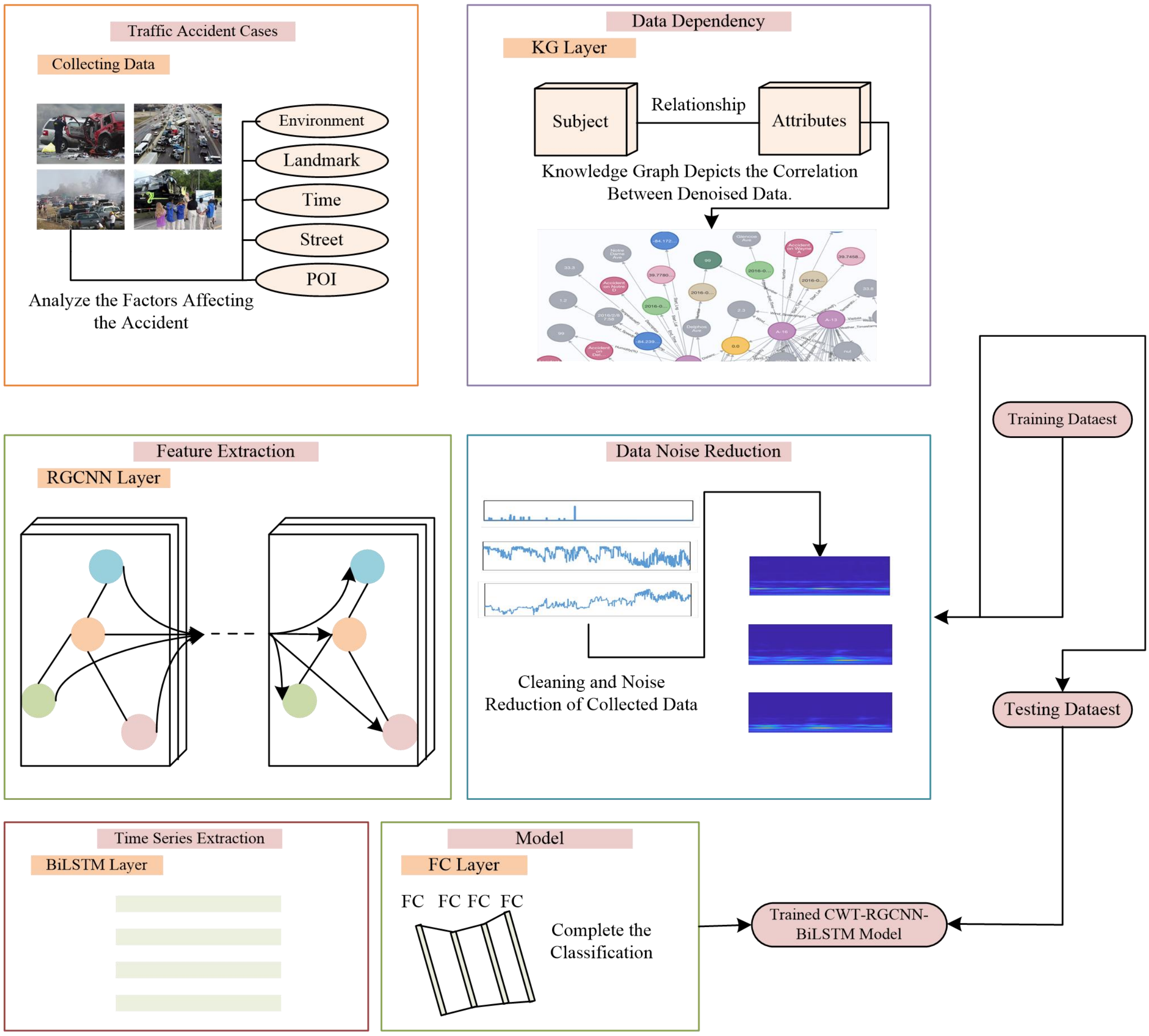


Fig. 1. Traffic accident prediction structure diagram based on CWT-KG-RGCNN-BiLSTM model

traffic accident data are obtained simultaneously. Firstly, the value of a is determined to change b, so as to obtain the time domain $\phi_{a,b}(t)$, and then the value of a is changed to obtain the frequency domain $wt(a,b)$. Among them,

$$\phi_{a,b}(t)=|a|^{-\frac{1}{2}}\phi\left(\frac{t-b}{a}\right) \tag{4}$$

$$wt(a,b)=|a|^{-\frac{1}{2}}\int_{-\infty}^{\infty}x(t)\phi^*\left(\frac{t-b}{a}\right)dt \tag{5}$$

where a represents the scale parameter of the expansion or contraction wavelet of the traffic accident data in the morlet function ; b represents the shift parameters of the traffic

C. Feature Extraction Based on RGCNN

The traffic accident data is denoised at the CWT layer, and the time-frequency diagram is generated and input to the RGCNN layer. CNN can automatically extract the features contained and can avoid the influence of other factors on it. The relationship between traffic accident data belongs to the topological structure, which makes it difficult to extract CNN, and GCNN can better extract features from the topological structure of the graph, which solves this problem. However, the traditional GCNN ignores the relationship between traffic accidents, and R-GCNN can stitch the features between traffic accident data. Therefore, the data after noise reduction is input into R-GCNN, which can better extract the spatial characteristics of traffic accidents. The R-GCNN-based aggregation process is shown in Fig.4.

On the basis of GCNN, R-GCNN also extracts the relationship features between its data, among which

$$h_i^{(l+1)}=\sigma\left(\sum_{r\in R}\sum_{j\in N_i^r}\frac{1}{c_{i,r}}W_r^{(l)}h_j^{(l)}+W_0^{(l)}h_i^{(l)}\right) \tag{6}$$

In the formula, $h_i^{(l+1)}$ represents the characteristics of the nodes after the traffic accident data is updated, $\sigma(\ )$ represents the activation function, $c_{i,r}$ represents the regularization constant, $W_r^{(l)}$ represents the weight matrix of the corresponding relationship characteristics between the traffic accident data, $h_j^{(l)}$ represents the characteristics of the nodes related to the data in the input traffic accident knowledge graph, $W_0^{(l)}$ represents the characteristic weight matrix of the traffic accident data itself, and $h_i^{(l)}$ represents the characteristics of the nodes in the input traffic accident knowledge graph. The propagation formula is to calculate the characteristics of the corresponding relationship between traffic accident data through the feature weight matrix corresponding to the relationship matrix.

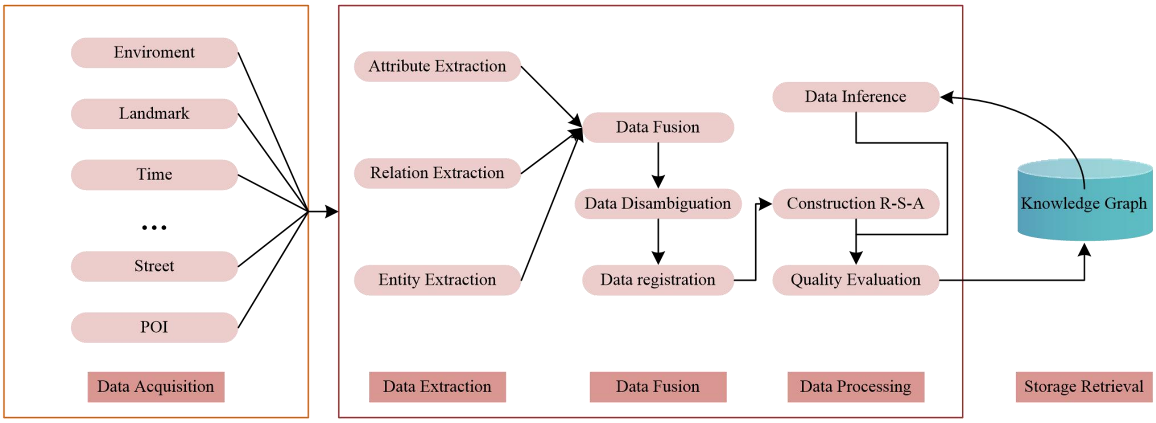


Fig. 2. Knowledge graph depicts the flow chart of the relationship between traffic accident data.

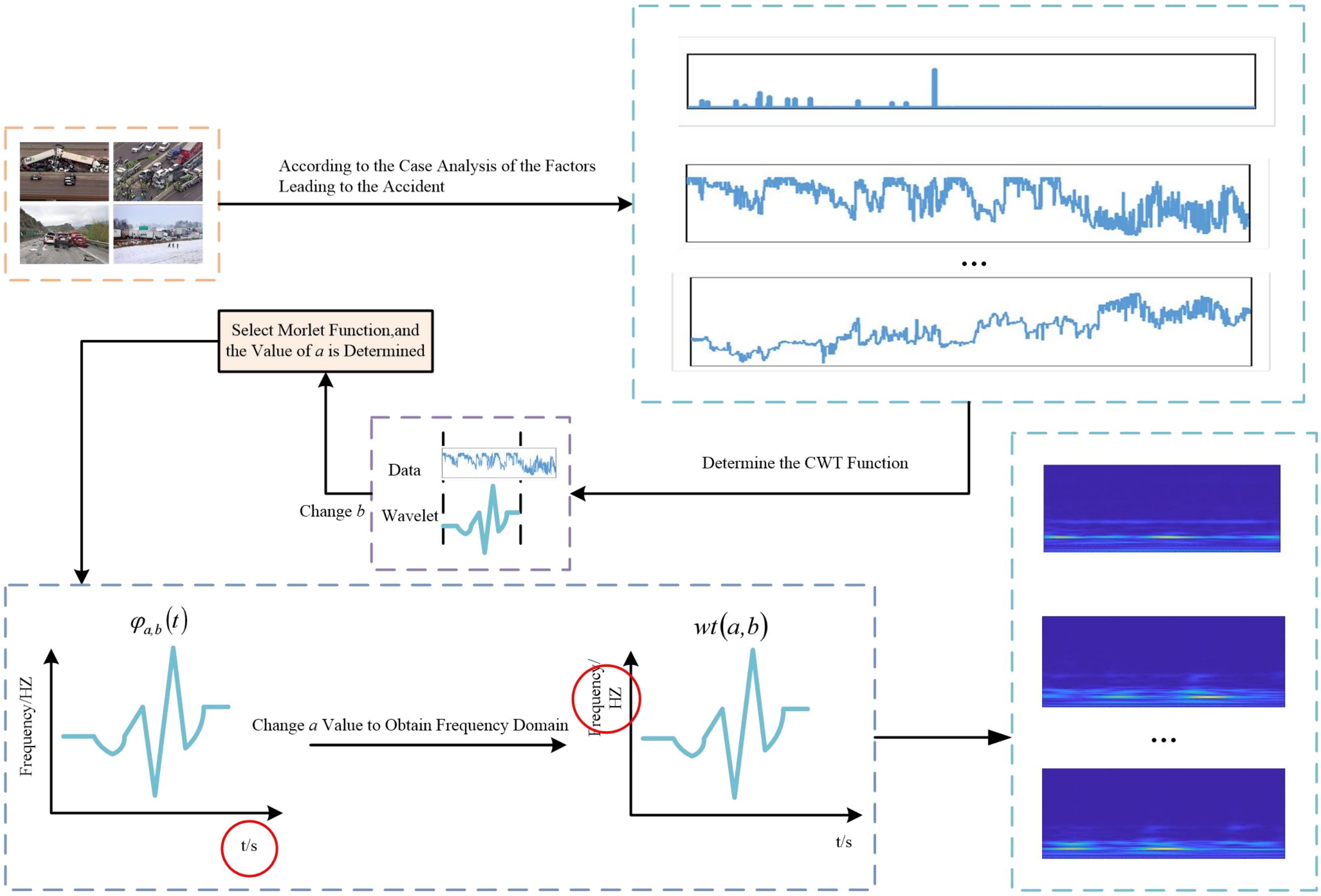


Fig. 3. Noise reduction process of traffic accident data based on CWT

In model training, the larger the picture, the more the relationship between traffic accident data, which will lead to the difficulty of model training, or may lead to over-fitting. Regularization solves two possible problems by sharing parameters. The decomposition of W_r is as follows

$$W_r^{(l)} = \sum_{b=1}^{B} a_{rb}^{(l)} V_b^{(l)} \tag{7}$$

In the formula, r represents the correspondence between each traffic accident, B represents the hyper parameter, which controls the number of V_b , a_{rb} represents the decomposition coefficient of W_r on V_b , and V_b represents the shared parameter.

D. Time Series Extraction Based on BiLSTM

The RGCNN layer completes the spatial feature extraction of traffic accident data. In order to further extract the data time series features, the data is input into the BiLSTM layer. Traditional recurrent neural networks cannot preserve memory for a long time. On this basis, LSTM introduces gate functions to preserve data for a long time, but it will cause data to extract features only in a single direction. On the basis of the two, BiLSTM can retain memory for a long time and can also extract features from the data in two directions, so as to complete the classification and finally complete the prediction of traffic accidents. Therefore, the data processed by RGCNN is input into the BiLSTM layer, and the BiLSTM traffic accident data structure diagram is shown in Fig.5.

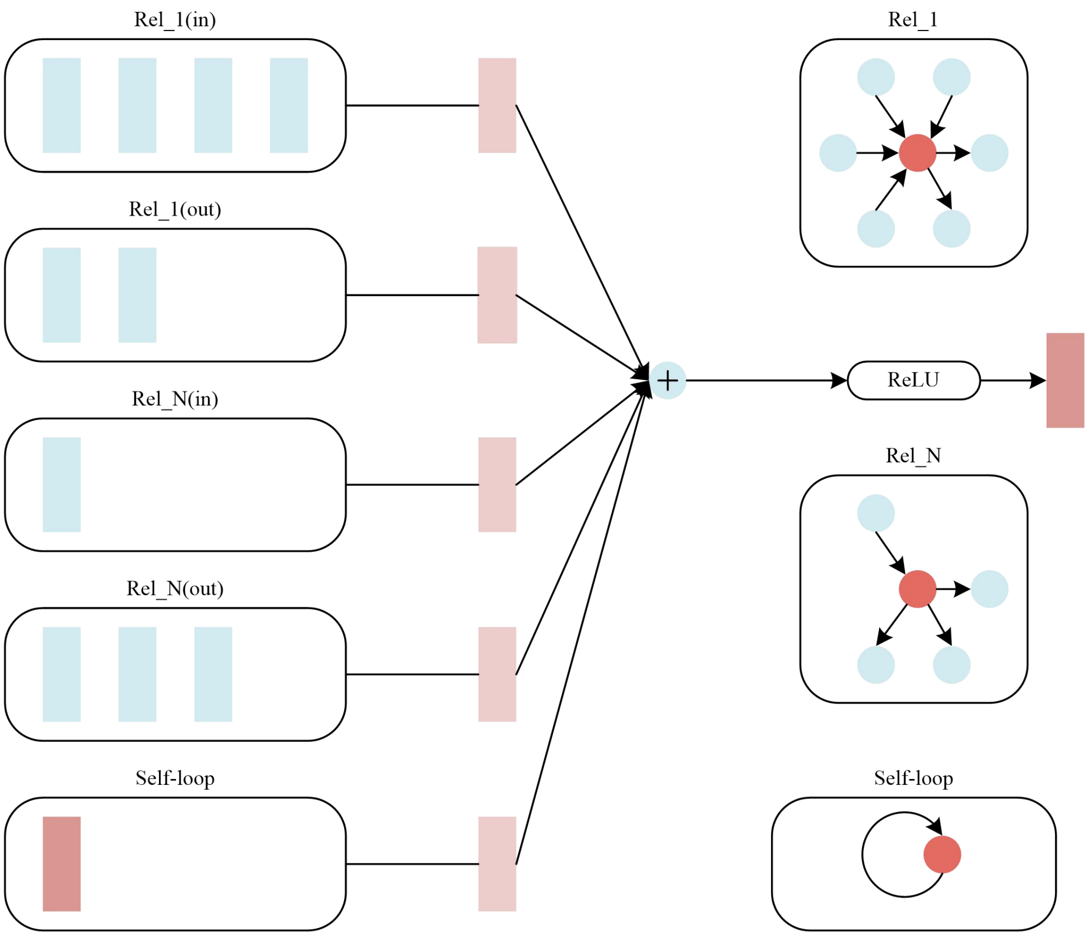


Fig. 4. Based on R-GCNN aggregation process

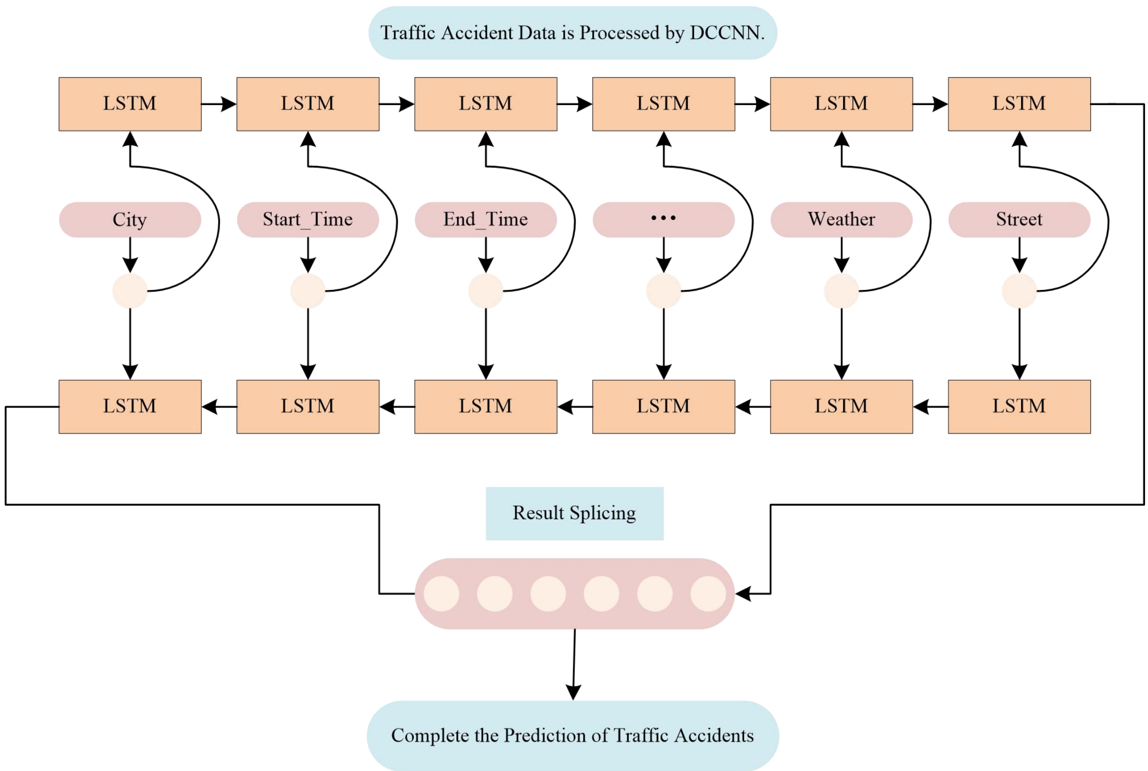


Fig. 5. Structure diagram of BiLSTM processing traffic accident data

As shown in Fig.5, in the BiLSTM layer, the traffic accident data processed by the RGCNN layer simultaneously performs forward and backward propagation, and finally combines the processed data. In LSTM, data first enters the forgetting door and filters useless data ; after that, the data retained by the forgetting gate enters the input gate, mainly to update the old unit state. After the traffic accident data is processed by the forgetting gate and the input gate, the output data time series characteristics, the model

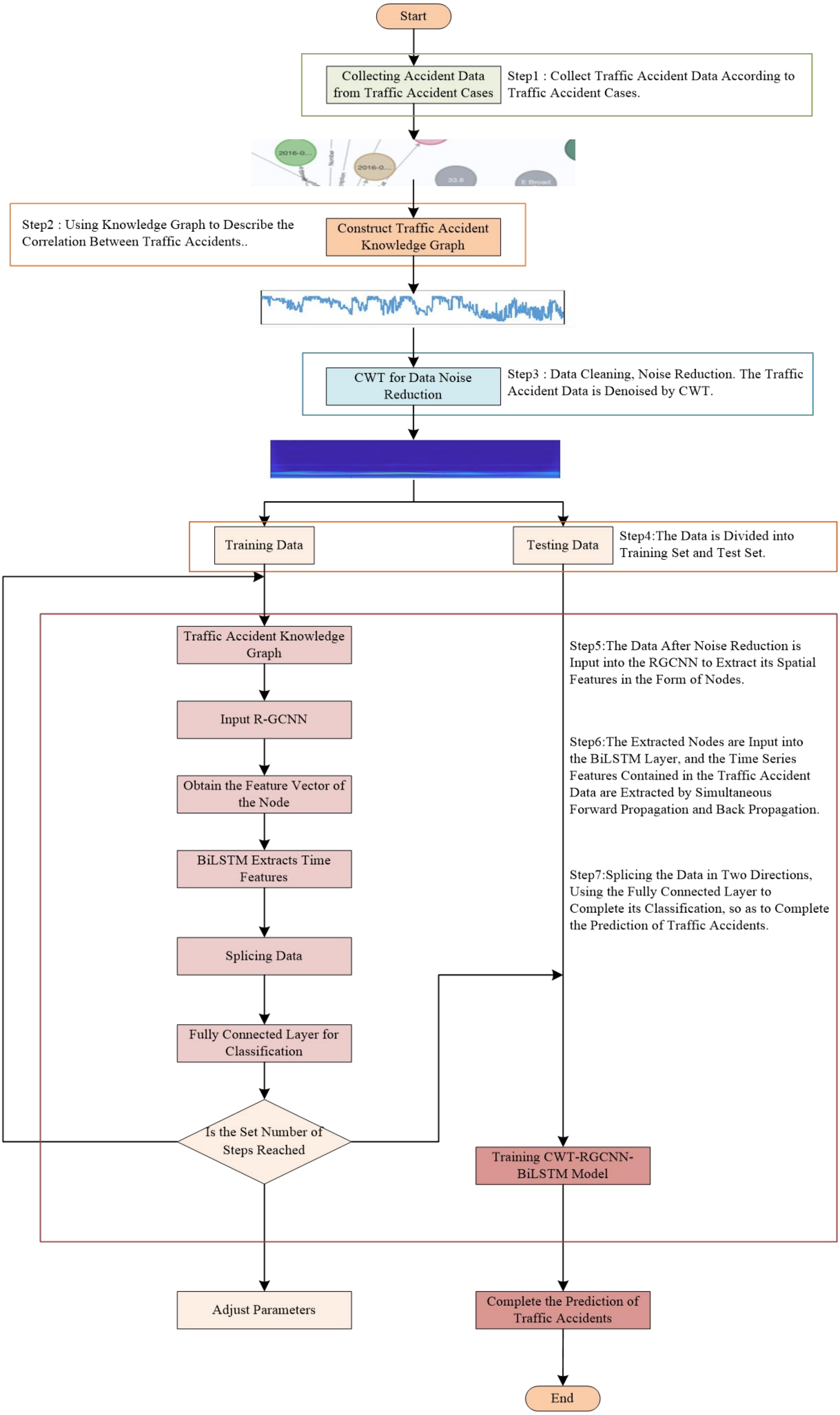


Fig.6. Model structure flow chart of traffic accident prediction based on GK-CWT-DCCNN-LSTM

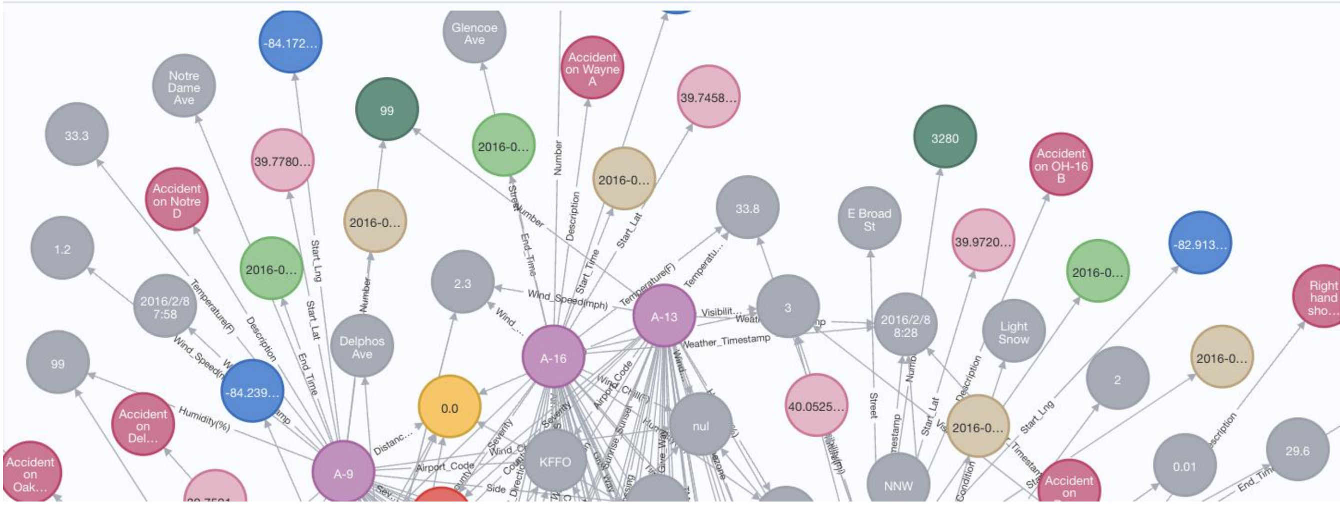


Fig.7. Traffic accident knowledge map visualization

TABLE I
TRAFFIC ACCIDENT DATA TYPE

Classification	Attributes	type	unit	Description
ID		object		Accident identifier
Source		object		Accident data sources
TMC		Float 64		Describe event information in detail
Environment	Temperature ,	Float 64	F	
	Wind Chill	Float 64	F	
	Humidity	Float 64	%	
	Visibility	Float 64	mi	
	Wind_Direction	Object		
	Wind_Speed	Float 64	mp	
	Precipitation	Float 64	h	
	Weather_Conditions	Object	in	Display weather conditions (rain, snow, thunderstorms, fog, etc.)
Landmarks	Start_Lat,Start_Lng,End_Lat, End_Lng	Float 64		
Time	Start_Time, End_Time	Object		
	Sunrise_Sunset	Object		
	Civil_Twilight	Object		The daytime period of folk twilight
	Nautical_Twilight	Object		The Nautical Twilight Daytime
	Astronomical_Twilight	Object		Show the day according to the astronomical twilight
Streets	Street, Number, Side, Zipcode, City, County, State	Object		
POI	Amenity, Bump,Crossing, Give_Way, Junction, No_Exit, Railway, Roundabout, Station, Stop, Traffic_Calming,Traffic_Signal,Turning_Loop	Bool		Is There Amenity, Bump,Crossing, Give_Way, Junction, No_Exit, Railway, Roundabout, Station, Stop, Traffic_Calming, Traffic_Signal and Turning_Loop Nearby
Severity	1-4	Int64		1 the least, 4 the most serious impact

completes the classification and prediction of traffic accidents. Among them

$$Z^{f}=\sigma\left(W_{f}*\left[h_{t-1},x_{t}\right]+b_{f}\right) \quad (8)$$

$$i_{t}=\sigma\left(W_{i}*\left[h_{t-1},x_{t}\right]+b_{i}\right) \quad (9)$$

$$\tilde{C}_{t}=\tanh\left(W_{C}*\left[h_{t-1},x_{t}\right]+b_{c}\right) \quad (10)$$

$$o_{t}=\left(W_{o}\left[h_{t-1},x_{t}\right]+b_{o}\right) \quad (11)$$

$$h_{t}=o_{t}*\tanh\left(C_{t}\right) \quad (12)$$

After processing by BiLSTM, the traffic accident data is input into the fully connected layer to classify it and complete the prediction of traffic accidents. The full connection layer calculation formula is as follows

$$z^{l+1(j)}=W_{ij}^{l}d^{l(i)}+b_{j}^{l} \quad (13)$$

In the formula, $z^{l+1(j)}$ represents the logtis value of the jth output neuron in the $l+1$ layer , W_{ij}^{l} denotes the weight between the i-th neuron and the $l+1$-th neuron in layer l. This layer is mainly to complete the classification of traffic accident data.

In summary, the model structure flow chart of traffic accident prediction based on GK-CWT-DCCNN-BiLSTM is shown in Fig.6.

III. DATA ANALYSIS

A. Data source

The data used in this article is kaggle's public dataset of traffic accidents in 49 states of the United States from 2016 to 2019 [34]. The factors that lead to traffic accidents in the

TABLE II
DATA DISPLAY OF ENTITY-RELATIONSHIP-ATTRIBUTE IN TRAFFIC ACCIDENTS

Amenity is an Entity		accident is an Entity		Weather is an Entity	
Amenity	FALSE	Airport_Code	KFFO	Humidity(%)	100
Astronomical_Twilight	Day	City	Dayton	Pressure(in)	29.63
Bump	FALSE	Country	US	Temperature(F)	33.8
Civil_Twilight	Day	County	Montgomery	Visibility(mi)	3
Crossing	FALSE	Description	Accident on Wayne Ave at glencoe Ave.Expect delays	Weather_Condition	Overcast
Give_Way	FALSE	Distance	0.01	Wind_Direction	SW
Junction	FALSE	End_Time	2016/2/8 9:13	Wind_Speed(mph)	2.3
No_Exit	FALSE	Number	100.0	weather	C-13
Railway	FALSE	Severity	2	weather_Timestamp	2016/2/8 8:28
Roundabout	FALSE	Side	R		
Station	FALSE	Source	MapQuest		
Stop	FALSE	Start_Lat	39.745888		
Sunrice_Sunset	FALSE	Start_Lng	-84.17041		
Traffic_Calmingg	FALSE	Start_Time	2016/2/8 8:43		
Traffic_Signal	FALSE	State	OH		
Traffic_Loop	FALSE	Street	Glencoe Ave		
		TMC	201		
		Timezone	US.Eastern		
		Zipcode	45410-1721		
		accident	A-16		

United States are divided into environment, landmarks, time, streets and points of interest. The environment includes temperature, wind chill, humidity, visibility, wind direction, wind speed, precipitation and weather conditions. Landmarks include starting point dimension, starting point longitude, end point dimension and end point longitude. Time includes start time, end time, day and night. The street includes name, number, opposite side and street postcode. Points of interest include amenities, deceleration points, intersections, junctions, railways, roundabouts, car shows and stop signs. The whole data classification is shown in Table I.

B. Knowledge Graph Visualization

The knowledge graph constructed by K = { R, S, A } is mainly divided into three categories : entity, relationship and attribute. Entities are mainly divided into three categories, namely Amenity, accident and Weather, which contain a total of 3145725 entities. The relationship is mainly divided into two categories, namely accident_a and accident_b, a total of 2097148 relationships. Attributes includ Start_Time Airport_Code, Astronomical_Twilight, Bump, City, Civil_Twilight, County, Crossing, Humidity, State and Wind_chill. The knowledge graph depicted by triples is shown in Fig.7. Due to the variety, the data of the three main entities are shown in Table II.

C. Comparative Analysis

The model based on KG-CWT-RGCNN-BiLSTM mainly considers data denoising, graph feature extraction and time series extraction. In order to verify the performance of the model, KG-CWT-RGCNN-BiLSTM is compared with KG-RGCNN-BiLSTM, KG-CWT-CNN-BiLSTM is compared with KG-CWT-RGCNN-BiLSTM, and KG-CWT-RGCNN-LSTM is compared with KG-CWT-RGCNN-BiLSTM. The performance of the model is evaluated by three values : accuracy, recall and F1-score. The experimental results are shown in Table III.

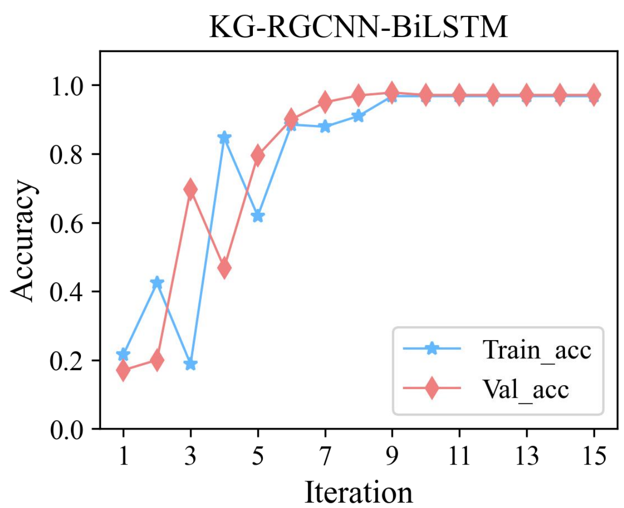
The accuracy of the models based on KG-RGCNN-BiLSTM and KG-CWT-RGCNN-BiLSTM is shown in Fig.8, where (a) represents the accuracy of the model KG-RGCNN-BiLSTM and (b) represents the accuracy of the model KG-CWT-RGCNN-BiLSTM.

As shown in Fig.8, the accuracy value of (a) is 0.9711, and the accuracy value of (b) is 0.9932. Because the preliminary traffic accident data contains noise, (b) the preliminary traffic accident data contains noise, (b) the model first denoises and then characterizes the knowledge graph. Combined with Table III, (b) the overall performance

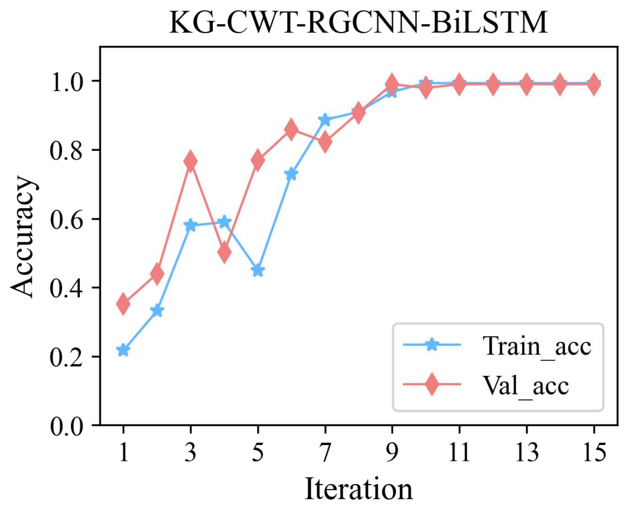
TABLE III
MODEL COMPARISON EXPERIMENTAL RESULTS

Model	Precision	Recall	F1-score
KG-RGCNN-BiLSTM	0.9711	0.9589	0.9565
KG-CWT-RGCNN-BiLSTM	0.9932	0.9872	0.9914
KG-CWT-CNN-BiLSTM	0.9735	0.9604	0.9702
KG-CWT-RGCNN-BiLSTM	0.9932	0.9872	0.9914
KG-CWT-RGCNN-LSTM	0.9781	0.9632	0.9765
KG-CWT-RGCNN-BiLSTM	0.9932	0.9872	0.9914

of the model is higher than that without noise reduction, and the three performance of the evaluation model is also higher. That is, before data analysis, it is necessary to reduce noise and reduce certain errors. Therefore, compared with KG-RGCNN-BiLSTM and KG-CWT-RGCNN-BiLSTM,

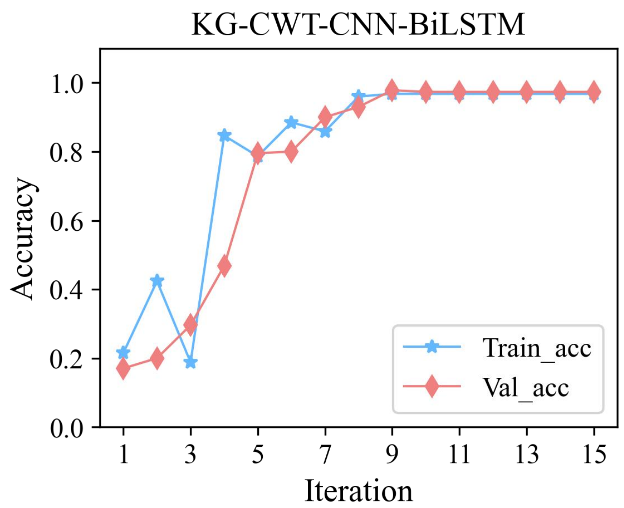


(a)

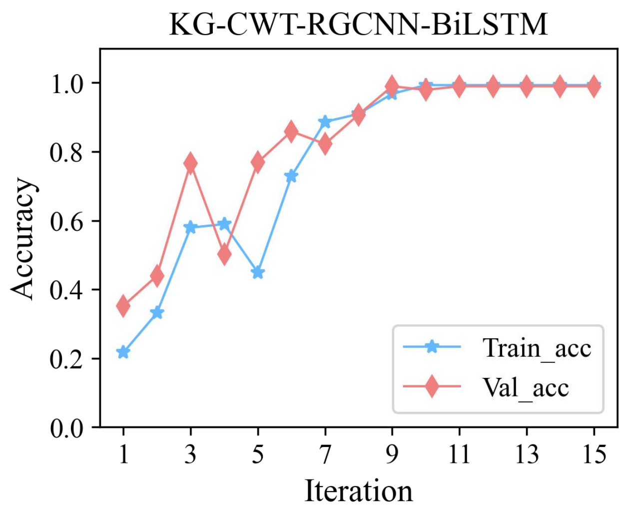


(b)

Fig.8. Model accuracy based on KG-RGCNN-BiLSTM and KG-CWT-RGCNN-BiLSTM



(a)



(b)

Fig.9. Model accuracy based on KG-CWT-CNN-BiLSTM and KG-CWT-RGCNN-BiLSTM

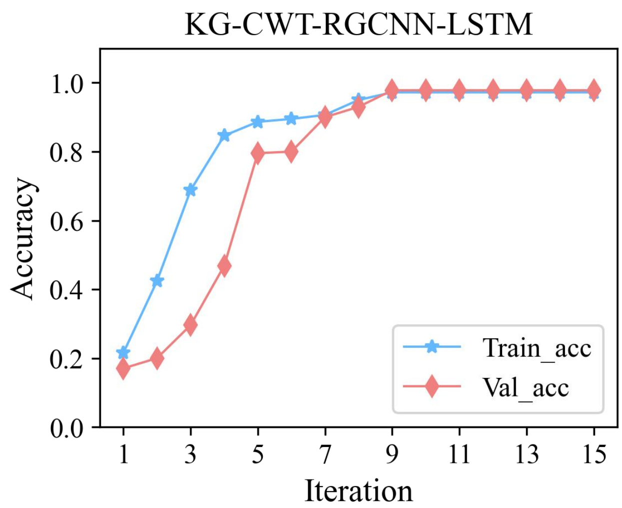
the model of KG-CWT-RGCNN-BiLSTM has high accuracy.

The accuracy of the model based on KG-CWT-CNN-BiLSTM and KG-CWT-RGCNN-BiLSTM is shown in Fig.8, where (a) represents the accuracy of the model KG-CWT-CNN-BiLSTM, (b) represents the accuracy of the model KG-CWT-RGCNN-BiLSTM.

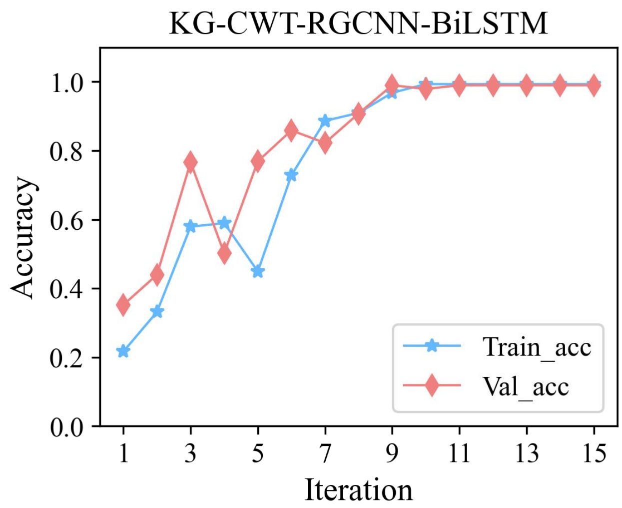
As shown in Fig.9, the accuracy value of (a) is 0.9735, and the accuracy value of (b) is 0.9932. CNN can automatically extract the features contained in the data, and use the knowledge graph to describe the traffic accident data.The generated knowledge graph input into CNN can also automatically extract features. Due to the large number, GCNN has better feature extraction performance for this topological structure, and GCNN is more stable than CNN for graph feature extraction. Therefore, compared with KG-CWT-CNN-BiLSTM and KG-CWT-RGCNN-BiLSTM, the KG-CWT-RGCNN-BiLSTM model has high accuracy.

The accuracy of the model based on KG-CWT-RGCNN-LSTM and the model based on KG-CWT-RGCNN-BiLSTM is shown in Fig.10, where (a) represents the accuracy of the model KG-CWT-RGCNN-LSTM and (b) represents the accuracy of the model KG-CWT-RGCNN-BiLSTM.

As shown in Fig.10, the accuracy value of (a) is 0.9781, and the accuracy value of (b) is 0.9932. The extraction of time features by LSTM is better than that of general RNN, and its ability to classify data as a classifier is higher.



(a)



(b)

Fig.10. Model accuracy based on KG-CWT-RGCNN-BiLSTM and KG-CWT-RGCNN-BiLSTM

However, LSTM cannot predict traffic accident data forward and backward at the same time, while BiLSTM can predict forward and backward at the same time, and then stitch when data output, that is, the extraction of time series features by BiLSTM is more accurate than LSTM.Combined with Table III, the recall rate of BiLSTM and the value of F1-score are higher than LSTM. So, compared with KG-CWT-RGCNN-LSTM and KG-CWT-RGCNN-BiLSTM, the accuracy of KG-CWT-RGCNN-BiLSTM model is high.

In summary, from the comparison of data noise reduction, graph feature extraction and time series feature extraction, the accuracy, recall and F1-score values of the KG-CWT-RGCNN-BiLSTM model are higher than other models. The model performance is better than other models, so the model is used to predict traffic accidents.

D. Parameter analysis

1) The number of convolution kernels

The size of the convolution kernel is an important parameter in the whole model, which has a great correlation with the accuracy of the model. The size of the convolution kernel is set to 8,16,32,64 and 128 to evaluate the influence of the model accuracy. The effect of convolution kernel size on model accuracy is shown in Fig.11.

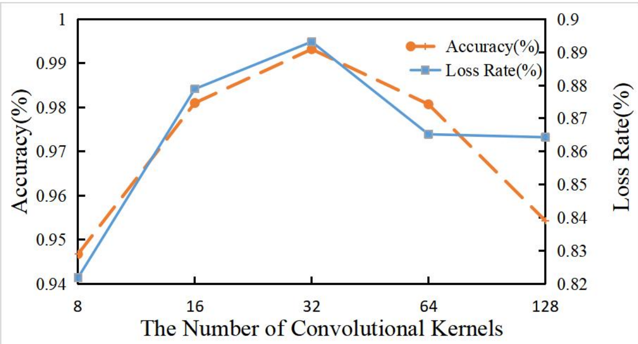


Fig.11 Effect of convolution kernel size on model accuracy

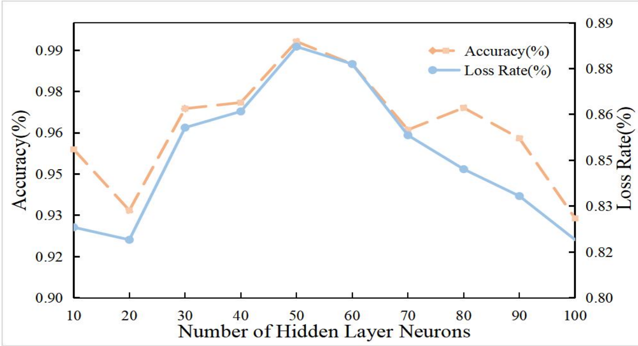
As shown in Fig.11, when the number of convolution kernels is 32, the model accuracy is higher than other parameters. The convolution kernel is mainly to increase or decrease the original data volume. When the number of convolution kernels is small, it will lead to a decrease in the classification degree and evaluation degree of the traffic accident prediction model. When the number of convolution kernels is large, it will lead to over-fitting of the traffic accident prediction model. Therefore, when the number of convolution kernels is set to 32, the model reaches the optimal value.

When the number of convolution kernels is relatively small, the accuracy of fault classification and degree evaluation is relatively low, and the overall diagnosis result of the model is not ideal. When the number of convolution kernels is too large, the two tasks appear serious over-fitting phenomenon at the same time, and the accuracy of model test decreases sharply.

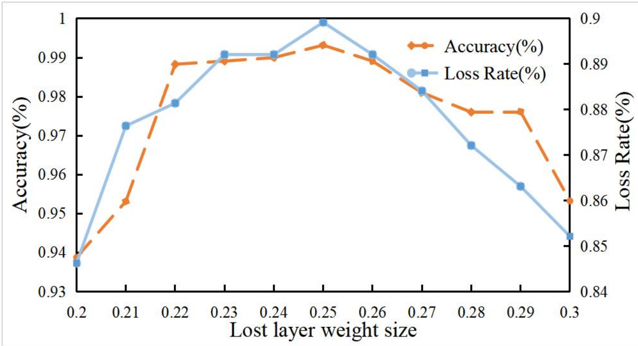
2) The number of hidden layer neurons in BiLSTM

In the BiLSTM layer, the number of hidden layer neurons and the weight of the lost layer will affect the accuracy of the model. The value of the neurons is evaluated from 10-100, and the weight of the lost layer is evaluated from 0.2-0.3. The influence of the number of hidden layer neurons and the weight of the lost layer on the accuracy is shown in Fig.12.

As shown in Fig.12, (a) represents the influence of the number of hidden layer neurons on the accuracy of the model, and (b) represents the influence of the number of hidden layer neurons on the accuracy of the model. Fig.12 shows that the greater the number of neurons is not, the better the prediction effect is. The value of the neuron is set to 50, and the weight of the lost layer is set to 0.25.



(a)



(b)

Fig.12 The influence of the number of hidden layer neurons and the weight of lost layer on the accuracy.

In summary, the comparative analysis shows that the model performance based on KG-CWT-RGCNN-BiLSTM is better than other models, and the model accuracy is optimal through parameter analysis. The main network parameters of the whole model are shown in Table IV.

TABLE IV
THE MAIN NETWORK PARAMETERS OF THE MODEL

Convolution2dLayer	32	Learningrate	0.01
BiLSTMlayer	50	LearnRateDropPeriod	20
DroupoutLayer	0.25	LearnRateDropFactor	0.8

IV. CONCLUSION

In order to predict massive traffic accidents, combined with the characteristics of KG, CWT, RGCNN and BiLSTM, this paper constructs a prediction method based on KG-CWT-RGCNN-BiLSTM model. The model considers traffic accident prediction comprehensively, and the accuracy rate reaches 99.32 %. After comparative analysis, the performance is higher than other networks. The model has the following characteristics.

1) Using KG theory to analyze the correlation of traffic accident data, the generated topology can more clearly identify the factors leading to the accident. The CWT theory is used to denoise the data, which solves the noise problem

contained in the manually acquired data.Using RGCNN theory to extract data features is a further improvement of CNN. BiLSTM theory is used to extract the characteristics of data time series and complete the classification and prediction, which improves the LSTM method that can not propagate forward and backward at the same time. Finally, a model based on KG-CWT-RGCNN-BiLSTM was constructed.

2) By comparing the model of KG-CWT-RGCNN-BiLSTM with the model of KG-RGCNN-BiLSTM, the model of KG-CWT-CNN-BiLSTM with the model of KG-CWT-RGCNN-BiLSTM and the model of KG-CWT-RGCNN-LSTM with the model of KG-CWT-RGCNN-BiLSTM, the three factors of noise reduction, RGCNN feature extraction and BiLSTM feature extraction are analyzed respectively. The model evaluation contains three criteria : accuracy rate, recall rate and F1-score. The results show that the model based on KG-CWT-RGCNN-BiLSTM, which deals with the above three factors at the same time, has the highest value among the three evaluation values.

3) Determine the model based on KG-CWT-RGCNN-BiLSTM to predict traffic accidents, and analyze the parameters contained in the model. The results show that when the number of convolution kernels is 32, the number of BiLSTM hidden layer neurons is 50, the weight of loss layer is 0.25 and the learning rate is set to 0.01, the performance of the model is the best, and there is no over-fitting phenomenon, and the accuracy of prediction results is higher.

However, because the data analyzed only contains road and environmental factors, and lacks human and vehicle-related factors, the knowledge graph constructed has certain limitations and low universality. The constructed model has many levels. Although the model has high accuracy, the running speed is slow, and further research is needed.

References

[1] A. García-Ferrer, A. de Juan and P. Poncela, “Forecasting traffic accidents using disaggregated data,” *International Journal of Forecasting*, vol.22, no.2, pp.203-222, 2006.

[2] Mehdi Hosseinpour, Ahmad Shukri Yahaya, Seyed Mohammadreza Ghadiri, Joewono Prasetijo, “Application of Adaptive Neuro-Fuzzy Inference System for Road Accident Prediction,” *KSCE journal of civil engineering*, vol.17, no.7, pp.1761-1772, 2013.

[3] Markus Deublein, Matthias Schubert, Bryan T. Adey, “Prediction of road accidents: comparison of two Bayesian methods,” *Structure and Infrastructure Engineering*, vol.10.no.11, pp.1394-1416, 2014.

[4] W. Li, X.J. Zhao and S.Y. Liu, “Traffic Accident Prediction Based on Multivariable Grey Model,” *Information*, vol.11, no.4, pp.184-184, 2020.

[5] X.H. Jin, J.Y. Zheng and X.Y. Geng, “Prediction of Road Traffic Accidents Based on Grey System Theory and Grey Markov Model,” *International Journal of Safety and Security Engineering*, vol.10, no.2, pp.263-268, 2020.

[6] Z.Q. Yuan, D. Huang, W.P. Tong, Z.Y. Liu, “Characteristic Analysis and Prediction of Traffic Accidents in the Multiethnic Plateau Mountain Area,” *Journal of Transportation Engineering, Part A: Systems*, vol.146,no.8, pp.04020068-04020068, 2020.

[7] Li Haitao, Bai Qiaowen, Zhao Yonghua, Qu Zhaowei, Xin Wang, “TSDCN: Traffic safety state deep clustering network for real‐time traffic crash‐prediction,” *IET Intelligent Transport Systems*, vol.15, no.1, pp.132-146, 2020.

[8] OspinaMateus Holman, Quintana Jiménez Leonardo Augusto, LopezValdes Francisco J., Sana Shib Sankar, “Prediction of motorcyclist traffic crashes in Cartagena (Colombia): development of a safety performance function,” *RAIRO - Operations Research*, vol.55, no.3, pp.1257-1278, 2021.

[9] Muhammad Babar Ali Rabbani, Muhammad Ali Musarat, Wesam Salah Alaloul, Muhammad Shoaib Rabbani, Ahsen Maqsoom,Saba Ayub, Hamna Bukhari, et al, “A Comparison Between Seasonal Autoregressive Integrated Moving Average (SARIMA) and Exponential Smoothing (ES) Based on Time Series Model for Forecasting Road Accidents,” *Arabian Journal for Science and Engineering*, vol.46, no.11, pp.1-26, 2021.

[10] Deretić Nemanja, Stanimirović Dragan, Awadh Mohammed Al, Vujanović Nikola, Djukić Aleksandar, “SARIMA Modelling Approach for Forecasting of Traffic Accidents,” *Sustainability*, vol.14,no.8, pp.4403-4403, 2022.

[11] Marcillo Pablo, Valdivieso Caraguay Ángel Leonardo and HernándezÁlvarez Myriam, “A Systematic Literature Review of Learning-Based Traffic Accident Prediction Models Based on Heterogeneous Sources,” *Applied Sciences*, vol.12, no.9, pp.4529-4529, 2022.

[12] Mamoudou Sangare, Sharut Gupta, Samia Bouzefrane, Soumya Banerjee, Paul Muhlethaler, “Exploring the forecasting approach for road accidents: Analytical measures with hybrid machine learning,” *Expert Systems With Applications*, vol.167, pp.113855, 2020.

[13] Alqatawna Ali, Rivas Álvarez Ana María and García-Moreno Santos Sánchez-Cambronero, “Comparison of Multivariate Regression Models and Artificial Neural Networks for Prediction Highway Traffic Accidents in Spain: A Case Study,” *Transportation Research Procedia*, vol.58, pp.277-284, 2021.

[14] Wakatsuki Yuki, Tatebe Jumpei and Xing Jian, “Improving the Accuracy of Traffic Accident Prediction Models on Expressways by Considering Additional Information,” *International Journal of Intelligent Transportation Systems Research*, vol.20,no.1, pp.309-319, 2022.

[15] Z. Tian and S.R, Zhang, “Deep learning method for traffic accident prediction security,” *Soft Computing*, vol.26, no.11, pp.5363-5375, 2022.

[16] Lida Barba, Nibaldo Rodríguez and Cecilia Montt, “Smoothing Strategies Combined with ARIMA and Neural Networks to Improve the Forecasting of Traffic Accidents,” *The Scientific World Journal*,vol.2014, pp.152375, 2014.

[17] Gyanendra Singh, Mahesh Pal, Yogender Yadav, Tushar Singla, “Deep neural network-based predictive modeling of road accidents,” *Neural Computing and Applications*, vol.32,no.16, pp.1-10, 2020.

[18] Li Duowei, Wu Jianping and Peng Depin, “Online Traffic Accident Spatial-Temporal Post-Impact Prediction Model on Highways Based on Spiking Neural Networks,” *Journal of Advanced Transportation*, vol.2021, pp.1-20, 2021.

[19] X.R. Sun, Y. Meng and W.L. Wang, “Identifying Traffic Events from Weibo with Knowledge Graph and Target Detection,” *Data Analysis and Knowledge Discovery*, vol.4, no.12, pp.136-147, 2020.

[20] Q.Q. Qiu, “The Analysis and Application of Urban Traffic Trip Data Based on Knowledge Graph,” *North China University of Technology*, 2021. DOI:10.26926/d.cnki.gbfgu.2021.000079.

[21] L.Y. Zhang, M. Zhang, J.Z. Tang, J. Ma, X.K. Duan, J. Sun, et al, “Analysis of Traffic Accident Based on Knowledge Graph,” *Journal of Advanced Transportation*, vol.2022, Article ID.3915467, 2022.

[22] R.H. Liu and X.H. Wei, “Application of Improved Convolutional Neural Network in Text Classification,” *IAENG International Journal of Computer Science*,vol.49,no.3, pp.762-767, 2022.

[23] Z.H. Wu, S.R. Pan,F.W. Chen,G.D. Long, C.Q. Zhang, Philip S. Yu. “A Comprehensive Survey on Graph Neural Networks.,” *CoRR*, vol.abs/1901.00596, vol.32,no.1, pp.4-24, 2019.

[24] G.Y. Zhu, M. Zhang and Y. Yi, “Prediction of Evolution Results of Urban Rail Transit Emergencies Based on Knowledge Graph,” *Journal of Electronics & Information Technology*, vol45.no.03, pp.949-957, 2023.

[25] L. Yu, B.W Du, X. Hu, L.L. Sun, L.Z. Han, W.F. Lv, “Deep Spatio-Temporal Graph Convolutional Network for Traffic Accident Prediction,” *Neurocomputing*, vol.423, pp. 135-147., 2020.

[26] Z.H. Zhang, W.Z. Yang and Wushour Silamu, “Traffic Accident Prediction Based on LSTM-GBRT Model,”. *Journal of Control Science and Engineering*, vol.2020, pp.1-10, 2020.

[27] Z.H. Zhang, W.Z. Yang , T.T Yuan, D.H. Li, X.Y. Wang, “Traffic Accident Prediction Based on LSTM Neural Network Model”, *Computer Engineering and Applications*, vol.55,no.14, pp.249-253+259, 2019.

[28] J.W. Liu and Z.Y. Song, “Overview of recurrent neural networks,” *Control and Decision*, vol.37, no.11, pp.2753-2768, 2022.

[29] B.Y. Chen, “Research on Forecast of Road Traffic Accidents Based on Bi-LSTM,” *Chang’an University*, 2021. DOI:10.26976/d.cnki. gchau. 2021.001367.
[30] X.P. Luo, X. Xu and Herschel Rabitz, “On the fundamental conjecture of HDMR: a Fourier analysis approach,” *Journal of Mathematical Chemistry*, vol.55,no.2, 2017.
[31] H.H. Pan, P. Gao, S.L. Zhou, D. Xie, Prediction of Traffic Based on Wavelet Neural Network,” *Journal of Luoyang Institute of Science and Technology(Social Science Edition)*, vol.28, no.04, pp.53-57, 2018.
[32] J.P. Zhao and J.L. Ding, “Road Accidents Prediction based on Wavelet Analysis of the Grey GM(1,1) Model,” *Mathematics in Practice and Theory*,vol.45,no.12, pp.119-124, 2015.
[33] H. Yi, “The Ferquency Domain Algorithm of Continuous Wavelet Transform Based on the Discretization of Fourier Transform,”*Journal of Jinggangshan University (Natural Sciences Edition)*, vol.40.no.03, pp.1-8, 2019.
[34] https://www.kaggle.com/datasets/sobhanmoosavi/us-accidents

Changfeng Zhu was born in Gansu, China , in 1972. In 2003 and 2013, he obtained master's and doctor's degrees in the Planning and Management of Traffic and Transportation in Lanzhou Jiaotong University. He is now a professor and doctoral supervisor of the School of Traffic and Transportation of Lanzhou Jiaotong University. His main research direction is transportation planning and management. He has made fruitful research achievements, presided over a number of national and local level fund projects, and published more than 50 papers that were retrieved by databases EI and SCI.