

Article

Multi-Source Data-Driven Local-Global Dynamic Multi-Graph Convolutional Network for Bike-Sharing Demands Prediction

Juan Chen ^{1,2}  and Rui Huang ^{1,*}¹ SHU-UTS SILC Business School, Shanghai University, Shanghai 201899, China; chenjuan82@shu.edu.cn² Smart City Research Institute, Shanghai University, Shanghai 201899, China

* Correspondence: 20723326@shu.edu.cn

Abstract: The prediction of bike-sharing demand plays a pivotal role in the optimization of intelligent transportation systems, particularly amidst the COVID-19 pandemic, which has significantly altered travel behaviors and demand dynamics. In this study, we examine various spatiotemporal influencing factors associated with bike-sharing and propose the Local-Global Dynamic Multi-Graph Convolutional Network (LGDMGCN) model, driven by multi-source data, for multi-step prediction of station-level bike-sharing demand. In the temporal dimension, we dynamically model temporal dependencies by incorporating multiple sources of time semantic features such as confirmed COVID-19 cases, weather conditions, and holidays. Additionally, we integrate a time attention mechanism to better capture variations over time. In the spatial dimension, we consider factors related to the addition or removal of stations and utilize spatial semantic features, such as urban points of interest and station locations, to construct dynamic multi-graphs. The model utilizes a local-global structure to capture spatial dependencies among individual bike-sharing stations and all stations collectively. Experimental results, obtained through comparisons with baseline models on the same dataset and conducting ablation studies, demonstrate the feasibility and effectiveness of the proposed model in predicting bike-sharing demand.

Keywords: bike-sharing demand prediction; graph convolutional network; multi-source data



Citation: Chen, J.; Huang, R.

Multi-Source Data-Driven
Local-Global Dynamic Multi-Graph
Convolutional Network for
Bike-Sharing Demands Prediction.
Algorithms **2024**, *17*, 384.

<https://doi.org/10.3390/a17090384>

Academic Editor: Frank Werner

Received: 25 May 2024

Revised: 27 August 2024

Accepted: 29 August 2024

Published: 1 September 2024



Copyright: © 2024 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Within the urban transportation framework, the Bike-Sharing System (BSS) functions as the essential “capillary” transport network, primarily addressing the challenge of the “last mile” commute from subway or bus stations to the final destination. The BSS has witnessed rapid growth and expansion since its inception, owing to its inherent convenience, environmental friendliness, and capacity to alleviate pressure on public transportation systems to a certain extent. However, due to indiscriminate deployment and managerial oversights, issues such as excessive proliferation, haphazard parking, uneven distribution, as well as imbalanced supply have emerged within the BSS.

Addressing these issues requires operators to comprehend users’ cycling demands, enabling proactive planning for scheduling and deployment. Bike-sharing demand prediction is a fundamental problem in managing transportation systems. However, achieving accurate and long-term demand forecasts is challenging for operators due to the multifaceted factors influencing bike-sharing demand.

We categorize numerous factors influencing bike-sharing demand into four main classes across spatial and temporal dimensions:

1. **Spatial Correlation Factors:** Spatial correlation refers to the association between the demand for bike-sharing and spatial locations at the same time. On one hand, the demand for bike-sharing varies depending on the different Points of Interest (POIs) near bike-sharing stations. For example, during the city’s evening peak hours, the demand for bike-sharing near office buildings is higher compared to residential areas.

On the other hand, there exists mutual influence in demand between bike-sharing stations, and this influence is correlated with the distance between stations. The increase in demand at station A, as depicted in Figure 1, leads to a corresponding rise in demand at neighboring station B, whereas the demand at more distant station C remains relatively unchanged.

2. **Spatial Heterogeneity Factors:** Spatial heterogeneity refers to sudden, irregular spatial factors influencing bike-sharing demand. For instance, the addition or closure of stations (referred to collectively as dynamic changes in stations) impacts the demand at surrounding stations. As illustrated in Figure 2, the blue dots represent regular stations, yellow dots represent newly added stations, and purple dots represent stations that have disappeared.
3. **Temporal Correlation Factors:** Temporal correlation refers to fluctuations in demand at the same station over time. For instance, people's travel needs differ on weekdays and weekends, leading to varying demand at the same station at different times. Additionally, since the outbreak of the COVID-19 pandemic, people's travel patterns have been significantly affected, resulting in fluctuations in bike-sharing demand following the pattern of each pandemic wave, as shown in Figure 3.
4. **Temporal Heterogeneity Factors:** Temporal heterogeneity factors refer to the influence of various unforeseen elements at different points in time on the demand for the same station, including holidays and adverse weather conditions. As illustrated in Figure 4, on days affected by rainy or snowy weather, the demand decreases.

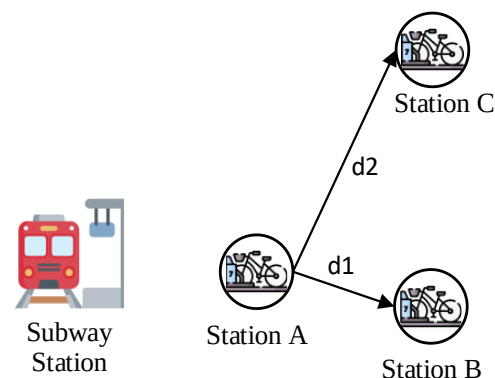


Figure 1. The mutual impact of demand between stations.



Figure 2. The dynamic changes in station locations within specific regions of Chicago.

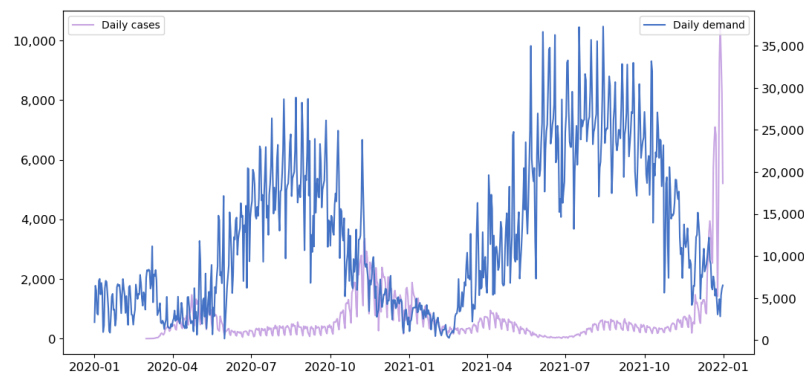


Figure 3. The relationship between daily demand for shared bicycles and daily confirmed COVID-19 cases.

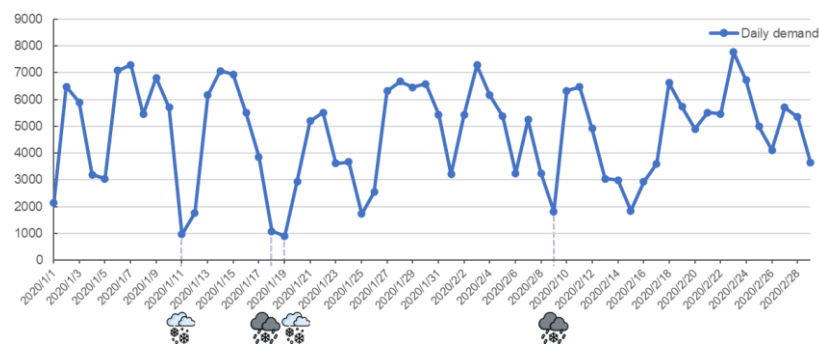


Figure 4. The impact of rainy or snowy weather on the demand for shared bicycles.

Existing research has predominantly utilized deep learning models to learn the spatiotemporal correlation features of bike-sharing demand. Zi et al. divided the bike-sharing demand time series into hourly, daily, and weekly levels, merging them into the model to capture different temporal dependencies [1]. Ma et al. proposed a GC-LSTM model with a spatiotemporal attention matrix, utilizing attention mechanisms to learn spatial correlations [2]. Lin and Peeta introduced a data-driven graph convolutional model (GCNN-DDGF) to learn spatial correlations between bike-sharing stations [3]. Yi et al. proposed the FGST model, leveraging graph generation and spatiotemporal embedding to model spatiotemporal dependencies [4]. Some studies have also incorporated temporal heterogeneity factors such as weather and holidays for more accurate predictions [2,5]. However, certain factors remain underexplored, leading to suboptimal model performance in real-world applications. Firstly, most existing multi-step forecasting models are limited to capturing local spatiotemporal correlations, focusing on nearby temporal and spatial features while ignoring globally similar spatiotemporal patterns. Secondly, the significant impact of the COVID-19 pandemic on travel patterns has not been well-modeled by existing models. Thirdly, for docked bike-sharing systems, the dynamic changes in stations, i.e., station additions and closures, cannot be predicted based on historical demand alone. Yet, these changes significantly affect long-term demand forecasting, and there is currently a lack of research addressing this influential factor.

To address the aforementioned challenges, this paper proposes a Local-Global Dynamic Multi-Graph Convolutional Neural Network (LGDMGCN) model based on the multi-source data-driven approach. This model, operating in the context of the impact of COVID-19 and dynamic changes in stations, not only captures local-global multiscale spatiotemporal correlations but also incorporates diverse information to capture irregular spatiotemporal heterogeneities and dynamics, leading to more accurate multi-step predictions. Specifically, the LGDMGCN model employs a dynamic multi-graph generator to leverage various spatial semantic features and the temporal dynamic semantic information

of stations to create dynamic multi-graphs. Subsequently, based on spatial semantics, these dynamic multi-graphs are separately fed into local or global spatiotemporal multi-graph convolution modules to learn spatiotemporal features on different scales of bike-sharing demand. Further, a temporal attention mechanism is employed to capture the global dependencies of multi-step time. Finally, the output sequence is concatenated with multi-source temporal semantic features such as COVID-19 pandemic data, holidays, and weather, and input into a Gated Recurrent Unit (GRU) for prediction. The contributions of this paper are summarized as follows:

1. In the proposed LGDMGCN model, a local-global spatiotemporal multi-graph convolution module is designed to learn multiscale spatiotemporal correlation features. It fully utilizes the spatial semantic information of stations to construct multiple graphs and employs multi-graph convolution to capture spatial features. Time gate units are used to capture temporal features, and the local-global structure incorporates features both around individual stations and across all stations globally. Additionally, a time attention mechanism is applied to focus on the long-term dependencies in the time series, capturing global relationships.
2. In the context of the COVID-19 pandemic, the model integrates multi-source data, including confirmed COVID-19 cases, holidays, and weather conditions, to learn the heterogeneity features of time. Considering the actual scenario of dynamic station changes, a dynamic multi-graph generator is employed to create dynamic multi-graphs. Subsequently, a spatiotemporal dynamic multi-graph convolution module is used to learn spatiotemporal correlation features with spatial heterogeneity.
3. The proposed prediction algorithm demonstrated its applicability to real-world scenarios by effectively integrating external factors such as holidays, weather, and the COVID-19 pandemic. Experimental results on real datasets showed that the LGDMGCN model exhibited superior performance in multi-step prediction, considering temporal and spatial heterogeneity factors. The model's ability to adapt to varying external conditions and capture complex spatiotemporal dependencies makes it a valuable tool for optimizing intelligent transportation systems. Moreover, through methods like virtual station construction, the algorithm holds potential applicability to dockless bike-sharing and regional-level predictions, providing insights for broader intelligent transportation systems.

2. Related Works

2.1. Analysis and Prediction Methods for the BSS

Early scholars employed traditional statistical and regression methods to analyze and predict bike-sharing traffic. The widely used method for time series applications like bike sharing is the Autoregressive Integrated Moving Average model (ARIMA). Gallop et al. proposed a seasonal ARIMA model based on ARIMA and weather data to analyze bike-sharing numbers in Vancouver [6]. Yoon et al. developed an application system capable of estimating the availability of bicycles at each station, employing an enhanced ARIMA model to forecast the number of shared bicycles at each station. [7]. Zheng et al. introduced a multi-factor regression model based on ARIMA to predict bike-sharing traffic [8]. However, traditional statistical and regression methods require data to undergo a sliding operation to obtain stationary sequences before conducting linear regression. Such methods rely on manual feature engineering by professionals and are not suitable for handling complex nonlinear relationships.

With the advancement of computer processing power and the widespread use of terminal devices generating vast amounts of passive data [9], the era of big data has arrived. Scholars began to explore machine learning and deep learning methods for bike-sharing predictions. Common machine-learning models for bike-sharing system predictions include Random Forest, Gradient Boosting Tree, and Bayesian networks. Yang et al. proposed a random forest model to predict the number of shared bicycles that considers meteorological data, capable of simultaneously handling categorical and numerical variables without

requiring normalization and possessing interpretability [10]. Li et al. introduced a Gradient Boosting Regression Tree (GBRT) model for bike-sharing predictions [11]. Froehlich et al. used Bayesian networks to predict bike availability within 10–20 min [12]. However, machine learning prediction methods fell short of fully modeling the complex spatiotemporal relationships in bike-sharing time series.

In recent years, driven by the development of deep learning, many deep learning models have been applied to bike-sharing system predictions. Simultaneously, bike-sharing has experienced rapid development in many cities worldwide, easing transportation for people but also posing challenges to urban traffic management. Managers and scholars have focused on researching more accurate bike-sharing traffic prediction methods to optimize management, leading to numerous relevant studies. Research on bike-sharing traffic prediction based on deep learning models generally falls into three categories: region-based, cluster-based, and station-based. Region-based studies often involve dividing the area to be predicted into a grid and forecasting the traffic for each grid separately. Bao et al. divided the area into a 5×5 grid, utilized LSTM to capture temporal features, and employed ConvLSTM to capture spatiotemporal features, demonstrating good predictive performance during morning peak hours [13]. However, direct grid division may sever some complete functional areas, rendering the prediction results unsuitable for management applications. Some scholars adopted clustering methods based on different rules to irregularly divide the city into areas or clusters for prediction [14]. Gu et al. used the DBSCAN algorithm [15] to cluster bike parking points, achieving meaningful area division [16]. Mehdizadeh Dastjerdi and Morency employed the Louvain heuristic clustering algorithm [17] to hierarchically cluster bike-sharing stations, conducting community-level demand predictions based on clustering results [18]. While cluster-based methods tend to overlook some fine-grained spatial features [2], station-based predictions are beneficial for the management of docked bike-sharing systems. These predictions can be directly used for operational decision-making and user experience improvement. Lin et al. proposed the GCNN-DDGF model for predicting the demand for each bike-sharing station in New York City, automatically learning hidden relationships between stations [3]. Chai et al. introduced a multi-graph convolutional network for station-level traffic prediction in New York and Chicago [19]. Lee and Ku developed a dual-attention-based recurrent neural network model for demand prediction in Taipei City's YouBike system [20]. Station-level predictions, applicable to both docked and dockless bike-sharing systems, offer the finest granularity and can be adapted to different scenarios, such as constructing stations as needed, clustering to form group stations and more.

2.2. Modeling Spatiotemporal Dependency Relationships

Various deep learning models have been employed to model spatiotemporal dependencies in bike-sharing systems. Foundational models for time-series prediction, such as Recurrent Neural Network (RNN) models [21], Long Short-Term Memory (LSTM) models [22–24], and Gated Recurrent Unit [22], were initially applied to bike-sharing prediction problems. However, these models cannot capture spatial features. In response to this issue, Convolutional Neural Networks (CNNs), commonly used in the image domain, have been applied to bike-sharing systems [25,26]. Nevertheless, CNN models still cannot model non-Euclidean topological features of bike-sharing stations or clusters. Graph Convolutional Networks (GCNs) have a natural advantage in modeling non-Euclidean structures and have been widely applied in various fields. Many scholars have introduced GCNs into bike-sharing systems. Guo et al. constructed a Spatiotemporal Graph Neural Network (ST-GNN) model based on GCN and GRU for bike-sharing demand prediction [27]. To further model complex spatial features, some scholars designed different graphs, constructing Multi-Graph Convolutional (MGCN) models that fuse multiple graphs [19,28]. Others have fused different spatial features on various scales through model structure design, such as creating local and global model structures [29,30]. In addition to spatiotemporal correlations, there exists spatiotemporal dynamics that prevent prediction results from

achieving the desired practical effects. Therefore, some Dynamic Graph Convolutional Networks (DGCNs) have been proposed and applied to the field of traffic flow prediction. F. Li et al. introduced a Dynamic Graph Convolutional Recurrent Network (DGCRN) model for traffic prediction, featuring a hypernetwork designed to auto-generate dynamic graphs, and defined the computation method for dynamic graph convolution [31]. Wang et al. proposed a deep learning model called STSeq2Seq for predicting traffic state multiple steps ahead, where the PAM module can dynamically generate graphs based on input traffic sequence data. While Recurrent Neural Networks are suitable for modeling time dependencies, Graph Convolutional Networks have an advantage in modeling topological spatial relationships of stations. The design of spatiotemporal neural network models and the use of spatiotemporal attention mechanisms can better capture spatiotemporal correlations. However, research on spatiotemporal dynamics in the field of bike-sharing is still limited, and there has been no research considering the impact of station additions and closures on demand predictions.

2.3. External Influencing Factors

While model design can capture the spatiotemporal complexity of bike-sharing, practical applications are affected by various external factors. Common external influencing factors include weather, holidays, and Points of Interest (POI). Liang et al. introduced an external factor fusion module in their model, integrating POI information, weather forecast data, and holiday data for more accurate predictions [29]. Liu et al. designed the NFE module to extract features related to traffic flow and holidays, proposing a dynamic feature fusion strategy for external factors [32]. Ma et al. considered weather information, POI data, and user personal information in their bike-sharing demand prediction model [2]. Z. Pan et al. argued that metadata are the fundamental reason influencing spatiotemporal complex dependency relationships, and model parameters should be generated by a meta-learner. Metadata includes features of graph nodes, such as POI and geographic location, as well as edge attributes, such as distances between nodes. The model extracts metadata for edges and nodes using MetaGAT and MetaRNN, respectively [33].

Moreover, some unpredictable events can significantly impact demand, such as natural disasters, major events, and viral pandemics. Wang et al. proposed the EAST_NET model, a time-aware spatiotemporal network designed to learn dynamic patterns in time series data during unexpected events [34]. Essien et al. extracted information about popular events by crawling Twitter tweets and incorporated weather and traffic accident information to help the model make more efficient and accurate predictions [35]. Jiao et al. used the STL-LSTM model to predict bus passenger flow during the COVID-19 pandemic in Beijing, considering daily passenger flow as the research object [36]. Tu et al. extracted information about the COVID-19 pandemic from search engines and proposed the DST-FNN model, combining deep learning and fuzzy learning for passenger flow prediction [37]. Mehdizadeh Dastjerdi and Morency analyzed the impact of the COVID-19 pandemic on bike-sharing, using the CNN-LSTM model to predict bike demand during the pandemic [18].

3. Symbol Definitions and Problem Description

3.1. Symbols and Relevant Definitions

In this section, we provide an overview of the symbols and relevant definitions used in this paper. The symbols and their meanings are listed in Table 1.

Time Interval: The temporal scope within the research domain is partitioned into T time intervals, with minutes as the smallest unit of time.

Bike-Sharing Static Multi-Graph: Assuming there are N stations in the research area, the static network of bike-sharing is represented by a graph $G(V, E, A)$. Nodes $V = \{v_1, v_2, \dots, v_N\}$ represent the set of bike-sharing stations, edges between two nodes $E = \{e(i, j) | i, j = 1, 2, \dots, N\}$ represent the connections between stations, and the adjacency matrix $A = \{a(i, j) | i, j = 1, 2, \dots, N\} \in \mathbb{R}^{N \times N}$ represents the degree of connection between stations. Based on three different location semantic information, the graph G is assigned

different weights, resulting in the static multi-graph $\mathbf{G} = \{G_{dis}, G_{his}, G_{poi}\}$. It includes three types of graphs: distance graph $G_{dis}(V, E, A_{dis})$, historical connectivity graph $G_{his}(V, E, A_{his})$, and POI Similarity graph $G_{poi}(V, E, A_{poi})$.

Table 1. Symbol explanations.

Symbol	Definition
m	Time step
T	Total time length
D	Total number of days
N	Total number of stations
G	Static graph
A	Adjacency matrix
$\mathbf{G}$	Static multi-graph
$\mathbf{G}^t$	Dynamic multi-graph
X	Bike-sharing flow
COV	Daily confirmed COVID-19 cases
TF	Time features for each day
W	Weather conditions for each time period
h	Historical time steps
p	Future time steps to be predicted
d	Distance threshold
N_F	Number of features

Bike-Sharing Dynamic Multi-Graph: Due to resource optimization, bike-sharing managers make irregular adjustments to stations, leading to station additions or closures. Therefore, based on the static multi-graph $\mathbf{G}$, dynamic graphs are created for each time period to characterize station dynamics. $\mathbf{G}^t$ is used to represent the dynamic multi-graph in each time period, where $\mathbf{G}^t$ includes G_{dis}^t , G_{his}^t and G_{poi}^t . $G_{dis}^t = \{G_{dis}^1, G_{dis}^2, \dots, G_{dis}^T\}$, $G_{his}^t = \{G_{his}^1, G_{his}^2, \dots, G_{his}^T\}$, $G_{poi}^t = \{G_{poi}^1, G_{poi}^2, \dots, G_{poi}^T\}$, $t = 1, 2, \dots, T$.

Bike-Sharing Riding Record: A record of a bike-sharing journey from one station to another, denoted as $R(SS, ES, ST, ET)$, where SS represents the start station, ES represents the end station, ST represents the start time, and ET represents the end time.

Bike-Sharing Flow: The flow of bike-sharing within a time, divided into output flow and input flow, represented by $X^O = \{x_1^O, x_2^O, \dots, x_T^O\}$ and $X^I = \{x_1^I, x_2^I, \dots, x_T^I\}$, respectively. In this research, we focus on predicting bike-sharing demand (output flow), denoted as $X = \{x_1, x_2, \dots, x_T\} \in \mathbb{R}^{T \times N}$. The demand for shared bicycles within a specific time interval t can be represented by $x_t = \{x_t^1, x_t^2, \dots, x_t^N\} \in \mathbb{R}^N$, which can be calculated by Formula (1).

$$x_t = \sum_{n=1}^N | \{ R(SS, ES, ST, ET) | SS = n \wedge ST = t \} | \quad (1)$$

Confirmed COVID-19 Cases: This study also considers the important external influencing factor of confirmed COVID-19 cases, denoted as $COV = \{cov_1, cov_2, \dots, cov_D\} \in \mathbb{R}^D$, where D represents the total number of days in the research time frame.

Time Features: Includes time semantic features for weekdays, weekends, and holidays, represented by $TF = \{tf_1, tf_2, \dots, tf_D\} \in \mathbb{R}^D$.

Weather Condition: Includes temperature, humidity, wind speed, and atmospheric pressure, represented by $W = \{w_1, w_2, \dots, w_T\} \in \mathbb{R}^T$.

3.2. Problem Description

Assuming the current time index is t_0 , given the dynamic multigraphs of the past $h\mathbf{G}^{(t_0-h+1):t_0}$, the demand for shared bicycles $X^{(t_0-h+1):t_0}$, and other external features

$E = [\text{COV} : \text{TF} : W]$, we aim to learn a mapping function $f(\cdot)$, that can predict the future shared bicycle traffic for the next p time steps $Y^{(t_0+1):(t_0+p)}$, as indicated by Formula (2):

$$Y^{(t_0+1):(t_0+p)} = f(G^{(t_0-h+1):t_0}, X^{(t_0-h+1):t_0}, E) \quad (2)$$

4. Local-Global Dynamic Multi-Graph Convolutional Neural Network Model

This study refers to multi-step prediction models [31,38,39] and adopts the encoder-decoder structure. The ST-Conv block proposed by Yu et al. effectively extracts spatiotemporal features. The combination of GCN and Gated TCN in this block allows for prediction without relying on recurrent neural networks [40]. However, the STGCN model is limited to single-step prediction. Inspired by this research, our study employs an improved “GCN + Gated TCN” structure in the encoder to capture spatiotemporal features. Considering station dynamics and multiscale spatial dependencies, a Global-Local Dynamic Spatiotemporal Graph Convolutional Module is proposed. Wang et al.’s study directly uses GCN + Gated TCN to learn spatiotemporal features and introduces a Pattern-Aware Adjacency Matrix (PAM) module to consider pattern similarity features in input time series [38]. In contrast, the dynamic adjacency matrix in this study primarily considers the impact of station dynamic changes on the demand of surrounding stations. Therefore, a dynamic multi-graph generator is designed to generate dynamic multi-graphs containing multiple spatial semantic features and information about station changes. In the decoder, this study not only incorporates conventional external time features but also considers the severity of COVID-19’s impact on bike-sharing demand in the post-pandemic era.

The overall architecture of the proposed model is shown in Figure 5. The architecture consists of an encoder and a decoder. The encoder is composed of a Global-Local Dynamic Spatiotemporal Graph Convolution Network (Global-local Dynamic ST-GCN), capturing dynamic spatiotemporal features to improve prediction accuracy. The decoder comprises Temporal Attention (TA), Time Feature Fusion (TFF), and Gated Recurrent Unit (GRU) to further optimize prediction results by focusing on the dynamic nature of time. Additionally, rich spatial and temporal multi-source features are, respectively, incorporated into the encoder and decoder, providing abundant spatiotemporal semantic information for predictions.

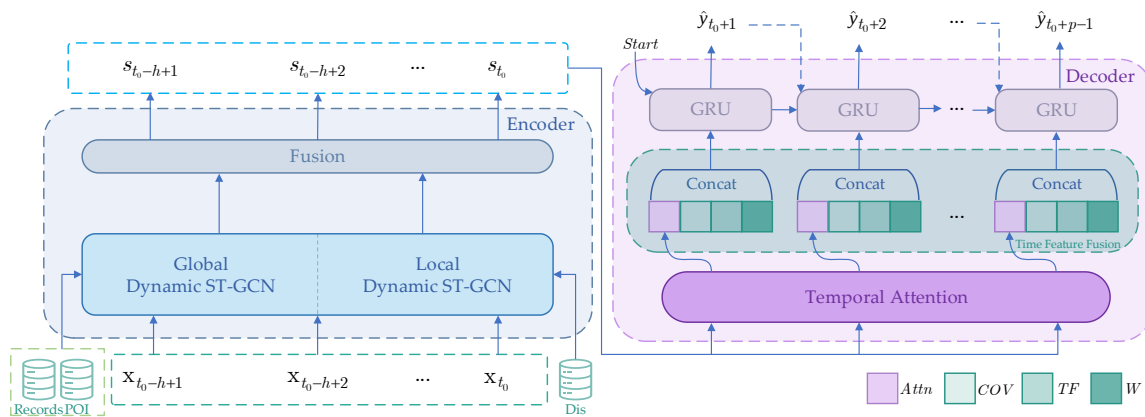


Figure 5. The architecture of LGDMGCN. $\{x_{t_0-h+1}, x_{t_0-h+2}, \dots, x_{t_0}\}$: the input time series. Records: historical ride records between stations. POI: POIs of stations. Dis: distance between stations. $\{s_{t_0-h+1}, s_{t_0-h+2}, \dots, s_{t_0}\}$: the feature sequence produced by the encoder. *Attn*: the output of Temporal Attention. $\{\hat{y}_{t_0-h+1}, \hat{y}_{t_0-h+2}, \dots, \hat{y}_{t_0}\}$: the predictive sequence.

4.1. Local-Global Dynamic Spatiotemporal Graph Convolution

The time series of historical bike-sharing demand $X = \{x_{t_0-h+1}, x_{t_0-h+2}, \dots, x_{t_0}\}$, the station graph network $G(V, E, A)$, and spatial semantic feature data (station distances, POI, etc.) are input to the encoder. The dynamic spatiotemporal graph convolution module captures dynamic features in both time and space dimensions. Gated Temporal

Convolution layers are applied in the time dimension, and a Graph Convolutional Neural Network is used in the spatial dimension.

As shown in Figure 6, the Local-Global Dynamic Spatiotemporal Graph Convolution Module first constructs dynamic multi-graphs with different semantics using the Dynamic Multigraph Generator. These semantics include station distances, historical riding records, station POI information, and information about station dynamic changes. The multiple graphs encompass diverse aspects and varying perspectives of dynamic location semantic information. Merging or concatenating these images into a single model input in a simplistic manner would lead to information loss, thereby failing to capture features on different scales. In the middle section of Figure 6, an example of dynamic multi-graphs can be observed. During time period t , there are three stations (represented by blue solid circles). In time period $t + 1$, the number of stations increases to four (the newly added station on the right changes from a hollow circle to a blue solid circle). In time period $t + 2$, another station is added on the left, making a total of five stations. Secondly, to preserve the diversity and integrity of information to the maximum extent, the dynamic spatiotemporal graph convolution model is divided into two parts: global and local. The Local Spatiotemporal Module focuses on the feature relationships between each station and its adjacent stations, while the Global Spatiotemporal Module focuses on the feature relationships between each station and all other stations. In the upper-middle part of Figure 6, the dynamic distance graph shows connections only between nearby stations (connections between stations are represented by black solid lines, while the connections of absence stations are shown by gray dashed lines). From time period t to $t + 2$, as the number of stations increases, gray dashed lines become black solid lines, but there will still be no connections between distant stations, such as the leftmost and rightmost stations. Therefore, the dynamic distance graph contains local spatial information, which will be input into the local spatiotemporal module. In contrast, the lower part of Figure 6 shows the dynamic historical connection graph and the dynamic POI similarity graph, which are not limited by distance and thus contain global spatial information. Consequently, these two dynamic multi-graphs are input into the global spatiotemporal module together.

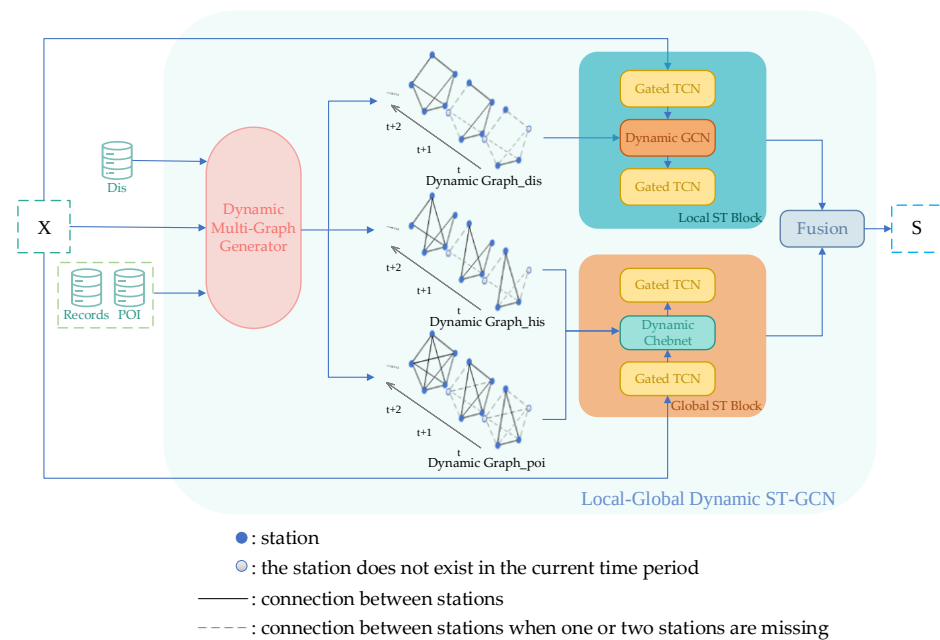


Figure 6. Local-Global Dynamic Spatiotemporal Graph Convolution Module.

Here is the overall algorithm flowchart for Local-Global Dynamic Spatiotemporal Graph Convolution Module, as shown in Algorithm 1.

Algorithm 1. Local-global dynamic ST-GCN algorithm flow

Input: X: historical bike – sharing demand, $R(SS, ES, ST, ET)$: historical ride records, Dis : station distance, and POI

Output: S: fused feature

1. Calculate A_{dis} , A_{his} , and A_{poi} based on Dis , $R(SS, ES, ST, ET)$, and POI
2. Construct multi – graph G_{dis} , G_{his} , and G_{poi}
3. **For** $t \in T$ **Do**
4. **If** station change **Then**
5. Update A_{dis}^t , A_{his}^t , and A_{poi}^t
6. **End If**
6. Generate G_{dis}^t , G_{his}^t , and G_{poi}^t
7. $A_{loc}^t \leftarrow A_{dis}^t$
8. $LDST(X) \leftarrow \Gamma *_{\mathcal{T}} \text{ReLU}(\tilde{D}^{-\frac{1}{2}} \tilde{A}_{loc}^t \tilde{D}^{-\frac{1}{2}} (\Gamma *_{\mathcal{T}} X) \Theta_{gcn})$
9. $A_{glo}^t \leftarrow \alpha A_{his}^t + (1 - \alpha) A_{poi}^t$
10. $GDST(X) \leftarrow \Gamma *_{\mathcal{T}} \text{ReLU}(\Theta_{Cheb} *_{G_{cheb}} (\Gamma *_{\mathcal{T}} X))$
11. $S \leftarrow \text{Concatenate } GDST(X) \text{ and } GDST(X)$
- End For**
12. **Return** fused feature S

4.1.1. Dynamic Multi-Graph Generation Module

The efficacy of Multi-Graph Convolutional Network (MGCN) models in leveraging diverse semantic information of nodes has been substantiated in the context of shared bicycle traffic prediction [19]. Consequently, inspired by the construction of multiple graphs in MGCN models, we extend this approach to the establishment of dynamic multiple graphs.

1. Multi-Graph Construction

The multi-graph construction module involves building three graphs: the distance graph G_{dis} , the historical contact graph G_{his} , and the POI similarity graph G_{poi} . All three graphs have stations as nodes, and the edges represent connections between stations. The key difference lies in the calculation of adjacency matrix weights, explained as follows:

- Distance Graph G_{dis}

Incorporating local positional semantic information, the calculation formula is given by Equation (3):

$$a_{dis}(i, j) = \begin{cases} \frac{1}{Dis(i, j)}, & Dis(i, j) \leq d \\ 0, & \text{others} \end{cases} \quad (3)$$

where a_{dis} represents the weight value of the adjacency matrix A_{dis} for the distance graph, $Dis(i, j)$ denotes the straight-line distance between stations, and d is the set distance threshold. In most cases, bike-sharing stations with closer distances exhibit more similar demands. As the calculation formula for the distance graph follows a threshold structure, with a weight value of 0 for the adjacency matrix when the distance between two stations is beyond the threshold ($Dis(i, j) > d$), the graph constructed based on historical ride records captures the correlation features of nearby stations.

- Historical Connectivity Graph G_{his}

The adjacency matrix of the historical connectivity graph A_{his} is computed based on historical ride records, as per Equation (4):

$$a_{his}(i, j) = \sum_{t_{start}}^{t_{end}} |\{R(SS, ES, ST, ET) | SS = i \wedge ES = j\}| \quad (4)$$

where $a_{his}(i, j)$ denotes the number of bike-sharing ride records from station i to station j . G_{his} encompasses global positional semantic information. In G_{his} , nodes with larger edge weights often indicate diverse functionalities between the corresponding two nodes. For instance, frequent records might exist between residential areas and office buildings during

people's commuting. Thus, the graph established based on historical ride records captures the close relationship between distant stations at a global level.

- POI Similarity Graph G_{poi}

The POI difference graph G_{poi} includes global positional semantic information and is calculated using the TF-IDF method [41] for the similarity of POIs between stations, as shown in Equation (5):

$$a_{poi}(i, j) = \frac{TF_IDF_i \cdot TF_IDF_j}{\|TF_IDF_i\| \|TF_IDF_j\|} \quad (5)$$

where $\cdot$ denotes the dot product of vectors, and $\| \cdot \|$ represents the magnitude of a vector.

In G_{poi} , larger edge weights indicate higher POI similarity between nodes, signifying that stations with similar functional attributes have similar demand patterns. Thus, the graph established based on POI similarity captures the similarity of demand patterns across stations on a global scale. Although both graphs incorporate global positional semantic information, the distinction lies in G_{poi} , where larger edge weights indicate nodes with similar or identical functional attributes.

2. Dynamic Multi-Graph Generation

At each time step, stations undergo dynamic changes. As illustrated in Figure 6, at time t , only three stations exist; at time $t + 1$, an additional station is added; and at time $t + 2$, the count increases to five stations. Therefore, dynamic multi-graphs are generated for each time interval: $G_{dis}^t(V^t, E^t, A_{dis}^t)$, $G_{his}^t(V^t, E^t, A_{his}^t)$ and $G_{poi}^t(V^t, E^t, A_{poi}^t)$.

4.1.2. Local Spatiotemporal Module

Following the dynamic graph generator, three dynamic multi-graphs are generated for each time interval. Since the dynamic distance graph G_{dis}^t reflects local positional semantic information, it is employed in the Local Spatiotemporal Module (Local ST Block). The Local Spatiotemporal Module comprises two Gated Temporal Convolutions (Gated TCN) and one Dynamic Graph Convolution (Dynamic GCN).

The specific structure of the Gated TCN is illustrated in Figure 7 [40]. Dynamic Graph Convolution employs a variant of graph convolution proposed by Thomas et al. [42].

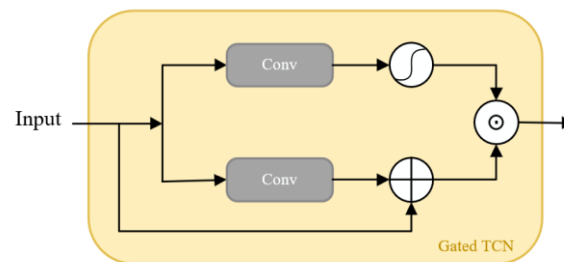


Figure 7. Gated Temporal Convolution Module.

Assuming there are N_F feature sequences, extended to dynamic multidimensional input $X \in \mathbb{R}^{N \times N_F}$. The structure of the Local Spatiotemporal (LDST) Module is as follows: Gated Temporal Convolution-Dynamic Graph Convolution-Gated Temporal Convolution. The calculation formula is given by Equation (6):

$$LDST(X) = \Gamma *_{\mathcal{T}} \text{ReLU}(\tilde{D}^{-\frac{1}{2}} \tilde{A}_{loc}^t \tilde{D}^{-\frac{1}{2}} (\Gamma *_{\mathcal{T}} X) \Theta_{gcn}) \quad (6)$$

where Γ is the convolutional kernel, $*_{\mathcal{T}}$ denotes the convolution operation, A_{loc}^t represents the adjacency matrix of the dynamic distance graph G_{dis}^t , $\tilde{D}$ is the degree matrix, and Θ_{gcn} represents learnable parameters.

4.1.3. Global Spatiotemporal Module

The dynamic historical connectivity graph G_{his}^t and the POI Similarity graph G_{poi}^t , which capture global positional semantic information, are utilized in the Global Spatiotemporal Module. Similar to the Local Spatiotemporal Module, the Global Spatiotemporal Module also comprises two Gated Temporal Convolutions (Gated TCN). However, in the graph convolution part, the Global Spatiotemporal Module employs Chebnet [43]. Given that G_{poi}^t is a fully connected graph, the computational approach of Chebnet is better suited for capturing global positional semantic information. Let x denote the input of the Chebnet, g_θ be the convolutional kernel, and $*_{G_{cheb}}$ represent the Chebnet graph convolution operation, and the computation is as follows:

$$\begin{aligned} A_{glo}^t &= \alpha A_{his}^t + (1 - \alpha) A_{poi}^t \\ g_\theta *_{G_{cheb}} x &= \sum_{k=0}^{K-1} \theta_k T_k(\tilde{L}_{glo}) x \end{aligned} \quad (7)$$

where α is the fusion factor of A_{his}^t and A_{poi}^t , L_{glo} is the global Laplacian matrix, and θ are polynomial coefficients.

The structure of the Global Spatiotemporal (GDST) Module is as follows: Gated Temporal Convolution-Dynamic Chebnet-Gated Temporal Convolution, and the calculation formula is given by Equation (8):

$$GDST(X) = \Gamma *_{\mathcal{T}} \text{ReLU}(\Theta_{Cheb} *_{G_{cheb}} (\Gamma *_{\mathcal{T}} X)) \quad (8)$$

4.1.4. Fusion Module

The final step of the dynamic spatiotemporal graph convolution involving local and global information is the fusion of outputs from the Local Spatiotemporal Module and the Global Spatiotemporal Module:

$$S = [\text{LDST}(X); \text{GDST}(X)] \quad (9)$$

where $S = \{s_{t_0-h+1}, s_{t_0-h+2}, \dots, s_{t_0}\}$ represents the fused features, i.e., the output sequence of the encoder, and $[\ ;]$ denotes the concatenation operation.

4.2. Temporal Attention Decoder

After the encoding phase, the model extracted the spatiotemporal features during dynamic station changes. The decoding process now employs a structured temporal attention prediction algorithm. In practical applications, the demand for bike-sharing is also influenced by external factors. To further model temporal dynamics, a time attention mechanism is introduced in the decoder. This mechanism utilizes rich temporal semantic features to model relationships between different time steps and focuses on the impact of exceptional factors, such as adverse weather conditions or epidemic outbreaks, on the demand for bike-sharing.

4.2.1. Temporal Attention Module

The decoder initially receives the output sequence from the encoder and inputs it into the temporal attention mechanism, computed according to the following Formula (10):

$$attn_t = \sum_{i=1}^T \frac{\exp(e_i)}{\sum_{j=1}^T \exp(e_j)} s_i \quad (10)$$

where e_t represents the input vector for the attention mechanism, and $attn_t$ is the final attention-weighted representation vector [29].

4.2.2. Multi-Source Feature Fusion

To ensure the prediction algorithm's applicability in real-world scenarios, the model integrates additional factors like holidays, weather conditions, and fluctuations due to events like epidemics. These factors are processed through the Multi-Source Feature Fusion, allowing the model to adapt to varying external conditions and providing robust predictions under different circumstances. Let $Attn = \{attn_{t_0+1}, attn_{t_0+2}, \dots, attn_{t_0+p-1}\}$ represent the output of the temporal attention module, which is concatenated with external features on the time dimension, including COVID-19 confirmed case data, time-specific features, and weather conditions:

$$c_t = [attn_t; cov_t; tf; w_t] \quad (11)$$

where c_t represents the fused feature at the time t .

4.2.3. Gated Recurrent Unit (GRU)

The fused feature output is then passed through a Gated Recurrent Unit [44], enabling the model to learn and capture long-term dependencies in bike-sharing demand and producing accurate multi-step predictions.

5. Experiment

5.1. Multi-Source Data Acquisition and Preprocessing

The data preprocessing stage encompasses comprehensive processing of both bike-sharing demand data and multi-source feature data.

Regarding the bike-sharing demand data, the Divvy bike-sharing ride records data for Chicago from the year 2020 are utilized (<https://divvy-tripdata.s3.amazonaws.com/index.html> (accessed on 16 February 2023)). The initial steps involve data cleansing, including the removal of null and outlier records, alongside filtering out records outside the research area to ensure data accuracy and consistency. Subsequently, the raw ride records are transformed into demand data, with ride counts aggregated based on designated time intervals and station identifiers, thus generating spatiotemporal demand data for bike-sharing. For time series forecasting tasks, the demand data are further segmented into time windows to create historical observation data and actual label values for forecasting.

Concerning the preprocessing of multi-source feature data, initial steps involve statistical analysis of historical ride records and their standardization for subsequent construction of multi-graph models. Subsequently, processing of POI data (<https://www.openstreetmap.org> (accessed on 1 April 2023)) surrounding stations entails steps such as data acquisition, cleansing, classification, and matching to acquire information on points of interest around each station. Distance between stations is calculated using latitude and longitude information and applying spherical distance formulas, followed by standardization. Additionally, considering the influence of time, time features including weekdays, weekends, and holidays are extracted. Climate data (<https://www.wunderground.com> (accessed on 22 April 2023)) and COVID-19 confirmed cases data (www.chicago.gov (accessed on 22 April 2023)) are also obtained, processed, and standardized to provide environmental factors relevant to bike-sharing demand. Through these steps, comprehensive preprocessing of multi-source feature data is conducted, furnishing a reliable data foundation for subsequent modeling and analytical tasks.

5.2. Experiment Settings and Parameters

The experiment was conducted on a personal computer equipped with an AMD Ryzen 54600H processor (manufactured by Advanced Micro Devices, Inc., Santa Clara, CA, USA) and an NVIDIA GeForce GTX 1650 graphics card (manufactured by NVIDIA Corporation, Santa Clara, CA, USA). Python (version 3.6) programming language and the PyTorch (version 1.8) deep learning framework were utilized, running on the Windows operating system.

For the proposed LGDMGCN model, each spatiotemporal graph convolutional component consisted of 2 spatiotemporal graph convolution modules, with a hidden size of 64 for all layers. The maximum diffusion step was set to 1. During the training process of LGDMGCN and all the baseline models, the ADAM optimizer was employed with a learning rate of 0.01, which decayed by a factor of 0.5 every 10 epochs. All models were subjected to early stopping to prevent overfitting. The training set, validation set, and test set account for 60%, 20%, and 20%, respectively. The batch size was set to 16.

5.3. Evaluation Metrics

To demonstrate the accuracy of the prediction method, this study employs two evaluation metrics: Mean Absolute Error (MAE) and Root Mean Square Error (RMSE) to assess the prediction results. MAE provides a better reflection of the actual error in the predicted values, while RMSE is more sensitive to outliers compared to MAE. Smaller values of MAE and RMSE indicate that the predictive model accurately describes the experimental data. The formulas for both evaluation metrics are as follows:

$$\text{MAE} = \frac{1}{n} \sum_{i=1}^n |Y_i - \hat{Y}_i| \quad (12)$$

$$\text{RMSE} = \sqrt{\text{MSE}} = \sqrt{\frac{1}{n} \sum_{i=1}^n (Y_i - \hat{Y}_i)^2} \quad (13)$$

5.4. Experimental Results and Analysis

Considering the practical applicability of bike-sharing demand prediction, time is divided into 2 h intervals as a time step. Based on previous research and observations of the data, it has been found that bike-sharing demand exhibits noticeable periodicity within the same time intervals every day. Therefore, utilizing historical data spanning one day (12-time steps) enables capturing this periodicity effectively.

5.4.1. Comparative Experiment

In the comparative experiment, the following models are selected as baseline methods:

- LSTM (Long Short-Term Memory): An improved RNN model commonly used for time series prediction [45].
- STGCN (Spatiotemporal Graph Convolutional Network): A spatiotemporal prediction model with a “gated TCN + GCN+ gated TCN” structure of spatiotemporal graph convolution modules. This model innovatively employs a combination of Graph Convolution and Gated Causal Convolution, eliminating the reliance on RNN for prediction. It is particularly effective in modeling relationships between stations based on the road network structure and has shown good prediction results in road speed prediction problems [40].
- DGCRN (Dynamic Graph Convolutional Recurrent Network): This model is based on an encoder-decoder structure, considering the impact of dynamics on traffic prediction, and defines the computation of dynamic graph convolution. It uses a hyper-network to dynamically generate an adjacency matrix at each step of the RNN, allowing the model to adaptively capture changing spatial relationships. The generated dynamic matrix is merged with the original road network matrix to provide a richer spatial representation [31].
- AttConvLSTM (Attention Convolutional LSTM): An encoder-decoder framework, utilizing convolutional and ConvLSTM units, is employed to identify complex features, capturing spatiotemporal characteristics and the impact of pick-up and drop-off interactions on citywide passenger demand. The model embeds an attention mechanism to depict the influence of underlying citywide travel patterns [46].
- STSeq2Seq (Spatiotemporal Sequence-to-Sequence): A model that captures dynamic relationships between time series data points by modeling non-local spatial correlations

between traffic sections. It utilizes an encoder-decoder architecture where the encoder processes the input sequence and the decoder generates the output sequence. It also introduces an attention mechanism in the decoder to model dependencies between different time steps [38].

Using historical demand data from the past 12 time intervals to predict the future multi-step demand for shared bicycles at the station level, the LGDMGCN model was compared with other baseline models using the Chicago shared bicycle dataset. Experimental results are presented in Table 2:

Table 2. Comparative experimental results.

Pred_Step	4 h		8 h		12 h		16 h	
Model	MAE	RMSE	MAE	RMSE	MAE	RMSE	MAE	RMSE
LSTM	0.4099	0.4525	0.4141	0.5410	0.4534	0.5983	0.5167	0.7246
STGCN	0.0380	0.2342	0.0506	0.2564	0.0560	0.2675	0.0765	0.2760
DGCRN	0.0915	0.2679	0.0951	0.3322	0.0976	0.3520	0.0985	0.3718
AttConvLSTM	0.0959	0.2548	0.1016	0.2682	0.1044	0.2690	0.1069	0.2702
STSeq2Seq	0.0398	0.2474	0.0408	0.2629	0.0432	0.2686	0.0498	0.2761
LGDMGCN	0.0350	0.2476	0.0367	0.2628	0.0370	0.2629	0.0371	0.2679

Comprehensive comparative experimental results indicate that various time series forecasting models exhibit certain differences in performance across different prediction time steps, as evaluated by the MAE and RMSE metrics. Specifically, it can be observed that:

1. The LSTM model, which solely considers temporal relationships between time series, lacks spatial correlations, resulting in overall inferior performance compared to models that incorporate spatiotemporal relationships.
2. The AttConvLSTM model, by combining Convolutional and LSTM layers and incorporating an attention mechanism, demonstrates the effectiveness of modeling spatiotemporal dependencies and attention mechanisms for bike-sharing station prediction compared to LSTM.
3. The STGCN model, which constructs stations into graphs and utilizes graph convolution for spatial modeling, performs well in capturing the spatial structure of bike-sharing stations. By employing TCN modules for temporal modeling, it achieves good performance in 4 h predictions. However, as STGCN is a single-step prediction model, its performance decreases significantly in multi-step predictions.
4. Both the STSeq2Seq and DGCRN models capture the spatiotemporal characteristics of bike-sharing station demand through sequence-to-sequence model structures, exhibiting good performance in multi-step predictions. With increasing prediction steps, their performance remains relatively stable. Additionally, both models consider spatiotemporal dynamics. The STSeq2Seq model dynamically adjusts the weights of adjacency matrices in graph convolution based on the dynamic temporal information of input sequences, while the DGCRN model generates dynamic graphs through Hyper Network modules and node embedding features. Experimental results indicate that the STSeq2Seq model outperforms DGCRN, possibly because of its adoption of spatiotemporal graph convolution modules in the encoder, benefiting from the good performance of STGCN in single-step prediction, allowing it to better capture spatiotemporal correlations. It is also possible that the dynamic graph generation approach of the STSeq2Seq model is more suitable for bike-sharing station demand prediction than that of the DGCRN model.
5. The LGDMGCN model exhibits excellent performance across all prediction time steps, with lower MAE and RMSE compared to other models. This indicates the effectiveness of the LGDMGCN model in multi-step bike-sharing demand prediction. Moreover, the LGDMGCN model demonstrates relative stability in performance across different prediction time steps, indicating good robustness. It effectively captures complex rela-

tionships between spatiotemporal data, thereby improving prediction accuracy and reliability. In addition to considering spatiotemporal correlations, the LGDMGCN model also accounts for spatiotemporal heterogeneity, resulting in further performance improvements compared to the STSeq2Seq model.

5.4.2. Ablation Experiment

To further validate the effectiveness of each module in the LGDMGCN model, ablation experiments were conducted on different components, including the Dynamic Multi-Graph Generation module (LGDMGCN *w/o* At), Global Spatiotemporal Graph Convolution module (LGDMGCN-Global), Local Spatiotemporal Graph Convolution module (LGDMGCN-Local), Time Attention Mechanism (LGDMGCN *w/o* Attn), and External Time Feature (LGDMGCN *w/o* Ext). The results are summarized in Table 3.

Table 3. Ablation experiment results.

Model	MAE (16 h)	RMSE (16 h)
LGDMGCN <i>w/o</i> At	0.0483	0.2896
LGDMGCN-Global	0.0373	0.2957
LGDMGCN-Local	0.0395	0.2841
LGDMGCN <i>w/o</i> Attn	0.0410	0.2754
LGDMGCN <i>w/o</i> Ext	0.0378	0.2849
LGDMGCN	0.0371	0.2679

According to the results shown in Table 3, removing the dynamic graph generation module from the LGDMGCN model (LGDMGCN *w/o* At model) led to an increase in MAE and RMSE, indicating that the dynamic graph generation module positively impacts the model's performance by utilizing multi-source data in both time and space. This suggests that the introduction of this module plays a crucial role in more accurately modeling the spatiotemporal graph. Secondly, removing the local spatiotemporal graph convolution module in the LGDMGCN-Global model resulted in a slight increase in MAE and RMSE. This confirms the effectiveness of local spatial features in capturing local spatial correlations at the station level. Removing the global spatiotemporal graph convolution module in the LGDMGCN-Local model led to an increase in MAE and RMSE, indicating that the global spatiotemporal graph convolution module positively influences the model's performance by utilizing global spatial features of the stations. Comparing the LGDMGCN-Global and LGDMGCN-Local models, it is observed that the global features of station space have a more significant impact on the model's performance compared to local features. The combination of both provides the model with different levels of spatial semantic information, which is beneficial for modeling spatial correlations on different scales.

On the other hand, removing the time attention mechanism in the LGDMGCN *w/o* Attn model and the external time feature in the LGDMGCN *w/o* Ext model both resulted in an increase in MAE and RMSE. This indicates that these two modules play a crucial role in modeling temporal relationships and utilizing external features. In the end, the complete LGDMGCN model achieved the best performance in terms of MAE and RMSE, demonstrating the importance of integrating all modules to enhance the overall model performance and further emphasizing the effectiveness of the LGDMGCN model.

5.4.3. Time Granularity Analysis

Time granularity refers to the time interval or duration covered by each data point during data collection or analysis. The size of temporal granularity significantly influences the capture of information, the complexity of models, and the accuracy of predictions. As shown in Figure 8, the impact of different time granularities on the prediction average metrics of the LGDMGCN model.

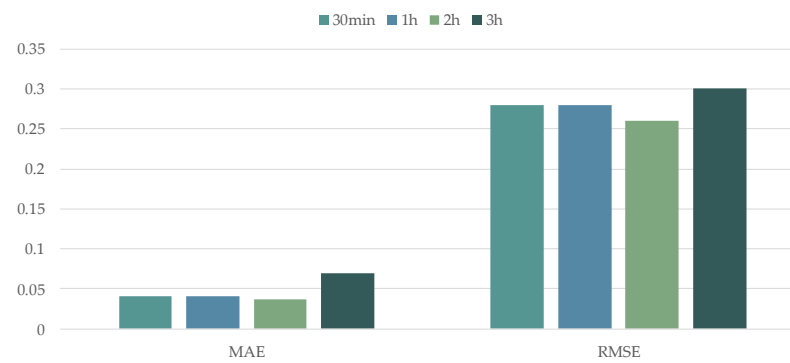


Figure 8. The metrics of different time granularities.

It can be observed that the model performs best when the time granularity is 2 h. Results for 30 min and 1 h intervals are slightly inferior to the 2 h interval, while performance deteriorates the most with a 3 h interval. Generally, smaller temporal granularities can provide more temporal features, making the model more complex and detailed. This can help the model capture more patterns and regularities during training, thereby enhancing prediction accuracy. The time granularities of 30 min and 1 h are frequently adopted in research concerning shared bicycle demand prediction [1,19]. However, our study focuses on the Chicago area during the COVID-19 pandemic, where some stations were closed due to minimal demand. These factors lead to sparse time series data compared to normal periods. In such cases, increasing the time granularity allows each data point to represent a longer period, reducing the proportion of missing data points. Although this approach may sacrifice some analytical precision, we are studying multi-step and relatively long-term forecasting, where some disregard for detailed information is acceptable. Therefore, the key lies in selecting the appropriate temporal granularity based on specific requirements and data characteristics to achieve the optimal balance.

5.4.4. Spatial Correlation and Heterogeneity Effects Analysis

In the selected research scope of the experiment, comprising 90 stations, the geographical features of each station and the spatial characteristics between stations influence the demand for each station as well as the correlation between the demands of different stations. Within the same period, if a station has a higher number of Points of Interest (POIs) or if the POIs are more popular during that period, the demand for shared bicycles at that station tends to be higher. Similarly, if the distance between two stations is shorter, the correlation in demand for shared bicycles between those two stations is higher. The weighted heatmap of the distance adjacency matrix for 40 selected stations on 1 January 2020 is shown in subplot (a) on the left side of Figure 9a. The horizontal and vertical axes represent the chosen 40 stations, and each small square in the figure represents the weight between station i and station j , which is solely related to the distance between the stations. The darker color indicates a smaller distance between stations, implying a higher correlation in demand for shared bicycles between the two stations.

When spatial heterogeneity factors are present, the described relationship between a station's own demand and the correlation of demand between stations can change. For instance, considering the scenario of a station disappearing, after the disappearance of a station, the demand for shared bicycles tends to increase for stations closer to the vanished station. In the proposed LGDMGCN model, the dynamic graph generation module dynamically adjusts the weights of the adjacency matrix between stations based on the appearance or disappearance of stations, enabling the model to capture spatial heterogeneity changes. Figure 9b illustrates the weighted heatmap of the distance adjacency matrix on 17 March 2020, where station 7 has been removed. Comparing the two sub-figures of (a) and (b), it can be observed that the color of the small squares (7, 6) and (6, 7) has become lighter, indicating a decrease in the weight between stations 6 and 7; meanwhile,

the color of (6, 4) has become darker, indicating an increase in the weight between stations 6 and 4. After the removal of station 7, the dynamic graph generation module decreases the weight of that station while increasing the weight of stations closer to it. When spatial heterogeneity factors occur, the model, through such dynamic adjustments, focuses on the changes in heterogeneity of existing stations.

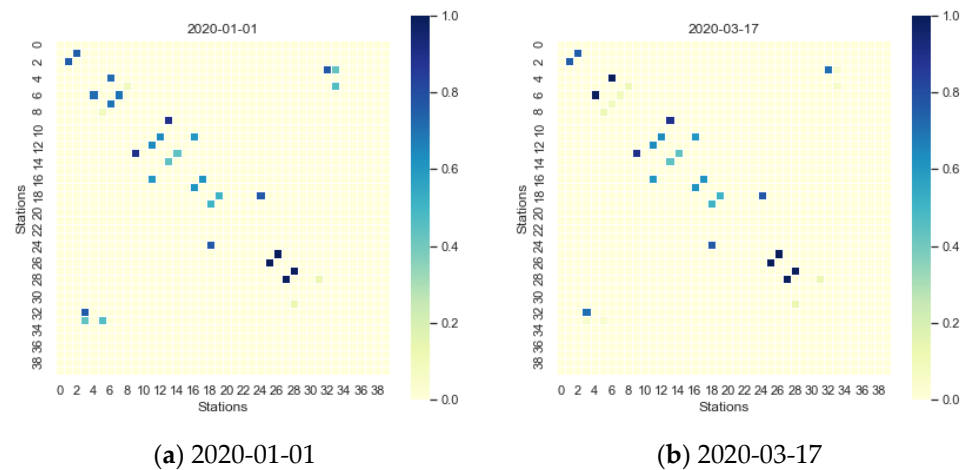


Figure 9. Dynamic adjacency matrix weighted heatmaps at different times.

5.4.5. Analysis of Temporal Correlation and Heterogeneity Effects

The demand for shared bicycles is correlated with temporal changes, and whenever factors with periodicity and seasonality vary, the demand also changes accordingly. However, when temporal heterogeneity factors come into play, it can have a significant impact on demand. In the decoder part of the LGDMGCN model, through the combined design of multiple time feature data sources and temporal attention, the model can dynamically adjust the corresponding changes in demand based on external features. Considering the backdrop of the COVID-19 pandemic, shared bicycle demand fluctuates in response to the variations in pandemic cases. Figure 10 illustrates the correlation between the daily confirmed COVID-19 cases and the demand for bike-sharing predicted by the LGDMGCN model. The solid purple line represents the daily confirmed COVID-19 cases, while the dashed blue line represents the corresponding predicted demand for shared bicycles. Each data point on the blue line corresponds to the model's forecast for the day, demonstrating how demand reacts to changes in pandemic severity. The chart clearly shows the trend of predicted reduction in demand during periods of increased COVID-19 cases, and vice versa. Although there is a slight observable lag, it essentially reflects the trend of temporal heterogeneity effects on bike-sharing demand during the pandemic.

We also computed the MAE for each model during both the severe periods (November) and the relatively mild periods (June) of the COVID-19 pandemic.

As shown in Figure 11, overall, the performance of the models in November, when the COVID-19 pandemic was severe, was worse than in June, when the pandemic was relatively stable. Specifically, the LSTM, STGCN, DGCRN, and AttConvLSTM models exhibited significant performance discrepancies between these two periods, indicating that these models struggled to capture demand fluctuations during severe pandemic conditions. In contrast, the STSeq2Seq model demonstrated smaller performance differences across the periods, possibly due to its PAM module effectively capturing dynamic changes in the time series. The LGDMGCN model, benefiting from the inclusion of COVID-19 case data and temporal attention mechanisms, achieved the best performance in both periods.

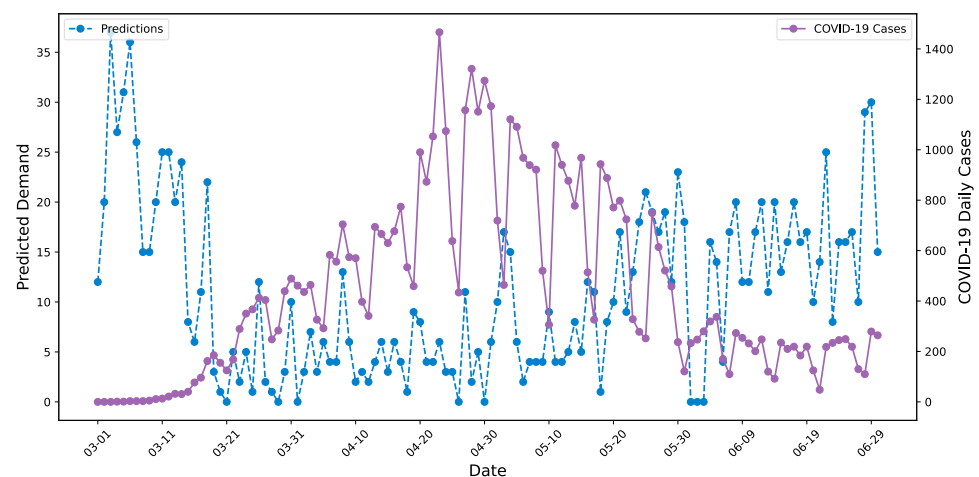


Figure 10. Relationship between daily confirmed COVID-19 cases and predicted demand for bike-sharing.

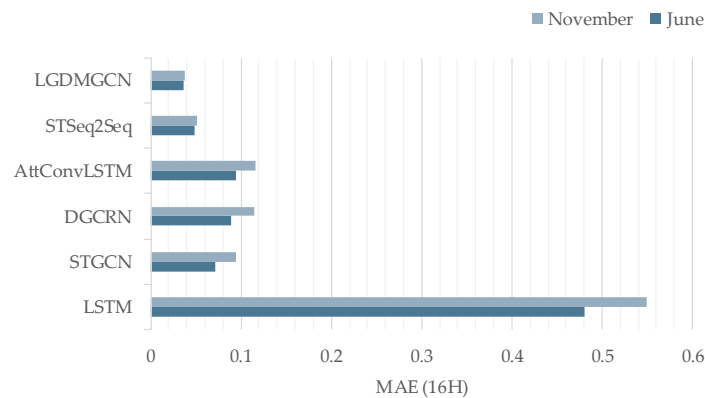


Figure 11. Comparison of MAE for models across different periods.

6. Conclusions

In this study, we aimed to address the fundamental issue of predicting shared bicycle demand in transportation system management. By summarizing various factors influencing shared bicycle demand in the spatial and temporal dimensions, we proposed an innovative LGDMGCN (Local-Global Dynamic Multigraph Convolutional Neural Network) model to more accurately predict shared bicycle demand. Experimental results demonstrate that the model performs exceptionally well in real-world scenarios, especially when considering complex factors such as the COVID-19 pandemic, dynamic changes in station availability, and spatiotemporal heterogeneity, achieving significant performance improvements.

The innovations of this paper are primarily reflected in several aspects. Firstly, we designed a local-global spatiotemporal multigraph convolutional module, leveraging the spatial semantic information of stations to construct multiple graphs. Through multigraph convolution, it captures spatial features, while utilizing time-gated units captures temporal features. This structure allows the model to simultaneously focus on local and global spatiotemporal features, enabling a more comprehensive learning of multiscale correlation features in shared bicycle demand. Secondly, against the backdrop of the COVID-19 pandemic and dynamic changes in station availability, we introduced a dynamic multigraph generator and spatiotemporal dynamic multigraph convolutional module to adapt to changes in demand during pandemic periods and dynamic station availability. This effectively enhances the model's adaptability and predictive accuracy. Lastly, through the introduction of a temporal attention mechanism, we successfully captured global dependencies in time series, thereby enhancing the accuracy of multi-step predictions.

During the experimental phase, we thoroughly validated the LGDMGCN model using real datasets and compared it with existing models. The results indicate that, compared to other models, the LGDMGCN model performs exceptionally well in predicting shared bicycle demand, particularly in complex real-world scenarios. This demonstrates the high practicality and performance of the proposed model in solving the shared bicycle demand prediction problem.

However, predicting shared bicycle demand remains a complex and multidimensional problem. Future research could further deepen the consideration of other factors, such as traffic flow and social media information. And the multi-modal prediction with interactions between different modes of transportation explores the interrelationships among shared bicycles, buses, subways, taxis, and shared scooters. The model developed in this study provides shared bicycle operators with a more accurate and long-term demand prediction tool, offering valuable insights for efficient management and resource allocation in urban transportation systems.

Author Contributions: Funding acquisition, J.C.; methodology, J.C.; project administration, J.C.; software, R.H.; supervision, J.C.; validation, R.H.; visualization, R.H.; writing—original draft, J.C.; writing—review and editing, R.H. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the National Natural Science Foundation of China, grant number 61104166.

Data Availability Statement: The data presented in this study are available on request from the corresponding author. The datasets used during the current study are available at <https://divvy-tripdata.s3.amazonaws.com/index.html>, accessed on 16 February 2023.

Acknowledgments: The authors would like to thank the reviewers for useful suggestions.

Conflicts of Interest: The authors declare there are no conflicts of interest regarding the publication of this paper. The authors have no financial and personal relationships with other people or organizations that could inappropriately influence our work.

Abbreviations

Abbreviation	Full Term
LGDMGCN	Local-Global Dynamic Multi-Graph Convolutional Neural Network
COVID-19	Coronavirus Disease 2019
POI	Points of Interest
GRU	Gated Recurrent Unit
BSS	Bike-Sharing System
Global-local Dynamic ST-GCN	Global-Local Dynamic Spatiotemporal Graph Convolution Network
TA	Temporal Attention
TFF	Time Feature Fusion
Local ST Block	Local Spatiotemporal Module
Gated TCN	Gated Temporal Convolutions
Dynamic GCN	Dynamic Graph Convolution
LDGCN	Local Dynamic Graph Convolutional Neural
LDST	Local Dynamic Spatiotemporal
GDST	Global Spatiotemporal

References

1. Zi, W.; Xiong, W.; Chen, H.; Chen, L. TAGCN: Station-Level Demand Prediction for Bike-Sharing System via a Temporal Attention Graph Convolution Network. *Inf. Sci.* **2021**, *561*, 274–285. [CrossRef]
2. Ma, X.; Yin, Y.; Jin, Y.; He, M.; Zhu, M. Short-Term Prediction of Bike-Sharing Demand Using Multi-Source Data: A Spatial-Temporal Graph Attentional LSTM Approach. *Appl. Sci.* **2022**, *12*, 1161. [CrossRef]
3. Lin, L.; He, Z.; Peeta, S. Predicting Station-Level Hourly Demand in a Large-Scale Bike-Sharing Network: A Graph Convolutional Neural Network Approach. *Transp. Res. Part C Emerg. Technol.* **2018**, *97*, 258–276. [CrossRef]

4. Yi, P.; Huang, F.; Peng, J. A Fine-Grained Graph-Based Spatiotemporal Network for Bike Flow Prediction in Bike-Sharing Systems. In Proceedings of the 2021 SIAM International Conference on Data Mining (SDM), SIAM, Virtual, 29 April–1 May 2021; pp. 513–521.
5. Chen, L.; Zhang, D.; Wang, L.; Yang, D.; Ma, X.; Li, S.; Wu, Z.; Pan, G.; Nguyen, T.-M.-T.; Jakubowicz, J. Dynamic Cluster-Based over-Demand Prediction in Bike Sharing Systems. In Proceedings of the 2016 ACM International Joint Conference on Pervasive and Ubiquitous Computing, Heidelberg, Germany, 12–16 September 2016; pp. 841–852.
6. Gallop, C.; Tse, C.; Zhao, J. A Seasonal Autoregressive Model of Vancouver Bicycle Traffic Using Weather Variables. *I-Manag. J. Civ. Eng.* **2011**, *1*, 9. [\[CrossRef\]](#)
7. Yoon, J.W.; Pinelli, F.; Calabrese, F. Cityride: A Predictive Bike Sharing Journey Advisor. In Proceedings of the 2012 IEEE 13th International Conference on Mobile Data Management, Bengaluru, India, 23–26 July 2012; IEEE: New York, NY, USA, 2012; pp. 306–311.
8. Zheng, Z.; Zhou, Y.; Sun, L. A Multiple Factor Bike Usage Prediction Model in Bike-Sharing System. In Proceedings of the Green, Pervasive, and Cloud Computing: 13th International Conference, GPC 2018, Hangzhou, China, 11–13 May 2018; Revised Selected Papers 13. Springer: Berlin/Heidelberg, Germany, 2019; pp. 390–405.
9. Chen, C.; Ma, J.; Susilo, Y.; Liu, Y.; Wang, M. The Promises of Big Data and Small Data for Travel Behavior (Aka Human Mobility) Analysis. *Transp. Res. Part C Emerg. Technol.* **2016**, *68*, 285–299. [\[CrossRef\]](#)
10. Yang, Z.; Hu, J.; Shu, Y.; Cheng, P.; Chen, J.; Moscibroda, T. Mobility Modeling and Prediction in Bike-Sharing Systems. In Proceedings of the 14th Annual International Conference on Mobile Systems, Applications, and Services, Singapore, 26–30 June 2016; pp. 165–178.
11. Li, Y.; Zheng, Y.; Zhang, H.; Chen, L. Traffic Prediction in a Bike-Sharing System. In Proceedings of the 23rd SIGSPATIAL International Conference on Advances in Geographic Information Systems, Seattle, WA, USA, 3–6 November 2015; pp. 1–10.
12. Froehlich, J.E.; Neumann, J.; Oliver, N. Sensing and Predicting the Pulse of the City through Shared Bicycling. In Proceedings of the Twenty-First International Joint Conference on Artificial Intelligence, Pasadena, CA, USA, 11–17 July 2009.
13. Bao, J.; Yu, H.; Wu, J. Short-Term FFBS Demand Prediction with Multi-Source Data in a Hybrid Deep Learning Framework. *IET Intell. Transp. Syst.* **2019**, *13*, 1340–1347. [\[CrossRef\]](#)
14. Feng, S.; Chen, H.; Du, C.; Li, J.; Jing, N. A Hierarchical Demand Prediction Method with Station Clustering for Bike Sharing System. In Proceedings of the 2018 IEEE Third International Conference on Data Science in Cyberspace (DSC), Guangzhou, China, 18–21 June 2018; IEEE: New York, NY, USA, 2018; pp. 829–836.
15. Erman, J.; Arlitt, M.; Mahanti, A. Traffic Classification Using Clustering Algorithms. In Proceedings of the 2006 SIGCOMM Workshop on Mining Network Data, Pisa, Italy, 15 September 2006; pp. 281–286.
16. Gu, J.; Zhou, Q.; Yang, J.; Liu, Y.; Zhuang, F.; Zhao, Y.; Xiong, H. Exploiting Interpretable Patterns for Flow Prediction in Dockless Bike Sharing Systems. *IEEE Trans. Knowl. Data Eng.* **2020**, *34*, 640–652. [\[CrossRef\]](#)
17. Blondel, V.D.; Guillaume, J.-L.; Lambiotte, R.; Lefebvre, E. Fast Unfolding of Communities in Large Networks. *J. Stat. Mech. Theory Exp.* **2008**, *2008*, P10008. [\[CrossRef\]](#)
18. Mehdizadeh Dastjerdi, A.; Morency, C. Bike-Sharing Demand Prediction at Community Level under COVID-19 Using Deep Learning. *Sensors* **2022**, *22*, 1060. [\[CrossRef\]](#)
19. Chai, D.; Wang, L.; Yang, Q. Bike Flow Prediction with Multi-Graph Convolutional Networks. In Proceedings of the 26th ACM SIGSPATIAL International Conference on Advances in Geographic Information Systems, Seattle, WA, USA, 6–9 November 2018; pp. 397–400.
20. Lee, S.-H.; Ku, H.-C. A Dual Attention-Based Recurrent Neural Network for Short-Term Bike Sharing Usage Demand Prediction. *IEEE Trans. Intell. Transp. Syst.* **2022**, *24*, 4621–4630. [\[CrossRef\]](#)
21. Chen, P.-C.; Hsieh, H.-Y.; Su, K.-W.; Sigalingging, X.K.; Chen, Y.-R.; Leu, J.-S. Predicting Station Level Demand in a Bike-Sharing System Using Recurrent Neural Networks. *IET Intell. Transp. Syst.* **2020**, *14*, 554–561. [\[CrossRef\]](#)
22. Wang, B.; Kim, I. Short-Term Prediction for Bike-Sharing Service Using Machine Learning. *Transp. Res. Procedia* **2018**, *34*, 171–178. [\[CrossRef\]](#)
23. Zhang, C.; Zhang, L.; Liu, Y.; Yang, X. Short-Term Prediction of Bike-Sharing Usage Considering Public Transport: A LSTM Approach. In Proceedings of the 2018 21st International Conference on Intelligent Transportation Systems (ITSC), Maui, HI, USA, 4–7 November 2018; IEEE: New York, NY, USA, 2018; pp. 1564–1571.
24. Pan, Y.; Zheng, R.C.; Zhang, J.; Yao, X. Predicting Bike Sharing Demand Using Recurrent Neural Networks. *Procedia Comput. Sci.* **2019**, *147*, 562–566. [\[CrossRef\]](#)
25. Yang, H.; Xie, K.; Ozbay, K.; Ma, Y.; Wang, Z. Use of Deep Learning to Predict Daily Usage of Bike Sharing Systems. *Transp. Res. Rec.* **2018**, *2672*, 92–102. [\[CrossRef\]](#)
26. Li, X.; Xu, Y.; Chen, Q.; Wang, L.; Zhang, X.; Shi, W. Short-Term Forecast of Bicycle Usage in Bike Sharing Systems: A Spatial-Temporal Memory Network. *IEEE Trans. Intell. Transp. Syst.* **2021**, *23*, 10923–10934. [\[CrossRef\]](#)
27. Guo, R.; Jiang, Z.; Huang, J.; Tao, J.; Wang, C.; Li, J.; Chen, L. BikeNet: Accurate Bike Demand Prediction Using Graph Neural Networks for Station Rebalancing. In Proceedings of the 2019 IEEE SmartWorld, Ubiquitous Intelligence & Computing, Advanced & Trusted Computing, Scalable Computing & Communications, Cloud & Big Data Computing, Internet of People and Smart City Innovation (SmartWorld/SCALCOM/UIC/ATC/CBDCom/IOP/SCI), Leicester, UK, 19–23 August 2019; IEEE: New York, NY, USA, 2019; pp. 686–693.

28. Lv, M.; Hong, Z.; Chen, L.; Chen, T.; Zhu, T.; Ji, S. Temporal Multi-Graph Convolutional Network for Traffic Flow Prediction. *IEEE Trans. Intell. Transport. Syst.* **2021**, *22*, 3337–3348. [[CrossRef](#)]
29. Liang, Y.; Ke, S.; Zhang, J.; Yi, X.; Zheng, Y. GeoMAN: Multi-Level Attention Networks for Geo-Sensory Time Series Prediction. In Proceedings of the Twenty-Seventh International Joint Conference on Artificial Intelligence; International Joint Conferences on Artificial Intelligence Organization: Stockholm, Sweden, July 2018; pp. 3428–3434.
30. Liu, L.; Qiu, Z.; Li, G.; Wang, Q.; Ouyang, W.; Lin, L. Contextualized Spatial–Temporal Network for Taxi Origin–Destination Demand Prediction. *IEEE Trans. Intell. Transp. Syst.* **2019**, *20*, 3875–3887. [[CrossRef](#)]
31. Li, F.; Feng, J.; Yan, H.; Jin, G.; Jin, D.; Li, Y. Dynamic Graph Convolutional Recurrent Network for Traffic Prediction: Benchmark and Solution. *arXiv* **2021**, arXiv:2104.14917. [[CrossRef](#)]
32. Liu, L.; Zhen, J.; Li, G.; Zhan, G.; He, Z.; Du, B.; Lin, L. Dynamic Spatial–Temporal Representation Learning for Traffic Flow Prediction. *IEEE Trans. Intell. Transp. Syst.* **2021**, *22*, 7169–7183. [[CrossRef](#)]
33. Pan, Z.; Liang, Y.; Wang, W.; Yu, Y.; Zheng, Y.; Zhang, J. Urban Traffic Prediction from Spatio–Temporal Data Using Deep Meta Learning. In Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining, Anchorage, AK, USA, 4–8 August 2019; ACM: Anchorage AK USA, 2019; pp. 1720–1730.
34. Wang, Z.; Jiang, R.; Xue, H.; Salim, F.D.; Song, X.; Shibasaki, R. Event-Aware Multimodal Mobility Nowcasting. In Proceedings of the AAAI Conference on Artificial Intelligence, Virtual, 22 February–1 March 2022.
35. Essien, A.; Petrounias, I.; Sampaio, P.; Sampaio, S. A Deep–Learning Model for Urban Traffic Flow Prediction with Traffic Events Mined from Twitter. *World Wide Web* **2021**, *24*, 1345–1368. [[CrossRef](#)]
36. Jiao, F.; Huang, L.; Song, R.; Huang, H. An Improved STL–LSTM Model for Daily Bus Passenger Flow Prediction during the COVID–19 Pandemic. *Sensors* **2021**, *21*, 5950. [[CrossRef](#)] [[PubMed](#)]
37. Tu, Q.; Zhang, Q.; Zhang, Z.J.; Gong, D.; Tang, M.C. A Deep Spatio–Temporal Fuzzy Neural Network for Subway Passenger Flow Prediction with COVID–19 Search Engine Data. *IEEE Trans. Fuzzy Syst.* **2022**, *31*, 394–406. [[CrossRef](#)]
38. Wang, X.; Guan, X.; Cao, J.; Zhang, N.; Wu, H. Forecast Network–Wide Traffic States for Multiple Steps Ahead: A Deep Learning Approach Considering Dynamic Non–Local Spatial Correlation and Non–Stationary Temporal Dependency. *Transp. Res. Part C: Emerg. Technol.* **2020**, *119*, 102763. [[CrossRef](#)]
39. Xu, H.; Zou, T.; Liu, M.; Qiao, Y.; Wang, J.; Li, X. Adaptive Spatiotemporal Dependence Learning for Multi–Mode Transportation Demand Prediction. *IEEE Trans. Intell. Transp. Syst.* **2022**, *23*, 18632–18642. [[CrossRef](#)]
40. Yu, B.; Yin, H.; Zhu, Z. Spatio–Temporal Graph Convolutional Networks: A Deep Learning Framework for Traffic Forecasting. *arXiv* **2017**, arXiv:1709.04875.
41. Sparck Jones, K. A Statistical Interpretation of Term Specificity and Its Application in Retrieval. *J. Doc.* **1972**, *28*, 11–21. [[CrossRef](#)]
42. Kipf, T.N.; Welling, M. Semi–Supervised Classification with Graph Convolutional Networks. *arXiv* **2016**, arXiv:1609.02907.
43. Defferrard, M.; Bresson, X.; Vandergheynst, P. Convolutional Neural Networks on Graphs with Fast Localized Spectral Filtering. *Adv. Neural Inf. Process. Syst.* **2016**, *29*, 3844–3852.
44. Chung, J.; Gulcehre, C.; Cho, K.; Bengio, Y. Empirical Evaluation of Gated Recurrent Neural Networks on Sequence Modeling. *arXiv* **2014**, arXiv:1412.3555.
45. Tian, Y.; Pan, L. Predicting Short–Term Traffic Flow by Long Short–Term Memory Recurrent Neural Network. In Proceedings of the 2015 IEEE International Conference on Smart city/SocialCom/SustainCom (SmartCity), Chengdu, China, 19–21 December 2015; IEEE: New York, NY, USA, 2015; pp. 153–158.
46. Zhou, X.; Shen, Y.; Zhu, Y.; Huang, L. Predicting Multi–Step Citywide Passenger Demands Using Attention–Based Neural Networks. In Proceedings of the Eleventh ACM International Conference on Web Search and Data Mining, Los Angeles, CA, USA, 5–9 February 2018; pp. 736–744.

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.