IEEE Access
Multidisciplinary : Rapid Review : Open Access Journal

Received 8 September 2023, accepted 3 November 2023, date of publication 7 November 2023, date of current version 15 November 2023.

Digital Object Identifier 10.1109/ACCESS.2023.3330909

RESEARCH ARTICLE

Research on Regional Traffic Flow Prediction Based on MGCN-WOALSTM

KERANG CAO[1], YARU LIU[1], LINI DUAN[2], SEN XU[1], AND HOEKYUNG JUNG[3]
[1]Key Laboratory of Intelligent Technology of Chemical Process Industry in Liaoning Province, College of Computer Science and Technology, Shenyang University of Chemical Technology, Shenyang 110142, China
[2]Department of Big Data Management and Application, Shenyang University of Chemical Technology, Shenyang 110142, China
[3]Department of Computer Engineering, Pai Chai University, gwangyeoksi 35345, South Korea

Corresponding author: Hoekyung Jung (hkjung@pcu.ac.kr)

This work was supported by the Ministry of Science and Information and Communications Technology (MSIT), South Korea, through the Innovative Human Resource Development for Local Intellectualization Support Program, supervised by the Institute for Information and Communications Technology Planning and Evaluation (IITP), under Grant IITP-2023-RS-2022-00156334.

ABSTRACT Regional traffic flow forecasting is the key to the realization of intelligent transportation system. The existing traffic flow forecasting methods have problems such as insufficient Spatio-temporal Correlation modeling and ignoring the impact of weather factors, which lead to high prediction errors. Therefore, in this study, a Multi-channel Graph Convolutional Neural network (MGCN) was first established to analyze and express the spatial correlation of traffic flow on different dimensions, and a self-attention mechanism was used to weight the spatial correlation features of MGCN output. Next, an LSTM is built after the MGCN network layer to obtain temporal features, an Embedding layer is added to embed traffic flow temporal periodic features, and the Whale Optimization Algorithm (WOA) is introduced to find the global optimal LSTM network parameter combination, which is then applied to the prediction model. The performance of the model was tested using the public dataset PeMSD4 and corresponding weather data. Compared with prediction models that were not optimized by WOA and did not consider the influence of weather factors (GCN, LSTM, ASTGCN, etc.), the prediction errors RMSE, MAE, and MAPE of the final constructed prediction model were reduced, indicating that the MGCN-WOALSTM model has better traffic flow prediction performance.

INDEX TERMS MGCN, regional traffic flow prediction, LSTM, WOA, spatio-temporal correlation.

I. INTRODUCTION

With the rapid development of technology, people's economic level has reached new heights and the number of private cars has increased dramatically. This in turn has led to increasingly serious traffic congestion and frequent traffic accidents, which have a great impact on residents' living standards and road safety. Therefore the optimization of intelligent transportation systems [1] has become more and more important. Efficient prediction algorithms for more accurate traffic flow prediction are also indispensable for the realization of intelligent transportation systems. The goal of traffic forecasting is to estimate traffic conditions at future moments based on historical traffic flow data, which can balance traffic resources and reduce traffic congestion rates. It also plays an important role in traffic applications such as path optimization and traffic light control. Traffic flow data is not only random and uncertain due to weather and traffic accidents, but also has complex spatio-temporal Correlation. Traffic flow prediction should consider both the spatial correlation of traffic flow and the temporal correlation of the time series itself, while current prediction models are difficult to fully extract its spatio-temporal characteristics. It is of strong research significance how to fully extract the spatio-temporal characteristic relationships of traffic flow from the traffic road network topology and the temporal dimension, and construct an accurate and efficient regional traffic flow prediction model in real time [2].

Traffic flow prediction has always been a popular research problem. It started with time series analysis models to

The associate editor coordinating the review of this manuscript and approving it for publication was N. Ramesh Babu.

predict traffic flow information in future time steps, such as Kalman Filter Model (KFM), Autoregressive Integrated Moving Average (ARIMA) model, etc. Lu et al. [3] proposed a combined short-term traffic flow prediction method based on ARIMA and LSTM neural network to predict future traffic flow based on historical traffic flow data. Then comes the machine learning based prediction model [4] which is heavily used nowadays. Yang [5] proposed a traffic flow uncertainty prediction method based on KNN and Kalman filter to solve the problem of low fit between actual and predicted values. Ma et al. [6] proposed a GS-SVM prediction model, which models the mapping relationship between the traffic flow and spatio-temporal information by supporting vector machine algorithm, and uses grid search algorithm to dynamically optimize the SVM parameters. The above machine learning methods are less suitable for regions with complex traffic road network structures, and it is difficult to fully obtain the spatial correlation of traffic flow data on different roads.

Deep learning methods can tap deeper levels of complex spatio-temporal features of traffic data. Therefore, many deep learning models have been proposed and used to solve traffic flow prediction problems. Shu et al. [7] proposed to study the time series of traffic flows using an improved Gated Recurrent Unit (GRU) with bidirectional positive and negative feedback, and optimize the learning rate using a modified adaptive (RADAM) model to improve the accuracy of model prediction. Chen et al. [8] construct a dynamic spatio-temporal graph-based DST-GCNN dual-flow network to predict the dynamic graph structure using graph prediction flow and input the predicted structure into a traffic prediction flow consisting of a stack of graph-based spatio-temporal convolutional layers to predict traffic flows at future times. Cheng et al. [9] proposed to evaluate the intrinsic correlation between traffic variables using an econometric theory-based VAR model, then a CNN-LSTM hybrid neural network based on deep learning was used for multi-feature short-term traffic flow prediction. Zhang et al. [10] proposed a gated adaptive graph convolutional network (MSTA-GCN) model, which introduced a multi-headed spatio-temporal attention mechanism to capture the dynamic spatio-temporal correlation of traffic flows at different moments and in different spaces. An et al. [11] proposed a spatio-temporal graph convolutional network model (IGAGCN) for urban road network traffic flow prediction by fully considering the combination of information geometry method and attention mechanism to capture the spatial dependence of traffic flow. Zhang et al. [12] captured the spatial correlation of traffic road networks by GCN and introduced the Soft-Attention mechanism, based on which they established a traffic flow prediction model GRU-GCN-Soft Attention (GGCN-SA). Chen et al. [13] proposed a parallel structured deep learning model consisting of a GCN and a stacked bi-directional LSTM network (GCN-SBULSTM). In addition, severe weather conditions are an important factor affecting road traffic safety and traffic congestion status. The analysis of the correlation between traffic flow and weather factors at different frequencies and time intervals can help to select the most important weather factors affecting the magnitude of traffic flow [14]. Yang et al. [15] studied by example the severe weather related traffic safety and traffic flow problems, and the results showed that severe weather events such as rainfall and fog can affect traffic flow to some extent. Hou et al. [16] completed traffic flow prediction based on deep learning algorithms and data fusion, considering the impact of weather changes on traffic flow, using Stacked Auto-Encoder (SAE) and Radial Basis Function (RBF) to process traffic flow data and capture the correlation between weather perturbations and traffic flow periodicity. Miao et al. [17] proposed a Queue Hybrid Neural Network (QHNN) model based on LSTM and GRU to extract traffic flow features while taking weather weighting factors into account to predict traffic flow.

In summary, the existing traffic flow prediction models are still inadequate in extracting the spatio-temporal correlation of traffic flow data, and it is difficult to fully obtain the tighter spatial correlation features in the traffic road network topology. There is room for improvement in the extraction of time-dependent relationships. To address the problem of temporal traffic flow prediction with multiple traffic flow influencing factors and to take into account the complex topology of the traffic road network, this paper proposes the MGCN-WOALSTM traffic flow prediction model based on the whale optimization algorithm. The model adopts the construction of multi-channel GCN layers to capture the spatial correlation of traffic flow, and introduces a self-attention mechanism in the multi-channel GCN layers to further explore the dependencies of different road nodes in spatial locations. We build the LSTM neural network on the GCN layer of each channel for learning the dynamic changes in the temporal dimension of traffic data to obtain the temporal correlation. Then the WOA is used to optimize the LSTM network hyperparameters, the temporal periodic features are added to the LSTM in the form of embedding layer vectors, and the spatio-temporal feature vectors are stitched by the fully connected layer as the basis for prediction. Finally, on the basis of considering weather factors and using WOA to optimize network hyper-parameter, through testing on the public data set PeMSD4, and using RMSE, MAE, MAPE to evaluate model performance, it is verified that MGCN-WOALSTM model achieves the best prediction effect.

II. RELATED WORK

A. GCN: GRAPH CONVOLUTIONAL NETWORK

Since its innovation, GCN has been used in many applications, such as node classification, connection prediction, whole graph classification, etc. It is usually divided into Spectral domain graph convolution (converting data from the spatial domain to the spectral domain for processing according to the graph theory and convolution theorem) and Spatial graph convolution (defining convolution operations directly on the space without relying on the graph convolution theory).

GCN can use both structural information and node features to accomplish classification (regression) tasks. According to previous studies, it is found that GCN has superior performance when 2-3 neural network layers are set up. The GCN can be regarded as the same as the Convolutional Neural Network (CNN) to some extent, which is a feature extractor in practical applications. The difference between the two above mentioned is that the objects of action are different. The object of action in GCN is graph data, which can extract feature information of topological data more comprehensively. It makes up for the shortcoming of CNN, in which the object of action is spatial features that cannot be exacted adequately. The regional traffic road network is essentially a topological structure, which can be modeled perfectly by graph. Therefore, GCN can be well applied to the field of intelligent transportation.

Suppose now there is a weighted graph G = (V, A, X): Where V is the vertex set, consisting of a total of N nodes; A is the adjacency matrix (0-1), representing the relationship between the nodes; X is defined as a matrix of size m∗|V|, representing the node features, and each node has m-dimensional features.

GCN mathematical representation:

$$h_v^0 = x_v \tag{1}$$

$$h_v^k = \sigma(W_k \sum\nolimits_{u \in N(v)} \frac{h_u^{k-1}}{|N(v)|} + B_k h_v^{k-1}), \quad \forall k \in \{1, \ldots, K\} \tag{2}$$

$$z_v = h_v^K \tag{3}$$

GCN vector representation:

$$H^{(l+1)} = \sigma(H^{(l)} W_0^{(l)} + \tilde{A} H^{(l)} W_1^{(l)}) \tag{4}$$

The key idea of GCN is to generate the embedding representation of the current node based on the neighbor nodes in the graph structure. Each node has an independent computational graph based on the neighbor nodes, and the nodes use a neural network based approach to aggregate the information of the neighbors. The basic method of GCN aggregation is to average the information from the neighbors and apply it to the neural network layer. The GCN model can be of any depth, and nodes have embedding representation at each layer. The embedding representation at layer 0 is the node feature X, and the embedding representation of the nodes at layer K is computed by aggregating the node information of the neighbors at layer K-1. The nodes in the GCN model share common learnable parameters W and B at each layer, and because of the shared parameters, the model can cope with new nodes added to the graph data, which is not possible with traditional graph machine learning methods. However, it is not enough for the GCN to accomplish temporal data prediction by simply embedding the feature information of neighboring nodes in non-Euclidean space. When the number of layers of the graph convolutional network increases to a certain number, the feature vector expression of the GCN will weaken as the number of layers increases. Moreover, the node feature quantity will tend to converge to a certain value, and then the phenomenon of over-smoothing will occur.

B. ATTENTION MECHANISM

The attention mechanism also has many applications in the field of traffic flow prediction. The core idea of the attention mechanism is to globally scan the input information, and obtain the information that is more critical to the current task goal from a large amount of information. It adaptively captures the feature with the highest relevance, assigns it the highest weight, and assigns different weights to the parts with different degrees of relevance in the model training process. The attention mechanism operation process can be described as follows: Specifically, the attention matrix is calculated based on the query matrix Q and the key matrix K; the attention matrix is calculated using the attention scoring function; the relevance of the two is calculated based on the Query and a certain Key; different functions and computational mechanisms can be introduced to obtain the attention score, and the higher the score the higher the relevance. This can be done by finding the vector dot product of the two, introducing a neural network to find the value, etc. Query is an artificially defined hyper-parameter, which can be a dynamically generated vector or a vector of learnable model parameters. Then it is processed by the SoftMax normalization function to obtain the attention probability distribution vector, which gives the weight coefficients between 0 and 1 for all keys. Finally, the input matrix is multiplied with the probability distribution to obtain the attention weight matrix.

Firstly, we define the attention correlation matrices Q, K and V. Let N denotes the feature dimension of the input information, and the specific operation of the weighted attention matrix Att is shown in equation (5):

$$\begin{cases} Sim_i = Sim(Q, K) = Query \cdot Key \\ a_i = SoftMax(Sim_i) = \dfrac{e^{Sim_i}}{\sum\limits_{j=1}^{N} e^{Sim_j}} \\ Att(Q, K, V) = \sum\limits_{i=1}^{N} a_i \cdot V_i \end{cases} \tag{5}$$

C. LSTM: LONG SHORT TERM MEMORY

Ordinary Recurrent Neural Networks (RNN) are trying to remember all the information, no matter it is useful or not. LSTM usually works on sequential data. It is an improvement of RNN in sense of better handling the temporal dependencies of long sequences, and it can predict the time series variables more accurately. While in recurrent neural networks due to the problem of gradient explosion, or disappearance due to high powers of the matrix during the training of long sequences, in fact only short-period dependencies can be learned, which is called the long-range dependency problem.

The difference between LSTM and RNN is that LSTM is used to alleviate the gradient disappearance problem by controlling the information flow through three gating units (input gate, output gate and forgetting gate), and more importantly, LSTM has four times the number of parameters to control the model than RNN. The core idea is to selectively process the input information and design a memory cell with the same shape as the hidden state, which has the function of selective memory. It can choose to remember important information, filter out useless information, and reduce the memory burden. The input of LSTM gate is the current time step input X_t and the previous time step hidden state h_{t-1}. The output is calculated by the fully connected layer whose activation function is sigmoid function. And the forgetting gate is used to control which part of the previous state should be kept and which part should be forgotten.

There are three main stages within the LSTM. The first stage is the forgetting stage, in which f_t, is calculated to be used as the forgetting gating, as shown in Eq. (6):

$$f_t = \sigma(W_f X_t + U_f h_{t-1} + b_f) \tag{6}$$

The second stage is the selection memory stage, where the input information X_t is selected and memorized. Let i_t be the input gating, and the results obtained from the first and second stages are summed to obtain the c_t, which will be transmitted to the next state, calculated as follows:

$$i_t = \sigma(W_i X_t + U_i h_{t-1} + b_i) \tag{7}$$

$$\tilde{c}_t = \tanh(W_c X_t + U_c h_{t-1} + b_c) \tag{8}$$

$$c_t = f_t \times c_{t-1} + i_t + \tilde{c}_t \tag{9}$$

The third stage is the output stage, which is mainly controlled by o_t. And the results c_t of the previous stage are deflated by a tanh activation function to obtain the final output h_t, calculated as follows:

$$O_t = \sigma(W_O X_t + U_O h_{t-1} + b_O) \tag{10}$$

$$h_t = O_t \times \tanh(c_t) \tag{11}$$

where f_t, i_t, o_t denote the values of the forgetting gate, input gate, and output gate of the LSTM at the t time step, c_t and ct-1 respectively denote the cell states at the t and t-1 time step, W_f, W_i, W_c, W_O respectively represents the weight coefficient matrices to be trained for the forgetting gate, input gate, memory cell, and output gate, b_f, b_i, b_c, b_O respectively represents the bias parameters to be trained for the forgetting gate, input gate, memory cell, and output gate.

D. WHALE OPTIMIZATION ALGORITHM

The Whale Optimization Algorithm (WOA) is an intelligent optimization algorithm that simulates bubble net hunting behaviors of the humpback whale, and achieves the optimal solution through the process of searching, encircling, pursuing and attacking the prey by the whale population. It mainly updates the individual whale positions through the three stages of encircling the prey, bubble net hunting and searching for the prey, and has the advantages of simple operation and strong optimization-seeking ability with fewer parameters to be adjusted, etc. The WOA algorithm first initializes the whale population positions in the feasible solution. Each whale represents the potential Initial Optimal Solution of the optimization problem, and the position represents the characteristics of the whale. Each whale will search in the solution space according to certain rules, and the fitness function will be recalculated every time the whale moves to update the current optimal fitness value. When all whales have finished moving, the algorithm updates all whale positions and performs multiple iterations until the global optimal solution is found. The flow of the WOA is shown in Figure 1 below:

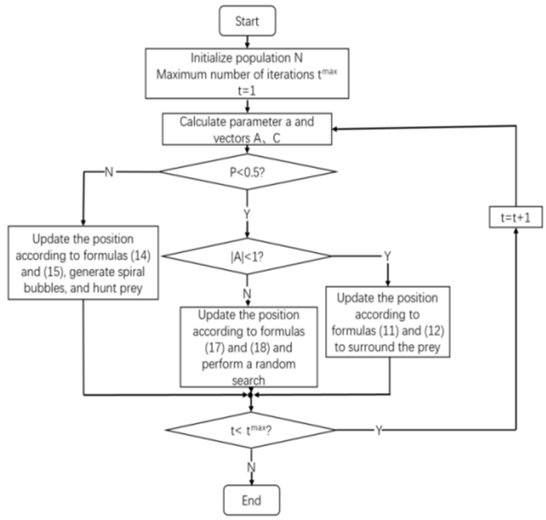


FIGURE 1. **Flow chart of whale optimization algorithm.**

Assume in a D-dimensional space, a population of N whales $X = (X_1, X_2, \ldots, X_N)$, where D denotes the number of variables in the solution of the optimization search problem. The WOA algorithm is implemented through three main stages:

1) SURROUNDING PREY

The WOA assumes that the current best candidate solution is the target prey position. After determining the target prey position, other whales will update their own position by encircling the prey, and this behavior can be expressed by the following equation:

$$D = |C \cdot X^*(t) - X(t)| \tag{12}$$

$$X(t+1) = X^*(t) - A \cdot D \tag{13}$$

where t denotes the current number of iterations, A and C are the coefficient vectors, X^* (t) are the position vectors of the current optimal solution, X (t) are the whale position vectors in the current number of iterations. Vector A and C are calculated as follows:

$$A = 2a \cdot r_1 - a; C = 2 \cdot r_2 \tag{14}$$

where the value of a decreases linearly from 2 to 0, and the r_1 and r_2 are random vectors in [0, 1].

2) BUBBLE NET PREDATION

The first contraction envelope mechanism is to calculate the distance between the current whale position and the target prey position, and the position update between the whale and the prey is expressed by the logarithmic spiral equation, as follows:

$$X(t+1) = D' \cdot e^{bl} \cdot \cos(2\pi l) + X^*(t) \tag{15}$$

$$D' = |X^*(t) - X(t)| \tag{16}$$

where D' denotes the distance between the individual whale and the target prey, and b is a constant, l is a random vector in [−1,1].

The WOA selects bubble net predation or contraction enclosure according to the probability p. Assuming that the probability of each of the two behaviors is 0.5 and p takes values in the range [0, 1], the mathematical formula is expressed as follows:

$$X(t+1) = \begin{cases} X^*(t) - A \cdot D, & p < 0.5 \\ D' \cdot e^{bl} \cdot \cos(2\pi l) + X^*(t), & p \geq 0.5 \end{cases} \tag{17}$$

Through the changes of probability p, parameter a and fluctuation range A, the whales will update their positions in different ways and move toward the more optimal position. As the number of iterations t increases, the parameter A and the convergence factor a gradually decrease. And if $|A| < 1$, each whale will gradually surround the current optimal solution, which belongs to the local optimal-seeking stage in WOA.

3) PREY SEARCH

WOA searches randomly according to the distance between each pairs of the whales, and updates the position for the purpose of finding the prey.

$$D = |C \cdot X_{rand}(t) - X(t)| \tag{18}$$

$$X(t+1) = X_{rand}(t) - A \cdot D \tag{19}$$

where X_{rand} (t) is a randomly selected whale position vector. When the absolute value of A is greater than 1, a whale is randomly selected and the whale population position is updated, forcing the whales to move away from the target prey, thereby exploring the global optimal prey position.

III. MGCN-WOALSTM TRAFFIC FLOW PREDICTION MODEL

A. TRAFFIC FLOW PREDICTION PROBLEM

Traffic flow prediction problem description: Define the undirected graph G = (V, E, A) to represent the complex topology of the traffic road network, where V represents the set of road network nodes and E is the set of edges. The adjacency matrix A represents the connection relationship between roads, which is obtained by adaptive learning. Additionally, let the feature matrix X represent the attribute features of road nodes in the road network, including traffic flow, average speed, rainfall, and snowfall. Thus the traffic flow prediction problem can be defined as a mapping function F, which represents the MGCN-WOALSTM model. The prediction could be obtained under a series of transformations of the traffic road network G and the traffic flow feature matrix X for the first T time intervals that are known. Let Y represent the predicted traffic flow, then the problem can be expressed as the following equation:

$$Y = F(G, (X_1, X_2, \ldots, X_T)) \tag{20}$$

Based on the above problem description, the steps for implementing the MGCN-WOALSTM traffic flow prediction model are as follows:

Step1: Select the region, obtain the distance between the target road section and other road sections in the road network region, obtain the historical traffic flow and speed data of other road sections in the road network region, plus the corresponding weather data set, divide the training set and the testing set, and pre-process the used data (data cleaning, normalization, etc.).

Step2: Set the initial structure and parameters of the multi-channel GCN model, pre-process the traffic flow data into the form of topological graph structure, and get the adjacency matrix by adaptive learning during the training process which will be used as the data input of the GCN model. Then the initialization parameters of the LSTM model are set, and the initial network structures such as input layer, hidden layer and output layer are established.

Step3: The global optimal hyperparameters of the LSTM network are optimized by using the WOA to derive the optimal solutions of the 2- to 4-layer LSTM network structure.

Step4: Construct a regional traffic flow prediction model based on MGCN-WOALSTM; determine the parameters such as the number of GCN layers and channels; use the Self-attention mechanism to weight the spatial correlation features of MGCN output, which will be used to calculate the degree of influence of different factors on the spatial feature vector; and use the weighted spatial feature vector as the input of the LSTM network. The WOA-optimized LSTM extracted traffic flow temporal correlation features and embedding layer vector-time-period features are spliced through the fully connected layer to obtain the final output predicted traffic flow results.

Step5: The MGCN-WOALSTM traffic flow prediction model is trained using the training set, and the trained model is tested with the test set.

Step6: Finally, the output data are back-normalized, and the prediction results are analyzed with the real values; the performance of the prediction model is evaluated with MAPE, RMSE, and MAE; and if the results are not satisfactory, the network parameters are re-adjusted back to Step2.

B. TRAFFIC FLOW PREDICTION MODEL MODULE

1) MULTI-CHANNEL GCN MODULE

GCN is mainly constructed by the adjacency matrix as well as the degree matrix, together with the feature matrix of the graph network data itself, which can extract the scale of the influence on the target node by other nodes, most notably by establishing the adjacency matrix (indicating the relationship between traffic road networks) and the graph feature data (traffic flow data converted into graph structure data). The adjacency matrix A of the traffic road network is usually determined by the location and number of sensors in the traffic dataset. But as the basis for GCN analysis of spatial dimensional features, this method is bound to lose some important spatial association information. To address this problem, this paper proposes to use a self-learning approach to obtain an adaptive road network adjacency matrix A^*, which is a shared parameter matrix in GCN, by training with traffic flow data.

This module is mainly used to express the spatial characteristics of traffic flow data in a more comprehensive analysis. The structure of the MGCN module is shown in Figure 2, where the input information is $X = (X_1, X_2, X_3 \ldots, X_T)$ and adjacency matrix $A*$, and X is the traffic flow feature data at T consecutive time steps. The module consists of two layers of GCN. The number of input and output channels in each layer are C_{11}, C_{12} and C_{21}, C_{22} respectively, where $C_{12} = C_{21}$. So the spatial feature vectors output from the two layers of GCN have C_{12} and C_{22} channels. The spatial features extracted from different channels in their respective dimensions are stitched together into a complete spatial feature vector, which is used as input information for the next stage of the WOALSTM module.

The first layer of GCN extracts the spatial correlation feature vector $G^{(1)} \in R^{V\times T\times C12}$ of traffic flow, which contains the GCN results of the C_{12} channels. It is shown in equation (21):

$$\begin{cases} G^{(l+1)} = f(G^{(l)}, A^*) \\ G^{(1)} = f(G^{(0)}, A^*) = \sigma(A^* X W_j^{(0)}) \end{cases} \tag{21}$$

where l represents the number of network layers, $G^{(l)}$ represents the features of the lth layer of the network, σ represents the ReLU activation function, and $W_j^{(0)}$, $j = \{1, 2, \ldots, C_{12}\}$ denotes the weight parameter matrix of the jth channel in the first layer GCN.

The second layer of GCN takes $G^{(1)}$ as input information and goes to analyze deeper spatial correlation based on the already obtained spatial feature vector to calculate new vector $G^{(2)} \in R^{V\times T\times C22}$, as shown in equation (22):

$$G^{(2)} = \mathrm{f}(G^{(1)}, A^*) = \sigma(A^* G^{(1)} W_Z^{(1)}) \tag{22}$$

where $W_Z^{(1)}$, $Z = \{1, 2, \ldots, C_{22}\}$ denotes the weight parameter matrix of the Zth channel in the second layer GCN.

2) WOALSTM MODULE

The structure of the initial LSTM module is shown in Figure 3. The multi-channel GCN module obtains the vector $G^{(2)}$, which describes the spatial correlation of the traffic flow at T consecutive time steps from C_{22} different dimensions respectively, and uses the self-attention mechanism to weight

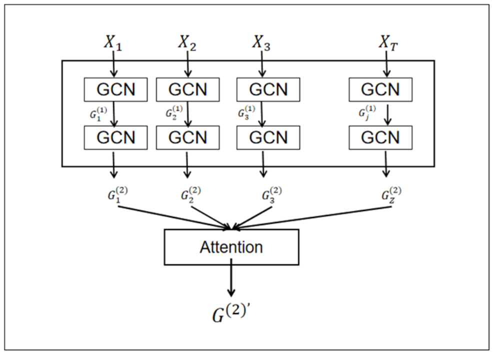


FIGURE 2. MGCN module structure.

the spatial correlation features of the MGCN output, as shown in equation (23).

$$\begin{cases} Q = W_Q G^{(2)} \\ K = W_K G^{(2)} \\ V = W_V G^{(2)} \\ G^{(2)'} = SoftMax(\dfrac{QK^T}{\sqrt{d_K}}) \cdot V \end{cases} \tag{23}$$

where W_Q, W_K, and W_V are the weight matrices that can be learned during the training process, d_K denotes the dimensionality of the input vector $G^{(2)}$, and the resulting vector $G^{(2)'}$ is used as the input to the WOALSTM.

WOALSTM constructs an independent LSTM network for the output of each channel of $G^{(2)'}$, and analyzes the temporal correlation of traffic flow on the basis of the spatial correlation feature vector $G^{(2)'}$ to obtain the spatial and temporal correlation features of traffic flow data.

$$\begin{cases} H_Z^{(1)} = LSTM_Z^{(1)}(G^{(2)'}) \\ H_Z^{(2)} = LSTM_Z^{(1)}(H_Z^{(1)}) \\ \ldots\ldots \\ H_Z^{(n)} = LSTM_Z^{(n)}(H_Z^{(n-1)}) \end{cases} \tag{24}$$

As shown in equation (24), $H_Z^{(n)}$ represents the final output of vector $G^{(2)'}$ through the LSTM network layer corresponding to channel Z. The feature vector H, obtained by concatenating the output feature vectors H_Z, $Z = (1, 2, \ldots, C_{22})$ of all channels, is used as the spatio-temporal feature vector of traffic flow.

In the training process of LSTM for temporal prediction task, the network structure and how to choose the optimal hyper-parameters such as the number of hidden layer neurons and learning rate have a direct impact on the prediction results. Based on the above reasons, this paper combines the intelligent optimization algorithm-WOA with LSTM, and uses the WOA to find the optimal combination of four parameters, namely, the number of LSTM layers n, the maximum number of iterations m, the number of hidden layer neurons unit, and the learning rate LR, to determine the optimal combination of LSTM parameters. Then the parameters will

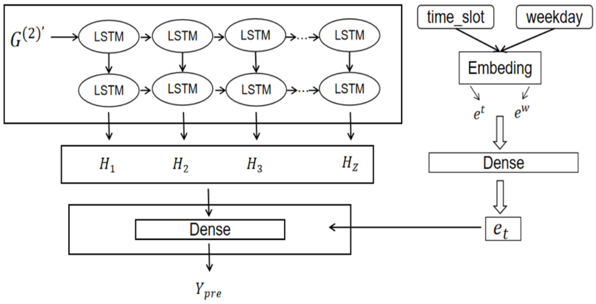


FIGURE 3. Initial LSTM module structure.

be applied to MGCN-LSTM for traffic flow prediction training and testing, and the traffic flow spatio-temporal feature vector H is obtained. Meanwhile the traffic flow data are analyzed in time dimension to obtain time_slot and weekday information, which are converted to vector form through the embedding layer and then spliced to obtain the time period feature vector et through the fully connected layer. Finally, the spatio-temporal feature vector H and the time-period feature vector et are input to a fully connected layer to obtain Y_{pre}, which is the regional traffic flow prediction result.

The detailed steps for optimizing the hyperparameters of the LSTM network based on WOA are as follows:

Step1: Initialization of the LSTM neural network parameters, setting the number of LSTM layers to 2 to 4 layers.

Step2: Initialization of WOA parameters, determining the maximum number of iterations M, the dimension of whales D, the number of whale populations N, and the upper limit UB and the lower limit LB of the parameters.

Step3: Initialize the location of the whales and generate the initial population. A certain whale individual location $X_i =$ (n, m, unit, LR) is randomly selected, where n represents the number of LSTM layers, m represents the maximum number of iterations, unit represents the number of hidden layer neurons, and LR represents the learning rate. Then X_i could be regarded as the combination of parameters to be optimized, which will be input to the WOA. In our experiment, the value range of n is [2, 4], the value range of m is [200,800], the value range of unit is [20,100], and the value range of LR is set to [0.001,0.01].

Step4: Calculation of whale population fitness values. The fitness value is calculated for the current whale population whenever the whales move. The Root Mean Square Error (RMSE) between the actual traffic flow value and the predicted value of the testing set is used as the fitness function and calculated as follows:

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(X_{true} - y_{pre})^2} \tag{25}$$

where n denotes the number of samples in the testing set, X_{true} denotes the actual traffic flow value, and y_{pre} denotes the model prediction value.

Step5: Update the whale individual seeking position. The starting parameter values randomly selected in step 3, are used as the historical optimal values to assign and train the parameters of the LSTM. All whale positions will be updated when each whale has completed the calculation of movement and fitness values. The search will stop when the optimal solution is found. Then the combination of parameters at this time will be used as the optimal hyperparameters of the MGCN-LSTM model. Otherwise, the individual whale search will be performed again.

Step6: The MGCN-LSTM neural network is trained and predicted using the hyper-parameters derived from the WOA algorithm.

3) FLOW CHART OF MGCN-WOALSTM TRAFFIC FLOW PREDICTION MODEL

The flow of MGCN-WOALSTM traffic flow prediction model is shown in Figure 4. The data set is divided after pre-processing the traffic flow data such as data cleaning, filling the null and normalization. The input data $\{X_1, X_2, \ldots .X_T\}$ is updated after the two-layer multi-channel GCN network, and feature extraction of the surrounding traffic network information is completed, which at this time now contains spatial feature information. Then the output spatial features are weighted by the Self-attention mechanism. The weighted spatial feature vector is used as the input information of the WOALSTM network, and the final output $Y_{pre} = \{y_1, y_2, \ldots .y_T\}$ is obtained through Dropout optimization and a fully connected layer, which is a mapping between the spatio-temporal feature vector H and the traffic flow prediction value Y_{pre}. For the model training process, the input data is the training set, and the output Y_{pre} and the actual traffic flow X_{true} are subjected to Root Mean Square Error (RMSE) loss calculation. The model is continuously optimized by using the WOA in order to find the optimal LSTM parameters. For the traffic flow prediction task, the input data is the testing set, and no further model optimization is performed. The output value $Y_{pre} = \{y_1, y_2, \ldots .y_T\}$ is the traffic flow prediction value of the model.

IV. RESULTS

A. TRAFFIC FLOW DATA PRE-PROCESSING

The dataset used in this paper is obtained from the official PeMS website (http://pems.dot.ca.gov) [18]. The selected PeMSD4 dataset is the traffic data of San Francisco Bay Area, which contains the traffic flow data collected by 307 sensor nodes of the California highway network in the United States for a total of 59 days from January 1, 2018 to February 28, 2018. The data is recorded every five minutes, indicating that each sensor contains 288 data nodes per day. The weather data for the corresponding dates were obtained from MesoWest, and the weather data were collected at the same time interval as PeMSD4.

The accuracy of the weather data and traffic flow data will be directly related to the performance of the prediction

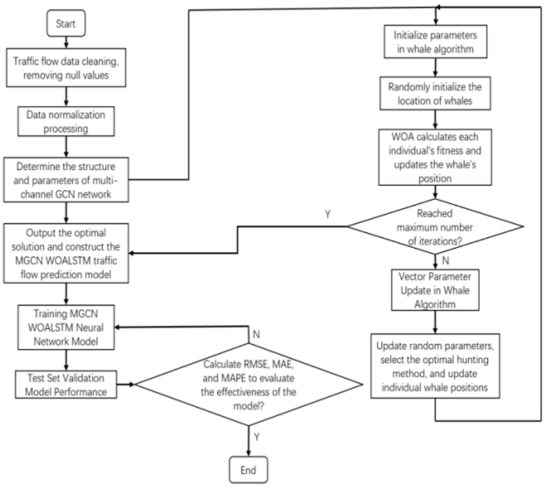


FIGURE 4. **Flow chart of MGCN-WOALSTM traffic flow prediction model.**

model. In order to avoid the influence of abnormal data on the prediction performance, the data are preprocessed first, and the abnormal data are replaced by the average of the traffic and speed values of the neighboring time periods at the corresponding time points. In addition, Min-Max normalization was used to scale the data to the range [0, 1]; for subsequent training and verification of the prediction model performance, the traffic flow data set was divided into training and testing sets in the ratio of 5:1. As a consquence data of the first 50 days were taken as the training set and the data of the last 9 days were taken as the test set.

The input of MGCN-WOALSTM prediction model is the pre-processed traffic flow and weather data, and the output is the predicted traffic flow value, and the input training set and testing set are constructed according to the characteristics of traffic flow and weather data. Assuming that the traffic flow data input is x, the weather data input is w, let X be the input data set of the prediction model, then we have

$$\begin{aligned} X &= (X_1, X_2, X_3, \ldots, X_T) \\ &= \{(x_1, w_1), (x_2, w_2), (x_3, w_3), \ldots, (x_t, w_t)\} \\ w &= \{w_1, w_2, w_3, \ldots, w_t\} \end{aligned} \tag{26}$$

where x_i (i = 1 to t) represents the traffic flow data at moment i, including traffic flow and average speed; w_j (j = 1 to t) represents the weather data at moment j, w_k (k = 1 to t) indicating the kth weather factor, such as visibility, rainfall, snowfall, etc. After taking into account the influence of each weather factor on traffic flow, only two weather factors, that is rainfall and snowfall, are selected in this paper.

B. MODEL EVALUATION METRICS

The degree of merit of a time series forecasting task is usually evaluated using three performance metrics: Root Mean Square Error (RMSE), Mean Absolute Error (MAE), and Mean Absolute Percentage Error (MAPE). Smaller values of the three evaluation metrics represent higher model prediction accuracy. Metrics RMSE is expressed by formula (25). The other two metrics can be expressed by formulas (27) and (28) respectively, as following:

$$MAPE = \frac{1}{n}\sum_{n}^{1}\left|\frac{y_{pre} - x_{true}}{y_{pre}}\right| \times 100 \tag{27}$$

$$MAE = \frac{1}{n}\sum_{i=1}^{n}|X_{true} - y_{pre}| \tag{28}$$

where n denotes the number of samples in the testing set, X_{true} denotes the actual traffic flow value, and y_{pre} denotes the model prediction value.

C. MODEL PARAMETER SETTINGS

To prevent over-fitting during the training of the LSTM neural network, a Dropout layer is added after each LSTM layer. In this model, the loss rate is set to 0.2, the batch_size is set to 32, the epoch is set to 300, the initial whale population in WOA is 30, and the initial number of iterations is 500. After predictive training, the optimal MGCN-WOALSTM model parameters are set as shown in Table 1:

TABLE 1. **Model optimal parameter settings.**

Parameter Name	Parameter Values
GCN Layer1(C_{11}, C_{12})	（4，8）
GCN Layer2 (C_{21}, C_{22})	（8，16）
LSTM layers n	3
Number of hidden layer neurons unit	（39，86，72）
Maximum iterations m	400
Learning Rate LR	0.001

D. MODEL PERFORMANCE COMPARISON

1) PERFORMANCE ANALYSIS OF DIFFERENT TRAFFIC FLOW FORECASTING MODELS

The performance of the regional traffic flow prediction model proposed in this paper is validated by the publicly available dataset PeMSD4 and evaluated using common performance metrics in three time-series prediction problems. In order to more intuitively reflect the superior results of the MGCN-LSTM model compared with other forecasting models such as GRU, LSTM, GCN and ASTGCN, the error evaluation metrics are introduced to test the model performance. The error results are shown in Table 2, where shows the error values of each model for predicting traffic flow within one day on the PeMSD4 dataset. It can be seen that the prediction effect of GRU and LSTM models that extract only the correlation on the temporal dimension of traffic flow is basically the same. And the effect of extracting spatial dependence using only GCN is not very satisfactory, because GCN only considers the spatial characteristics and ignores the temporal characteristics of traffic data. ASTGCN adds spatio-temporal attention mechanism to GCN to obtain dynamic spatio-temporal correlations among traffic

TABLE 2. Error comparison table between MGCN-LSTM model and other models.

Models / Metrics	RMSE	MAE	MAPE
GRU	9.75	7.01	17.01
LSTM	9.69	7.12	17.50
GCN	9.95	7.43	20.84
ASTGCN	8.84	6.89	16.22
MGCN-LSTM	8.01	6.24	16.61

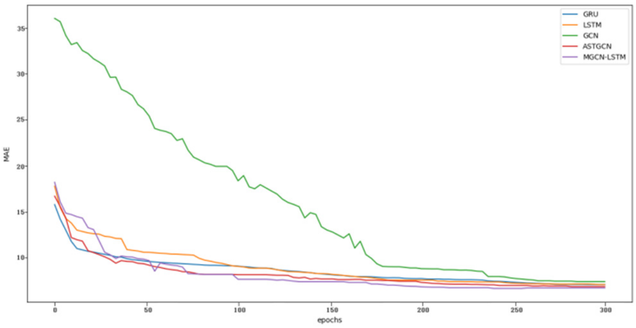


FIGURE 5. MAE for each model.

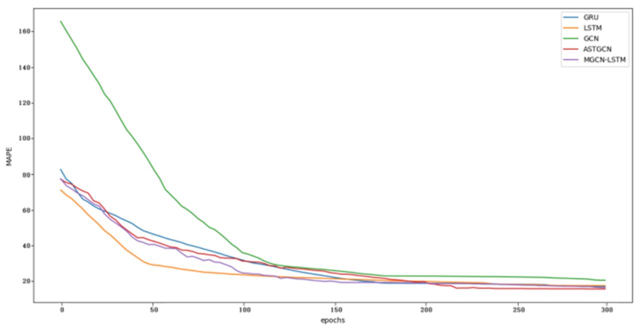


FIGURE 6. MAPE for each model.

road networks, which is a more accurate model in traffic flow prediction at present. The MGCN-LSTM constructed in this paper also captures the spatio-temporal characteristics of traffic flow from both temporal and spatial dimensions. In our model, the prediction index RMSE error decreases by 0.83, MAE error decreases by 0.65, and MAPE error increases by 0.39 compared with ASTGCN; but the overall trend of the three errors of MAE, MAPE, and RMSE gradually decreases with the increase of the number of training rounds, as shown in Figure 5. Figures 6 and 7 show that, overall the prediction performance of MGCN-LSTM model is better than other models, which proves the effectiveness of MGCN-LSTM model in spatio-temporal traffic prediction.

Figure 8 shows the comparison between the traffic flow values predicted by each model in the testing set for one day and the actual traffic flow collected by sensor node 1. It can be seen from the figure that the prediction results of the MGCN-LSTM constructed in this paper are in better agreement with the actual traffic values, and the trend of traffic flow changes within a day is more closely matched.

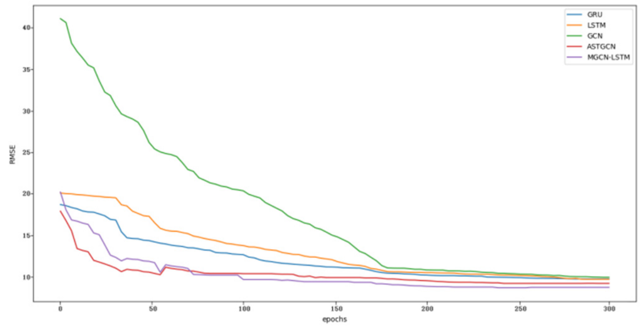


FIGURE 7. RMSE for each model.

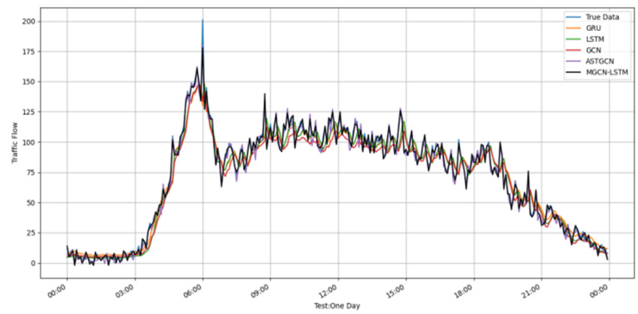


FIGURE 8. Traffic prediction results of each model.

TABLE 3. Comparison of MGCN-LSTM and MGCN-WOALSTM errors.

Models / Metrics	RMSE	MAE	MAPE
MGCN-LSTM	8.01	6.24	16.61
MGCN-WOALSTM	6.84	5.15	13.47

2) COMPARISON ANALYSIS OF BEFORE AND AFTER WOA OPTIMIZATION OF MGCN-LSTM

Each neural network in MGCN uses the same parameters and network structure, and optimizes the LSTM network parameters through WOA. The error results of the original MGCN-LSTM and the optimized prediction model are shown in Table 3, which shows that the MGCN-LSTM prediction model optimized by the whale optimization algorithm further reduces the prediction error than the original prediction model, and achieves better results than the initial MGCN-LSTM prediction, proving that the prediction model with optimized LSTM neural network structure parameters using WOA has better performance.

3) COMPARATIVE ANALYSIS OF TRAFFIC FLOW PREDICTION MODELS INCORPORATING WEATHER FACTORS

On the basis of MGCN-WOALSTM using the optimal combination of parameters, the weather factor is introduced. As can be seen in Table 4, the RMSE, MAE and MAPE errors of the MGCN-WOALSTM prediction model are reduced after considering rain and snow, two external factors affecting

TABLE 4. Comparison of traffic flow prediction model errors with and without integrated weather factors.

Metrics \ Models	RMSE	MAE	MAPE
MGCN-WOALSTM	6.84	5.15	13.47
MGCN-WOALSTM with Weather Factor	6.19	4.07	10.86

traffic flow. The experiment result indicates that the proposed MGCN-WOALSTM prediction model considering weather factors has better prediction performance, proving that the incorporation of weather features helps to improve the traffic flow prediction accuracy.

V. CONCLUSION

This paper proposes to build a regional traffic flow prediction model based on MGCN-WOALSTM, using historical traffic flow data and weather data, analyzing the spatial and temporal correlation of traffic flow from multiple dimensions, considering the correlation strength between weather factors and traffic flow, integrating weather factors into the traffic flow prediction model, and using the WOA to optimize the LSTM network parameters, so as to achieve the effect of reducing the prediction model error. Moreover, experiments were conducted on the real data set, in which the error metrics RMSE was reduced by 2.65, MAE by 2.82, and MAPE by 5.36 compared to ASTGCN, indicating that the prediction model proposed in this paper can better obtain the dynamic change pattern of traffic flow and has better prediction performance. In summary, MGCN-WOALSTM can capture deeper spatial and temporal characteristics from traffic data, and its prediction accuracy is verified to be better than other models. In future traffic flow prediction research, the impact of traffic accidents or more weather influencing factors (such as visibility, temperature, etc.) on traffic flow can be considered, and the optimal parameters of the model under the demand of long-term traffic flow prediction tasks can be further optimized and studied.

REFERENCES

[1] J. Tanimoto and X. An, "Improvement of traffic flux with introduction of a new lane-change protocol supported by intelligent traffic system," *Chaos, Solitons Fractals*, vol. 122, pp. 1–5, May 2019, doi: 10.1016/j.chaos.2019.03.007.

[2] L. Rui and Q. Li, "Short-term traffic flow prediction algorithm based on combined model," *J. Electron. Inf.*, vol. 38, no. 5, pp. 1227–1233, May 2016, doi: 10.11999/JEIT150846.

[3] S. Lu, Q. Zhang, G. Chen, and D. Seng, "A combined method for short-term traffic flow prediction based on recurrent neural network," *Alexandria Eng. J.*, vol. 60, no. 1, pp. 87–94, Feb. 2021, doi: 10.1016/j.aej.2020.06.008.

[4] H. Wang, X. Wei, J. Yao, and Y. Zhang, "Traffic flow prediction using machine learning methods," in *Proc. 3rd Int. Conf. Mach. Learn., Big Data Bus. Intell. (MLBDBI)*, Taiyuan, China, Dec. 2021, pp. 30–35.

[5] L. Yang, "Uncertainty prediction method for traffic flow based on K-nearest neighbor algorithm," *J. Intell. Fuzzy Syst.*, vol. 39, no. 2, pp. 1489–1499, Aug. 2020, doi: 10.3233/jifs-179923.

[6] C. Ma, S. Qin, J. Hu, and L. Yan, "Subway flow prediction based on improved support vector machine," in *Proc. 8th IEEE Int. Conf. Cyber Secur. Cloud Comput. (CSCloud)/7th IEEE Int. Conf. Edge Comput. Scalable Cloud (EdgeCom)*, Washington, DC, USA, Jun. 2021, pp. 115–120, doi: 10.1109/CSCloud-EdgeCom52276.2021.00030.

[7] W. Shu, K. Cai, and N. N. Xiong, "A short-term traffic flow prediction model based on an improved gate recurrent unit neural network," *IEEE Trans. Intell. Transp. Syst.*, vol. 23, no. 9, pp. 16654–16665, Sep. 2022, doi: 10.1109/TITS.2021.3094659.

[8] K. Chen, F. Chen, B. Lai, Z. Jin, Y. Liu, K. Li, L. Wei, P. Wang, Y. Tang, J. Huang, and X.-S. Hua, "Dynamic spatio-temporal graph-based CNNs for traffic flow prediction," *IEEE Access*, vol. 8, pp. 185136–185145, 2020, doi: 10.1109/ACCESS.2020.3027375.

[9] Z. Cheng, J. Lu, H. Zhou, Y. Zhang, and L. Zhang, "Short-term traffic flow prediction: An integrated method of econometrics and hybrid deep learning," *IEEE Trans. Intell. Transp. Syst.*, vol. 23, no. 6, pp. 5231–5244, Jun. 2022, doi: 10.1109/TITS.2021.3052796.

[10] H. Zhang, S. Kan, J. Cao, L. Chen, and T. Zhao, "A traffic flow-forecasting model based on multi-head spatio-temporal attention and adaptive graph convolutional networks," *Int. J. Modern Phys. C*, vol. 33, no. 10, Oct. 2022, Art. no. 2250137, doi: 10.1142/s0129183122501376.

[11] J. An, L. Guo, W. Liu, Z. Fu, P. Ren, X. Liu, and T. Li, "IGAGCN: Information geometry and attention-based spatiotemporal graph convolutional networks for traffic flow prediction," *Neural Netw.*, vol. 143, pp. 355–367, Nov. 2021, doi: 10.1016/j.neunet.2021.05.035.

[12] H. Zhang, L. Chen, J. Cao, X. Zhang, and S. Kan, "A combined traffic flow forecasting model based on graph convolutional network and attention mechanism," *Int. J. Modern Phys. C*, vol. 32, no. 12, Dec. 2021, Art. no. 2150158, doi: 10.1142/s0129183121501588.

[13] P. Chen, X. Fu, and X. Wang, "A graph convolutional stacked bidirectional unidirectional-LSTM neural network for metro ridership prediction," *IEEE Trans. Intell. Transp. Syst.*, vol. 23, no. 7, pp. 6950–6962, Jul. 2022, doi: 10.1109/TITS.2021.3065404.

[14] A. Nasser and V. Simon, "A novel method for analyzing weather effect on smart city traffic," in *Proc. IEEE 22nd Int. Symp. World Wireless, Mobile Multimedia Netw. (WoWMoM)*, Pisa, Pisa, Italy, Jun. 2021, pp. 335–340, doi: 10.1109/WoWMoM51794.2021.00061.

[15] W. Yang, L. Li, C. Hu, L. Zhang, and R. Fang, "Summary of research on impact of adverse weather on traffic safety and traffic flow," *J. Wuhan Univ. Technol., Transp. Sci. Eng.*, vol. 43, no. 5, pp. 844–854, Oct. 2019.

[16] Y. Hou, Z. Deng, and H. Cui, "Short-term traffic flow prediction with weather conditions: Based on deep learning algorithms and data fusion," *Complexity*, vol. 2021, Jan. 2021, Art. no. 6662959, doi: 10.1155/2021/6662959.

[17] F. Miao, L. Tao, J. Xue, and X. Zhang, "A queue hybrid neural network with weather weighted factor for traffic flow prediction," in *Proc. IEEE 24th Int. Conf. Comput. Supported Cooperat. Work Design (CSCWD)*, Dalian, China, May 2021, pp. 788–793, doi: 10.1109/CSCWD49262.2021.9437626.

[18] M. M. Ghazi, A. Ramezani, M. Siahi, and M. M. Ghazi, "Learning spatiotemporal features from incomplete data for traffic flow prediction using hybrid deep neural networks," 2022, *arXiv:2204.10222*.

[19] Y. Li, S. Chai, Z. Ma, and G. Wang, "A hybrid deep learning framework for long-term traffic flow prediction," *IEEE Access*, vol. 9, pp. 11264–11271, 2021, doi: 10.1109/ACCESS.2021.3050836.

[20] Y. Zhang and D. Xin, "Dynamic optimization long short-term memory model based on data preprocessing for short-term traffic flow prediction," *IEEE Access*, vol. 8, pp. 91510–91520, 2020, doi: 10.1109/access.2020.2994655.

[21] H. Ruan, B. Wu, B. Li, Z. Chen, and W. Yun, "Expressway exit station short-term traffic flow prediction with split traffic flows according originating entry stations," *IEEE Access*, vol. 9, pp. 86285–86299, 2021, doi: 10.1109/ACCESS.2021.3087658.

[22] T. N. Kipf and M. Welling, "Semi-supervised classification with graph convolutional networks," 2016, *arXiv:1609.02907*.

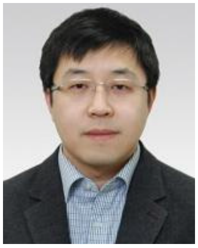
[23] L. Xu, J. Peng, and Y. Shao, "Prediction of expressway traffic flow on holidays based on improved GA-BP algorithm," *J. Comput. Theor. Nanoscience*, vol. 13, no. 5, pp. 3290–3294, May 2016, doi: 10.1166/jctn.2016.4988.

[24] J. Ye, S. Xue, and A. Jiang, "Attention-based spatio-temporal graph convolutional network considering external factors for multi-step traffic flow prediction," *Digit. Commun. Netw.*, vol. 8, no. 3, pp. 343–350, Jun. 2022, doi: 10.1016/j.dcan.2021.09.007.

[25] 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, doi: 10.1109/TITS.2021.3098461.

[26] C. Tang, J. Sun, Y. Sun, M. Peng, and N. Gan, "A general traffic flow prediction approach based on spatial–temporal graph attention," *IEEE Access*, vol. 8, pp. 153731–153741, 2020, doi: 10.1109/ACCESS.2020.3018452.

[27] S. Guo, Y. Lin, N. Feng, C. Song, and H. Wan, "Attention based spatial–temporal graph convolutional networks for traffic flow forecasting," in *Proc. Nat. Conf. Artif. Intell. Assoc. Advancement Artif. Intell. (AAAI)*, vol. 33, 2019, pp. 922–929, doi: 10.1609/aaai.v33i01.3301922.



KERANG CAO received the Ph.D. degree in computer engineering from Pai Chai University, South Korea, in 2011. He is currently an Associate Professor and a Postgraduate Supervisor with the College of Computer Science and Technology, Shenyang University of Chemical Technology, China. He chaired and participated in four projects of the International Cooperation Program of the Ministry of Science and Technology and the National Natural Science Foundation of China. He has published more than 20 articles in national and international journals and conferences, many of which have been indexed in SCI and EI journals. He has engaged in research works with the National Natural Science Foundation of China and the Korea Institute of Science and Technology Information. His major research interests include the analysis and design of intelligent optimization algorithms and the IoT technology.

YARU LIU is currently pursuing the M.S. degree in computer science and technology with the Shenyang University of Chemical Technology, Shenyang, China. Her research interests include deep learning and intelligent optimization algorithms.

LINI DUAN received the Ph.D. degree in production management and MIS from Chungnam National University, South Korea, in 2013. She is currently a Lecturer with the School of Economics and Management, Shenyang University of Chemical Technology, China. Her research interests include management information systems, logistics information management, and scheduling theory. Her research works have been published in many international journals.

SEN XU received the master's degree in computer science and engineering from the Harbin Institute of Technology, Harbin, China, and the Ph.D. degree in computer science and engineering from the University of South Carolina, Columbia, SC, USA. In 2015, he joined the School of Computer Science and Technology, Shenyang University of Chemical Technology, as a Faculty Member. He has published multiple highly cited papers in international conferences and journals. The first paper published during the Ph.D. studies was one of the earliest research literature on WiMAX. He has been cited more than 200 times by scholars from around the world. His research interests include information security, the design and analysis of security protocols, formal analysis and verification of protocols, secure multicast, and cryptography. As a reviewer of multiple international conferences and journals, he has been awarded the Exemplary Reviewer Award (on IEEE COMMUNICATIONS LETTERS) by the IEEE Communications Society.

HOEKYUNG JUNG received the Ph.D. degree in computer engineering from Kwangwoon University, South Korea, in 1993. He is currently a Professor with the College of AI/SW Creative Convergence, Pai Chai University, South Korea. His research interests include AI, deep learning, big data, embedded systems, and the IoT. His research works have been published in many international journals.