SAGE-GSAN: A graph-based method for estimating urban taxi CO emissions using street view images

Article *in* Journal of Cleaner Production · September 2024

DOI: 10.1016/j.jclepro.2024.143543

5 authors, including:

Zeqiang Chen
Wuhan University
82 PUBLICATIONS **1,676** CITATIONS
SEE PROFILE

Yan Zhang
The Chinese University of Hong Kong
30 PUBLICATIONS **429** CITATIONS
SEE PROFILE

Nengcheng Chen
Wuhan University
236 PUBLICATIONS **4,191** CITATIONS
SEE PROFILE

SAGE-GSAN: A Graph-Based Method for Estimating Urban Taxi CO Emissions Using Street View Images

Zeqiang Chen[a], Tongxu Zou[a], Zheng Xu [a], Yan Zhang[b c *]and Nengcheng Chen[a]

[a] *National Engineering Research Center of Geographic Information System, China University of Geosciences, Wuhan, 430074, China;* [b] *Institute of Space and Earth Information Science, The Chinese University of Hong Kong, 999077, China;* [c] *State Key Laboratory of Information Engineering in Surveying, Mapping, and Remote Sensing, Wuhan University, Wuhan 430079, China*

* **Corresponding author**: Yan Zhang, ORCID: 0000-0002-2059-4171, Email: yanzhang@cuhk.edu.hk, sggzhang@whu.edu.cn

Abstract

Accurately predicting the carbon emissions of urban street vehicles is a current challenge in the field of urban transportation. This study proposed a new SAGE-GSAN model (Graph SAmple and aggreGatE - Graph Spatial Attention Network) to solve this problem. It combines graph neural networks with streetscape features and road network structure to predict the traffic carbon monoxide emissions at the street level. The method takes the street view images, the 5,075 roads network structure in Wuhan and 19 street visual elements as the input features, and the carbon monoxide emissions obtained from the driving trajectories as the prediction data. The method achieved an experimental accuracy of 81.4% in predicting carbon monoxide emissions from street

cabs. This study also compares the prediction results of traditional neural networks and analyzes the effects of different street-level features and graph convolution layers on the prediction accuracy. The results of this study show that the graph neural network and attention mechanism techniques could solve the fine-grained carbon emission prediction problem at the urban street level effectively. The model code is shared at the: https://github.com/zou9229/CO_Predict_Code.

Keywords: Graph neural networks, Street view, Road network, Taxi, Carbon monoxide emissions

1. Introduction

The Paris Agreement signed in 2015 proposed to limit global warming to within 2°C above pre-industrial levels, and to strive to limit the temperature rise to within 1.5°C (Lawrence and Schafer 2019), to ensure the thriving growth of children (Howard et al. 2020). On 22 September 2020, at the 75th United Nations General Assembly, China proposed the goal of reaching peak carbon emissions before 2030 (Zeng et al. 2022) and achieving carbon neutrality before 2060 (Yang et al. 2022), and positioned it as a national development strategy (Zhang et al. 2023). Achieving these goals requires reductions in carbon emissions in various aspects. City carbon emissions are an important part of the national total (Ma et al. 2022), and transportation-related emissions are a significant factor contributing to urban carbon emissions (Liu et al. 2020). Taxis, as one of the most important modes of transportation in cities, play a crucial role in urban carbon emissions (Zhang et al. 2018). Although there are various types of vehicles in urban areas, this study focuses on taxis due to their crucial role in

city transportation. Taxis not only have the largest number, but also usually cover larger areas and operate for extended periods throughout the day, making their exhaust emissions a significant contributor to urban air quality. Therefore, accurately predicting taxis' carbon emissions is of great significance for achieving urban carbon neutrality (Wang et al. 2022). By predicting the carbon emissions of taxis on urban streets, we can better understand the carbon emissions in urban transportation(Zhao et al. 2017), which would be beneficial for governments in developing targeted policies and measures to reduce carbon emissions (Yang et al. 2020). Carbon monoxide (CO) pollutant is a common major carbon emission pollutant on the road, and motor vehicles emit the largest amount of it among all vehicles (Xue et al. 2013). CO poses harm to the environment and human health, and it undergoes oxidation reactions in the air, transforming into carbon dioxide (CO_2), which will exacerbate the greenhouse effect and affect the process of reaching peak carbon emissions and achieving carbon neutrality. Thus, it is significant to realize the prediction of the CO emission level of street city taxis by combining urban street view features and road network structure in this study.

The existing methods can be roughly categorized into micro and macro methods based on the object on which the forecasting process acts. From a micro-level perspective, some physical models are already available for calculating the related CO emissions of individual motor vehicles. This category of methods takes full account of the influence of multiple micro-factors on the emission factor and measures CO emissions in the dimension of the vehicle itself, so these methods are categorized as

micro-level methods. At present, micro-level models for motor vehicles are relatively mature(Saharidis and Konstantzos 2018). Models of motor vehicles' micro-emissions are mainly used to determine motor vehicles' emissions by simulating the characteristics of motor vehicles' emissions (Nocera et al. 2018) and substituting them with the usage characteristics of the motor vehicles themselves (Pathak et al. 2016). Kwiecień and Szopińska (2020) used the computer program to calculate emissions from road transport (COPERT) model to calculate the CO emissions on individual roads near residential areas. Pérez et al. (2019) found that passenger cars (mainly diesel cars) accounted for 80.7% of the total mileage in the city and were the main contributors to the pollution emissions of road transportation. Turkensteen (2017) studied whether green transportation would be feasible by using the comprehensive modal emission model (CMEM) to calculate fuel consumption and the carbon emissions of vehicles under given conditions of speed and load. Tao et al. (2019) used the CMEM model to invert the carbon dioxide emissions of trajectory points in taxis' travel trajectories. Gao et al. (2018) used the mobile source emission factor model (MOBILE) for localized research on motor vehicles' pollutant emissions in Wuhan and calculated the emission factors of different vehicle models. Ibarra-Espinosa et al. (2020) used the Vehicular Emissions INventory (VEIN) model in conjunction with travel demand to predict CO emissions in the São Paulo metropolitan area. In addition, some scholars have tried to predict the transient emissions of vehicles by deep learning and achieved better results (Howlader et al. 2023).

From a macro-level perspective, some researchers have used traditional models or

deep learning methods to predict the overall CO emissions at a larger scale, such as at the city or country level. Gao et al. (2021) proposed a novel fractional grey Riccati model to predict the carbon emission levels of the United States, China and Japan, and obtained better results. Wu et al. (2018) used the extended STIRPAT model to determine the relationship between CO2 emissions and different driving factors to predict peak carbon emissions in Qingdao. Acheampong and Boateng (2019) developed a model using artificial neural networks (ANN) to predict the intensity of carbon emissions in Australia, Brazil, China, India, and the United States. Huang et al. (2019) used a firefly algorithm (FA) optimized with Elman's neural network to predict China's carbon dioxide emissions. Lu et al. (2017) predicted real-time and fine-grained traffic carbon emissions for an entire city on the basis of a three-layer perceptron neural network (three-layer PNN) using a spatiotemporal dataset observed in Zhuhai City. Niu et al. (2020) used a generic regression neural network (GRNN) optimized by an improved fireworks algorithm (IFWA) to predict that China would be able to achieve its commitment to carbon peaking by 2030. Alfaseeh et al. (2020) used deep sequence learning to predict the greenhouse gas emissions of streets and achieved better results. Although they carried out their predictions at the street scale, they did not consider the road network structure. Zhou et al. (2023) proposed a method of identifying carbon emissions based on remote sensing images. They extracted the features of carbon emissions of the region using a convolutional neural network model based on a regional dataset of satellite-based remote sensing images.

While various methods can be used to predict the carbon emissions of

transportation (Seo and Park 2023), simple yet effective prediction of CO emissions at the level of road segments is 9still rare. Additionally, most previous studies have focused on predicting carbon emissions via traditional emission factor-based method (Sun and Huang 2022) or traditional neural network method (Fan et al. 2022). Rare methods have combined the structure of the road network to predict street-level CO emissions. The basic reason for this may be that it is relatively difficult to combine carbon emission data with a variety of information such as the road's topology and street view images, which requires complex data processing and feature extraction tools.

The urban environment is complex, with significant spatial heterogeneity in factors related to carbon emissions. The topology of road networks influences the direction of traffic flow to a certain extent, making this information crucial for predicting carbon emissions from traffic volume. emission-related factors pose significant challenges. Urban areas exhibit diverse factors influencing carbon emissions, such as varying traffic density, land use, and transportation infrastructure. These factors contribute to spatial variations in carbon emissions, emphasizing the need for a nuanced understanding of the urban environment in predictive modeling. Furthermore, the topology of road networks significantly influences traffic flow and emission patterns in cities. Incorporating road network topology is essential for accurate carbon emission predictions, as it shapes traffic flow and emission distribution. Additionally, due to the importance of the street-level perspective, street view images can authentically capture the urban environment, providing the model with richer contextual information. This research seeks to seamlessly integrate road network topology and street-level

perspectives. By harnessing technologies like graph neural networks, the research model aims to offer a holistic understanding, capturing the complex interplay between these factors. This comprehensive strategy is crucial for achieving precise street-level carbon emission predictions in urban environments.

In recent years, due to the support of graph neural networks (GNNs) for topological structures, there has been huge potential for geographic information in the field of spatial analysis, such as social functions classification (Hu et al. 2021), traffic forecasting (Jiang and Luo 2022), and geographic knowledge representation and reasoning(Zhu et al. 2020). GNNs could discover the spatial correlations and features contained in urban traffic's carbon emissions (Wang et al. 2020), and thus predict and optimize the carbon emissions and efficiency of urban traffic. Currently, many scholars have already used GNN in the field of urban transportation, and the common ones are traffic flow prediction (Bai et al. 2021) and passenger demand prediction (Bai et al. 2019), but the application of GNN in the research of CO emission remains relatively limited.

This research provides a reliable method for fine-grained predictions of CO emissions on road segments in urban areas by using GNN and a method of extracting the visual features from street view images. It constructs the input data on the basis of features extracted from urban street view data and the road network's structure and classifies road CO emissions into different levels for each street. Moreover, this research provides an important reference for environmental protection, urban planning, and traffic management, and may support efforts to achieve carbon neutrality and peak

carbon emissions, promoting the application and development of deep learning in the field of carbon emissions by urban traffic. In summary, the contributions of this research are the following.

1) This research proposes a fine-grained CO emission prediction method at the road level rather than on a city or regional basis. It incorporates features extracted from street view images and the road network's structure to estimate the CO emission level for each road.
2) This research designs a new SAGE-GSAN model, which can better consider the influence of the road network's topology and the correlation between streets on CO emissions, adaptively learn the relationships between street features, encoding these relationships to improve the accuracy of the predictions.
3) This research analyzes the visual features of streets related to high CO emissions from an environmental perspective. It describes the street environment in a quantitative way, which can help us understand the relationship between carbon emissions and the urban environment better.

2. Data and Preparation

This research selected 5075 urban streets in Wuhan City as the research object. The data included panoramic street view image (SVI) data from Tencent Maps, the structure of the road network from Wuhan City's open street map (OSM), and taxi global positioning system (GPS) data from the Wuhan City Traffic Management Bureau for 3 July 2017 in Wuhan City.

2.1 Structure of the Road Network

In this research, the OSM structure of Wuhan city's road network was used as the graph structure of GNN, where each road was a node in the graph, with a total of 5075 nodes labeled, starting from 0. The road network's structure is shown in Figure 2.1, which was generated in ArcMap with a 25 m buffer.

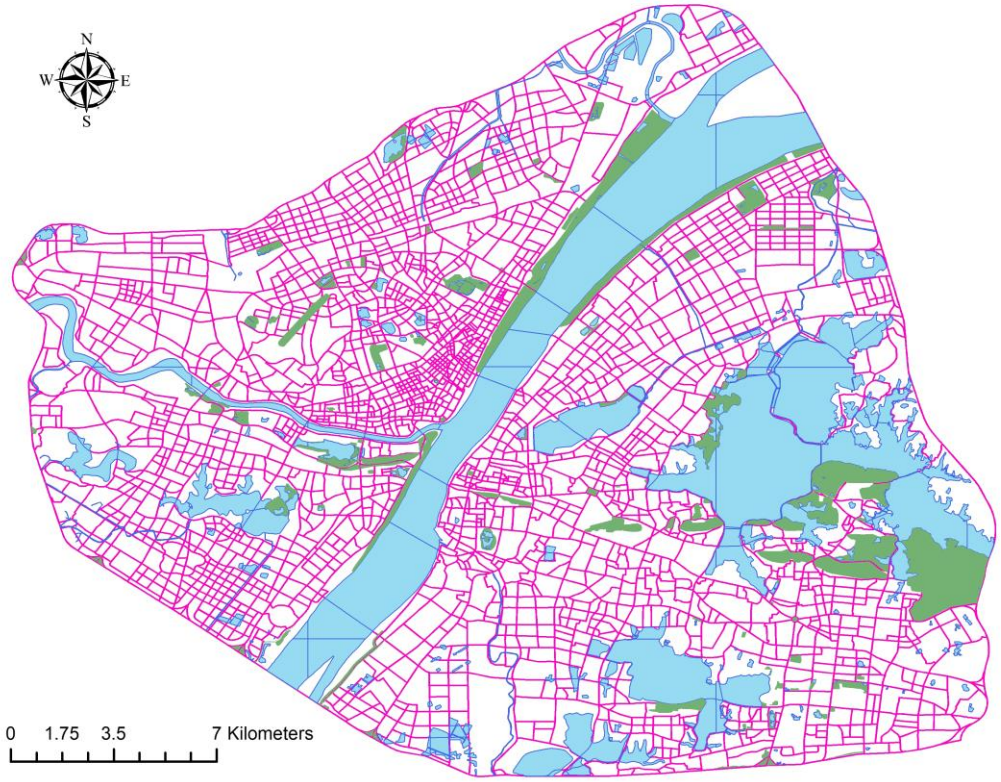


Figure 2.1 Structure of Wuhan City's road network.

2.2 Features of street view images

Street views are urban close-up images that are captured, acquired, processed, stored, and presented by mounting a camera on a car, bicycle, or pedestrian. These images are taken on a street-by-street basis and allow computers to view phenomena in urban streets from a human perspective. Widely acclaimed, street views provide an extremely detailed dataset of urban environments that can be used for a variety of

studies (Zhang et al. 2021), including in the fields of urban planning, transportation, and the environment (Hankey et al. 2021), as well as for validating and training various computer vision algorithms, such as object detection, image classification, and semantic segmentation (Biljecki and Ito 2021).

In this research, street views were used to analyze the visual characteristics of streets with high CO emissions. Through the acquisition and processing of street view images, this research found that there may be some landscape features of streets with high CO emissions, such as the width of the road, the number and location of motor vehicles on the road, the size of the buildings along the street, and the rate of greening at the roadside, etc. These environmental features can reflect traffic flows, pedestrian activity, and other factors related to the CO emissions of the street, and these features may be closely related to the intensity of the road's CO emission, so they can provide strong support and a basis for fine-grained predictions of CO emissions.

In this research, the features of street views were used as part of the input to the GNN. Street view images were downloaded from Tencent Maps, totaling 56,461 panoramic images, and the objects in the street view images were classified through semantic segmentation and used as the features of the streets. To achieve better feature extraction, this research used the pre-trained DeepLabV3+ model , previously trained by Chen et al. (2018), for semantic segmentation and feature extraction. The model segments the images into 19 categories of features, including cars, roads, buildings, walls, fences, trees, lawns, signs, buses, motorcycles, trucks, sky, poles, lights, pedestrians, cyclists, trains, bicycles, and sidewalks. Since each road had more than one

street view image, this research summarized and normalized the average value of each object category for each road as the features of the road.

2.3 Carbon Emission Ratings of Street Taxis

Since CO is the main component of motor vehicles' road-related carbon emissions and can represent the level of road-related carbon emissions, this research used the CO emission levels of street taxis as the label data of the GNN network. The CO emissions of taxis on the streets were obtained through a micro-factor method; here, the COPERT model was chosen for this calculation, as it is able to obtain relatively accurate results (Gräbe and Joubert 2022). Given a vehicle's speed V(km/h) and type, the CO emission factor of motor vehicles $E_{HOT(CO)}$(g/(km·vehicle)) can be calculated using Equation (1) (Shang et al. 2014).

$$E_{HOT(CO)} = \frac{71.7 + 11.4V}{1 + 35.4V - 0.248V^2} \quad (1)$$

Based on the GPS data of taxis in Wuhan City for one day, the taxi trajectory points were spatially connected with the 25 m buffer zone of the road network to obtain the taxi trajectory data of 5075 roads in Wuhan City, as shown in Figure 2.2.

Figure 2.2 Taxi trajectory points after spatial connection.

The traffic flow information of each road was obtained by combining the trajectory point data with the information on the road network. The CO emission factor was calculated for each trajectory point, as the vehicle speeds are not the same at each trajectory point. Then the average CO emission factor for each road was computed by summing the emission values of all trajectory points corresponding to this road and dividing by the number of points. Then the CO emissions of each street over one day were calculated by multiplying the CO emission factor by the corresponding road's traffic flow and length, as shown in Equation (2):

$$E_{CO} = E_{HOT(CO)} \times Cab\ \ traffic\ \ flow \times Road\ \ length \quad (2)$$

where E_{CO} is the CO emissions of the street in one day, $E_{HOT(CO)}$ is the CO emission factor of the road, $Cab\ \ traffic\ \ flow$ is the daily traffic flow of the road, and $Road\ \ length$ is the length of the road.

The total daily CO emissions of each road are multiplied by the number of

passenger cars and taxis in the *2018 Wuhan Statistical Yearbook*[1]. After multiplying it by 365 (days in one year), the annual CO emissions of passenger cars in Wuhan can be estimated to be approximately 32,000 tons. This result has an error of about 50% less than the 61,514 tons compared with the annual CO emissions data for passenger cars in the *2017 Wuhan Motor Vehicle Pollution Prevention and Control Annual Report*[2]. This is reasonable, because passenger cars include large vehicles such as gasoline and diesel vehicles, and there are also issues such as incomplete data on taxis. Furthermore, about 70% of the data showed that the difference between the average CO emissions factor calculated by Amap and that calculated by the GPS trajectory data was within 0.15, which confirmed the reasonableness of the data.

Finally, the CO emissions obtained for each road were divided into seven levels using natural breaks (Chen et al. 2013) and used as the label data for the GNN model. A schematic diagram of the CO emission levels of the divided roads is shown in Figure 2.3. Most of the streets have CO emission Level 1 (low CO emissions), and a few roads with high traffic volumes have higher CO emission levels.

[1]Wuhan Bureau of Statistics,2018 Wuhan Statistical Yearbook, http://tjj.wuhan.gov.cn/tjfw/tjnj/202004/P020200426461240969401.pdf

[2]Wuhan Ecological Environment Bureau, 2017 Wuhan Motor Vehicle Pollution Prevention and Control Annual Report, http://hbj.wuhan.gov.cn/fbjd_19/xxgkml/zwgk/wrfz/jdchjgl/202001/t20200107_576406.html

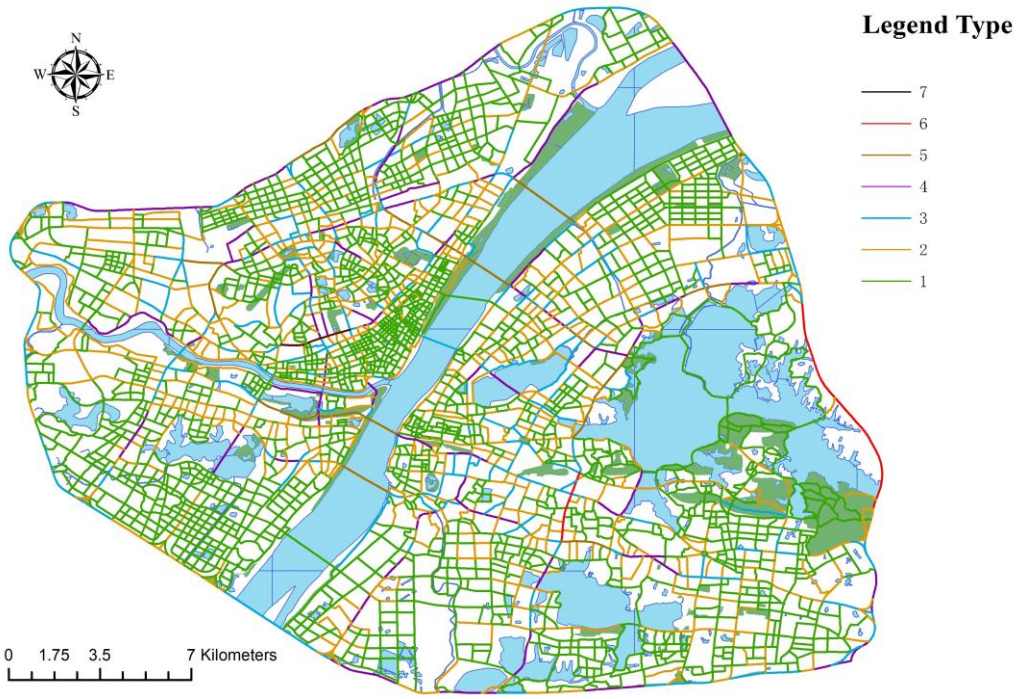


Figure 2.3 CO emission levels of the roads classified by the natural breaks method.

The street view images, carbon emission, and road network structure data used in this study may contain certain uncertainties. These uncertainties primarily arise from the following aspects: 1) The road network structure data may be outdated or miss certain small roads. 2) The time of capture and weather conditions of the street view images may affect the image quality and the accuracy of feature extraction. 3) Since the carbon emission of each vehicle on this road is $\int_0^l E_{HOT(CO)} dl$, where l is the road length, and $E_{HOT(CO)}$ is the carbon emission factor of the track point, since the speed v varies with the location, according to Equation (1), $E_{HOT(CO)}$ also varies with the location, and the integral curve is not smooth, so carbon emission cannot be accurately obtained. Corresponding solutions include:

1) Cleaning and processing the road network data obtained from OSM, removing unnecessary information and anomalous data while retaining the required road network structure information.

2) Conducting quality control on the obtained street view data, including checking the image resolution, clarity, and the accuracy of the geographic location. Aggregating and normalizing the average value of each type of object in the feature vectors of each road segment as the road's feature. This method captures the diversity and complexity of the road environment more effectively. Averaging can reduce the impact of noise or outliers in the images to a certain extent, improving the stability and reliability of the feature vectors. Normalization confines the numerical range of the feature vectors within a certain interval, facilitating model training and optimization and preventing weight imbalance issues caused by numerical differences among different features.

3) Speed v will not increase or drop sharply, so the carbon emission factor will not fluctuate greatly. Therefore, the average carbon emission factor $\bar{E}$ of each vehicle on the road (the average sum of $E_{HOT(CO)}$ of each point of the vehicle on the road) is taken, and the result obtained by multiplying the road length l is approximated as the carbon emission E_{co} of the vehicle on the road in one day. Then, the total carbon emission of the road in one day can be obtained by Equation (3), where $i = 1,2,3,\dots,N$, $\bar{\bar{E}}$ is the average carbon emission factor

of all taxis on the road (i.e. the average sum of all $\bar{E}$ on the road), and c is the traffic flow of the road.

$$E_{co} = \sum_{i=1}^{N} \bar{E}_i * l = \bar{\bar{E}} * l * c \#(3)$$

3. Methodology

This research methodology aims to integrate street view images, road network structure, and taxi GPS data, leveraging a Graph Convolutional Neural Network (GCNN) to predict carbon emissions (2023; Patent No. CN116663724A). The entire process can be divided into two main stages: data processing and model development. As shown in Figure 3.1, in the previous section, the researchers collected three types of data and completed the first stage of data processing, which involved semantic segmentation for feature extraction from street view images, conversion of the road network structure into an undirected graph, and calculation of taxi CO emissions using the COPERT model.

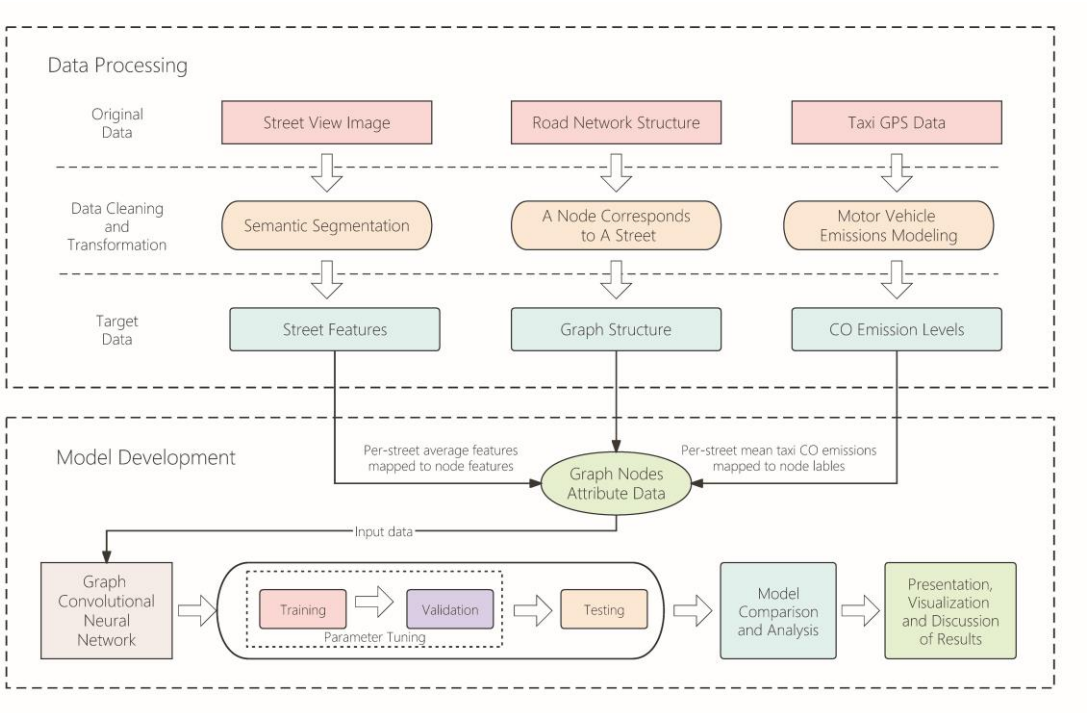


Figure 3.1 Research framework diagram.

GNNs are deep learning algorithms that have been widely applied in various fields (Zhou et al. 2020), such as geography (Zhang et al. 2023). GNNs are particularly suitable for handling spatial and network data because they can operate on graph-structured data (Z. et al. 2021). One of the advantages of using GNNs is that they can capture complex spatial relationships that are difficult to model with other deep learning algorithms (A. 2020). GNNs have been used in urban planning and transportation research to model and optimize transportation networks and predict traffic flows, and also in environmental modeling to predict and map environmental variables such as air pollution, water quality, and land cover (Zhou et al. 2020).

There are various types of GNN models, including graph convolutional networks (GCNs) (Fu et al. 2017), graph attention networks (GATs) (Veli V C Kovi C et al. 2017), and graph sample and aggregate (SAGE) (Hamilton et al. 2017). GCN is one of the earliest types of GCNNs. It aggregates the features of neighboring nodes and uses a fixed weight matrix to perform convolutional operations on the graph data. GAT is a type of attention-based graph neural network that can dynamically learn the importance of each node in the network and better handle the graph data. SAGE is a type of sampled-based graph neural network that constructs subgraphs by randomly sampling nodes and edges from the local neighborhood. This makes SAGE efficient for processing large-scale graph data while still achieving high accuracy. In this research, SAGE (Hamilton et al. 2017) was used to predict the street-level CO emissions from taxis. The road network's structure was converted into a large-scale graph structure, as

shown in Figure 3.2.

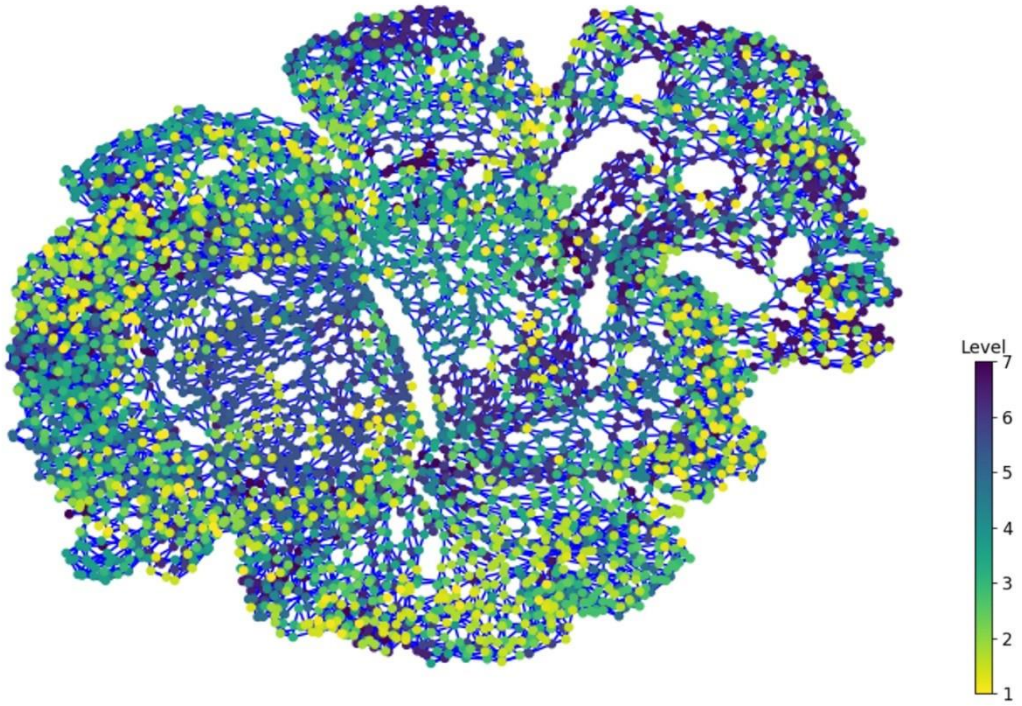


Figure 3.2 The undirected graph structure corresponding to the road network's structure.

One advantage of SAGE is its ability to capture the structural information of a graph by sampling the neighbors, which enables better generalization to new nodes and graphs. Moreover, SAGE is computationally efficient and can handle much larger graphs than other GNNs, which is particularly useful for large-scale geographic datasets, such as that of Wuhan (see Figure 3.2).

This study has designed a graph neural network model—SAGE-GSAN—to estimate carbon emissions from motor vehicles on urban streets. The model is constructed using the DGL framework and incorporates two layers of SAGEConv (GraphSAGE) to extract features from graph data. These layers update each node's representation by aggregating the features of neighboring nodes. Typically, a two-layer

SAGE model is a common choice as it maintains adequate expressive power while preventing overfitting and conserving computational resources. A spatial attention module (GSAN) is introduced between the two layers to incorporate the degree of attention to positional information within the node representations. This module dynamically adjusts the weights of node features using both positional and feature information, enabling the model to better utilize the spatial relationships between nodes. During forward propagation, the first SAGEConv layer extracts features and applies the ReLU activation function; if positional information (pos) is provided, it is processed through the spatial attention module for feature adjustment; this is followed by a second feature extraction via the SAGEConv layer and another application of the ReLU activation function. Finally, a global pooling operation aggregates the node features to produce the final node representation, as detailed in the model structure shown in Figure 3.3.

In the SAGE-GSAN model, the graph structure input is derived from the road network obtained in an earlier phase, and the graph node attribute data come from street features and taxi CO emission levels collected in the same phase. Specifically, the average features of each street extracted through semantic segmentation serve as node features, while the total CO emissions from taxis on each street are used as node labels. This study employs the aforementioned input data to train, validate, and test the GCNN, adjusting parameters during the training process. The model is then subjected to comparison and analysis, followed by the presentation, visualization, and discussion of the results.

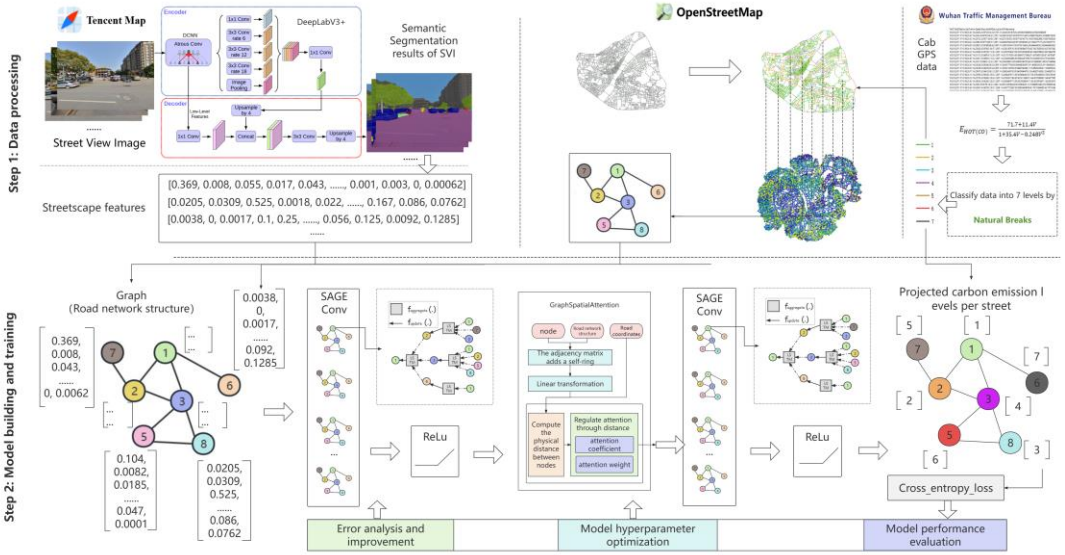


Figure 3.3 Method of predicting CO emission levels. Step 1: Data processing. Step 2: Building and training the SAGE-GSAN model.

In terms of the aggregation functions selected for the model, the commonly used aggregation functions in SAGE models include mean aggregation, GCN aggregation, pooling aggregation, and long short-term memory (LSTM) aggregation (Hamilton et al. 2017). Since the CO emissions of different roads may be influenced by multiple factors, there may be some underlying time-related correlations in the data. For example, the CO emissions of some roads may be affected by the weather or traffic conditions of the previous road, and CO from the previous road may bleed into the next road over time. LSTM allows for the exchange of information between nodes, and this information propagation process can be applied at each node using LSTM to update node representations based on information from neighboring nodes (Graves 2012). In graph data, the order of nodes may have practical significance, such as in time series data or sentence structures in language models (Yu et al. 2019). In our experiment, the streets are interconnected, so the order of streets may also hold practical meaning. In these

cases, LSTM can preserve this ordered information and be beneficial in a GNN (Van Houdt et al. 2020). Therefore, in this research, LSTM was selected as the aggregation function of SAGE for the experiment. Furthermore, the model was designed to aggregate information from the neighbors of each node up to a distance of two hops, meaning that for each node, it considered the information of its neighbors and the information from the neighbors of these neighbor nodes, as well as information of the node itself. This approach allowed the model to have a more comprehensive understanding of the local information around each node while avoiding the aggregation of too much information from distant nodes, reducing the impact of noise and improving the accuracy of the predictions. The rectified linear unit (ReLU) activation function (Yarotsky 2017) was used to enhance the expressive power and computational efficiency of the model .

This research develops a Graph Spatial Attention Network (GSAN) that leverages an efficient graph message-passing mechanism combining location and feature information to enhance graph representation learning. By incorporating self-loops and linear transformations, the model effectively captures the fundamental characteristics of nodes. In the GSAN, the message-passing method, named consistently with traditional neural network forward propagation, emphasizes computational interactions among nodes. This model utilizes a spatial attention mechanism that calculates physical distances between nodes and dynamically adjusts the weights of neighboring node features. This focus on significant neighbors allows the model to capture spatial correlations more accurately and adapt to various graph

structures and task demands.

To utilize spatial information effectively, the research integrates the physical distances between nodes as a critical factor in message passing, calculating Euclidean distances to capture absolute spatial relationships even in high-dimensional spaces. Given the position information posi and posj of the current node i and neighbor node j, as shown in Equation (4) is used to calculate the Euclidean physical distance between nodes in space. This enables the model to focus on proximate nodes, enhancing the capture of local structures and spatial relationships within the graph. Additionally, an attention mechanism on node features adjusts the importance of neighbor features dynamically using a learnable parameter matrix, allowing the model to adaptively focus on different node information throughout the learning process.

$$dist = \left\| pos_i - pos_j \right\|_2 \#(4)$$

In order to introduce the attention mechanism for node features, this study uses a learnable parameter matrix att to adjust the neighbor node feature x_j, as shown in Equation (5), where $\odot$ symbol represents the multiplication of the attention coefficients of neighbor node features one by one, aiming to dynamically adjust the importance of node features through the attention coefficients. If the element in att is large, then the feature x_j of the corresponding node j will be enhanced, indicating that the node is more important in the current task. If the element in att is small, then the feature x_j of the corresponding node j will be suppressed, indicating that the node is relatively less important in the current task. This mechanism enables the model to pay attention to the information of different nodes in the graph adaptively during the

learning process, which is very effective for the task based on graph neural network and attention mechanism in this study.

$$x_j' = x_j \odot att \#(5)$$

Non-linear relationships are further captured through the adjustment of neighboring node features processed by a LeakyReLU activation function, enhancing the model's expressive capability. To emphasize critical relationships between nodes, the study calculates attention weights based on physical distances using a softmax function, where negative distances prioritize closer nodes, facilitating more effective message passing.

Finally, this study applied the calculated attention weight to the neighbor node features, multiplied each element one by one, and obtained the weighted neighbor node features, that is, the final weighted message result, as shown in Equation (6). In this way, when considering the relationship between node j and other nodes, the weight of node features is dynamically adjusted according to its global importance (through $attention_{weights}$) and the importance of different parts of node features to the task (through $attention_{coefficents}$). This allows the model to learn how to focus on different parts of the node in different contexts, and to better adapt to different tasks and graph structures.

$$message_{result} = attention_{weights} \odot attention_{coefficents} \odot x_j' \#(6)$$

Ultimately, the computed attention weights are applied element-wise to the neighbor features, yielding a weighted message output. This process enables the model to dynamically adjust feature weights based on global importance and the relevance of

specific feature parts to the task. Through these mechanisms, the GSAN effectively combines location and feature information in graph message passing, adapting well to node spatial relationships while maintaining structural integrity and focusing on proximate, significant neighbor nodes for efficient information transfer.

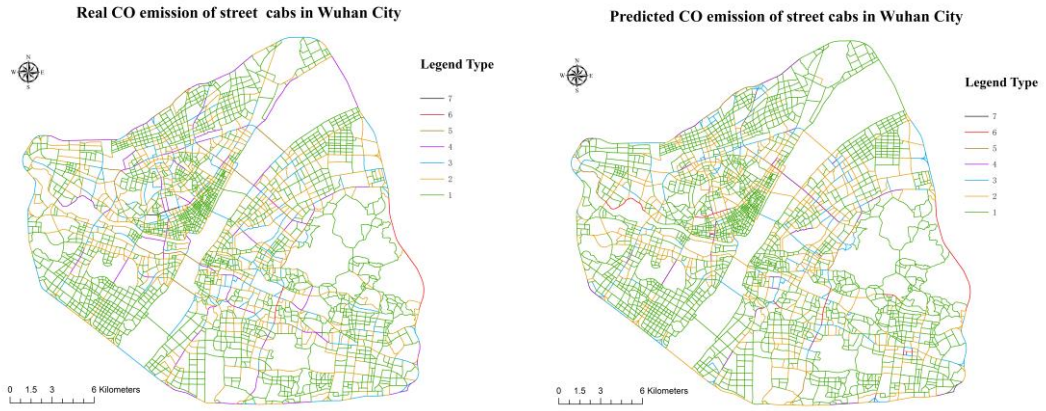
4. Experiments and Results

In this research, the task of predicting taxis' CO emissions in the streets of Wuhan was treated as a supervised node classification task. The data in Section 2 were split into the training, validation, and testing sets at a ratio of 6:2:2. Specifically, 60% of the data was used to train the model's parameters, 20% was used to adjust the model's hyperparameters to prevent overfitting on the test set, and the remaining 20% was used to evaluate the model's performance and ability to generalize.

4.1 Experimental results

The present research demonstrated the feasibility of utilizing the SAGE-GSAN model for predicting the CO emissions of taxis in the streets of Wuhan, achieving a maximum prediction accuracy of 81.4%[3]. A comparison between the true and predicted values is illustrated in Figure 4.1.

[3] The experimental part of the data and the core code have been uploaded to github: https://github.com/zou9229/CO_Predict_Code/tree/master



a) Real CO emissions of taxis b) Predicted CO emissions of taxis

Figure 4.1 Predictions of the SAGE-GSAN model regarding the street-level CO emissions of taxis. In the legend, Types 1–7 represent low to high levels of CO emissions.

It can be observed that the majority of CO emission levels in the streets predicted by the model were consistent with the original results. Furthermore, according to the analysis in Figure 2.2, it can be inferred that streets with a higher volume of taxi traffic tended to have higher CO emission levels. This suggests that among the 19 types of street features, those with a higher proportion of automobile-type features have higher CO emission levels, indicating that automobile-related features are important predictors in the model.

In Figure 4.1, it can be seen that the deviation in the predictions for the road-dense area was relatively low, and its error range was often within one level, while the deviation in the predictions for areas with sparse roads was relatively large, which may be due to the nature of the graph's structure, the nodes (i.e., streets) of the sparse area had fewer neighboring nodes and contained relatively fewer features of the street environment in the area and features of individual roads. The presence of missing

features will have some influence on the predictions. In addition, it can be seen that the longer the road, the more deviations in the predictions it produces, which may be due to the lack of environmental features, as this research only focused on street view features and the road network's structure, ignoring the roads' length, traffic flows, and other realistic features, resulting in unreasonable predictions of the CO emissions of longer streets, which could be increased to improve the accuracy in terms of the roads' length and other features (Mohajeri et al. 2015).

It can also be seen that the CO emission levels on urban ring roads are relatively high, which is because ring roads usually connect important areas such as core areas, commercial centers, residential areas, and industrial areas of the city, which have higher traffic demands, and, therefore, the CO emission levels on ring roads are correspondingly higher. In addition, the CO emission levels of the outer ring road of the city were higher than those of the inner ring road, which is mainly because the outer ring road is usually one of the main traffic arteries of the city that is responsible for connecting different areas of the city and the associated traffic with the surrounding areas. The outer ring road often has greater road capacity and high-speed capacity, and with the continuous expansion and development of the city, the population and economic activities gradually spread to the peripheral areas of the city. The outer ring roads usually run through these rapidly developing urban fringe areas.

Specifically, some main roads and Bridges have high CO emission levels. For example, Xiongchu Avenue, as an east-west main road, crosses Wuchang District and Hongshan District, connecting many urban areas and important transportation hubs.

The CO emission level of each section is 3-5, but the predicted result is between 1-6, with a large fluctuation, which may be caused by its geographical structure. As one of the important traffic arteries in Wuhan, Heping Avenue runs through the riverside zone of Wuhan, connecting the downtown area and major urban areas. Due to its geographical location and convenience, the CO emission level of each section is 2-4, and the deviation of the predicted results is relatively small. Hanyang Avenue, as an east-west main road, connects a number of important areas and traffic arteries in Wuhan and bears a large amount of traffic flow. The CO emission level of each section is about level 2, and the deviation of the predicted results is relatively small. As one of the most important Bridges in Wuhan City, the Yangtze River Bridge connects two major urban areas, Hanyang and Wuchang, which are densely populated. Therefore, the Yangtze River Bridge carries a large amount of vehicle flow and pedestrian flow, and its CO emission level reaches level 5, but the predicted result is only level 2, which may be due to the fact that the bridge is a bridge road with few adjacent roads, caused by not getting enough characteristics. The second Wuhan Yangtze River Bridge also has level 5 CO emission, but its prediction results are relatively accurate, reaching level 4. The Baishazhou Bridge and Yangsigang Bridge are not densely populated and the surrounding road network is relatively undeveloped. Their CO emission levels are only level 1 and Level 3, and their prediction results are only level 1 and level 2, with relatively low CO emission levels and relatively accurate prediction.

The bias in the predictions in other cases may be due to poor data quality or insufficient data, or it may be due to the high complexity of the environmental features

themselves, which are difficult to quantify or obtain. In addition to the features selected in this research, the types of buildings around the streets, urban green coverage, the particular topography of certain regions, and the influence of weather conditions on motor vehicle emissions may all lead to bias in the results of the predictions.

In summary, prediction bias can be caused by a variety of reasons. Various considerations are needed to improve the accuracy of the results.

4.2 Model comparison

This research compares the performance differences between the traditional neural network architecture and the SAGE-GSAN model architecture in the task of predicting road emission levels. A Multi-Layer Perceptron (MLP) model (Rumelhart et al. 1986), illustrated in Figure 4.2, was constructed for the research. This model was trained using preprocessed road semantic segmentation features and road CO emission level labels without incorporating road network structure. The model was employed for a task involving 19 input features (street features) and 7 outputs (road CO emission levels). It consisted of two hidden layers with 64 and 32 neurons, respectively, employing ReLU activation functions for non-linear transformations, ultimately providing CO emission level predictions.

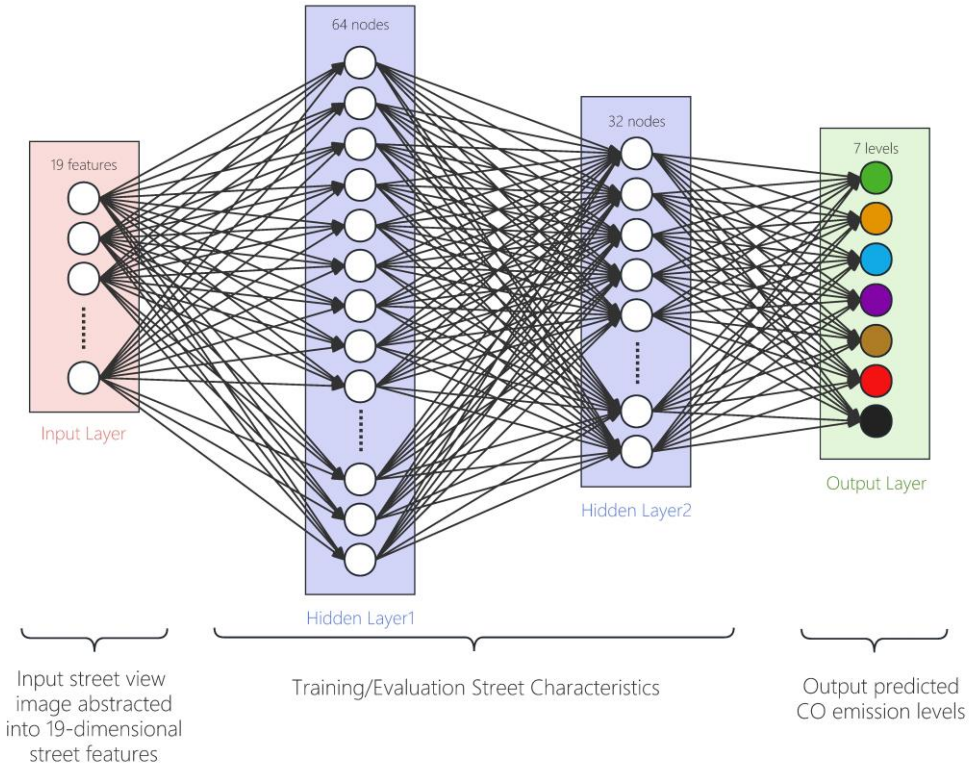


Figure 4.2 MLP predicts street-level CO emissions.

The experimental results of MLP and SAGE-GSAN are presented in Table 3.1. The analysis of the experimental results reveals that the MLP model achieved a classification accuracy of 71.03%, with a notably high MAPE value, suggesting significant prediction errors under certain conditions. Additionally, the maximum error (ME) on the test set was 6, indicating significant discrepancies between predicted and actual labels for individual samples. On the other hand, SAGE-GSAN, incorporating road network structure information for predictions, exhibited performance data for the best model. The SAGE-GSAN model achieved an optimal accuracy of 81.4%, representing an approximate 10% improvement over the MLP model on the test set. Furthermore, SAGE-GSAN demonstrated significant reductions in error metrics (MSE, RMSE, MAE), particularly a substantial decrease in MAPE from 0.8277 in the MLP model to 0.167, showcasing the superior fitting and more accurate predictive

capabilities of the SAGE-GSAN model when dealing with data enriched with topological structure information.

Table 4.1 Comparison of prediction accuracy between MLP and SAGE-GSAN.

Model Name	Best Accuracy	MSE	RMSE	MAE	MAPE	ME	MSL
SAGE-GSAN	81.4%	0.336	0.519	0.180	0.167	4.0	0.042
MLP	71.0%	0.535	0.731	0.360	0.828	6.0	0.166

4.3 Graph convolution layer accuracy comparison

This article compared SAGE-GSAN with other common GNN models (GCN、graph convolutional network via initial residual and identity mapping(GCNII) (Chen et al. 2020)、GAT、graph attention networks version 2(GATv2) (Brody et al. 2021)、Edge(dynamic graph CNN for learning on point clouds) (Wang et al. 2019)、simplifying graph convolutional networks (SG)(Wu et al. 2019)、chebyshev spectral graph convolution networks (Cheb) (Defferrard et al. 2016)、principal neighborhood aggregation for graph networks(PNA) (Corso et al. 2020)、directional graph networks (DGN) (Beaini et al. 2020) and topology adaptive graph convolutional networks(TAG) (Du et al. 2017), with a learning rate of 0.004 and a weight decay coefficient of $5e^{-5}$. After the models had run for 1000 epochs, the highest accuracy was selected, and the results of various evaluation indicators(including mean square error(MSE)、root mean square error(RMSE)、mean absolute error(MAE)、mean absolute percentage error(MAPE)、mean error(ME)、mean squared log error(MSL)) are shown in Table

4.1.

Table 4.1 Comparison of accuracy among different GNN models.

Model Name	Best Accuracy	MSE	RMSE	MAE	MAPE	ME	MSL
SAGE-GSAN	81.4%	0.336	0.519	0.180	0.167	4.0	0.042
GCN	81.2%	0.413	0.643	0.286	0.240	4.0	0.051
GCNII	81.1%	0.453	0.673	0.293	0.251	6.0	0.054
GAT	81.5%	0.456	0.676	0.296	0.268	6.0	0.055
GATv2	81.1%	0.547	0.740	0.323	0.323	6.0	0.067
Edge	81.2%	0.405	0.636	0.272	0.199	5.0	0.048
SG	81.4%	0.459	0.677	0.301	0.257	5.0	0.056
Cheb	81.3%	0.408	0.639	0.271	0.215	6.0	0.048
PNA	79.5%	0.278	0.527	0.182	0.115	5.0	0.032
DGN	81.1%	0.547	0.739	0.323	0.322	6.0	0.067
TAG	81.2%	0.356	0.597	0.254	0.201	4.0	0.044

The results in Table 4.1 indicate that all models achieved similar accuracy, with the best accuracy being around 81%. However, based on the MSE, RMSE, MAE, MAPE, ME, and MSL, it can be seen that the SAGE-GSAN model outperformed all the other models in most indicators. More intuitive results can be observed in Figure 4.3 (to facilitate the graphical representation, the ME metrics have all been multiplied by 0.1).

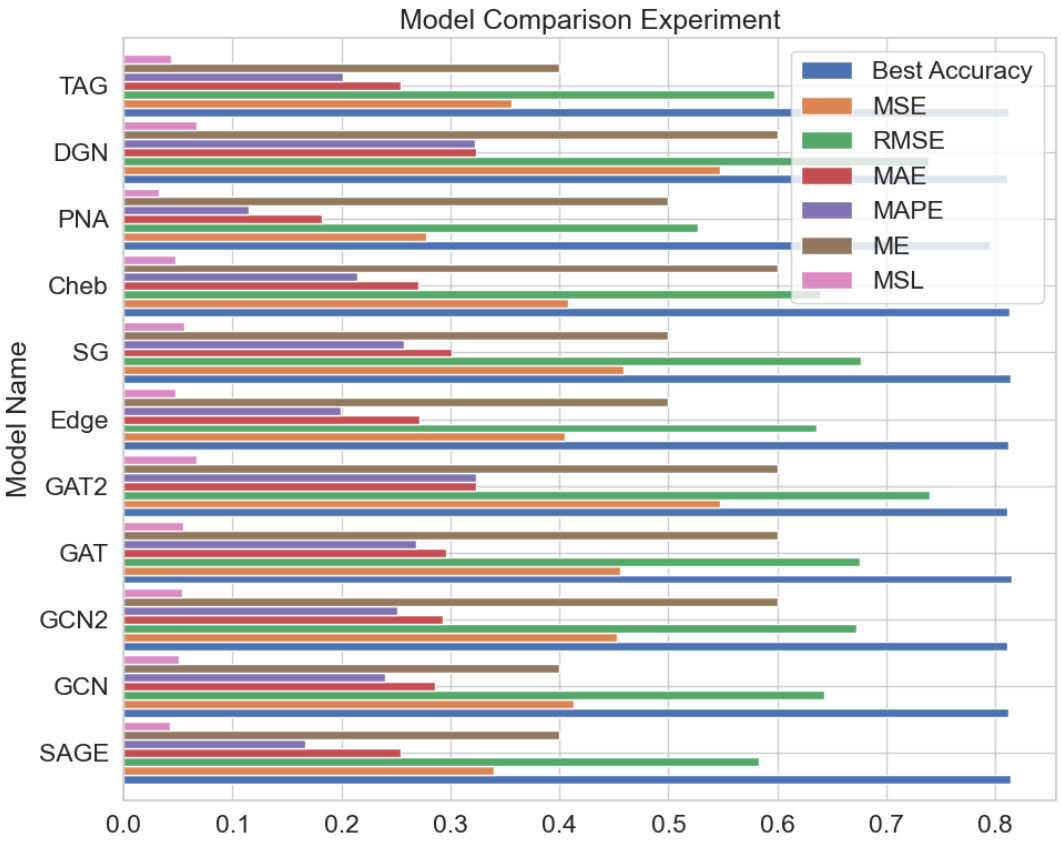


Figure 4.3 Bar chart comparing the models' metrics.

Specifically, although the PNA model had the lowest scores for all indicators except for ME, indicating better predictive performance in terms of error metrics, it had the worst accuracy in predicting CO emissions and was therefore not included in the comparison with other models. Apart from the PNA model, the SAGE-GSAN model had the lowest MSE, RMSE, MAE, MAPE, ME, and MSL values, indicating the smallest overall error compared with the other models.

Moving on, the performance of the GCN, GCNII, GAT, GATV2, Edge, SG, Cheb, DGN, and TAG models was relatively similar, with their best accuracy ranging from 81.1% to 81.5%, and their error metrics being relatively close. Among them, the TAG model's MSE, RMSE, MAE, MAPE, and MSL values were second only to those of the

SAGE-GSAN model, but its best accuracy was lower than that of the SAGE-GSAN model. In addition, this research ran all 11 models for 1000 epochs to test the accuracy of their predictions, and the change in the training loss is shown in Figure 4.4.

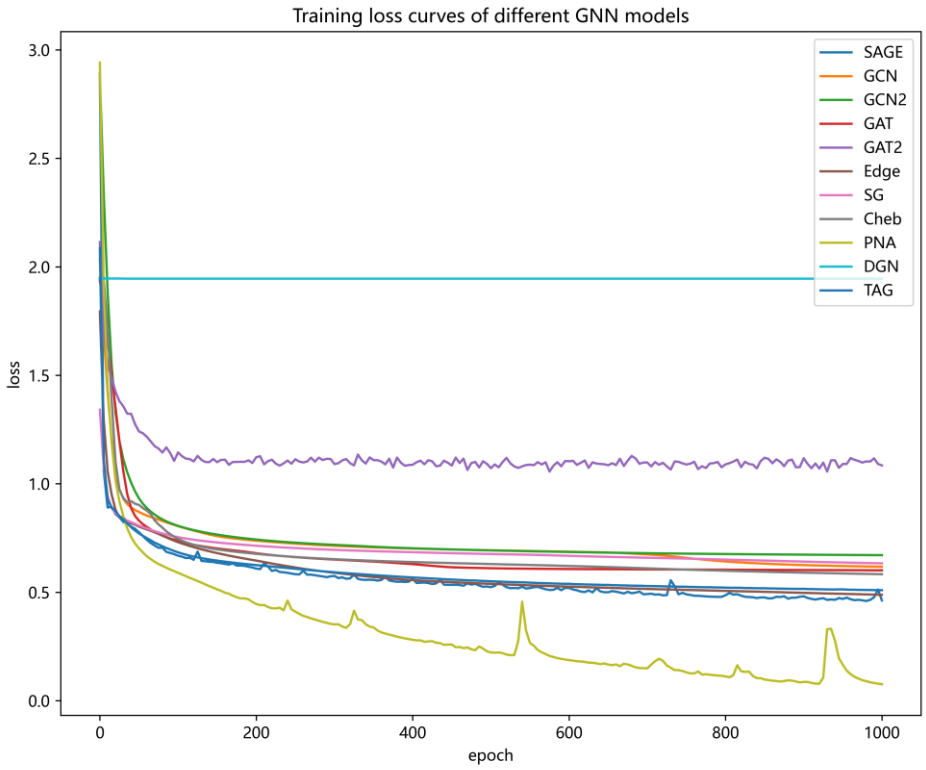


Figure 4.4 Training loss curves for different models.

As shown in Figure 4.4, when the SG model was used, the training loss was low at the beginning, which indicated that the model already had a good ability to extract features without training. PNA had the fastest convergence speed and the best results in terms of the evaluation index, but the lowest accuracy. Some models consistently outperformed the other models; for example, SAGE-GSAN, Edge, and TAG consistently had lower loss values than the others, suggesting that they were better at narrowing the difference between the predicted and actual values in the experiments in

this research. At the same time, these models converged faster than the other models. The models converged after about 20 durations, while the GCNII and Cheb models even converged after about 100 durations. In general, most of the GNN models had low loss values and high accuracy. The results showed that the GNN model has good generalization ability and can accurately predict the CO emission levels of taxis in urban streets.

Overall, according to these results, it can be concluded that the SAGE-GSAN model performs the best in the task of predicting CO emissions on Wuhan's streets, achieving high prediction accuracy while maintaining low evaluation metrics and a good fit. This demonstrated that in our experiment, the SAGE-GSAN model with the LSTM aggregation function was able to capture the correlation and dynamics among the nodes better, improving the predictive performance of the model. Therefore, the SAGE-GSAN-based prediction model was an effective method for predicting CO emissions on Wuhan's streets. Additionally, it can be seen that other GNNs also achieved high accuracy, indicating that GNNs are effective for predicting carbon emissions on urban streets.

4.4 Ablation experiment

In order to investigate which category of objects in the road features had the most significant impact on the predictions, we conducted ablation experiments by removing one category of features from the 19 categories. The SAGE-GSAN model used in the experiment served as the standard, and the results are presented in Table 4.3.

Table 4.3 Results of the ablation study.

Feature class	Best accuracy	MSE	RMSE	MAE	MAPE	ME	MSL
Car	81.4%	0.320	0.565	0.242	0.177	4.0	0.040
Road	81.4%	0.318	0.564	0.240	0.171	4.0	0.040
Building	81.4%	0.323	0.568	0.237	0.183	4.0	0.040
Wall	81.6%	0.312	0.558	0.233	0.184	4.0	0.039
Fence	81.5%	0.317	0.563	0.233	0.172	4.0	0.039
Trees	81.7%	0.322	0.567	0.247	0.169	4.0	0.040
Grass	81.4%	0.300	0.548	0.224	0.167	4.0	0.037
Traffic signs	81.4%	0.292	0.54	0.232	0.152	4.0	0.037
Bus	81.7%	0.317	0.563	0.238	0.186	4.0	0.040
Motorcycle	82.0%	0.305	0.552	0.234	0.158	4.0	0.038
Trucks	81.4%	0.299	0.547	0.231	0.168	4.0	0.038
Sky	81.3%	0.310	0.557	0.236	0.169	5.0	0.039
Poles	81.4%	0.313	0.560	0.235	0.178	4.0	0.0389
Lights	81.4%	0.316	0.563	0.237	0.177	4.0	0.039
Pedestrians	81.4%	0.313	0.560	0.233	0.181	4.0	0.039
Cyclist	81.4%	0.318	0.564	0.236	0.179	4.0	0.039
Trains	81.4%	0.310	0.557	0.233	0.180	4.0	0.039
Bicycles	81.4%	0.334	0.578	0.244	0.200	4.0	0.041
Sidewalk	81.4%	0.319	0.565	0.239	0.193	4.0	0.040

This research removed each category of features individually, including cars, roads,

buildings, walls, fences, trees, lawns, signs, buses, motorcycles, trucks, sky, poles, lights, pedestrians, cyclists, trains, bicycles, and sidewalks. By analyzing the changes in the accuracy of the predictions, we could identify which category of features had the greatest impact on the model's performance.

The results indicated a relatively consistent accuracy among the different features in street view images, as the values of all metrics were relatively small, indicating the good performance of the model for predicting the CO emissions of taxis in Wuhan, based on the features provided. The best accuracy achieved by all the feature categories ranged from 81.4% to 82.0%, with the MSE ranging from 0.292 to 0.334, the RMSE ranging from 0.54 to 0.578, the MAE ranging from 0.224 to 0.247, the MAPE ranging from 0.152 to 0.2, and the MSL ranging from 0.037 to 0.041. The ME of all features except for "sky" was 4.0. Among them, the "traffic signs" feature had the smallest MSE, RMSE, MAPE, and MSL, indicating that the model fit better to the data after losing this feature. This suggests that the "traffic signs" feature has a negative correlation with the accuracy of the model. This observation can be easily justified by the fact that the number of traffic signs reflects the complexity of the traffic situation in the area, with more traffic signs indicating more sophisticated road traffic, as well as higher traffic volumes, which logically causes a rise in CO emissions. However, when the proportion of the "car" features is low, this may lead to incorrect judgments and a decrease in the model’s accuracy. Conversely, the "bicycles" feature had the largest MSE, RMSE, MAPE, and MSL, indicating a lower fit between the model and the data after this feature was lost. As bicycles are non-motorized vehicles and do not produce CO

emissions, the higher the prevalence of bicycles in the traffic flow, the lower the CO emissions tend to be. Based on the experiments and analysis conducted above, it can be deduced that to reduce CO emissions in urban areas, it is advisable to promote bicycle use, utilize public transportation, or purchase new energy vehicles that do not produce CO to replace fossil fuel vehicles. Furthermore, it is crucial to undertake sensible urban road traffic planning to ensure a balanced distribution of traffic flow across different roadways.

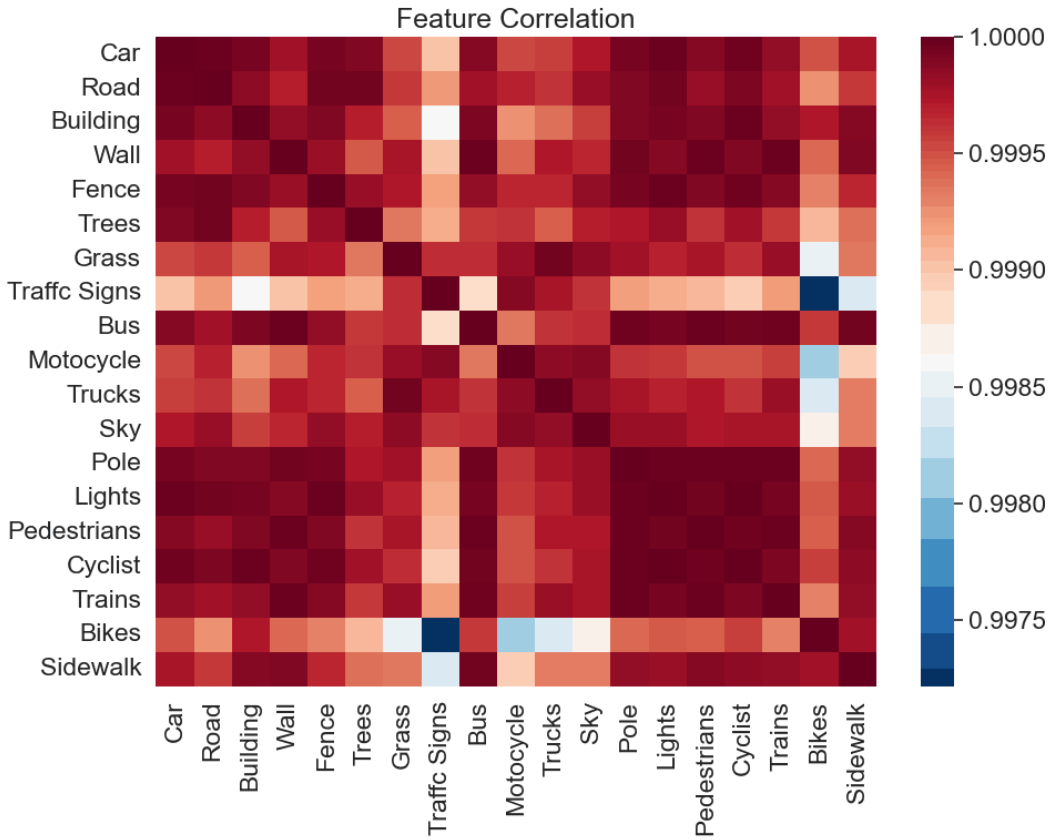


Figure 4.5 Heat map of the features' correlations.

This research also analyzed the correlations between different features using the Pearson correlation coefficient based on Table 4.3, and the results are shown in Figure 4.5. From the heatmap of the features' correlation, it can be observed that the feature of

"bicycles" feature had the highest negative correlation with the "traffic signs" feature, which confirmed the previous analysis of the accuracy of the "traffic signs" feature. Meanwhile, the "bicycles" feature did not exhibit strong positive correlations with other features, but the loss of the "bicycles" feature could reduce the accuracy of the results. Therefore, it can be inferred that since bicycles do not emit CO, the "bicycles" feature may reflect the intensity of CO emissions in the road to some extent, and roads with a higher proportion of bicycles have lower CO emissions.

Through the ablation experiments, this research systematically excluded each feature of the streets and observed the corresponding effects on the accuracy of the predictions. Through the analysis using Pearson’s correlation coefficients, this research found that the exclusion of most features had little effect on the accuracy, while the exclusion of a few specific features had a certain impact on the accuracy. This indicated that these specific street features play an important role in predicting CO emissions. These findings may help us better understand the relationship between streets’ features and CO emissions, and provide valuable information for further optimizing the model used for predictions. Moreover, these findings provide some value for understanding the impact of different street features on CO emissions and can provide a reference for formulating targeted measures to reduce CO emissions and improve urban CO pollution.

4.5 Discussion

The present research introduces a novel application of GNN for predicting CO emission levels by extracting and quantifying street-level environmental attributes that influence their emission potential. Unlike previous overall regional research methods

that rely on macro statistics, this research focuses on fine-grained road-level CO emission prediction, perfectly utilizing the advantages of GNNs, and adopting the analytical perspective of the streetscape to conduct research on the impacts of CO emissions, unlike previous studies where it was not possible to analyze which types of objects in the streetscape had a large impact on CO emissions from the results. This methodology not only improves prediction precision and spatial granularity but also opens up a fresh avenue for micro-scale understanding and management of urban carbon emissions.

The findings of this research bear substantial relevance to the fields of environmental science, transportation planning, and sustainable urban development. By providing precise predictions of street-level carbon monoxide emissions, this research equips urban planners with a crucial tool for accurately identifying high-emission areas. This capability not only aids in optimizing urban traffic flow management but also enables relevant authorities to implement more targeted emission control strategies, significantly enhancing the efficiency and effectiveness of environmental management. The data and analysis provided by this study offer robust support for the formulation of environmental policies, particularly in establishing stricter air quality standards and specific emission reduction targets. With the support of scientific data, policymakers can more effectively evaluate the impact of existing policies and adjust and optimize measures based on real-time data, ensuring the adaptability and foresight of environmental policies. From a public health perspective, this research underscores the direct benefits of improved air quality on residents' health levels. Implementing

measures such as promoting the use of clean energy vehicles and optimizing public transportation systems can effectively reduce the emission of air pollutants, thereby directly lowering the incidence of respiratory diseases, cardiovascular illnesses, and other public health issues related to poor air quality. Furthermore, the integration of big data with advanced algorithmic techniques provided by this study offers a new set of analytical tools and conceptual frameworks for assessing urban carbon emissions globally.

The implementation of these research outcomes is expected to exert far-reaching societal effects. Against the backdrop of accelerating urbanization, the accurate prediction and control of street-level emissions are critical for meeting air quality standards and improving residents' overall quality of life. Additionally, it serves as a robust evidentiary basis for advancing green mobility models, thereby catalyzing the formation of low-carbon, efficient urban transportation systems that significantly contribute to national objectives of reaching peak carbon emissions and attaining carbon neutrality.

4.6 Limitations

There are some limitations and shortcomings in this research. First, our dataset may not be comprehensive enough, and more data are needed to train the model to improve the accuracy of the predictions. Secondly, it can be inferred from the comparison graphs that adding the features of road length and road traffic volume would produce more accurate predictions, and the method in this research still had bias in the predictions for some regions, which may be related to the environmental characteristics

of the region, which needs to be further explored. Finally, more experiments and comparisons are needed to prove the superiority of the method for predicting CO emissions.

5. Conclusion and Outlook

The purpose of this research was to investigate the application of graph neural network model for predicting CO emissions from motor vehicles in urban streets. Specifically, a two-layer SAGE-GSAN network was constructed in this research to predict the CO emissions of taxis in the streets of Wuhan City for one day, with the street view features extracted by DeepLabV3 and the road network's structure used as input. The introduction of road network structure incorporates interconnection information between roads into consideration in the estimation of CO emissions. Simultaneously, the guidance from street view images helps us better understand the degree of correlation between urban features and CO emissions. According to the experimental results of different models, this research found that the graph neural network model was highly accurate in predicting the CO emissions of the city's taxis and the SAGE-GSAN model has the best results in terms of the combined indices and had better convergence in the training loss. Furthermore, through the experimental results, this research found that taxis had a great impact on the results and are an important feature. The ablation experiment showed that bicycles and traffic signs will have a certain impact on the experimental results. This research demonstrates the potential of using machine learning and geospatial techniques to address complex environmental issues in urban areas. Moreover, it may provide a valuable tool for

monitoring and mitigating carbon emissions at the street level, useful insights basis for designing targeted measures for reducing emissions, and important implications for urban planning and air quality management, which is a critical step towards sustainable urban development and environmental management.

This research acknowledges much room for improvement. Firstly, a more comprehensive dataset is essential to enhance prediction accuracy. Incorporating road characteristics and traffic volume features could address biases in regional predictions. Secondly, further exploration of regional environmental factors is needed. Thirdly, conducting additional experiments and comparisons will strengthen the model's validation and demonstrate its effectiveness in predicting CO emissions. Addressing these aspects will refine our research and its utility in urban carbon emission management.

REFERENCES

A. A. 2020. Traffic Flow Prediction Using Graph Convolution Neural Networks. 2020 10th International Conference on Information Science and Technology (ICIST); p. 91-95.

Acheampong AO, Boateng EB. 2019. Modelling carbon emission intensity: Application of artificial neural network. J Clean Prod. 225:833-856.

Alfaseeh L, Tu R, Farooq B, Hatzopoulou M. 2020. Greenhouse gas emission prediction on road network using deep sequence learning. Transportation Research Part D: Transport and Environment. 88:102593.

Bai J, Zhu J, Song Y, Zhao L, Hou Z, Du R, Li H. 2021. A3T-GCN: Attention Temporal Graph Convolutional Network for Traffic Forecasting. ISPRS International Journal of Geo-Information.

Bai L, Yao L, Kanhere SS, Wang X, Liu W, Yang Z. 2019. Spatio-Temporal Graph Convolutional and Recurrent Networks for Citywide Passenger Demand Prediction. CIKM '19. New York, NY, USA; p. 2293-2296.

Beaini D, Passaro S, L E Tourneau V, Hamilton WL, Corso G, Li O P. 2020. Directional Graph Networks. Arxiv E-Prints:2010-2863.

Biljecki F, Ito K. 2021. Street view imagery in urban analytics and GIS: A review. Landscape Urban Plan. 215:104217.

Brody S, Alon U, Yahav E. 2021. How Attentive are Graph Attention Networks? Arxiv E-Prints:2105-14491.

Chen J, Yang ST, Li HW, Zhang B, Lv JR. 2013. Research on Geographical Environment Unit Division Based on the Method of Natural Breaks (Jenks). The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences. XL-4/W3:47-50.

Chen L, Zhu Y, Papandreou G, Schroff F, Adam H. 2018. Encoder-Decoder with Atrous Separable Convolution for Semantic Image Segmentation; p. 1802-2611.

Chen M, Wei Z, Huang Z, Ding B, Li Y. 2020. Simple and Deep Graph Convolutional Networks. Proceedings of Machine Learning Research; p. 1725-1735.

Corso G, Cavalleri L, Beaini D, Li O P, Veli V C Kovi C P. 2020. Principal Neighbourhood Aggregation for Graph Nets. Arxiv E-Prints:2004-5718.

Defferrard MEL, Bresson X, Vandergheynst P. 2016. Convolutional Neural Networks on Graphs with Fast Localized Spectral Filtering. Arxiv E-Prints:1606-9375.

Du J, Zhang S, Wu G, Moura JMF, Kar S. 2017. Topology Adaptive Graph Convolutional Networks. Arxiv E-Prints:1710-10370.

Fan R, Zhang X, Bizimana A, Zhou T, Liu J, Meng X. 2022. Achieving China's carbon neutrality : Predicting driving factors of CO 2 emission by artificial neural network. J Clean Prod. 362.

Fu S, Liu W, Zhang K, Zhou Y, Tao D. 2017. Semi-supervised classification by graph p-Laplacian convolutional networks. Inform Sciences. 560:92-106.

Gao J, Hui HU, Xing P, Gang LI, Wei GU, Liao L. 2018. Emission Characteristics of Pollutants from Motor Vehicles in Wuhan Based on MOBILE 6.2. Journal of Taiyuan University of Technology.

Gao M, Yang H, Xiao Q, Goh M. 2021. A novel fractional grey Riccati model for carbon emission prediction. J Clean Prod. 282.

Gräbe RJ, Joubert JW. 2022. Are we getting vehicle emissions estimation right? Transportation Research Part D: Transport and Environment. 112:103477.

Graves A. 2012. Long Short-Term Memory. In Graves, edito. Supervised Sequence Labelling with

Recurrent Neural Networks. Berlin, Heidelberg: Springer Berlin Heidelberg; p. 37-45.

Hamilton WL, Ying R, Leskovec J. 2017. Inductive Representation Learning on Large Graphs. Arxiv E-Prints:1706-2216.

Hankey S, Zhang W, Le HTK, Hystad P, James P. 2021. Predicting bicycling and walking traffic using street view imagery and destination data. Transportation Research Part D: Transport and Environment. 90:102651.

Howard C, Tcholakov Y, Holz C. 2020. The Paris agreement: charting a low-emissions path for a child born today. Lancet Planet Health. 4(1):e4-e6.

Howlader AM, Patel D, Gammariello R. 2023. Data-driven approach for instantaneous vehicle emission predicting using integrated deep neural network. Transportation Research Part D: Transport and Environment. 116:103654.

Hu S, Gao S, Wu L, Xu Y, Zhang Z, Cui H, Gong X. 2021. Urban function classification at road segment level using taxi trajectory data: A graph convolutional neural network approach. Computers, Environment and Urban Systems. 87:101619.

Huang Y, Wang H, Liu H, Liu S. 2019. Elman neural network optimized by firefly algorithm for forecasting China's carbon dioxide emissions. Systems Science \& Control Engineering. 7(2):8-15.

Ibarra-Espinosa S, Ynoue RY, Ropkins K, Zhang X, de Freitas ED. 2020. High spatial and temporal resolution vehicular emissions in south-east Brazil with traffic data from real-time GPS and travel demand models. Atmos Environ. 222:117136.

Jiang W, Luo J. 2022. Graph neural network for traffic forecasting: A survey. Expert Syst Appl. 207:117921.

Kwiecień J, Szopińska K. 2020. Mapping Carbon Monoxide Pollution of Residential Areas in a Polish City. Remote Sens-Basel. 12(18):2885.

Lawrence MG, Schafer S. 2019. Promises and perils of the Paris Agreement. Science. 364(6443):829-830.

Liu Z, Ciais P, Deng Z, Lei R, Davis SJ, Feng S, Zheng B, Cui D, Dou X, Zhu B et al. 2020. Near-real-time monitoring of global CO_2 emissions reveals the effects of the COVID-19 pandemic. Nat Commun. 11:5172.

Lu X, Ota K, Dong M, Yu C, Jin H. 2017. Predicting Transportation Carbon Emission with Urban Big Data. Ieee T Sust Comput. 2(4):333-344.

Ma L, Xiang L, Wang C, Chen N, Wang W. 2022. Spatiotemporal evolution of urban carbon balance and its response to new-type urbanization: A case of the middle reaches of the Yangtze River Urban Agglomerations, China. J Clean Prod. 380:135122.

Mohajeri N, Gudmundsson A, French JR. 2015. CO2 emissions in relation to street-network configuration and city size. Transportation Research Part D: Transport and Environment. 35:116-129.

Niu D, Wang K, Wu J, Sun L, Liang Y, Xu X, Yang X. 2020. Can China achieve its 2030 carbon emissions commitment? Scenario analysis based on an improved general regression neural network. J Clean Prod. 243:118558.

Nocera S, Ruiz-Alarcón-Quintero C, Cavallaro F. 2018. Assessing carbon emissions from road transport through traffic flow estimators. Transportation Research Part C: Emerging Technologies. 95:125-148.

Pathak SK, Sood V, Singh Y, Channiwala SA. 2016. Real world vehicle emissions: Their correlation with driving parameters. Transportation Research Part D: Transport and Environment. 44:157-176.

Pérez J, de Andrés JM, Borge R, de la Paz D, Lumbreras J, Rodríguez E. 2019. Vehicle fleet characterization study in the city of Madrid and its application as a support tool in urban transport and air quality policy development. Transport Policy. 74:114-126.

Rumelhart DE, Hinton GE, Williams RJ. 1986. Learning representations by back-propagating errors. Nature. 323(6088):533-536.

Saharidis GKD, Konstantzos GE. 2018. Critical overview of emission calculation models in order to evaluate their potential use in estimation of Greenhouse Gas emissions from in port truck operations. J Clean Prod. 185:1024-1031.

Seo J, Park S. 2023. Optimizing model parameters of artificial neural networks to predict vehicle emissions. Atmos Environ. 294:119508.

Shang J, Zheng Y, Tong W, Chang E, Yu Y. 2014. Inferring gas consumption and pollution emission of vehicles throughout a city: ACM; p. 1027-1036.

Sun W, Huang C. 2022. Predictions of carbon emission intensity based on factor analysis and an improved extreme learning machine from the perspective of carbon emission efficiency. J Clean Prod. 338.

Tao J, Qi L, Chu M, Yuqian L. 2019. Computing the CO_2 Emissions of Taxi Trajectories and Exploring Their Spatiotemporal Patterns in Wuhan City; p. 1115.

Turkensteen M. 2017. The accuracy of carbon emission and fuel consumption computations in green vehicle routing. Eur J Oper Res. 262(2):647-659.

Van Houdt G, Mosquera C, Nápoles G. 2020. A review on the long short-term memory model. Artif Intell Rev. 53(8):5929-5955.

Veli V C Kovi C P, Cucurull G, Casanova A, Romero A, Li O P, Bengio Y. 2017. Graph Attention Networks. Arxiv E-Prints:1710-10903.

Wang P, Zhang Y, Hu T, Zhang T. 2022. Urban traffic flow prediction: a dynamic temporal graph network considering missing values. Int J Geogr Inf Sci. 37:1-28.

Wang X, Ma Y, Wang Y, Jin W, Wang X, Tang J, Jia C, Yu J. 2020. Traffic Flow Prediction via Spatial Temporal Graph Neural Network. WWW '20. New York, NY, USA; p. 1082-1092.

Wang Y, Sun Y, Liu Z, Sarma SE, Bronstein MM, Solomon JM. 2019. Dynamic Graph CNN for Learning on Point Clouds. Acm T Graphic. 38(5).

Wu CB, Huang GH, Xin BG, Chen JK. 2018. Scenario analysis of carbon emissions ' anti-driving effect on Qingdao's energy structure adjustment with an optimization model , Part I: Carbon emissions peak value prediction. J Clean Prod. 172:466-474.

Wu F, Souza A, Zhang T, Fifty C, Yu T, Weinberger K. 2019. Simplifying Graph Convolutional Networks. Proceedings of Machine Learning Research; p. 6861-6871.

Xue H, Jiang S, Liang B, Wang W. 2013. A Study on the Model of Traffic Flow and Vehicle Exhaust Emission. Math Probl Eng. 2013:736285.

Yang H, Huang X, Hu J, Thompson JR, Flower RJ. 2022. Achievements, challenges and global implications of China's carbon neutral pledge. Front Env Sci Eng. 16(8):111.

Yang L, Wang Y, Lian Y, Han S. 2020. Factors and scenario analysis of transport carbon dioxide emissions in rapidly-developing cities. Transportation Research Part D: Transport and Environment. 80:102252.

Yarotsky D. 2017. Error bounds for approximations with deep ReLU networks. Neural Networks. 94:103-114.

Yu Y, Si X, Hu C, Zhang J. 2019. A Review of Recurrent Neural Networks: LSTM Cells and Network

Architectures. Neural Comput. 31(7):1235-1270.

Z. W, S. P, F. C, G. L, C. Z, P. SY. 2021. A Comprehensive Survey on Graph Neural Networks. Ieee T Neur Net Lear. 32(1):4-24.

Zeng N, Jiang K, Han P, Hausfather Z, Cao J, Kirk-Davidoff D, Ali S, Zhou S. 2022. The Chinese Carbon-Neutral Goal: Challenges and Prospects. Adv Atmos Sci. 39(8):1229-1238.

Zhang J, Chen F, Wang Z, Wang R, Shi S. 2018. Spatiotemporal Patterns of Carbon Emissions and Taxi Travel Using GPS Data in Beijing. Energies.

Zhang Y, Chen N, Du W, Li Y, Zheng X. 2021. Multi-source sensor based urban habitat and resident health sensing: A case study of Wuhan, China. Build Environ. 198:107883.

Zhang Y, Chen N, Wang S, Wen M, Chen Z. 2023. Will carbon trading reduce spatial inequality? A spatial analysis of 200 cities in China. J Environ Manage. 325(Pt A):116402.

Zhang Y, Liu P, Biljecki F. 2023. Knowledge and topology: A two layer spatially dependent graph neural networks to identify urban functions with time-series street view image. Isprs J Photogramm. 198:153-168.

Zhao P, Kwan M, Qin K. 2017. Uncovering the spatiotemporal patterns of CO2 emissions by taxis based on Individuals' daily travel. J Transp Geogr. 62:122-135.

Zhou J, Cui G, Hu S, Zhang Z, Yang C, Liu Z, Wang L, Li C, Sun M. 2020. Graph neural networks: A review of methods and applications. Ai Open. 1:57-81.

Zhou S, Zhang X, Chu S, Zhang T, Wang J. 2023. Research on remote sensing image carbon emission monitoring based on deep learning. Signal Process. 207:108943.

Zhu D, Zhang F, Wang S, Wang Y, Cheng X, Huang Z, Liu Y. 2020. Understanding Place Characteristics in Geographic Contexts through Graph Convolutional Neural Networks. Ann Am Assoc Geogr. 110(2):408-420.

Z. Chen, T. Zou, Y. Zhang, and N. Chen, "A method and device for predicting and visualizing carbon emissions of motor vehicles on urban streets," Patent CN116663724A, issued 29 Aug. 2023.