

Graph Neural Networks in Intelligent Transportation Systems: Advances, Applications and Trends

Hourun Li, Yusheng Zhao, Zhengyang Mao, Yifang Qin, Zhiping Xiao, Jiaqi Feng, Yiyang Gu,
Wei Ju, *Member, IEEE*, Xiao Luo, and Ming Zhang

Abstract—Intelligent Transportation System (ITS) is crucial for improving traffic congestion, reducing accidents, optimizing urban planning, and more. However, the complexity of traffic networks has rendered traditional machine learning and statistical methods less effective. With the advent of artificial intelligence, deep learning frameworks have achieved remarkable progress across various fields and are now considered highly effective in many areas. Since 2019, Graph Neural Networks (GNNs) have emerged as a particularly promising deep learning approach within the ITS domain, owing to their robust ability to model graph-structured data and address complex problems. Consequently, there has been increasing scholarly attention to the applications of GNNs in transportation, which have demonstrated excellent performance. Nevertheless, current research predominantly focuses on traffic forecasting, with other ITS domains, such as autonomous vehicles and demand prediction, receiving less attention. This paper aims to review the applications of GNNs across six representative and emerging ITS research areas: traffic forecasting, vehicle control system, traffic signal control, transportation safety, demand prediction, and parking management. We have examined a wide range of graph-related studies from 2018 to 2023, summarizing their methodologies, features, and contributions in detailed tables and lists. Additionally, we identify the challenges of applying GNNs in ITS and propose potential future research directions.

Index Terms—Intelligent Transportation System, Graph Neural Network, Spatio-temporal Analysis

I. INTRODUCTION

As cities grow and transportation systems evolve, several issues have become increasingly apparent, such as traffic congestion, environmental pollution, and a rising number of traffic accidents. To address these challenges and improve traffic flow, route planning, and transportation safety, the Intelligent Transportation System (ITS) was introduced over five decades ago in the U.S. [1]. Today, ITS applications are integral to everyday life, including Electronic Toll Collection (ETC), Traffic Management Systems (TMS), Global Positioning Systems (GPS), and Commercial Vehicle Operations (CVO). ITS encompasses a broad range of areas, including traffic forecasting, autonomous vehicles, traffic signal control, and more. Notably, traffic forecasting has emerged as a prominent

Hourun Li, Yusheng Zhao, Zhengyang Mao, Yifang Qin, Jiaqi Feng, Yiyang Gu, Wei Ju, and Ming Zhang are with School of Computer Science, National Key Laboratory for Multimedia Information Processing, Peking University-Anker Embodied AI Lab, Peking University, Beijing, China. (e-mail: lihourun@stu.pku.edu.cn, yusheng.zhao@stu.pku.edu.cn, zhengyang.mao@stu.pku.edu.cn, qinyifang@pku.edu.cn, sf77@stu.pku.edu.cn, yiyanggu@pku.edu.cn, juwei@pku.edu.cn, mzhang_cs@pku.edu.cn)

Zhiping Xiao and Xiao Luo are with Department of Computer Science, University of California, Los Angeles, USA. (e-mail: patricia.xiao@cs.ucla.edu, xiaoluo@cs.ucla.edu)

research area, garnering significant attention due to its critical applications in optimizing route planning, facilitating road traffic, and reducing traffic accidents.

According to Verses et al. [2], addressing practical challenges such as managing massive and noisy data, scalability, and generalization remains difficult. Over the past three decades, statistical methods, including simple linear time series models like Autoregressive Integrated Moving Average (ARIMA) [3], [4], and traditional machine learning methods such as Logistic Regression (LR), Support Vector Regression (SVR), and k-Nearest Neighbors (KNN) [5]–[7] have been proposed to tackle these issues. However, the increasing volume of data and the complexity of road conditions have made traditional methods less effective. Consequently, there is a need for more efficient algorithms and scalable models to fully leverage massive data and develop accurate and efficient ITS solutions. Additionally, advancements in computational techniques, such as graphical processing units, have propelled the effectiveness of deep learning models. Since 2015, deep-learning models for traffic forecasting have seen significant progress, with GNNs emerging as the most popular models after 2019 [8]. GNNs excel not only at modeling graph-structured problems but also at capturing temporal-spatial dependencies and representing relationships in non-Euclidean spaces [8]–[10].

After conducting a comprehensive survey of research in the field of ITS, we found that a significant portion of studies primarily focuses on traffic forecasting. However, we believe that other critical domains within ITS also warrant greater attention. Despite the recent shift toward promising techniques such as deep learning and reinforcement learning, GNNs still require more exploration and application. Considering the graph-based nature of traffic networks and the inherent advantages offered by GNNs, we argue that GNNs represent a highly promising and competitive solution for ITS. Our investigation centers on papers related to GNNs in ITS published between 2018 and 2023, providing a detailed summary of their contributions. We also identified key research challenges within ITS and proposed potential future directions for leveraging GNNs. Our main contribution can be summarized as follows:

- *Comprehensive Review.* Extensive research work and surveys from 2018 to 2023 for ITS are reviewed in detail, covering six distinct research domains rather than focusing solely on traffic forecasting. Moreover, we offer an in-depth analysis of the papers reviewed, summarize methods and challenges, and present informative tables and lists.
- *A Comprehensive Taxonomy.* We categorized the reviewed studies based on various criteria, including research

domain, graph methods, and domain-specific challenges. This approach helps readers comprehensively understand each research domain in ITS.

- *Challenges and Future Directions.* After a comprehensive review, we summarize the key challenges when applying GNNs in ITS and propose potential future directions, which provide valuable insights for researchers looking to explore and advance this field.

We organize the rest of the survey as follows: In section II, we quickly review the related surveys in transportation domains and briefly introduce them. In Section III, we present the foundational knowledge on ITS and GNNs, along with a discussion on problem formulation. In Section IV, we investigate and review extensive graph-based studies in six domains, including traffic forecasting, vehicle control system, traffic signal control, transportation safety, demand prediction, and parking management. In Section V, we summarize the challenges and potential future directions in GNNs for ITS based on the previous review results. Finally, we present the conclusions in Section VI.

II. RELATED SURVEYS

In this section, we review the most relevant and representative surveys in Intelligent Transportation Systems (ITS), primarily focusing on those published in the past five years. Special emphasis is given to approaches based on Graph Neural Networks (GNNs), for which we provide a thorough introduction.

As previously mentioned, ITS encompass a variety of research fields. However, recent surveys have predominantly concentrated on traffic forecasting [8], [9], [11]–[16], with only a few exploring other ITS domains [2], [10], [17], [18]. Traffic flow prediction has been the most prominent research topic in traffic forecasting since 2015, as highlighted by Liu et al. [19]. Surveys prior to 2015 primarily focused on statistical methods [3], [4], [20]–[23] and traditional machine learning models [5]–[7], [24]. Notable early surveys by Vlahogianni et al. [15], [25] in 2004 and 2014 concentrated on short-term traffic forecasting. However, these traditional methods have limitations in addressing complex transportation problems due to their shallow architectures. With advancements in theory and hardware, deep learning models have gained popularity since the mid-2010s, significantly advancing traffic forecasting [8], [12]. Since 2019, GNNs have become particularly prominent, underscoring their growing importance in ITS [9]–[11], [14]. In recent years, researchers have increasingly focused on the temporal and spatial dependencies of traffic data [26], [27], leading to the exploration of new research trends and directions. The detailed development of models and GNN applications in ITS is illustrated in Fig.1.

It is worth noting that Jiang et al. published a comprehensive review of GNNs for traffic forecasting [9], which summarizes the research on this topic. They reviewed 212 articles published between 2018 and 2020, provided a detailed taxonomy of problems and methods, and compiled information on open-source data and code resources. In the following year, they extended their work with another survey [14], which updates the previous review by describing the latest research developments

and trends up to 2022. This follow-up also highlighted specific challenges and proposed informative future directions. While their surveys offer an exhaustive overview of traffic forecasting, they do not cover other research areas within ITS.

The most relevant survey is the work by Rahmani et al. [10], which stands out as the most recent and comprehensive review of GNNs in general ITS research. This review covers fundamental concepts such as graphs and GNNs and spans seven ITS research domains: traffic forecasting, demand prediction, autonomous vehicles, intersection management, parking management, urban planning, and transportation safety. For each domain, it provides a brief introduction followed by an independent review of the relevant literature. In contrast, our survey offers a deeper examination of the application of graphs and GNNs in ITS. We provide an extensive exploration of graph construction, GNN customization for specific challenges, and performance assessment across various ITS domains, offering insights that are not covered in [10]. Furthermore, instead of introducing individual papers independently, our systematic approach to summarizing related literature offers readers a more reliable and comprehensive analysis. Our review encompasses six domains: traffic forecasting, autonomous vehicles, traffic signal control, transportation safety, demand prediction, and parking management. It offers a more thorough and detailed exploration of ITS compared to existing surveys.

III. BACKGROUND

In this section, we will provide background information on ITS, graphs, and GNNs. We will begin by introducing the key concepts of ITS and the associated research domains. Next, we will explain fundamental graph concepts, covering different types of graph data and their characteristics. Finally, we will provide an overview of GNN variants and basic GNN models, laying the foundation for the more detailed discussions in the subsequent sections.

A. Intelligent Transportation System (ITS)

Before 1980, ITS was primarily a forward-looking concept aimed at overcoming the limitations of surface transportation capacity [1], [28]. During this period, the focus was on improving road network efficiency through optimized traffic signals and in-vehicle navigation systems [28]–[30]. In the 21st century, technological advancements have made traffic data more accessible, leading to a surge in data-driven approaches [31]. Real-time traffic data is now used extensively for traffic management, including road condition prediction, congestion identification, and navigation [31]–[36]. Nowadays, ITS is a rapidly growing interdisciplinary field [37], offering innovative services to enhance performance, improve travel safety, and provide useful information to users. Zhang et al. [31] identify six primary subsystems in ITS. For our focus on applying GNNs to address specific issues, we have further categorized these subsystems into specific research fields: traffic forecasting, vehicle control system, traffic signal control, transportation safety, demand prediction, and parking management, as illustrated in Fig.2.

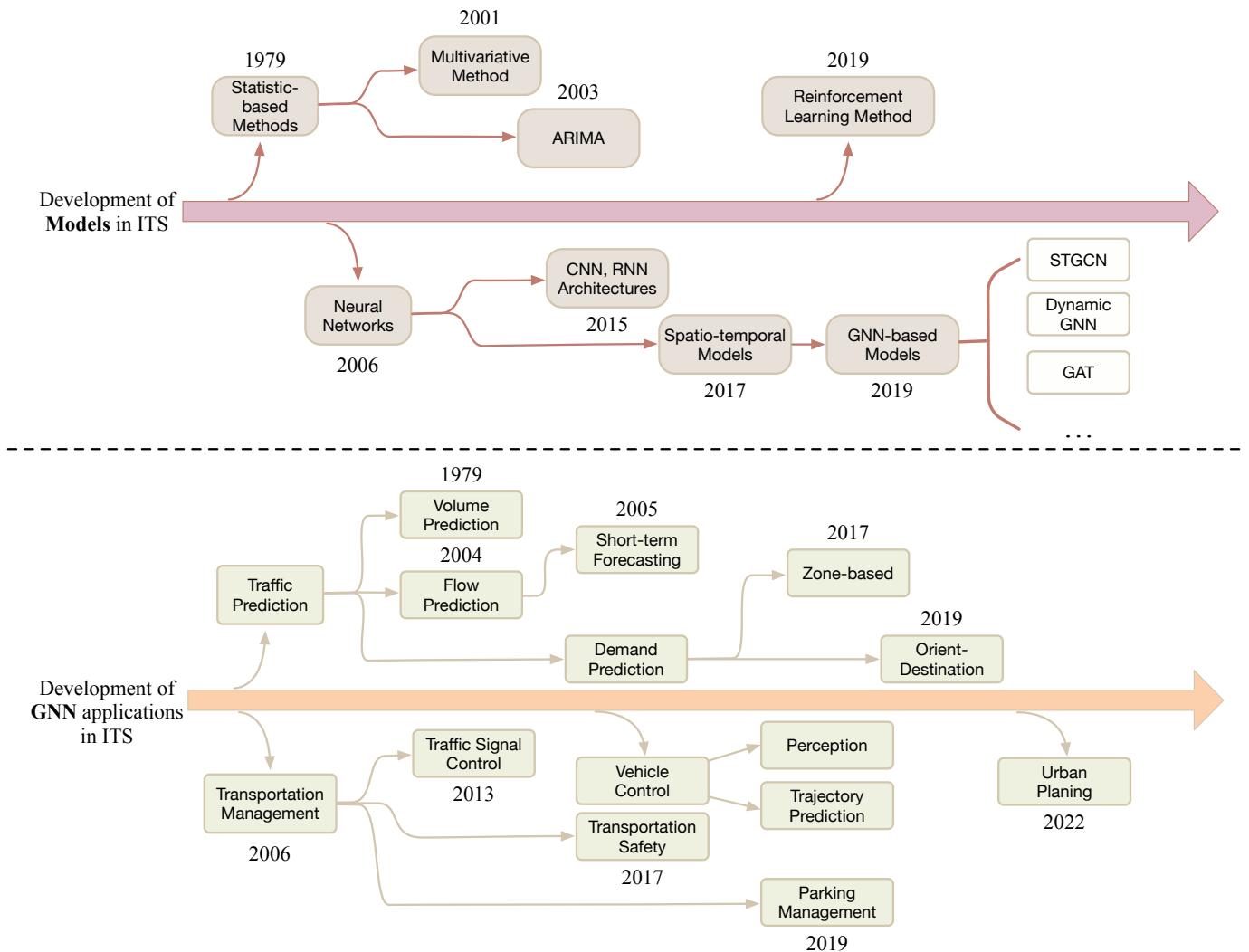


Fig. 1: Development of Models and Application in Intelligent Transportation System

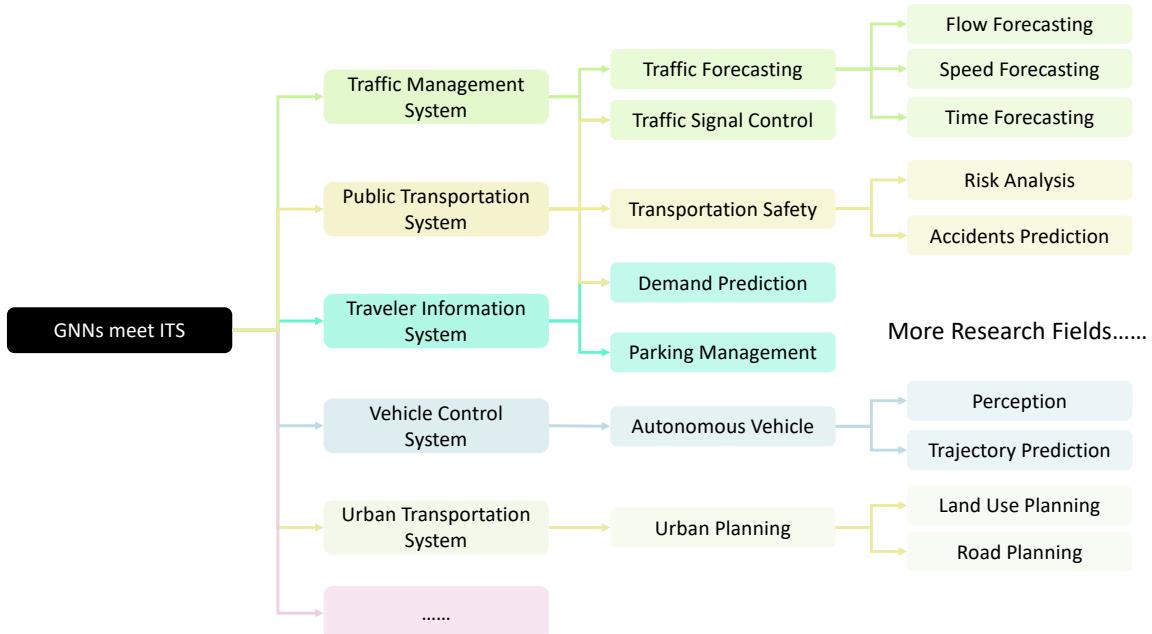


Fig. 2: Research Fields in Intelligent Transportation System

B. Graphs and Graph Neural Networks in ITS

In this section, we will briefly introduce graphs and GNNs, which are necessary for the following discussion.

1) **Graphs and Graph Types:** In general, a graph consists of a set of nodes $\mathcal{V} = \{v_1, v_2, \dots, v_{|\mathcal{V}|}\}$ and a corresponding set of edges \mathcal{E} , denoted as $\mathcal{G} = (\mathcal{V}, \mathcal{E})$. The graph \mathcal{G} can be represented by an adjacency matrix $\mathbf{A} \in \mathbb{R}^{|\mathcal{V}| \times |\mathcal{V}|}$, where $\mathbf{A}_{ij} = 1$ if $e_{ij} = (v_i, v_j) \in \mathcal{E}$, and $\mathbf{A}_{ij} = 0$ otherwise. In addition, each node v_i can be associated with node features $x_i \in \mathbb{R}^d$, and the feature matrix of the graph is denoted as $\mathbf{X} \in \mathbb{R}^{|\mathcal{V}| \times d}$, where d is the dimension of the features. Therefore, a graph can also be represented as $\mathcal{G} = \{\mathbf{X}, \mathbf{A}\}$ and the graph can be further categorized into several types:

- **Directed/Undirected Graphs.** In an undirected graph, if there is an edge $(v_i, v_j) \in \mathcal{E}$, then $(v_j, v_i) \in \mathcal{E}$ must also hold, and vice versa. While in a directed graph, this constraint does not apply.
- **Weighted/Unweighted Graphs.** In a weighted graph, edges are assigned weight values to indicate their varying importance or other necessary information. In an unweighted graph, all edges are considered equal.
- **Signed/Unsigned Graphs.** Signed graphs consist of edges with positive or negative signs, representing different types of relationships. In contrast, unsigned graphs have edges without sign distinctions, indicating neutral relationships.
- **Static/Dynamic Graphs.** Static graphs have fixed node and edge features that remain constant throughout. In contrast, dynamic graphs evolve, where new nodes or edges can emerge or disappear at any time step. Therefore, it is essential to model temporal information accurately to capture the changes over time. Normally, we represent a dynamic graph as a sequence of static graph screen-shots:

$$\mathcal{G} = \{\mathcal{G}_1, \mathcal{G}_2, \dots, \mathcal{G}_T\},$$

where $\mathcal{G}_t = (\mathcal{V}_t, \mathcal{E}_t)$, $t \in \{1, 2, \dots, T\}$ and T is the total number of time steps.

- **Homogeneous/Heterogeneous Graphs.** Homogeneous graphs consist of nodes and edges of a single type, representing uniform entities and relationships. In contrast, heterogeneous graphs involve multiple types of nodes and edges, capturing diverse entities and complex interactions.
- **Nested Graphs.** Nested graphs are hierarchical, with nodes that can be expanded into subgraphs containing additional nodes and links. Nested graphs are valuable in ITS for managing multi-level information flows and conducting multi-level prediction tasks.
- **Hypergraphs.** Hypergraphs enhance traditional graphs by allowing edges to connect any number of nodes, which enables the modeling of more complex relationships among groups of entities.

2) **Graph Construction in ITS:** Graph data is highly effective for representing traffic flow data. For instance, in a system with N road sections, we can model this as a graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$, where $\mathcal{V} = \{v_1, v_2, \dots, v_N\}$ denotes the set of nodes, each corresponding to a road section. The edges, represented by \mathcal{E} , capture the relationships between these nodes, such as connectivity. If two road sections are directly connected, an

edge is included in the edge set, denoted as $(v_i, v_j) \in \mathcal{E}$, where v_i and v_j are the connected road sections, and \mathcal{E} is the set of all such edges. This graph construction method provides a clear way to model and analyze traffic flow and connectivity.

However, graph construction varies across different ITS research domains and is highly problem-specific. For example, in urban traffic management, a dynamic graph is often used where road intersections are represented as nodes and roads as edges, capturing traffic connectivity and flow. Edges in this graph might be weighted to reflect parameters such as travel time, congestion levels, or distance. While in demand prediction, nodes might represent bus stops or train stations, while edges could indicate routes with attributes such as frequency, capacity, and schedule adherence. The diverse methods of graph construction highlight the need for tailored approaches that address the unique requirements and challenges of each application. Therefore, in the following sections, we will explore the typical graph construction methods used in each research domain in detail.

C. Graph Neural Networks (GNNs)

1) **Definition of GNNs:** Graph Neural Networks (GNNs) are designed specifically for graph data. A central mechanism of GNNs is the message-passing paradigm [38], which iteratively aggregates information from neighboring nodes to capture the structural characteristics of the graph. Let $h_v^{(l)}$ indicates the embedding of v at layer $l \in \{1, \dots, L\}$, the message-passing mechanism can be denoted as:

$$h_v^{(l)} = \mathcal{U}^{(l)} \left(h_v^{(l-1)}, \mathcal{A}^{(l)} \left(\{h_u^{(l-1)}\}_{u \in \mathcal{N}(v)} \right) \right), \quad (1)$$

where $\mathcal{A}^{(l)}$ and $\mathcal{U}^{(l)}$ represent the message aggregating and embedding updating function at layer l respectively, and $\mathcal{N}(v)$ denotes the neighbors of v . After L iterations, we can obtain the node-level representation $h_v^{(L)}$. The graph-level representation $h_{\mathcal{G}}$ can be obtained by aggregating all node representations at layer L with a readout function. Formally, it can be defined as:

$$h_{\mathcal{G}} = \text{READOUT}(\{h_v^{(L)}\}_{v \in \mathcal{V}}), \quad (2)$$

where READOUT could be averaging or other graph-level pooling functions depending on the model [39]–[41].

2) **Variants of GNNs:** In the rapidly evolving field of GNNs, a variety of variants have emerged as crucial tools for effectively utilizing graph-structured data and tackling the unique challenges across different domains. The existing literature offers numerous taxonomies to classify these variants based on structural, operational, and application-oriented characteristics. For the purpose of our survey, which specifically targets ITS, we will not present an exhaustive taxonomy of GNNs. Instead, we will introduce a few preliminary GNN variants that are highly relevant to our discussions on ITS applications. For those interested in a more thorough exploration of GNN taxonomies and a broader classification of these networks, we recommend referring to the comprehensive works by Wu et al. [42], Zhou et al. [43], and Zhang et al. [44].

- **Convolutional GNNs.** The underpinning of Convolutional GNNs is rooted in the concept of expanding traditional convolution operations from Euclidean to non-Euclidean

domains. Convolutional GNNs can be categorized into spectral and spatial methods. Spectral methods are based on spectral graph theory but suffer from computational inefficiencies and a lack of localization in the spatial domain. Spatial methods directly compute convolutions by aggregating and transforming features from a node's neighborhood, avoiding costly spectral transformations. In conclusion, convolutional GNNs adapt convolutional operations to graph-structured data, offering a robust framework for feature learning on graphs.

- **Recurrent-Based GNNs.** Recurrent-based GNNs combine the characteristics of GNNs and recurrent neural networks (RNNs) to effectively capture dependencies in sequential and graph-based data. By employing a message-passing mechanism and recurrent updates, recurrent-based GNNs can dynamically adapt to changes in the graph structure. This approach provides a flexible and adaptable framework that can effectively handle spatial and temporal complexities in real-world datasets, particularly in dynamic social and traffic networks.
- **Spatial-Temporal GNNs.** Spatial-temporal GNNs integrate the spatial structure of a graph using a spatial graph neural network and a time-series processing model such as RNN, LSTM, or Transformer. This integration allows for simultaneous data processing in both the temporal and spatial dimensions. Therefore, spatial-temporal GNNs are commonly employed to analyze graph data with temporal dynamics, particularly for traffic forecasting [9].
- **Graph Autoencoders.** Graph autoencoders (GAEs) are unsupervised generative models designed to encode graph-structured data into a low-dimensional representation using encoders such as GCN [45]. Subsequently, GAEs reconstruct the original graph from this representation. GAEs are commonly utilized for tasks such as node embedding, link prediction, and graph generation. An example of a well-known GAE is the Variational Graph Autoencoder (VGAЕ) [46].
- **Graph Reinforcement Learning.** Graph reinforcement learning (GRL) is an innovative approach that integrates reinforcement learning (RL) with graph-based representations to address challenges associated with graph-structured data where strategic decision-making is desirable. In GRL, agents develop optimal policies by interacting with graph-structured environments, leveraging the connections and relationships between nodes to inform their decisions. GRL is applied to optimize network routing and conduct molecular design on graphs.

IV. THE APPLICATIONS OF GNNs IN ITS

This section will review research across six critical domains within ITS: traffic forecasting, vehicle control system, traffic signal control, transportation safety, demand prediction, and parking management. For each domain, we define the core problems, evaluate their potential impact, and identify the associated challenges or requirements. Next, we present a logically organized review of the research, emphasizing key findings and trends rather than discussing each paper independently. Finally,

we discuss how graphs are constructed within each specific domain and examine the modifications made to standard GNN to address the unique needs and challenges of each area.

A. Traffic Forecasting

Traffic forecasting, also referred to as traffic prediction, has been a prominent research topic in ITS over the last few decades [47]. Its objective is to predict the future traffic state on a road network. Based on the observed metrics, traffic forecasting can be roughly categorized into two sub-types: flow forecasting and speed/time forecasting.

1) *Traffic Flow Forecasting:* Traffic flow forecasting, which uses flow as a metric, involves estimating the volume of vehicles during specific periods and across different segments of the transportation network. Accurate traffic flow forecasting is essential for managing congestion, planning routes, handling incidents, and improving overall transportation infrastructure efficiency. The challenges arise from the complex and constantly changing factors influencing traffic flow, such as daily commuting patterns, road conditions, weather, special events, accidents, and construction work [49].

Recent research has introduced various models that can effectively capture the complex interdependencies in spatial and temporal data. These models, such as STGCN by Song et al. [55], DyHSL by Zhao et al. [49], DCRNN by Li et al. [58], ASTGCN by Guo et al. [56], utilize advanced neural network architectures to handle dynamic relationships within traffic systems. They provide a comprehensive understanding of traffic behavior by addressing spatial aspects (e.g. road connectivity, road intersections) and temporal factors (e.g. traffic flow variations) over time.

One line of research involves utilizing GNNs in combination with RNNs [70], [71] to recursively capture both spatial and temporal information [53], [54], [58]. For instance, the Diffusion Convolutional Recurrent Neural Network (DCRNN) by Li et al. [58] replaces fully connected layers in the GRU [71] with diffusion convolution. Similarly, the Adaptive Graph Convolutional Recurrent Network (AGCRN) by Bai et al. [53] focuses on learning node-specific features and uncovering hidden inter-dependencies through an adaptive graph convolutional recurrent methodology. AGCRN reflects a trend toward tailoring models to understand complex network dynamics. Furthermore, the Hypergraph Convolutional Recurrent Neural Network (HGC-RNN) by Yi et al. [54] combines hypergraph convolution with RNNs, specifically targeting traffic flow forecasting. This combination highlights the potential of integrating different neural network architectures to improve prediction accuracy.

Another line of research [49], [55]–[57], [72] involves developing a large spatio-temporal graph and utilizing GNNs to capture spatio-temporal correlations. For example, the Spatial-temporal Synchronous Graph Convolutional Network (STGCN) by Song et al. [55] creates a spatio-temporal graph structure to perform localized graph convolution operations, thereby improving data processing capabilities. Attention-based Spatial-temporal Graph Convolutional Network (ASTGCN) by Guo et al. [56] incorporates an attention mechanism within the spatio-temporal graph context [55] to enhance performance

TABLE I: A Comprehensive Overview of Most Related Studies for Traffic Forecasting

Model	Article	Year	Task	Graph Construction	Spatial Module	Temporal Module	Summary
FPTN	[48]	2023	Flow	road network	transformer	transformer	FPTN improves traffic forecasting with sensor-based data division, triple types of embeddings, and an efficient Transformer encoder, reducing computational demands.
DyHSL	[49]	2023	Flow	learned hypergraph	HGNN	HGNN	DyHSL improves traffic forecasting with hypergraphs for dynamics and interactive convolutions for spatio-temporal relations, effective across multiple datasets.
DSTAGNN	[50]	2022	Flow	dynamic	GNN	GNN	DSTAGNN dynamically models spatial-temporal road network interactions by utilizing enhanced multi-head attention and multi-scale gated convolution.
Bi-STAT	[51]	2022	Flow	road network	transformer	transformer	Bi-STAT enhances traffic forecasting with adaptive spatial-temporal transformers, handling diverse task complexities and leveraging past data for improved prediction.
STFGNN	[52]	2021	Flow	road network	GNN	GNN	STFGNN enhances traffic forecasting by fusing data-driven temporal and spatial graphs and employing gated convolutions, effectively handling long sequences.
AGCRN	[53]	2020	Flow	generated	GCN	RNN	AGCRN enhances prediction by two adaptive modules, focusing on node-specific patterns and automatic inter-dependency learning without pre-defined graphs.
HGC-RNN	[54]	2020	Flow	road network	HGNN	RNN	HGC-RNN leverages hypergraph convolution and RNNs for structured time-series sensor network data, capturing complex structural and temporal dependencies.
STSGCN	[55]	2020	Flow	road network	GCN	GCN	STSGCN models localized spatial-temporal correlations and accounts for heterogeneities across different periods, simplifying spatial-temporal network data forecasting.
ASTGCN	[56]	2019	Flow	road network	GCN	attention	ASTGCN improve forecasting with a spatial-temporal attention mechanism and convolutions, focusing on dynamic correlations to make more accurate predictions.
LRGCN	[57]	2019	Flow	road network	RGCN	RGCN	LRGCN, designed for time-evolving graph path classification, integrating temporal dependencies and graph dynamics by relational GCN to process time-based relations.
DCRNN	[58]	2017	Flow	road network	GCN	RNN	DCRNN models forecasting as a diffusion process on directed graphs, using bidirectional random walks and an encoder-decoder architecture with scheduled sampling.
CAGRU	[59]	2021	Speed	road network	GAT	GRU	CAGRU predicts traffic speed and identifies patterns using a convolutional attention-based neural network based on traffic flow data without relying on historical speed data.
DMSTGCN	[60]	2021	Speed	learned	DGNN	DGNN	DMSTGCN learns dynamic spatial dependencies between road segments and incorporates multi-varied traffic data, capturing multifaceted spatio-temporal traffic features.
FASTGNN	[61]	2021	Speed	road network	ASTGCN	ASTGCN	FASTGNN, a federated learning framework, features a differential privacy-based method to protect topological information and an innovative aggregation approach.
-	[62]	2020	Speed	road network	GraphSAGE	-	This paper uses GraphSAGE to forecast spatially heterogeneous traffic speed and imputes missing data for segment networks with nonlinear spatial-temporal correlations.
ATT-LSTM	[63]	2020	Speed	road network	GAT	LSTM	Attention-based LSTM (ATT-LSTM), a short-term level prediction model, predicts traffic speed and imputes missing traffic data with a data preprocessing module.
GATCN	[64]	2020	Speed	road network	GAT	TCN	GATCN, a deep learning framework combining GAT and TCN, effectively learns spatio-temporal traffic flow characteristics and neighborhood information with multiple layers.
MTL-GRU	[65]	2020	Speed	road network	GNN	GRU	MTL-GRU, a multitask learning GRU model with residual mappings, selects the most informative features to enhance traffic flow and speed forecasting.
DSTL-GR	[66]	2023	Time	road network	GraphSAGE	LSTM	DSTL-GR enhances travel time prediction by considering spatial and temporal dependence, as well as exogenous variables, through a combination of GNNs and RNNs.
DeepTrans	[67]	2020	Time	road network	DCRNN	DCRNN	DeepTRANS enhances travel time estimation by incorporating traffic forecasting into an existing deep learning-based bus ETA model, improving congestion prediction.
SST-GNN	[68]	2020	Time	road network	SGNN	SGNN	SST-GNN predicts by encoding spatial correlations, using neighborhood aggregation and a spatio-temporal mechanism with position encoding for periodic patterns.
-	[69]	2019	Time	road segment	clustering	-	The model predicts bus travel times using real-time taxi and bus data, dividing routes into dwelling and transit segments with two tailored models for each.

by focusing on salient features. Additionally, the Long Short-Term Memory R-GCN (LRGCN) approach by Li et al. [57] is designed to encode spatio-temporal graphs more efficiently, addressing the inherent complexities of such data structures.

The third line of research involves utilizing GNNs in conjunction with transformers [48], [51], [73], [74], leveraging the success of transformers across various domains [75]–[77]. This approach effectively merges the transformers' ability to handle long-range dependencies with the spatial processing strengths of GNNs, resulting in two distinct types of methodologies for traffic flow forecasting. The first type features modular integration where spatial and temporal components are distinctly processed and then combined. For example, the Spatial-Temporal Transformer Networks (STTNs) introduced by Xu et al. [73] employ a spatial transformer to dynamically model directed spatial dependencies via self-attention, capturing real-time traffic conditions and directional flow. This is complemented by a temporal transformer that handles long-range bidirectional temporal dependencies, optimizing long-term forecasting accuracy. Similarly, the Multi-Spatial-Temporal Encoder-Decoder Model (MST-EDM) [78] processes different temporal scales with distinct encoders that are later fused, offering a granular analysis of temporal correlations.

The second type adopts a more unified or pure transformer approach, leveraging a streamlined architecture to enhance processing efficiency and scalability. The Fast Pure Transformer Network (FPTN) [48], for instance, restructures traffic flow data analysis by aligning it along the sensor dimension, applying multiple layers of Transformer encoders to simultaneously

capture complex spatio-temporal correlations. This method significantly reduces computational time and resource usage. Additionally, the Bidirectional Spatial-Temporal Adaptive Transformer (Bi-STAT) [51] uses an adaptive encoder-decoder architecture that dynamically adjusts to the complexity of traffic forecasting tasks. It integrates innovative features like a dynamic halting module (DHM) for iterative computation, enhancing the adaptability and accuracy of the forecasts.

The diversity of these models demonstrates the breadth of innovation in this area. Each approach provides unique insights and methodologies, contributing to an extensive and more diverse toolkit for traffic analysts and urban planners.

2) *Traffic Speed/Time Forecasting*: Traffic speed forecasting and time forecasting are closely intertwined within ITS, playing vital roles in improving navigational routing and estimating arrival times in various applications. Traffic speed forecasting involves estimating the average velocity of vehicles over a specific segment during a defined time interval. In contrast, traffic time forecasting predicts the duration required to travel through a particular route or segment. Due to the intrinsic relationship between travel speed and travel time, both forecasting types face similar challenges, such as information scarcity and heterogeneous traffic conditions. Recent research in this field [59]–[68], [79]–[81], has effectively utilized both spatial and temporal dimensions in traffic speed data, similar in traffic flow forecasting.

Information scarcity is a biggest problem in these fields, which highlights the difficulty in generating accurate predictions when faced with limited, incomplete, or sparse traffic data [63],

[82]. Such scarcity can stem from various reasons, including the lack of coverage by sensor networks, the high costs associated with the deployment and maintenance of extensive traffic monitoring systems, and the challenges in collecting data on roads with low traffic volumes or in remote areas. In response to these challenges, Liu et al. [62] have developed a technique that applies a data recovery algorithm based on identifying nonlinear spatial and temporal correlations within the road network. This algorithm helps impute missing speed data for different segments and enables traffic speed forecasting across a diverse and heterogeneous road network. Moreover, Huang et al. [82] have utilized Probabilistic Principal Component Analysis (PPCA) to model travel speeds reliably, even when data is missing from specific road segments. They have also employed spectral clustering to categorize roads with similar traffic conditions into clusters, which reduces the variability of traffic conditions within each group. This enhances predictive consistency and facilitates parallel computing to improve overall prediction performance.

In traffic speed forecasting, Liu et al. [62] employs the GraphSAGE model [83], a novel approach tailored for sparse network conditions, to enhance the accuracy of traffic speed predictions. This approach emphasizes the importance of spatial information in the context of sparse connectivity. Khodabandeh et al. [59] proposed CAGRU, combining graph convolution techniques with attention-based gated recurrent units [71] to capture both spatial and temporal relationships within traffic speed data. This fusion approach enriches the model's understanding of complex traffic dynamics. Zhang et al. [65] proposed a multi-task learning framework (MTL-GRU) that simultaneously processes traffic flow and speed forecasting. This approach enables the model to learn from the intertwined nature of traffic speed and flow, leading to a more nuanced representation of spatio-temporal data and enhancing the predictive accuracy for both metrics.

In traffic time forecasting, Tran et al. [67] have taken the lead in this field by incorporating traffic flow forecasting models into their travel time prediction system, called DeepTrans. Their methodology uses machine learning to examine vast datasets of historical traffic patterns, allowing for more precise travel time estimations. Diving deeper into the interplay between spatial and temporal factors, Kang et al. [80] introduced a novel spatio-temporal forecasting framework focused on the urban context. This approach can process and integrate multifaceted data streams, capturing the intricate dynamics of urban traffic. The model considers not only the physical layout of the transportation network but also the fluctuating congestion levels over time. By assimilating this spatio-temporal information, their model extracts essential representations that significantly improve the reliability of travel time forecasts.

Although traffic speed/time forecasting and traffic flow forecasting are related, they are still distinct areas in transportation domains. Traffic flow forecasting offers a macroscopic view, focusing on overall traffic conditions and trends across a broader area or network [49], [58], [84], which involves understanding traffic patterns, volume, and congestion across a network. On the other hand, traffic speed/time forecasting delves into the microscopic details, emphasizing the temporal elements of

travel. It provides detailed insights into the travel duration between specific locations, making it valuable for journey planning and management [85]–[87]. More details about the related work for traffic forecasting can be found in Table.I.

3) Discussion: How to Construct a Graph in Traffic Forecasting? Contemporary methods primarily align with two categories: *empirical graph construction* and *learning-based graph construction*. Empirically constructed graphs, such as road networks, are commonly used [48], [53], [56], [58]. Techniques like dynamic time wrapping also play a role in capturing dependencies among time series [52], [88]. Recent advancements have introduced methods to learn underlying graph structures from spatio-temporal data [49], [50], [52], [89], [90]. Some approaches [52], [89] focus on dissecting and understanding spatial and temporal structures individually, then integrating them for advanced spatio-temporal forecasting. For instance, Li et al. [52] proposed Spatial-Temporal Fusion Graph Neural Networks (STFGNN), which employ a spatial fusion graph along with a temporal graph, showcasing the efficacy of multi-dimensional graph structures in analysis. In contrast, other studies [49], [50], [90] aim to simultaneously model spatio-temporal structures. Zhao et al. [49] introduced a dynamic hypergraph over the spatio-temporal graph to enhance forecasting accuracy. Meanwhile, the Dynamic Spatial-Temporal Aware Graph Neural Network (DSTAGNN) [50] focuses on learning integrated spatio-temporal graphs, utilizing multi-head attention [75] to capture dynamic spatial relevancies.

How standard GNNs are modified for traffic forecasting? Traffic forecasting involves predicting future traffic patterns by leveraging temporal information for accuracy. This process is divided into two sub-types, both of which utilize similar methodologies and share the fundamental principle that integrating spatio-temporal information is essential for enhancing model performance. To achieve this, a combination of GNNs [83], [91], [92] and sequential models such as RNNs, Long Short-Term Memory networks (LSTMs) [70], [93], and Gated Recurrent Units (GRUs) [71] is commonly employed. Sequential models excel at processing time-series data, making them particularly effective for capturing and forecasting time-dependent traffic patterns. When combined with GNNs, which model spatial relationships, this approach adeptly captures both temporal sequences and spatial dependencies within traffic data, resulting in a more accurate and comprehensive traffic forecasting model.

B. Vehicle Control System

Vehicle Control System (VCS) is an essential component of ITS, which is responsible for optimizing the operation, safety, and efficiency of autonomous vehicles in traffic networks. Recently, learning-based methods, especially GNNs, have demonstrated significant potential in addressing the complexities of urban environments. This section will delve into two key subsystems within VCS: perception, which involves interpreting sensory data, and trajectory prediction, which focuses on forecasting vehicle movement. Finally, we present an overview of the most related studies for VCS in Table.II.

1) Perception: Perception plays a vital role in vehicle control system in identifying and categorizing objects around a vehicle.

TABLE II: A Comprehensive Overview of Most Related Studies for Vehicle Control System

Model	Article	Year	Datasets	GNN Module	Summary
GTNet	[94]	2023	ModelNet40, ShapeNet part	Graph Transformer	GTNet uses a Local Transformer to calculate neighboring point weights through dynamic graph-based cross-attention within domains, and a Global Transformer to expand its range using global self-attention.
MHNet	[95]	2023	ModelNet40, NTU	Spectral GNN	MHNet introduces a polynomial hypergraph filter, which dynamically extracts multi-scale node features.
DiffConv	[96]	2022	ModelNet40, Toronto3D, ShapeNet part	Spectral GNN	DiffConv uses density-dilated neighborhoods where each point's radius depends on its kernel density. It also uses masked attention to introduce task-specific learned variations to the neighborhood.
DeltaConv	[97]	2022	ModelNet40, SHREC11, ScanObjectNN, ShapeNet	Spectral GNN	DeltaConv uses a graph-based anisotropic convolutional operator by combining a set of geometric operators defined on scalar and vector fields to encode the directional information of each surface point.
3DCTN	[98]	2022	ModelNet40, ScanObjectNN	Graph Transformer	3DCTN combines convolutions and transformers to learn local and global features. It uses a multi-scale local feature aggregation block and a global feature learning block to process downsampled point sets.
Point Transformer	[99]	2021	S3DIS, ModelNet40, ShapeNet part	Graph Transformer	Point Transformer introduces an expressive transformer layer tailored for point cloud processing. It employs local self-attention and integrates vector attention to achieve elevated accuracy levels.
PCT	[100]	2021	ModelNet40, ShapeNet	Graph Transformer	Point Cloud Transformer (PCT) improves capturing local context capture within the point cloud by using coordinate-based input embedding with the help of farthest point sampling and nearest neighbor search.
CurveNet	[101]	2021	ModelNet40, ModelNet10, ShapeNet part	Spatial GNN	CurveNet enhances point cloud shape descriptors by organizing connected points through guided walks within point clouds and aggregating them to enhance their individual point-wise features.
LDGCNN	[102]	2021	ModelNet40, ShapeNet	Spatial GNN	LDGCNN is a linked dynamic graph CNN created for direct classification and segmentation of point clouds, addressing sparsity and unstructured nature. It also includes theoretical analysis and model visualization.
3D-GCN	[103]	2020	ModelNet40, ModelNet10, ShapeNet part	Spatial GNN	3D-GCN is a novel approach for processing 3D point clouds in computer vision that offers scale and shift invariance by utilizing learnable kernels and a graph max-pooling mechanism to extract robust features.
DHGNN	[104]	2019	ModelNet40, NTU	Spectral GNN	DHGNN addresses limitations in graph/hypergraph-based deep learning by dynamically updating hypergraph structures and encoding high-order data relations through vertex and hyperedge convolutions.
DGCNN	[105]	2019	ModelNet40	Spatial GNN	DGCNN, a novel neural network module dubbed EdgeConv suitable for point clouds, enhances CNN-based high-level tasks by incorporating local neighborhood information and adapting to topology.
RGCNN	[106]	2018	ShapeNet part	Spectral GNN	RGCNN directly processes point clouds, utilizing spectral graph theory and Chebyshev polynomial approximation to capture dynamic graph structures adaptively, enhancing point cloud understanding.
AGCN	[107]	2018	Sydney urban	Spectral GNN	AGCN, a flexible Graph CNN that takes data of arbitrary graph structure as input, enables task-driven adaptive graph and distance metric learning for diverse data such as molecular and social networks.
KCNet	[108]	2018	ModelNet40, ShapeNet	Spatial GNN	KCNet improves semantic learning efficiency for 3D point clouds by introducing a point-set kernel for 3D geometry and recursive feature aggregation on a nearest-neighbor graph that focuses on local structures.
Local-SpecGCN	[109]	2018	ModelNet40, McGill Shape, ShapeNet part, ScanNet Indoor Scene	Spectral GNN	Local-SpecGCN uses spectral graph convolution on local graphs and a graph pooling strategy for point cloud feature learning, enhancing feature descriptors by aggregating information from clustered nodes.
ECC	[110]	2017	Sydney Urban Objects, ModelNet10, ModelNet40	Spatial GNN	ECC adapts convolution operators for arbitrary graphs, avoiding the spectral domain, and uses specific edge labels in a vertex's neighborhood to condition filter weights, enabling diverse graph classification tasks.

Perception involves two critical tasks: semantic segmentation with classification and object detection with tracking [111], which includes clustering and assigning specific classes to pixels in an image. In perception, point cloud is a common method for representing 3D data. It can capture complex 3D shapes and their unique irregular structures. Traditional deep learning methods usually convert point clouds into 3D voxel grids or collections of images before feeding them into deep neural networks, potentially resulting in information loss and computational overhead [112]. An alternative approach leverages the graph-like nature of point clouds, leading to a growth in research efforts employing GNNs to enhance the efficiency and accuracy of 3D data analysis. In the following sections, we will discuss GNN-based methods for learning representations from point cloud data.

Graph-Based Methods in Spatial Domain. Spatial Convolutional Graph Neural Networks can be described as the process of spreading node features to neighboring nodes by utilizing a convolutional kernel. This is then followed by applying an activation function using a trainable weight matrix to transform these features into the subsequent hidden layer [113].

As a pioneering approach, Simonovsky et al. [110] introduced Edge-Conditioned Convolution (ECC) as the first graph-based method in a spatial domain. This method uses edge labels in vertex neighborhoods to calculate adaptive convolution kernel weights, thus enabling more effective utilization of edge information compared to traditional point-based convolutions. However, ECC [110] has a limitation in that it primarily depends on the inherent graph structure of the input point cloud, which inherently restricts its flexibility to model non-local relations. To address the limitation, several methods [102], [105], [114] have been proposed.

Dynamic Graph Convolutional Neural Network (DGCNN) [105] introduces an EdgeConv neural network architecture to segment point clouds and capture semantically related

structures. This approach learns a dynamic graph representation of the point cloud, which evolves across layers and during the training phase as learnable parameters are updated. Building upon earlier developments like ECC and DGCNN, the Linked Dynamic Graph Convolutional Neural Network (LDGCNN) [102] enhances the capabilities of DGCNN by establishing links between hierarchical features derived from various dynamic graphs. This linkage enables the computation of informative edge vectors while simultaneously reducing the model's size.

To capture the local neighborhood structural information of a point, kernel-based approaches have been extensively explored [101], [103], [108]. For instance, KCNet [108] introduced a point-set kernel of learnable 3D points. They utilized a kernel correlation layer to compute affinities between each data point's nearest neighbors and these point-set kernels. They also implemented recursive feature propagation and aggregation along the edges, effectively utilizing local high-dimensional feature structures. Similarly, 3D-GCN [103] proposed deformable kernels that were designed to extract shift and scale-invariant local 3D features from point clouds. Furthermore, Xiang et al. [101] introduced a method to arrange connected points through guided walks within the point clouds and subsequently aggregated them to enhance their point-wise features, effectively improving the representation of point cloud geometry.

Graph-Based Methods in Spectral Domain. Spectral Convolutional Graph Neural Networks are based on spectral graph theory [115]. In this framework, graph signals are filtered through the eigendecomposition of the graph Laplacian. Regularized Graph CNN (RGCNN) [106] perform graph convolution and feature learning based on spectral graph theory, which treats point cloud features as signals on a graph and employs Chebyshev polynomial approximation for graph convolution. RGCNN adapts to the corresponding learned features by updating the graph Laplacian matrix in each layer, effectively capturing evolving graph structures during

the learning process. Traditional spectral GCNs require the prior computation of graph Laplacians and pooling hierarchies for the entire graph, which can be computationally intensive. In dealing with the challenges of diverse graph topology in data, two promising approaches have been proposed. One approach is the Adaptive Graph Convolutional Neural Network (AGCN) [107], which enhances the generalization capacity of GCNs by incorporating a learnable distance metric to parameterize the similarity between two vertices within a graph, allowing for the dynamic construction of graphs. The other approach is Local-SpecGCN [109], which conducts spectral filtering on dynamically generated local graphs. It uses recursive clustering based on spectral coordinates to facilitate graph pooling, which enhances the learning process by mitigating point isolation. Instead of using conventional max pooling, Wang et al. introduced a recursive clustering and pooling strategy that enables the amalgamation of information from nodes within clusters defined by their spectral coordinates.

Hypergraphs attract the attention of researchers as a tool for capturing high-order data correlations. One notable example is Hypergraph Neural Networks (HGNN) [104], which uses a hyperedge convolution operation to capture high-order data correlations and represent complex structures within point clouds. This operation aggregates node features into hyperedge features and then updates node features through hyperedge feature aggregation. Hypergraph Gragh Convolutional Network (HyperGCN) [116] uses non-linear Laplacian operators [117] to convert hypergraphs into more straightforward graphs by breaking hyperedges down into subgraphs with edge weights that depend solely on their degrees. Hypergraph convolution relies on a predefined structure for propagation. To address this limitation, Bai et al. [118] introduced an attention mechanism for dynamic connection learning among hyperedges. This mechanism ensures that information propagates and gathers in graph regions relevant to specific tasks, leading to the learning of more discriminative node embeddings. Multi-modal Hypergraph Neural Network (MHNet) [95] uses hypergraph structures to model high-order and multi-modal data correlations effectively. It achieves this by employing a polynomial hypergraph filter that dynamically extracts multi-scale node features through parametric polynomial fitting.

Recent advancements have been made in convolution operations for point clouds. Conventional approaches commonly perform convolution operations to the irregular point clouds by imposing a fixed view, such as using fixed neighborhood sizes. To address this issue, DiffConv [96] introduced density-dilated neighborhoods, where the radius for each point depends on its kernel density. DiffConv uses masked attention, which introduces task-specific irregularity to the neighborhood, making the convolution more flexible and effective. Another approach, DeltaConv. [97] proposed a new method to construct anisotropic convolution layers for geometric CNNs. It designed a graph-based anisotropic convolutional operator by combining a set of geometric operators defined on scalar and vector fields to encode directional information for each surface point.

Graph Transformer-based Methods. While transformers have been extensively used in computer vision, graph-based transformers are explicitly tailored for learning 3D point cloud

representation. The transformer architecture is well-suited for point cloud analysis due to its self-attention operator, which functions as a set operator by preserving permutation and cardinality invariance of input elements [99]. As an example within this category, Point Transformer (PT) [99] introduces a transformer layer that is highly expressive and specifically designed for point cloud processing. The Point Transformer employs local self-attention that ensures scalability even in large scenes. Additionally, integrating vector attention is pivotal in achieving elevated accuracy levels. Another tailored transformer for point clouds is Point cloud transformer (PCT) [100]. PCT employs a coordinate-based input embedding module to learn distinctive features by combining raw positional encoding and input embedding, harnessing the individual spatial coordinates of each point. Furthermore, it enhances performance by substituting the original self-attention module with an offset-attention module. Unlike PT, PCT excels in capturing global interaction and local neighborhood information.

To improve efficiency in point cloud classification, 3D Convolution-Transformer Network (3DCTN) [98] by Lu et al. combines GNN convolutions with transformers, which helps effectively learn local and global features. To be more specific, 3DCTN processes downsampled point sets using a combination of a multi-scale local feature aggregating block and a global feature learning block, which are implemented by GNNs and Transformers, respectively. While most Transformer-based approaches primarily rely on global attention mechanisms to extract point cloud features, they often fail to adequately capture feature learning from local neighbors. To address this challenge, Graph Transformer Network (GTNet) [94] employs local and global Transformer modules. The local Transformer module calculates neighboring point weights through dynamic graph-based intra-domain cross-attention, which assigns different weights to each neighboring point's influence on the centroid's features. In contrast, the global Transformer module expands the local Transformer's reach by utilizing global self-attention, which enables a broader feature extraction.

2) *Trajectory Prediction:* Trajectory prediction involves anticipating the future paths of road users based on their past trajectories and the surrounding environment. Road users include vehicles, cyclists, and pedestrians. The environment includes static factors such as terrain and obstacles and dynamic factors such as the movements of nearby agents [119].

Several models have been developed to improve trajectory prediction by adopting the paradigm of spatial and temporal convolution through GNNs. For instance, GRIP [120] enhanced trajectory prediction by incorporating interactions among adjacent objects, represented as an undirected graph. It employs a GCN module to model the graph network, and the output of GCN is then input into an LSTM encoder-decoder for predicting the trajectories of surrounding vehicles. SCALE-Net [121] creates an efficient and scalable framework to maintain high prediction performance for numerous vehicles. It employs an Edge-Enhanced Graph Convolutional Network (EGCN) to update node features based on an attention mechanism influenced by edge features from neighboring nodes. Social-STGCNN [122] represents pedestrian trajectories as spatio-temporal graphs and employs GCN and TCN to operate on

these graphs, enabling the model to predict the entire sequence simultaneously. Chandra et al. [123] uses a two-layer Graph-LSTM architecture for trajectory prediction. The initial layer is applied to forecast future trajectories of traffic participants, while the second layer captures interaction-related factors among participants using a weighted dynamic geometric graph network (DGG). Additionally, they introduced a regularization algorithm based on spectral clustering to minimize errors in long-term predictions. GSTCN [124] uses a GCN to capture spatial interactions and a CNN to handle temporal correlations among neighboring vehicles. The spatial-temporal features are encoded and decoded using a GRU network in their framework.

Recently, the attention mechanism has gained popularity for various sequence-based tasks, including predicting the trajectory of autonomous vehicle systems. The Spatial-Temporal Graph Attention network (STGAT) [125] utilizes an LSTM encoder to encode trajectories, followed by GAT for attention-weighted interaction information, and an LSTM decoder for trajectory prediction. SCOUT [126] utilizes GAT to incorporate dynamic agent interactions, aiming to improve socially aware and consistent trajectory predictions. The Attention-based Spatio-Temporal Graph Neural Network (AST-GNN) [127] utilizes a dual-attention mechanism: the first for capturing spatial interactions among all agents and the second for considering the temporal movement patterns of each agent in the past. The Spatio-Temporal Graph Dual-Attention Network (STG-DAT) [128] employs a dual-attention mechanism to learn representations on spatio-temporal dynamic graphs, considering historical and future features from state, relation, and scene context information. The Triple Policies Fused Hierarchical Graph Networks (Tri-HGNN) [129] proposed triple policies fused hierarchical GNN for pedestrian trajectory prediction. More specifically, the extrinsic-level policy uses GAT for spatial and temporal embeddings, the intrinsic-level policy captures human intention with GCN, and the basic-level policy combines information for predictions through TCN. The Heterogeneous Driving Graph Transformer (HDGT) [130] models the driving scene as a heterogeneous graph, where agents, lanes, and traffic signs are considered as different types of nodes and edges. Besides, the transformer structure is applied hierarchically to accommodate the heterogeneous inputs.

3) Discussion: How to Construct a graph in Vehicle Control System? In vehicle perception, 3D representations of real-world objects can be obtained using 3D sensors or Light Detection and Ranging (LiDAR) technology. The individual points in a point cloud can be seen as nodes in a graph, while determining the graph structure is a complex process. Connecting all edges for the entire point cloud can require a lot of memory. Therefore, a simplified method involves using the K-Nearest Neighbors (KNN) approach to create a locally directed graph [102]. Moreover, there have been proposals for using kernel-based methods to calculate the affinities between each data point [101], [103], [108]. In trajectory prediction, both the agent (such as a vehicle or pedestrian) and static features in the environment (such as traffic signals and road signs) can be viewed as nodes in a graph. The connections between these nodes can represent their physical proximity or potential interaction. Node attributes may encompass the

entity's speed, acceleration, heading, and previous path, while edge properties may include relative speed, relative heading, and distance between the two entities. To account for temporal aspects, temporal encoding methods can be utilized, such as using time steps as node or edge attributes or employing recursive GNN variations for processing time series data.

How standard GNNs are modified for Vehicle Control System? When GNNs are deployed in real-time and high-stakes environments such as VCS, adapting standard GNN architectures is essential to meet specific operational demands. One key challenge is the dynamic nature of the traffic environment, where the graph structure frequently changes. To address this, methods like edge weighting are crucial for dynamically updating the graph structure without retraining the model from scratch. Another important consideration is incorporating both spatial and temporal awareness. Spatial-temporal GNNs are often employed to integrate spatial relationships and temporal dynamics into the model, which enhances its ability to understand and predict traffic patterns. Additionally, autonomous vehicles require real-time or near-real-time responses, making the computational efficiency of GNNs critical. Standard GNNs can be computationally intensive, especially with large graphs, so optimizing the network for low latency is essential. This can be achieved through model simplification, pruning, and reducing the number of layers to accelerate computation. Finally, in high-stakes scenarios, GNNs must account for uncertainty in their predictions. Techniques such as Bayesian GNNs [131] can provide a measure of confidence, which is vital for safe and reliable decision-making.

C. Traffic Signal Control

Traffic signal control (TSC) involves managing and coordinating traffic lights at intersections and road junctions to regulate traffic flow, improve road safety, and minimize delays. Effective TSC helps reduce the likelihood of accidents by managing the movement of vehicles and pedestrians. It also has the potential to reduce travel time and fuel consumption, leading to lower vehicle emissions and supporting environmental sustainability. However, the intricate interactions between intersections and the continuously changing traffic conditions make the real-world network of intersections highly complex, posing a significant challenge for adaptive traffic signal control.

1) Multi-agent Graph Reinforcement Learning: Single-agent reinforcement learning methods are constrained to managing traffic signals in a single intersection due to the curse of dimensionality when using a global single model for all intersections. As a result, single-agent RL is typically restricted to an isolated intersection without coordination with neighboring intersections [143]. In managing multiple intersections, a practical approach is to obtain neighborhood intersection information by combining the states of intersections and their adjacent areas [144]. However, this method encounters difficulties when dealing with increasing input dimensionality, leading to challenges in model convergence. Thankfully, multi-agent reinforcement learning allows for the individual control of each signal using a reinforcement learning (RL) agent and developing policies for each intersection, resulting in

TABLE III: A Comprehensive Overview of Most Related Studies for Traffic Signal Control

Model	Article	Year	Datasets	Simulator	Temporal Module	Spatial Module	Attention Based	Summary
AFMRL	[132]	2023	Simulated and real-world data (Jinan, Hangzhou, Manhattan)	CityFlow	-	GNN	✗	A multi-agent reinforcement learning approach for multi-intersection TSC. Adaptive partitioning is emphasized and feudal hierarchy is explored.
KeyLight	[133]	2023	Simulated and real-world data (Jinan, Hangzhou, New York)	CityFlow	-	GAT	✓	KeyLight integrates reinforcement learning and GNNs. NOV-LADLE state representation and residual connections are used in the model.
HG-M2I	[134]	2023	real world data (Chengdu)	SUMO	GRU	Bi-GRU	✓	The HG-M2I algorithm, spatial-temporal analysis and multi-agent RL based, optimizes TSC by hierarchical graph structures and input-output correlation.
MetaSTGAT	[135]	2022	Simulated and real-world data (Jinan, Hangzhou)	CityFlow	LSTM	GAT	✓	MetaSTGAT meta-learning based, merges GAT and LSTM to address spatial-temporal correlations and dynamic interaction of intersections.
PRCOLight	[136]	2022	Simulated and real-world data (Jinan, Hangzhou, New York)	CityFlow	-	GNN	✗	PRCOL uses lane capacity for the RL reward function and GNN modules to help RL decide the light phase and duration by predicting traffic flow.
DynSTGAT	[137]	2021	Simulated and real world data (Jinan, Hangzhou, New York)	CityFlow	TCN, LSTM, STGAT	STGAT	✓	DynSTGAT combines spatial-temporal graph attention networks and temporal convolutional network to enhance adaptive TSC.
IHG-MA	[138]	2021	Simulated and real-world data (Chengdu)	SUMO	Bi-GRU	Bi-GRU	✓	IHG-MA uses inductive heterogeneous GNNs to capture traffic features and a decentralized multi-agent actor-critic framework to optimize TSC.
GraphLight	[139]	2021	Simulated data	SUMO	-	GCNN	✗	GraphLight is a decentralized, graph-based, multi-agent system using actor-critic methods for TSC, distinguishing neighboring intersection impacts.
TSC-GNN	[140]	2021	real world data (Jinan, Hangzhou)	-	-	GAT	✓	TSC-GNN is a graph-based model for TSC utilizing probabilistic neural networks, to manage uncertainties and calculate Q-values.
STMARL	[141]	2020	Simulated and real-world data (Hefei, Hangzhou)	CityFlow	RNN	GNN	✓	STMARL applies spatial-temporal RL for TSC, using graphs, RNNs, GNNs, and deep Q-learning for distributed decision-making.
CoLight	[142]	2019	Simulated and real-world data (Jinan, Hangzhou, New York)	CityFlow	-	GAT	✓	CoLight uses graph attention networks for TSC, and captures spatial-temporal impacts from nearby intersections without indexing

significant advancements. Additionally, the application of multi-agent graph reinforcement learning has demonstrated promising progress [136], [137], [139], [145], [146].

Nishi et al. [147] are among the ones who first combine multi-agent reinforcement learning and graph neural networks to address the multi-intersection interaction problem and the spatial dependency. Their work employs GCNs to extract the geometric features. Zhong et al. [140] proposed a model named TSC-GNN to handle a problem that most studies model traffic state deterministically and to exploit the uncertainties of traffic conditions. Yoon et al. [148] claimed that the RL method encountered a restricted exploration problem, which means it cannot handle unseen conditions. They proposed a novel approach to obtain a transferable policy by using graph representation for the state and training it by GNNs. Based on Multi-Agent Reinforcement Learning (MARL), Saki et al. [149] used multi-objective reinforcement learning (MORL) to further improve the performance by determining the policy corresponding to each traffic flow ratio, which achieved the shortest average travel times in all environments compared with ruled based and single objective reinforcement learning. Some more similar literature [133], [134], [136]–[139], [150] based on graph reinforcement learning is listed in Table.III .

2) *Attention Mechanism*: Another issue is distinguishing the impact of neighboring traffic signals on the target intersection. For example, intersections on the main road may significantly affect the target intersection more than those on the side road. Most existing research does not differentiate the influence of surrounding intersections on the target intersection [147], [151]. The attention mechanism has proven valuable in adaptive signal light control. CoLight [142] advanced this by using GATs to identify the impact of neighboring intersections and effectively leverage joint intersections. This was achieved by developing an index-free model of neighboring intersections and averaging their influences using learned attention parameters. Sun et al. [133] introduced NOV-LADLE to address potential convergence failures in the attention mechanism, maintaining a concise state and emphasizing essential intersections. They also added a residual connection structure to GAT to accelerate convergence and enhance performance based on CoLight. Further studies, including DynSTGAT and TSC-GNN [137], [140], have also explored using the graph attention mechanism to tackle this issue. Table III summarizes these works.

3) *Spatial and Temporal Dependency*: Similar to the previous discussion, considering spatial and temporal relations is crucial. The historical states of surrounding intersections affect the prediction of future signals at a target intersection.

Wang et al. [141] pioneered the study of spatio-temporal dependencies among multiple traffic signals. They used graph structures to capture spatial features and recurrent neural networks to integrate historical traffic data, with decisions for each signal based on deep Q-learning. Li et al. [152] proposed a model using LSTM and GCN to extract spatial-temporal traffic features of intersection networks. LSTM processed variable-length inputs and extracted valid features from historical data, while GCN handled the LSTM output, linking interactions of intersections. They incorporated imitation learning instead of reinforcement learning. Yang [134] introduced the Hierarchical Graph Multi-agent Mutual Information (HG-M2I) algorithm to generate optimal embeddings of traffic networks. This algorithm fused multi-granularity information from each agent's current and historical states to develop optimal TSC policies, measuring the correlation between input states and output embeddings by maximizing mutual information.

Despite numerous studies attempting to incorporate temporal and spatial influences of surrounding intersections into the target intersection, they typically consider the spatial-temporal information separately. Wu et al. [137] proposed DynSTGAT, which uses TCN to simultaneously capture historical and current spatial-temporal information. To address dynamically changing traffic, Wang et al. [135] proposed MetaSTGAT, a meta-learning model based on GATs that adapts to dynamic traffic flow and fully utilizes the spatial-temporal characteristics of multi-intersections. Other literature also explores the exploitation of spatial and temporal information [135], [138].

4) *Discussion: How to construct a graph in traffic signal control?* In real-world scenarios, intersections often exhibit complicated structures. For instance, different roads can be connected to the same intersections, each with a varying number of lanes. To address this complexity, we can create a directed traffic light adjacency graph that reflects the spatial relationships among traffic lights. In this graph, nodes represent intersections, with additional states indicating the presence of traffic lights. The edges represent the roads connecting these intersections, with directional attributes specifying whether roads are one-way or two-way. Moreover, additional features, such as the number of lanes on each road, the queue length and the number of

vehicles in the lane, and the average speed of vehicles, should be considered. Models' performance can be assessed based on the average travel time of vehicles [141].

How standard GNNs are modified for traffic signal control? In the context of traffic signal control, the adaptation of standard GNNs involves addressing the unique challenges of dynamic traffic environments and the interdependencies between intersections. The first adaptation involves integrating GNNs with multi-agent reinforcement learning (RL) to effectively manage multiple intersections. Each intersection is controlled by an RL agent operating collaboratively within a Markov Decision Process. By employing multi-agent RL, GNNs enable these agents to model the spatial relationships and geographical structures of intersections, facilitating seamless information sharing and integration. The second adaptation involves combining GNNs with components that process sequential data, such as RNNs or LSTM networks. This approach captures temporal patterns by using GCNs or GATs to understand spatial relationships, while LSTMs handle the temporal sequencing of traffic data, which is crucial for predicting future states. Furthermore, the incorporation of attention mechanisms allows for differential weighting of the influence of various intersections and temporal phases. These adaptations render GNNs as a dynamic solution capable of handling the temporal and spatial complexities inherent in traffic signal control systems.

D. Transportation Safety

Transportation safety aims to identify high-risk areas for accidents and incidents, crucial for pinpointing hotspots for future improvements. By understanding the factors contributing to these risks, researchers can develop targeted interventions and policies to mitigate them, thereby enhancing overall transportation system safety.

1) Spatial and Temporal Granularity: In transportation safety, prediction models can be classified into four categories based on temporal and spatial granularity. Prediction periods divide models into long-term (day-level) [158], [162], [163] and mid-term (hour-level) [155], [160], [164]–[166]. Prediction regions distinguish models into link-level [156], [161] and region-level [153], [157], [167].

Zhou et al. [160] introduced a three-stage RiskOracle framework for minute-level citywide traffic accident prediction using a Multi-task Differential Time-varying GCN (Multi-task DTGN) to model dynamic subregion correlations. Zhang et al. [158] proposed GraphCast, a multi-modal sensing and GNN-based approach for regional accident prediction, utilizing social media and remote sensing data to handle noisy and heterogeneous multi-modal data. Tran et al. [153] developed a Multi-structured GNN (MSGNN) for predicting traffic incidents across entire networks, extracting area-wide features from various data sources for faster and more efficient incident prediction rather than link-level synchronization and map-matching. Huang et al. [168] noted that many machine learning techniques predict the number of traffic accidents in each cell of a discretized grid without considering the underlying graph structure of road networks. Accurate link-level prediction

requires a complex fusion of heterogeneous data resources and "map-matching" to represent all data with different granularity in the same map system.

2) Spatial-temporal Correlation: The work by Yu et al. [156] and Yuan et al. [162] both highlight that existing methods either ignore spatial-temporal correlations or make predictions at a coarse-grained level without considering the underlying graph structure of road networks.

Zhou et al. [159] proposed RiskSeq that addresses sporadic events with a self-adaptive ranking method. It employs a Differential Time-varying GCN (DT-GCN) enhanced with node-wise proximity and signal-wise differential operations to capture dynamic traffic and accident correlations. The framework also includes a Context-Guided LSTM for decoding risks across multiple spatial scales, focusing on spatiotemporal multi-granularity urban traffic risk prediction. Yu et al. [156] tackled link-level accident prediction with a spatio-temporal convolutional network framework. Their model predicts link-level incident risk by learning spatial-temporal features from a road network graph, using graph convolutional operations to capture dynamic spatial and temporal variations. Liu et al. [154] proposed a multi-task learning framework (TAP) based on edge computing, using spatio-temporal variational graph auto-encoders to enhance prediction accuracy by analyzing dynamic spatial-temporal traffic data correlations and integrating external factors. Wang et al. [157] introduced GSNet, a region-wide accident prediction model capturing geographical and semantic spatial-temporal correlations. The model features a weighted loss function to address the zero-inflation issue. Table.IV lists other spatial-temporal models [161], [169].

3) Discussion: How to construct a graph in transportation safety? A straightforward way to construct a graph in transportation safety is using zone-based methods, which models target area as several medium-sized regions or grids, represented as a vertex in the graph. Edges between vertices indicate the connectedness between subregions, existing when traffic elements within those subregions have strong correlations. More specifically, the traffic elements contributing to the vertex features include static road network features (e.g., road types and number of lanes) and dynamic traffic features (e.g., traffic volume and average speed). These elements help define the relationships or connectedness within the urban graph.

How standard GNNs are modified for transportation safety? When adapting GNN to transportation safety, a major challenge is dealing with the zero-inflated problem due to the rarity of accidents. As we create zone-based graphs for GNN, the frequency of zero values in the training data increases with higher spatio-temporal resolution, making it harder to make accurate predictions. Existing work addresses imbalanced anomaly data through various methods, including loss function handling [155], [157] and data preprocessing techniques like priori knowledge-based data enhancement [159], [160], negative sample undersampling [156], and graph augmentation [170]. Wang et al. [170] employed graph augmentation and contrastive loss to enhance latent representations in training, proposing an improved contrastive GNN-based framework for traffic anomaly analysis in ITS. Yu et al. [156] used negative sample undersampling to balance data by matching

TABLE IV: A Comprehensive Overview of Most Related Studies for Transportation Safety

Model	Article	Year	Datasets	Spatial Granularity	Temporal Granularity	Solution of zero-inflated	Summary
MSGNN	[153]	2023	loop detectors, GPS probe data (Brisbane, Gold Coast)	Region Level	Short-Term (1h)	clustering-based data imputation	MSGNN, as a sub-area level accident prediction model, captures spatio and temporal relations and uses a data imputation approach for sparse datasets.
TAP	[154]	2023	Real world dataset	Region Level	Short-Term (30min)	-	Multi-task learning framework (TAP) predicts traffic accidents using Spatio-temporal Variational Graph Auto-Encoders, accelerated by edge computing.
GGCMT	[155]	2022	Real world dataset (NYC)	Region Level	Mid-Term (3h)	prior risk data enhancement method.	Computer-vision-based GGCMT predicts accidents using a gated graph convolutional multi-task model, improved with prior risk data enhancement methods.
DSTGCN	[156]	2021	Real world dataset	Link Level	Short-Term	Negative sample undersampling method	DSTGCN predicts traffic accidents from heterogeneous data, capturing spatio-temporal correlations by Deep Spatio-Temporal Graph Convolutional Network.
GSNet	[157]	2021	Real world dataset (NYC, Chicago)	Region Level	Short-term(1h)	Weight loss function	GSNet predicts traffic accidents by analyzing spatio-temporal and heterogeneous data, using a specialized loss function for rare events.
GraphCast	[158]	2020	Real world dataset (NYC)	Region Level	Long-term (7day)	-	GraphCast, a multi-modal graph neural network, forecasts fine-grained traffic risks in cities by combining social media and remote sensing data.
RiskSeq	[159]	2020	Real world dataset (NYC, Suzhou)	Region Level	Short-term (10min)	priori knowledge-based data enhancement	RiskSeq, a fine-grained and multi-step accident prediction model, combines Differential Time-varying GCN and hierarchical sequence learning.
RiskOracle	[160]	2020	Real world dataset (NYC, Suzhou)	Region Level	Short-term (30min)	priori knowledge-based data enhancement	RiskOracle, a minute-level accident prediction model, combines Differential Time-varying Graph network, multi-task and region selection strategies.
TA-STAN	[161]	2019	Real world dataset (NYC)	Region Level	Mid-Term (12h)	-	TA-STAN predicts accidents by analyzing real-world traffic data, vehicle types, and external factors with a Spatial-Temporal Attention Network.

non-accident samples with accident samples, leveraging spatial-temporal GNNs for more accurate predictions. Wang et al. [167] explored a chain-like triggering mechanism connecting accident occurrences, utilizing Spatial-Temporal Categorical GNNs (STC-GNN) to manage multi-dimensional and chain effects for fine-grained temporal accident prediction.

E. Traffic Demand Forecasting

Traffic demand prediction involves predicting the volume of users needing transport to or from specific areas or locations, which is crucial for scheduling services efficiently and enhancing overall transport system responsiveness. However, the growth of modern cities has caused an increase in traffic-related issues, which puts much pressure on public transportation systems. To tackle this problem, ride-hailing services such as Uber, Lyft, and DiDi, as well as bike-sharing services like MoBike have emerged as potential solutions [171], [172]. As a result, there is now a pressing need for accurate traffic demand prediction that can precisely forecast future crowd demands, thus scheduling future transportation services and other downstream tasks. [173]. However, predicting traffic demand poses challenges. On the one hand, human behavior, influenced by weather and special events, introduces variability that makes accurate forecasting even more complicated. On the other hand, integrating diverse data sources for comprehensive analysis requires advanced models that are capable of processing complex, large-scale datasets in real-time.

1) *Traffic Zone-based Graph Methods*: One of the pioneering works in utilizing graph learning methods for demand prediction tasks is Spatio-Temporal Multi-Graph Convolution Network (ST-MGCN) [174]. It proposes to exploit graph structures from multiple perspectives to capture comprehensive information on the spatio-temporal characteristics of traffic systems. Specifically, ST-MGCN builds the graphs of zones from three angles: a neighborhood graph based on the spatial proximity, a functionality graph defined by the POI similarity, and a transportation connectivity graph induced by road networks such as motorways, highways, or public transportation systems like subways. A multi-graph convolution is then introduced to model the spatial dependencies between regions and provide informative representations for downstream demand prediction tasks. Similarly, PGDRT [175] builds the zone-wise relational graph using three types of temporal characteristics: adjacent visual characteristics, periodic characteristics, and representative characteristics, to provide a more

comprehensive view of temporal features in traffic systems. To fully exploit the rich information from multiple traffic systems, Multiview Spatio-Temporal Graph Neural Networks (MSTGNN) [176] proposes a multiview graph that jointly depicts the demand relationship between bus, metro, and taxi demands. The multiview graph enables MSTGNN to capture the interaction dependencies among the travel demands of different transportation systems. An auxiliary loss is used to encourage the consistency between graph features from multiple views and enhance the performance of TGCN modules.

2) *Spatio-temporal Graph-based Methods*: Classical GNN models for traffic demand prediction treat the spatial dependency as a static graph and cannot depict dynamic features. However, in reality, the spatial dependencies between most nodes change over time, while others remain relatively constant. To address this limitation, the Dynamical Spatio-Temporal Graph Neural Network (DSTGNN) [177] evaluates the stability of a node's spatial dependence based on the number of dissimilar neighbors and constructs a spatio-temporal graph that evolves over time. To encode the spatio-temporal information, the model uses a spatio-temporal embedding network that combines a Diffusion Convolution Neural Network (DCNN) with a modified transformer.

3) *Dynamic Graph-based Methods*: The traditional approach to modeling cities is to divide them into grid-like zones and construct graphs based on these divisions. However, this approach can lead to suboptimal solutions, and adapting to dynamic graph structures remains challenging. A new solution called Deep Multi-View Spatio-temporal Virtual Graph Neural Network (DMVST-VGNN) [183] improves learning capabilities related to spatial dynamics and long-term temporal dependencies. The DMVST-VGNN method proposes a graph generation process that provides a more flexible and fine-grained perspective on the spatio-temporal relationships between regions, as opposed to the simplistic grid-based division of the map. Another proposal by Nazzal et al. [180] extends the idea of dynamic and flexible graph structures to decentralized edge-computing scenarios and introduces a heterogeneous GNN-LSTM algorithm. This algorithm is designed to handle dynamic taxi graphs where taxis serve as nodes. The proposed heterogeneous GNN-LSTM structure has demonstrated the ability to capture dynamic decentralized graph structures and has shown promising results in taxi-level demand and supply forecasting.

4) *Improvements on Graph Encoders*: The traditional graph convolution network has limited capability to represent the complex information in traffic zone graphs. However, some

TABLE V: A Comprehensive Overview of Most Related Studies for Demand Prediction

Model	Article	Year	Prediction Task	Graph Views	GNN Module	Temporal Module	Summary
PGDRT	[175]	2023	Taxi Passenger	Neighborhood, Function, Connectivity	GCN	ConvLSTM	PGDRT considers a region's unique characteristics and the influence of regions on the model of the dependent relationship between regions.
MSTGNN	[176]	2023	Bus, Metro, Taxi	Neighborhood, Connectivity	GCN	Temporal GCN	MSTGNN uses a multiview graph consisting of bus, metro, and taxi views, with each view containing both local and global graphs.
STGMT	[178]	2023	Taxi & Highway	Traffic Network	Node2Vec	Multi-head Attention	STGMT combines Multi-head Temporal Attention (MTA) and Multi-head Temporal Interactive Attention (MTIA) for temporal features.
PAG-TSN	[179]	2023	Ride-hailing	Distance, POI relation	BAT-GCN	PA-GRU	PAG-TSN uses a bicomponent attention GCN and a periodic attentional GRU to integrate the extracted spatio-temporal information.
HetGNN-LSTM	[180]	2023	Taxi	Decentralized taxi graph	HetGNN	LSTM	HetGNN-LSTM proposes a semi-decentralized approach utilizing multiple cloudlets, moderately sized storage, and computation devices.
MFGCN	[181]	2023	Ride-hailing	OD network	MODGCN	TAS-LSTM	MFGCN is a multimodal fusion GCN that consists of a multimodal module to incorporate weather and temporal activity patterns.
SGCNPM	[182]	2023	Dockless Bike-Sharing	Distance, Function, Interconnectio	MGCN	LSTM	SGCNPM considers time, built environment, and weather to create a prediction method considering the influence of multiple factors.
DSTGNN	[177]	2022	Taxi & Bike	Spatial dependency	DCNN	Multi-head Attention	DSTGNN builds spatial graphs based on the stability of the node's spatial dependence to capture the dynamical relationship.
DMVST-VGN	[183]	2022	Ride-hailing	Multi-view Graph Generation	GAT	Multi-head Attention	The Model integrates 1D CNN, Multi-Graph Attention Neural Networks, and Transformer to construct multiview spatio-temporal information.
ST-MGCN	[174]	2019	Ride-hailing	Neighborhood, Function, Connectivity	ChebNet	RNN	ST-MGCN uses GNNs to model non-Euclidean pair-wise correlations between different regions by designing a spatio-temporal multi-graph.

works aim to enhance the expressiveness of graph encoders. STGMT [178] proposes the Sandwich-Transformer for processing spatio-temporal traffic graphs, which is composed of a Multi-head Temporal Attention (MTA) and a Multi-head Temporal Interactive Attention (MTIA). PAG-TSN [179] constructs a Bicomponent Attention Graph Convolution model (BAT-GCN) and a periodic attentional gated recurrent unit model to capture geographical relationships and temporal features of different periods, respectively. While previous research primarily concentrates on processing plain time-series traffic demand data for predictions, it is essential to recognize that contextual information and multimodal attributes, such as weather conditions, significantly impact ride-hailing and other public traffic systems. To tackle these challenges, Multimodal Fusion Graph Convolutional Network (MFGCN) [181] introduces an innovative Multimodal Fusion Graph Convolutional Network for traffic demand prediction. MFGCN incorporates a Multimodal Origin-Destination GCN (MOD-GCN) that comprises three GCNs to capture spatial patterns and a Multimodal Attribute Enhancement (MAE) module for integrating dynamic weather and metadata. SGCNPM [182] utilizes multiple modules that consist of GCN and LSTM operators to model the multiple factors in a dynamic traffic system, including time periods, built environment, and weather, to predict the short-term demand of a dockless bike-sharing system. A comprehensive overview of most related studies for demand prediction can be found in Table.V.

5) Discussion: How to construct a graph in demand prediction? In demand prediction, in addition to the zone-based construction approach discussed in the transportation safety section, we can also create a traffic graph based on network topology, where nodes represent stations, bus stops, etc., and edges can be constructed based on the correlation among the features of the nodes. Research has shown the effectiveness of location-based graphs, but constructing graphs from multiple perspectives has proven to introduce more informative relationships between geographical zones. For instance, the Spatio-Temporal Multi-Graph Convolution Network (ST-MGCN) [174] builds zone graphs from three perspectives: a neighborhood graph based on geographical distances between neighborhoods, a functionality graph derived from the similarity between POI vectors, and a transportation connectivity graph induced by road networks such as motorways, highways, or public transportation systems like subways.

How standard GNNs are modified for demand prediction?

In the field of demand prediction, it is essential to adapt GNNs to accommodate the spatio-temporal characteristics of the input flow graph. This involves incorporating temporal and spatial modules and other necessary modifications. Additionally, based on the enhancement of the multi-view approach we have explored [175], [176], [183], a multi-view fusion mechanism is required to aggregate from different representations. Furthermore, to capture the dynamic nature of demands, GNNs can be combined with other sequence models such as RNNs [175], [179], [180] and self-attention layers [177], [178].

F. Parking Management

Parking management has become a significant concern due to the limited resources of parking spaces and the increasing traffic pressure. One crucial research area of intelligent parking management is predicting parking availability, which involves forecasting future parking occupancy. Being able to predict parking availability on a city-wide scale is crucial for developing effective Parking Guidance and Information (PGI) systems, such as Baidu Map [184] and Google Map [185].

Predicting the availability of parking spaces faces several challenges, such as the non-Euclidean spatial autocorrelation between parking lots, the dynamic temporal autocorrelation within and between parking lots, and the information scarcity due to the lack of real-time data obtained from sensors to determine parking availability. GNNs have been recognized as an effective approach for processing spatial-temporal and graph-based structures, addressing the mentioned challenges, and providing more accurate prediction accuracy.

As one of the pioneering works on modeling the parking availability prediction with graph-based models, SHARE [186] and its variant SHARE-X [187] proposes a Semi-supervised Hierarchical Recurrent Graph Neural Network to analyze spatio-temporal parking data. Specifically, SHARE proposes a hierarchical graph convolution module that captures non-Euclidean spatial correlations between parking lots. It consists of two blocks: a contextual graph convolution block for local spatial dependencies and a soft clustering graph convolution block for global spatial dependencies. SHARE-X extends the idea of SHARE to address the lack of real-time sensors in real-world scenarios. Particularly, It leverages a parking availability approximation module to estimate parking availability for parking lots without sensor monitoring.

To better depict the strong spatiotemporal contextual autocorrelation between vacant parking spaces, dConvLSTM-DCN

[188] analyzed the historical zone-wise parking space data and found that there is both a temporal correlation within each parking lot and a spatial correlation among different parking spaces. Based on this observation, Feng et al. proposed a deep learning framework called dual Convolutional Long Short-Term Memory with Dense Convolutional Network (dConvLSTM-DCN) to predict the availability of vacant parking places in the short-term (within 30 minutes) and long-term (over 30 minutes) zone-wisely. The framework consists of two parallel ConvLSTM components that capture the spatial correlations among parking lots and provide an informative representation of the prediction process.

The traditional methods of obtaining real-time on-street parking information rely heavily on densely deployed sensors. Therefore, the high costs of current parking availability prediction models have made it difficult to use them in many cities and areas. To tackle this issue and avoid the high costs of deploying new sensors, MePark [189] aims to predict real-time on-street parking availability across a city by using existing infrastructure and easily accessible data without relying solely on specially deployed sensors. More specifically, MePark uses an iterative mechanism to combine the aggregated inflow and individual parking duration predictions to exploit the transaction data adequately. Additionally, it extracts distinctive features from multiple data sources, combining the MGCN and the LSTM network to capture complex spatio-temporal correlations.

1) Discussion: How to construct a graph in parking management? Constructing the graph in the parking management system follows a similar approach to previous research domains, such as transportation safety and demand prediction. The zone-based construction method can predict parking availability in a medium-sized grid or area, while the network topology-based method can predict availability in specific parking lots based on their geometric information.

How standard GNNs are modified for parking management? Firstly, the dynamic temporal autocorrelation and non-Euclidean relationship between parking lots is crucial for accurate short-term parking availability prediction. Many existing efforts suggest integrating recurrent network structures, such as LSTMs, into GNN architectures [186], [188], [189], while some methods directly reform GNNs with recurrent structures, such as SHARE and SHARE-X [187]. Secondly, information scarcity poses another challenge, prompting researchers to explore GNNs to exploit non-sensor data, like transaction data, to harness the prediction potential [189].

V. CHALLENGES AND FUTURE DIRECTIONS

After analyzing the current studies on GNNs in ITS, we discuss the challenges and future directions for applying GNNs in ITS. This is important to identify any gaps that need to be addressed and to provide insights for further research.

A. Research Challenges

1) Data: Constructing datasets is one of the primary challenges when applying GNNs in ITS. Data privacy is a major concern when collecting information from traffic sensors or GPS data. Currently, there are limited publicly available

data sources, such as Data.gov, The University of Sydney's Intelligent Vehicles and Safety Systems, and the Connected Vehicle DataSets from the Safety Pilot Model Deployment [18]. Some researchers have explored multi-modal models to access data from richer sources like social media [95], [181], [190], [191]. However, these sources often face issues related to credibility and a lack of valuable information. As a result, generating a large, high-quality, and comprehensive dataset for ITS remains a formidable task, requiring continued effort to address data privacy concerns and improve data sources.

2) Model: Domain-specific Model Design. The transportation network is a complex system that includes various nodes and edges, such as roads, intersections, and vehicles. Designing GNN models in ITS that can efficiently learn from such a heterogeneous and complex structure requires significant effort. The design of GNN applications in ITS heavily depends on the specific goals of the corresponding applications, as different goals require using different graph models and construction techniques. For instance, GNNs are commonly used in traffic forecasting and travel demand modeling to predict features or variables over graph nodes. While in areas such as traffic signal control, GNNs focus on learning control policies or unraveling agent interactions that involve learning or predicting over edges or the entire graph. Besides, GNNs face different challenges in various transportation domains. The pure GNN models can not effectively solve the problem, so some scholars have explored the potential of combining GNNs with other approaches. For instance, in decision-making problems, such as traffic signal control, reinforcement learning is an effective technique. When multiple intersections interact, multi-agent reinforcement learning methods combined with GNNs have been proposed [139], [139], [147]. In some particular scenarios, such as traffic accident prediction, positive samples like accidents can be rare when predicting within a fine-grained granularity. To improve accuracy, we can use data augmentation techniques like a priori knowledge-based data enhancement [159], [160] and negative sample undersampling methods [156]. Nearly every transportation domain has its own domain-specific problems and unique characteristics. Therefore, combining GNNs and other techniques requires nuanced graph construction, tailored problem analysis, and painstaking design.

Dynamic Spatio-temporal Dependency. Utilizing GNNs to model spatio-temporal dependencies in ITS presents a significant challenge [26]. This difficulty arises from the need to precisely capture both the complex and dynamic spatial interactions within the transportation network and the temporal dynamics that reflect the constantly changing traffic patterns. Transportation networks frequently exhibit dynamic spatial dependencies, as the graph structure can evolve over time due to the constantly changing urban environment. For example, in trajectory prediction, it is crucial to identify key agents and objects, such as vehicles, cyclists, and pedestrians, that can influence the predicted trajectory. Consequently, a graph framework capable of adapting to these real-time changes is necessary to maintain accurate predictions. Meanwhile, in terms of temporal dependencies, traffic conditions at any given moment are shaped by a multitude of past events. Accurately capturing these long-term dependencies is crucial for precise

forecasting, but it poses significant computational challenges and demands sophisticated memory mechanisms within the model [68], [192]. Furthermore, real-time data processing is vital for practical applications in ITS, adding another layer of complexity [193], [194]. The model must integrate and process this multifaceted data while adapting in an environment marked by continuous change and uncertainty. Therefore, effectively modeling spatio-temporal dependencies remains a pivotal yet challenging aspect of utilizing GNNs in ITS.

Robustness, Reliability, Interpretability. Deep learning has faced criticism for its lack of interpretability and black-box nature, which makes it difficult to assess the rationale behind recommendations made by graph-based deep learning approaches in transportation safety and other high-stakes fields. Furthermore, it is essential to ensure that GNN models remain reliable in large-scale real-world applications, such as during rush hours, sensor failures, or cyber-attacks [195]. As we strive to enhance model performance, we must also be vigilant about potential failures and undetected anomalies. Additionally, scalability remains a critical concern. Current GNN frameworks, including those based on TensorFlow, PyTorch, DGL, and PyG, face limitations in scalability, which restrict their application to large-scale graphs due to inadequate system support [27].

3) *Computation:* Processing, storing, and transmitting vast amounts of data has become increasingly critical in today's extensive data landscape, particularly within ITS. While GNNs and deep learning techniques are extensively utilized in ITS, they face substantial challenges due to their high computational demands. These challenges are amplified when performing real-time or near-real-time inference and processing large volumes of data from extensive camera networks. Additionally, the limited resources of IoT devices—such as constrained memory and computing power—complicate these issues further. To address these challenges, researchers have proposed various solutions, including edge computing, graph sampling, hardware acceleration, and optimized algorithms.

B. Future Directions

1) *More Integration of Advanced Techniques:* As noted, GNNs are highly effective for capturing spatial-temporal relationships and making inferences on graph-based data structures. However, since different problems have unique characteristics and challenges, it is crucial to design tailored models for each specific issue. Additionally, integrating other techniques into GNN frameworks can significantly enhance model performance and facilitate practical applications. For example, incorporating the edge learning paradigm [196] into GNN frameworks can address the storage, memory, and computational limitations of data-producing devices. This approach allows distributed edge devices to collaboratively train models and perform inferences, thus ensuring privacy and security [196]. Similarly, leveraging transfer learning [197] and meta-learning [135] can significantly improve a model's adaptability to different cities with varying traffic patterns. In conclusion, combining GNNs with advanced techniques such as reinforcement learning, transfer learning, meta-learning, generative adversarial networks (GANs), semi-supervised learning, and Bayesian networks opens new avenues

for addressing domain-specific problems and challenges. This synergistic approach not only leads to more robust and versatile solutions but also unveils exciting possibilities for solving complex, real-world problems across various domains. As research continues, it is crucial to explore these integrations further, pushing the boundaries of what can be achieved with GNNs and their allied technologies.

2) *More Expanding Applications of GNNs:* Currently, most research on GNNs in ITS focuses on traffic prediction. However, significant untapped potential remains for further developing GNNs and exploring their broader applications within this field. To enhance GNNs, efforts should be directed towards improving their efficiency, robustness, and generality. One approach is to implement multi-modal learning [95], [198], which allows models to integrate a richer set of contextual information. Additionally, incorporating more complex graph structures, such as heterogeneous graphs [180] and hypergraphs [116], can facilitate handling larger and more intricate graph structures. Furthermore, exploring GNN applications in other ITS research domains also holds promise. For example, the 3D structural understanding required for autonomous vehicles presents a unique opportunity for GNNs, especially compared to traditional transformer architectures [99], [100] used in point cloud processing, which often lack efficiency. By combining graph convolution with self-attention mechanisms, we can significantly enhance feature extraction and capture both local and global contexts effectively. This integration has the potential to greatly improve GNN performance in ITS applications [98]. While we have covered various domains, from traffic prediction to traffic safety, additional areas remain to explore, such as route planning, urban land-use planning, and traffic pattern recognition. Further investigation into these applications of GNNs could yield new insights and enhance their performance across a broader range of ITS scenarios.

3) *More Comprehensive Experiments:* Currently, some research experiments in the field of ITS rely on simulators. However, data generated by traffic simulation software may not always accurately reflect real-world conditions due to factors such as variations in driver behavior and alternative route planning [199]. Furthermore, even when models are evaluated using real-world data, such testing is often limited in scale, or the model's running time may not be fully reported. These limitations do not necessarily ensure the model's reliability, robustness, or generalization ability to real-world scenarios. According to work by Shi et al. [200], some models, such as those based on Deep Q-Networks (DQNs) for reinforcement learning, experience performance degradation when dealing with large-scale road networks or incomplete data, complicating their generalization. Therefore, it is essential to conduct more comprehensive experiments using large-scale real-world data to fully assess and validate the effectiveness of these models.

VI. CONCLUSION

With the rapid advancement of deep learning, GNNs have emerged as a promising tool in ITS. However, most existing research on GNNs in ITS has concentrated on traffic forecasting while overlooking other critical areas, such as autonomous

vehicles and transportation safety. In this work, we have reviewed and analyzed representative papers from 2018 to 2023 that explore various applications of GNNs across six domains within ITS. We have summarized and categorized these studies based on their research focus, the graph methods employed, and the domain-specific challenges encountered, and presented the findings in detailed tables and lists. Our analysis reveals that while many studies are limited to specific functionalities of GNNs, such as modeling graph-structured data and capturing spatio-temporal relationships, a vast potential is waiting to be fully realized. GNNs can be expanded to other areas of ITS, opening up new possibilities and avenues for research and development. Additionally, we have identified common challenges that need to be addressed when applying GNNs in ITS, including issues related to *data*, *model*, and *computation*. We also highlight future directions for GNN research in ITS, emphasizing the crucial role of integrating GNNs with other techniques. This approach can lead to more comprehensive solutions, broadening the application scope of GNNs and fostering more extensive experiments.

ACKNOWLEDGMENT

The authors are grateful to the anonymous reviewers for critically reading this article and for giving important suggestions to improve this article.

This paper is partially supported by the National Key Research and Development Program of China with Grant No. 2023YFC3341203, and the National Natural Science Foundation of China (NSFC Grant Numbers 62306014 and 62276002).

REFERENCES

- [1] J. Wootton, A. Garcia-Ortiz, and S. Amin, "Intelligent transportation systems: a global perspective," *Mathematical and computer modelling*, vol. 22, no. 4-7, 1995.
- [2] M. Véres and M. Moussa, "Deep learning for intelligent transportation systems: A survey of emerging trends," *TITS*, vol. 21, no. 8, 2019.
- [3] S. Lee and D. B. Fambro, "Application of subset autoregressive integrated moving average model for short-term freeway traffic volume forecasting," *TRR*, vol. 1678, no. 1, 1999.
- [4] B. M. Williams, "Multivariate vehicular traffic flow prediction: evaluation of arimax modeling," *TRR*, vol. 1776, no. 1, 2001.
- [5] P. Jian John Lu PHD and L. Rajaram, "Evaluation of intelligent transportation system operations using logistic regression models," *Institute of Transportation Engineers. ITE Journal*, vol. 83, no. 3, p. 40, 2013.
- [6] C.-H. Wu, J.-M. Ho, and D.-T. Lee, "Travel-time prediction with support vector regression," *T-ITS*, vol. 5, no. 4, pp. 276–281, 2004.
- [7] H. Chang, Y. Lee, B. Yoon, and S. Baek, "Dynamic near-term traffic flow prediction: system-oriented approach based on past experiences," *IET intelligent transport systems*, vol. 6, no. 3, pp. 292–305, 2012.
- [8] X. Fan, C. Xiang, L. Gong, X. He, Y. Qu, S. Amirgholipour, Y. Xi, P. Nanda, and X. He, "Deep learning for intelligent traffic sensing and prediction: recent advances and future challenges," *CCF Transactions on Pervasive Computing and Interaction*, vol. 2, 2020.
- [9] W. Jiang and J. Luo, "Graph neural network for traffic forecasting: A survey," *ESWA*, vol. 207, 2022.
- [10] S. Rahmani, A. Baghbani, N. Bouguila, and Z. Patterson, "Graph neural networks for intelligent transportation systems: A survey," *TITS*, 2023.
- [11] Z. Cui, K. Henrickson, R. Ke, and Y. Wang, "Traffic graph convolutional recurrent neural network: A deep learning framework for network-scale traffic learning and forecasting," *TITS*, vol. 21, no. 11, 2019.
- [12] D. A. Tedjopurnomo, Z. Bao, B. Zheng, F. M. Choudhury, and A. K. Qin, "A survey on modern deep neural network for traffic prediction: Trends, methods and challenges," *TKDE*, vol. 34, no. 4, 2020.
- [13] X. Yin, G. Wu, J. Wei, Y. Shen, H. Qi, and B. Yin, "Deep learning on traffic prediction: Methods, analysis, and future directions," *TITS*, vol. 23, no. 6, 2021.
- [14] W. Jiang, J. Luo, M. He, and W. Gu, "Graph neural network for traffic forecasting: The research progress," *ISPRS International Journal of Geo-Information*, vol. 12, no. 3, 2023.
- [15] E. I. Vlahogianni, M. G. Karlaftis, and J. C. Golias, "Short-term traffic forecasting: Where we are and where we're going," *Transportation Research Part C: Emerging Technologies*, vol. 43, 2014.
- [16] K. Lee, M. Eo, E. Jung, Y. Yoon, and W. Rhee, "Short-term traffic prediction with deep neural networks: A survey," *IEEE Access*, vol. 9, 2021.
- [17] A. K. Haghhighat, V. Ravichandra-Mouli, P. Chakraborty, Y. Esfandiari, S. Arabi, and A. Sharma, "Applications of deep learning in intelligent transportation systems," *Journal of Big Data Analytics in Transportation*, vol. 2, 2020.
- [18] J. Guerrero-Ibañez, J. Contreras-Castillo, and S. Zeadaaly, "Deep learning support for intelligent transportation systems," *Transactions on Emerging Telecommunications Technologies*, vol. 32, no. 3, 2021.
- [19] J. Liu, N. Wu, Y. Qiao, and Z. Li, "A scientometric review of research on traffic forecasting in transportation," *IET Intelligent Transport Systems*, vol. 15, no. 1, 2021.
- [20] M. S. Ahmed and A. R. Cook, *Analysis of freeway traffic time-series data by using Box-Jenkins techniques*, 1979, no. 722.
- [21] M. Levin and Y.-D. Tsao, "On forecasting freeway occupancies and volumes (abridgment)," *TRR*, no. 773, 1980.
- [22] B. M. Williams and L. A. Hoel, "Modeling and forecasting vehicular traffic flow as a seasonal arima process: Theoretical basis and empirical results," *Journal of Transportation Engineering*, vol. 129, no. 6, 2003.
- [23] Y. Kamarianakis and P. Prastacos, "Forecasting traffic flow conditions in an urban network: Comparison of multivariate and univariate approaches," *TRR*, vol. 1857, no. 1, 2003.
- [24] L. Vanajakshi and L. R. Rilett, "A comparison of the performance of artificial neural networks and support vector machines for the prediction of traffic speed," in *IV*, 2004, pp. 194–199.
- [25] E. I. Vlahogianni, J. C. Golias, and M. G. Karlaftis, "Short-term traffic forecasting: Overview of objectives and methods," *Transport reviews*, vol. 24, no. 5, 2004.
- [26] K.-H. N. Bui, J. Cho, and H. Yi, "Spatial-temporal graph neural network for traffic forecasting: An overview and open research issues," *Applied Intelligence*, vol. 52, no. 3, 2022.
- [27] Y. Li, D. Yu, Z. Liu, M. Zhang, X. Gong, and L. Zhao, "Graph neural network for spatiotemporal data: methods and applications," *arXiv preprint arXiv:2306.00012*, 2023.
- [28] A. Auer, S. Feese, S. Lockwood, and B. A. Hamilton, "History of intelligent transportation systems." Tech. Rep., 2016.
- [29] R. Mundy, *Management of Public Transportation Systems in the 1980s:(the Emergence of Parapublic Transportation)*. Department of Marketing and Transportation, College of Business, 1981.
- [30] R. C. Bushnell, J. T. Low, and J. B. Wiley, "Transportation network models: Past problems and prospects for the 1980s," *International Journal of Physical Distribution & Materials Management*, vol. 11, no. 8, 1981.
- [31] J. Zhang, F.-Y. Wang, K. Wang, W.-H. Lin, X. Xu, and C. Chen, "Data-driven intelligent transportation systems: A survey," *TITS*, vol. 12, no. 4, 2011.
- [32] R. Jia, P. Jiang, L. Liu, L. Cui, and Y. Shi, "Data driven congestion trends prediction of urban transportation," *IEEE Internet of Things Journal*, vol. 5, no. 2, 2017.
- [33] R. Günther, T. Wenzel, M. Wegner, and R. Rettig, "Big data driven dynamic driving cycle development for busses in urban public transportation," *Transportation Research Part D: Transport and Environment*, vol. 51, 2017.
- [34] A. Saroj, S. Roy, A. Guin, M. Hunter, and R. Fujimoto, "Smart city real-time data-driven transportation simulation," in *Winter Simulation Conference*, 2018.
- [35] Y. Wang and Z. Zeng, *Data-driven solutions to transportation problems*, 2018.
- [36] J. N. Njoku, C. I. Nwakanma, G. C. Amaizu, and D.-S. Kim, "Prospects and challenges of metaverse application in data-driven intelligent transportation systems," *IET Intelligent Transport Systems*, vol. 17, no. 1, 2023.
- [37] Y. Lin, P. Wang, and M. Ma, "Intelligent transportation system (its): Concept, challenge and opportunity," in *BigDataSecurity*, 2017.
- [38] J. Gilmer, S. S. Schoenholz, P. F. Riley, O. Vinyals, and G. E. Dahl, "Neural message passing for quantum chemistry," in *International conference on machine learning*. PMLR, 2017, pp. 1263–1272.

- [39] J. Lee, I. Lee, and J. Kang, "Self-attention graph pooling," in *International conference on machine learning*, 2019, pp. 3734–3743.
- [40] Z. Ying, J. You, C. Morris, X. Ren, W. Hamilton, and J. Leskovec, "Hierarchical graph representation learning with differentiable pooling," *Advances in neural information processing systems*, vol. 31, 2018.
- [41] M. Zhang, Z. Cui, M. Neumann, and Y. Chen, "An end-to-end deep learning architecture for graph classification," in *Proceedings of the AAAI conference on artificial intelligence*, vol. 32, no. 1, 2018.
- [42] Z. Wu, S. Pan, F. Chen, G. Long, C. Zhang, and S. Y. Philip, "A comprehensive survey on graph neural networks," *TNNLS*, vol. 32, no. 1, 2020.
- [43] J. Zhou, G. Cui, S. Hu, Z. Zhang, C. Yang, Z. Liu, L. Wang, C. Li, and M. Sun, "Graph neural networks: A review of methods and applications," *AI Open*, vol. 1, 2020.
- [44] Z. Zhang, P. Cui, and W. Zhu, "Deep learning on graphs: A survey," *TKDE*, vol. 34, no. 1, 2020.
- [45] S. Abadal, A. Jain, R. Guirado, J. López-Alonso, and E. Alarcón, "Computing graph neural networks: A survey from algorithms to accelerators," *CSUR*, vol. 54, no. 9, 2021.
- [46] T. N. Kipf and M. Welling, "Variational graph auto-encoders," *arXiv preprint arXiv:1611.07308*, 2016.
- [47] E. L. Manibardo, I. Laña, and J. Del Ser, "Deep learning for road traffic forecasting: Does it make a difference?" *IEEE Transactions on Intelligent Transportation Systems*, vol. 23, no. 7, pp. 6164–6188, 2021.
- [48] J. Zhang, J. Jin, J. Tang, and Z. Qu, "Fptn: Fast pure transformer network for traffic flow forecasting," in *ICANN*, 2023.
- [49] Y. Zhao, X. Luo, W. Ju, C. Chen, X.-S. Hua, and M. Zhang, "Dynamic hypergraph structure learning for traffic flow forecasting," in *ICDE*, 2023.
- [50] S. Lan, Y. Ma, W. Huang, W. Wang, H. Yang, and P. Li, "Dstagnn: Dynamic spatial-temporal aware graph neural network for traffic flow forecasting," in *ICML*, 2022.
- [51] C. Chen, Y. Liu, L. Chen, and C. Zhang, "Bidirectional spatial-temporal adaptive transformer for urban traffic flow forecasting," *TNNLS*, 2022.
- [52] M. Li and Z. Zhu, "Spatial-temporal fusion graph neural networks for traffic flow forecasting," in *AAAI*, 2021.
- [53] L. Bai, L. Yao, C. Li, X. Wang, and C. Wang, "Adaptive graph convolutional recurrent network for traffic forecasting," in *NeurIPS*, 2020.
- [54] J. Yi and J. Park, "Hypergraph convolutional recurrent neural network," in *KDD*, 2020.
- [55] C. Song, Y. Lin, S. Guo, and H. Wan, "Spatial-temporal synchronous graph convolutional networks: A new framework for spatial-temporal network data forecasting," in *AAAI*, 2020.
- [56] S. Guo, Y. Lin, N. Feng, C. Song, and H. Wan, "Attention based spatial-temporal graph convolutional networks for traffic flow forecasting," in *AAAI*, 2019.
- [57] J. Li, Z. Han, H. Cheng, J. Su, P. Wang, J. Zhang, and L. Pan, "Predicting path failure in time-evolving graphs," in *KDD*, 2019.
- [58] Y. Li, R. Yu, C. Shahabi, and Y. Liu, "Diffusion convolutional recurrent neural network: Data-driven traffic forecasting," in *ICLR*, 2018.
- [59] G. Khodabandeh, W. Kheriji, and F. H. Selem, "Link traffic speed forecasting using convolutional attention-based gated recurrent unit," *Applied Intelligence*, vol. 51, 2021.
- [60] L. Han, B. Du, L. Sun, Y. Fu, Y. Lv, and H. Xiong, "Dynamic and multi-faceted spatio-temporal deep learning for traffic speed forecasting," in *KDD*, 2021.
- [61] C. Zhang, S. Zhang, J. James, and S. Yu, "Fastggnn: A topological information protected federated learning approach for traffic speed forecasting," *IEEE Transactions on Industrial Informatics*, vol. 17, no. 12, 2021.
- [62] J. Liu, G. P. Ong, and X. Chen, "Graphsage-based traffic speed forecasting for segment network with sparse data," *TITS*, vol. 23, no. 3, 2020.
- [63] P. Wu, Z. Huang, Y. Pian, L. Xu, J. Li, and K. Chen, "A combined deep learning method with attention-based lstm model for short-term traffic speed forecasting," *JAT*, vol. 2020, 2020.
- [64] G. Guo and W. Yuan, "Short-term traffic speed forecasting based on graph attention temporal convolutional networks," *Neurocomputing*, vol. 410, 2020.
- [65] K. Zhang, L. Wu, Z. Zhu, and J. Deng, "A multitask learning model for traffic flow and speed forecasting," *IEEE Access*, vol. 8, 2020.
- [66] D. Wang, J. Zhu, Y. Yin, J. Ignatius, X. Wei, and A. Kumar, "Dynamic travel time prediction with spatiotemporal features: using a gnn-based deep learning method," *Annals of Operations Research*, 2023.
- [67] L. Tran, M. Y. Mun, M. Lim, J. Yamato, N. Huh, and C. Shahabi, "Deeptrans: a deep learning system for public bus travel time estimation using traffic forecasting," *VLDB*, vol. 13, no. 12, 2020.
- [68] A. Roy, K. K. Roy, A. Ahsan Ali, M. A. Amin, and A. M. Rahman, "Sst-gnn: simplified spatio-temporal traffic forecasting model using graph neural network," in *PAKDD*, 2021.
- [69] J. Ma, J. Chan, G. Ristanoski, S. Rajasegarar, and C. Leckie, "Bus travel time prediction with real-time traffic information," *Transportation Research Part C: Emerging Technologies*, vol. 105, 2019.
- [70] S. Hochreiter and J. Schmidhuber, "Long short-term memory," *Neural Computation*, vol. 9, no. 8, 1997.
- [71] K. Cho, B. van Merriënboer, D. Bahdanau, and Y. Bengio, "On the properties of neural machine translation: Encoder-decoder approaches," in *SSST@EMNLP*, 2014.
- [72] W. Ju, Y. Zhao, Y. Qin, S. Yi, J. Yuan, Z. Xiao, X. Luo, X. Yan, and M. Zhang, "Cool: A conjoint perspective on spatio-temporal graph neural network for traffic forecasting," *Information Fusion*, p. 102341, 2024.
- [73] M. Xu, W. Dai, C. Liu, X. Gao, W. Lin, G.-J. Qi, and H. Xiong, "Spatial-temporal transformer networks for traffic flow forecasting," *arXiv preprint arXiv:2001.02908*, 2020.
- [74] G. Huo, Y. Zhang, B. Wang, J. Gao, Y. Hu, and B. Yin, "Hierarchical spatio-temporal graph convolutional networks and transformer network for traffic flow forecasting," *TITS*, vol. 24, no. 4, 2023.
- [75] A. Vaswani, N. Shazeer, N. Parmar, J. Uszkoreit, L. Jones, A. N. Gomez, Ł. Kaiser, and I. Polosukhin, "Attention is all you need," in *NeurIPS*, 2017.
- [76] D. He, Y. Zhao, J. Luo, T. Hui, S. Huang, A. Zhang, and S. Liu, "Transrefer3d: Entity-and-relation aware transformer for fine-grained 3d visual grounding," in *Proceedings of the 29th ACM International Conference on Multimedia*, 2021.
- [77] Y. Zhao, J. Chen, C. Gao, W. Wang, L. Yang, H. Ren, H. Xia, and S. Liu, "Target-driven structured transformer planner for vision-language navigation," in *Proceedings of the 30th ACM International Conference on Multimedia*, 2022.
- [78] Y. Zhang and X. Gu, "Traffic flow prediction based on transformer and multi-spatial-temporal encoder-decoder," in *2023 4th International Conference on Computer Engineering and Application (ICCEA)*, 2023.
- [79] X. Mi, C. Yu, X. Liu, G. Yan, F. Yu, and P. Shang, "A dynamic ensemble deep deterministic policy gradient recursive network for spatiotemporal traffic speed forecasting in an urban road network," *Digital Signal Processing*, vol. 129, 2022.
- [80] L. Kang, G. Hu, H. Huang, W. Lu, and L. Liu, "Urban traffic travel time short-term prediction model based on spatio-temporal feature extraction," *JAT*, vol. 2020, 2020.
- [81] D. P. M. Abellana, "Multivariate travel time forecasting in a traffic network using fuzzy cognitive mapping," in *AIC*, 2023.
- [82] L. Huang, Y. Yang, X. Zhao, C. Ma, and H. Gao, "Sparse data-based urban road travel speed prediction using probabilistic principal component analysis," *IEEE Access*, vol. 6, 2018.
- [83] W. Hamilton, Z. Ying, and J. Leskovec, "Inductive representation learning on large graphs," in *NeurIPS*, 2017.
- [84] C. Zheng, X. Fan, C. Wang, and J. Qi, "Gman: A graph multi-attention network for traffic prediction," in *AAAI*, 2020.
- [85] C.-H. Wei and Y. Lee, "Development of freeway travel time forecasting models by integrating different sources of traffic data," *IEEE Transactions on Vehicular Technology*, vol. 56, no. 6, 2007.
- [86] J. Barceló, L. Montero, L. Marqués, and C. Carmona, "Travel time forecasting and dynamic origin-destination estimation for freeways based on bluetooth traffic monitoring," *TRR*, vol. 2175, no. 1, 2010.
- [87] A. Comi and A. Polimeni, "Bus travel time: Experimental evidence and forecasting," *Forecasting*, vol. 2, no. 3, 2020.
- [88] M. Li, Y. Zhu, T. Zhao, and M. Angelova, "Weighted dynamic time warping for traffic flow clustering," *Neurocomputing*, vol. 472, pp. 266–279, 2022.
- [89] Y. Jin, K. Chen, and Q. Yang, "Transferable graph structure learning for graph-based traffic forecasting across cities," in *Proceedings of the 29th ACM SIGKDD Conference on Knowledge Discovery and Data Mining*, 2023, pp. 1032–1043.
- [90] Q. Zhang, J. Chang, G. Meng, S. Xiang, and C. Pan, "Spatio-temporal graph structure learning for traffic forecasting," in *Proceedings of the AAAI conference on artificial intelligence*, vol. 34, no. 01, 2020, pp. 1177–1185.
- [91] T. N. Kipf and M. Welling, "Semi-supervised classification with graph convolutional networks," in *ICLR*, 2017.
- [92] K. Xu, W. Hu, J. Leskovec, and S. Jegelka, "How powerful are graph neural networks?" in *ICLR*, 2019.

- [93] I. Sutskever, O. Vinyals, and Q. V. Le, "Sequence to sequence learning with neural networks," in *NeurIPS*, 2014.
- [94] W. Zhou, Q. Wang, W. Jin, X. Shi, D. Wang, X. Hao, and Y. Yu, "Gtnet: Graph transformer network for 3d point cloud classification and semantic segmentation," *arXiv preprint arXiv:2305.15213*, 2023.
- [95] Z. Liu, Y. Luo, X. Pu, G. Min, and C. Luo, "A multi-modal hypergraph neural network via parametric filtering and feature sampling," *TBD*, 2023.
- [96] M. Lin and A. Feragen, "diffconv: Analyzing irregular point clouds with an irregular view," in *European Conference on Computer Vision*. Springer, 2022.
- [97] R. Wiersma, A. Nasikun, E. Eisemann, and K. Hildebrandt, "Deltaconv: anisotropic operators for geometric deep learning on point clouds," *TOG*, vol. 41, no. 4, 2022.
- [98] D. Lu, Q. Xie, K. Gao, L. Xu, and J. Li, "3dctn: 3d convolution-transformer network for point cloud classification," *TITS*, vol. 23, no. 12, 2022.
- [99] H. Zhao, L. Jiang, J. Jia, P. H. Torr, and V. Koltun, "Point transformer," in *ICCV*, 2021.
- [100] M.-H. Guo, J.-X. Cai, Z.-N. Liu, T.-J. Mu, R. R. Martin, and S.-M. Hu, "Pct: Point cloud transformer," *Computational Visual Media*, vol. 7, 2021.
- [101] T. Xiang, C. Zhang, Y. Song, J. Yu, and W. Cai, "Walk in the cloud: Learning curves for point clouds shape analysis," in *ICCV*, 2021.
- [102] K. Zhang, M. Hao, J. Wang, X. Chen, Y. Leng, C. W. de Silva, and C. Fu, "Linked dynamic graph cnn: Learning through point cloud by linking hierarchical features," in *M2VIP*, 2021.
- [103] Z.-H. Lin, S.-Y. Huang, and Y.-C. F. Wang, "Convolution in the cloud: Learning deformable kernels in 3d graph convolution networks for point cloud analysis," in *CVPR*, 2020.
- [104] Y. Feng, H. You, Z. Zhang, R. Ji, and Y. Gao, "Hypergraph neural networks," in *AAAI*, 2019.
- [105] Y. Wang, Y. Sun, Z. Liu, S. E. Sarma, M. M. Bronstein, and J. M. Solomon, "Dynamic graph cnn for learning on point clouds," *TOG*, vol. 38, no. 5, 2019.
- [106] G. Te, W. Hu, A. Zheng, and Z. Guo, "Rgcn: Regularized graph cnn for point cloud segmentation," in *ACMMM*, 2018.
- [107] R. Li, S. Wang, F. Zhu, and J. Huang, "Adaptive graph convolutional neural networks," in *AAAI*, 2018.
- [108] Y. Shen, C. Feng, Y. Yang, and D. Tian, "Mining point cloud local structures by kernel correlation and graph pooling," in *CVPR*, 2018.
- [109] C. Wang, B. Samari, and K. Siddiqi, "Local spectral graph convolution for point set feature learning," in *ECCV*, 2018.
- [110] M. Simonovsky and N. Komodakis, "Dynamic edge-conditioned filters in convolutional neural networks on graphs," in *CVPR*, 2017.
- [111] H.-H. Jebamikyous and R. Kashef, "Autonomous vehicles perception (avp) using deep learning: Modeling, assessment, and challenges," *IEEE Access*, vol. 10, 2022.
- [112] Y. Guo, H. Wang, Q. Hu, H. Liu, L. Liu, and M. Bennamoun, "Deep learning for 3d point clouds: A survey," *TPAMI*, vol. 43, no. 12, 2020.
- [113] M. Balcilar, G. Renton, P. Héroux, B. Gauzère, S. Adam, and P. Honeine, "Bridging the gap between spectral and spatial domains in graph neural networks," *arXiv preprint arXiv:2003.11702*, 2020.
- [114] Y. Wang and J. M. Solomon, "Object dgcn: 3d object detection using dynamic graphs," in *NeurIPS*, 2021.
- [115] F. R. Chung, *Spectral graph theory*. American Mathematical Soc., 1997, vol. 92.
- [116] N. Yadati, M. Nimishakavi, P. Yadav, V. Nitin, A. Louis, and P. Talukdar, "Hypergen: A new method for training graph convolutional networks on hypergraphs," in *NeurIPS*, 2019.
- [117] T.-H. H. Chan and Z. Liang, "Generalizing the hypergraph laplacian via a diffusion process with mediators," *TCS*, vol. 806, 2020.
- [118] S. Bai, F. Zhang, and P. H. Torr, "Hypergraph convolution and hypergraph attention," *Pattern Recognition*, vol. 110, 2021.
- [119] R. Huang, H. Xue, M. Pagnucco, F. Salim, and Y. Song, "Multimodal trajectory prediction: A survey," *arXiv preprint arXiv:2302.10463*, 2023.
- [120] X. Li, X. Ying, and M. C. Chuah, "Grip: Graph-based interaction-aware trajectory prediction," in *ITSC*, 2019.
- [121] H. Jeon, J. Choi, and D. Kum, "Scale-net: Scalable vehicle trajectory prediction network under random number of interacting vehicles via edge-enhanced graph convolutional neural network," in *IROS*, 2020.
- [122] A. Mohamed, K. Qian, M. Elhoseiny, and C. Claudel, "Social-stgcn: A social spatio-temporal graph convolutional neural network for human trajectory prediction," in *CVPR*, 2020.
- [123] R. Chandra, T. Guan, S. Panuganti, T. Mittal, U. Bhattacharya, A. Bera, and D. Manocha, "Forecasting trajectory and behavior of road-agents using spectral clustering in graph-lstms," *IEEE Robotics and Automation Letters*, vol. 5, no. 3, 2020.
- [124] Z. Sheng, Y. Xu, S. Xue, and D. Li, "Graph-based spatial-temporal convolutional network for vehicle trajectory prediction in autonomous driving," *TITS*, vol. 23, no. 10, 2022.
- [125] Y. Huang, H. Bi, Z. Li, T. Mao, and Z. Wang, "Stgat: Modeling spatial-temporal interactions for human trajectory prediction," in *ICCV*, 2019.
- [126] S. Carrasco, D. F. Llorca, and M. Sotelo, "Scout: Socially-consistent and understandable graph attention network for trajectory prediction of vehicles and vrus," in *IEEE Intelligent Vehicles Symposium*, 2021.
- [127] H. Zhou, D. Ren, H. Xia, M. Fan, X. Yang, and H. Huang, "Ast-gnn: An attention-based spatio-temporal graph neural network for interaction-aware pedestrian trajectory prediction," *Neurocomputing*, vol. 445, 2021.
- [128] J. Li, H. Ma, Z. Zhang, J. Li, and M. Tomizuka, "Spatio-temporal graph dual-attention network for multi-agent prediction and tracking," *TITS*, vol. 23, no. 8, 2022.
- [129] W. Zhu, Y. Liu, P. Wang, M. Zhang, T. Wang, and Y. Yi, "Tri-hggn: Learning triple policies fused hierarchical graph neural networks for pedestrian trajectory prediction," *PR*, 2023.
- [130] X. Jia, P. Wu, L. Chen, Y. Liu, H. Li, and J. Yan, "Hdgt: Heterogeneous driving graph transformer for multi-agent trajectory prediction via scene encoding," *TPAMI*, 2023.
- [131] F. Zhou, Q. Yang, T. Zhong, D. Chen, and N. Zhang, "Variational graph neural networks for road traffic prediction in intelligent transportation systems," *IEEE Transactions on Industrial Informatics*, vol. 17, no. 4, pp. 2802–2812, 2020.
- [132] J. Ma and F. Wu, "Feudal multi-agent reinforcement learning with adaptive network partition for traffic signal control," *arXiv preprint arXiv:2205.13836*, 2022.
- [133] Y. S. K. LIN and A. K. Bashir, "Keylight: Intelligent traffic signal control method based on improved graph neural network," *IEEE Transactions on Consumer Electronics*, 2023.
- [134] S. Yang, "Hierarchical graph multi-agent reinforcement learning for traffic signal control," *Information Sciences*, vol. 634, 2023.
- [135] M. Wang, L. Wu, M. Li, D. Wu, X. Shi, and C. Ma, "Meta-learning based spatial-temporal graph attention network for traffic signal control," *Knowledge-based Systems*, vol. 250, 2022.
- [136] C. Zhao and G. Wang, "Dynamic traffic light control with reinforcement learning based on gnn prediction," *Available at SSRN 4040526*, 2022.
- [137] L. Wu, M. Wang, D. Wu, and J. Wu, "Dynstgat: Dynamic spatial-temporal graph attention network for traffic signal control," in *CIKM*, 2021.
- [138] S. Yang, B. Yang, Z. Kang, and L. Deng, "Ihg-ma: Inductive heterogeneous graph multi-agent reinforcement learning for multi-intersection traffic signal control," *Neural Networks*, vol. 139, 2021.
- [139] Z. Zeng, "Graphlight: graph-based reinforcement learning for traffic signal control," in *ICCCS*, 2021.
- [140] T. Zhong, Z. Xu, and F. Zhou, "Probabilistic graph neural networks for traffic signal control," in *ICASSP*, 2021.
- [141] Y. Wang, T. Xu, X. Niu, C. Tan, E. Chen, and H. Xiong, "Stmarl: A spatio-temporal multi-agent reinforcement learning approach for cooperative traffic light control," *TMC*, vol. 21, no. 6, 2020.
- [142] H. Wei, N. Xu, H. Zhang, G. Zheng, X. Zang, C. Chen, W. Zhang, Y. Zhu, K. Xu, and Z. Li, "Colight: Learning network-level cooperation for traffic signal control," in *CIKM*, 2019.
- [143] F. I. Shashi, S. M. Sultan, A. Khatun, T. Sultana, and T. Alam, "A study on deep reinforcement learning based traffic signal control for mitigating traffic congestion," in *ECBIOS*, 2021.
- [144] Y. Huo, Q. Tao, and J. Hu, "Cooperative control for multi-intersection traffic signal based on deep reinforcement learning and imitation learning," *IEEE Access*, vol. 8, 2020.
- [145] T. Chu, J. Wang, L. Codicà, and Z. Li, "Multi-agent deep reinforcement learning for large-scale traffic signal control," *TITS*, vol. 21, no. 3, 2019.
- [146] J. K. Gupta, M. Egorov, and M. Kochenderfer, "Cooperative multi-agent control using deep reinforcement learning," in *Autonomous Agents and Multiagent Systems: AAMAS Workshops*, 2017.
- [147] T. Nishi, K. Otaki, K. Hayakawa, and T. Yoshimura, "Traffic signal control based on reinforcement learning with graph convolutional neural nets," in *ITSC*, 2018.
- [148] J. Yoon, K. Ahn, J. Park, and H. Yeo, "Transferable traffic signal control: Reinforcement learning with graph centric state representation," *Transportation Research Part C: Emerging Technologies*, vol. 130, 2021.
- [149] T. Saiki and S. Arai, "Flexible traffic signal control via multi-objective reinforcement learning," *IEEE Access*, 2023.
- [150] J. Ma and F. Wu, "Learning to coordinate traffic signals with adaptive network partition," *TITS*, 2023.

- [151] Z. Yu, S. Liang, L. Wei, Z. Jin, J. Huang, D. Cai, X. He, and X.-S. Hua, “Macar: Urban traffic light control via active multi-agent communication and action rectification,” in *IJCAI*, 2021.
- [152] X. Li, Z. Guo, X. Dai, Y. Lin, J. Jin, F. Zhu, and F.-Y. Wang, “Deep imitation learning for traffic signal control and operations based on graph convolutional neural networks,” in *ITSC*, 2020.
- [153] T. Tran, D. He, J. Kim, and M. Hickman, “Msgnn: A multi-structured graph neural network model for real-time incident prediction in large traffic networks,” *Transportation Research Part C: Emerging Technologies*, vol. 156, 2023.
- [154] Z. Liu, Y. Chen, F. Xia, J. Bian, B. Zhu, G. Shen, and X. Kong, “Tap: Traffic accident profiling via multi-task spatio-temporal graph representation learning,” *TKDD*, vol. 17, no. 4, 2023.
- [155] Y. Huang, F. Zhang, and J. Hu, “Deep spatial-temporal graph modeling of urban traffic accident prediction,” in *The International Conference on Image, Vision and Intelligent Systems*, 2022.
- [156] L. Yu, B. Du, X. Hu, L. Sun, L. Han, and W. Lv, “Deep spatio-temporal graph convolutional network for traffic accident prediction,” *Neurocomputing*, vol. 423, 2021.
- [157] B. Wang, Y. Lin, S. Guo, and H. Wan, “Gsnnet: learning spatial-temporal correlations from geographical and semantic aspects for traffic accident risk forecasting,” in *AAAI*, 2021.
- [158] Y. Zhang, X. Dong, L. Shang, D. Zhang, and D. Wang, “A multi-modal graph neural network approach to traffic risk forecasting in smart urban sensing,” in *SECON*, 2020.
- [159] Z. Zhou, Y. Wang, X. Xie, L. Chen, and C. Zhu, “Foresee urban sparse traffic accidents: A spatiotemporal multi-granularity perspective,” *TKDE*, vol. 34, no. 8, 2020.
- [160] Z. Zhou, Y. Wang, X. Xie, L. Chen, and H. Liu, “Riskoracle: A minute-level citywide traffic accident forecasting framework,” in *AAAI*, 2020.
- [161] L. Zhu, T. Li, and S. Du, “Ta-stan: A deep spatial-temporal attention learning framework for regional traffic accident risk prediction,” in *IJCNN*, 2019.
- [162] Z. Yuan, X. Zhou, and T. Yang, “Hetero-convlstm: A deep learning approach to traffic accident prediction on heterogeneous spatio-temporal data,” in *KDD*, 2018.
- [163] C. Huang, C. Zhang, P. Dai, and L. Bo, “Deep dynamic fusion network for traffic accident forecasting,” in *CIKM*, 2019.
- [164] J. Bao, P. Liu, and S. V. Ukkusuri, “A spatiotemporal deep learning approach for citywide short-term crash risk prediction with multi-source data,” *Accident Analysis & Prevention*, vol. 122, 2019.
- [165] C. Chen, X. Fan, C. Zheng, L. Xiao, M. Cheng, and C. Wang, “Sdcae: Stack denoising convolutional autoencoder model for accident risk prediction via traffic big data,” in *CBD*, 2018.
- [166] H. Ren, Y. Song, J. Wang, Y. Hu, and J. Lei, “A deep learning approach to the citywide traffic accident risk prediction,” in *ITSC*, 2018.
- [167] Z. Wang, R. Jiang, Z. Cai, Z. Fan, X. Liu, K.-S. Kim, X. Song, and R. Shibasaki, “Spatio-temporal-categorical graph neural networks for fine-grained multi-incident co-prediction,” in *CIKM*, 2021.
- [168] B. Huang and B. Hooi, “Traffic accident prediction using graph neural networks: New datasets and the travel model,” *Traffic*, vol. 27, no. 29, 2022.
- [169] Y. Zhang and T. Cheng, “Graph deep learning model for network-based predictive hotspot mapping of sparse spatio-temporal events,” *Computers, Environment and Urban Systems*, vol. 79, 2020.
- [170] Y. Wang, X. Lin, J. Wu, A. K. Bashir, W. Yang, J. Li, and M. Imran, “Contrastive gnn-based traffic anomaly analysis against imbalanced dataset in iot-based its,” in *GLOBECOM*, 2022.
- [171] Y. Tong, Y. Chen, Z. Zhou, L. Chen, J. Wang, Q. Yang, J. Ye, and W. Lv, “The simpler the better: a unified approach to predicting original taxi demands based on large-scale online platforms,” in *KDD*, 2017.
- [172] W. Zi, W. Xiong, H. Chen, and L. Chen, “Tagcn: Station-level demand prediction for bike-sharing system via a temporal attention graph convolution network,” *Information Sciences*, vol. 561, 2021.
- [173] J. L. Toole, S. Colak, B. Sturt, L. P. Alexander, A. Evsukoff, and M. C. González, “The path most traveled: Travel demand estimation using big data resources,” *Transportation Research Part C: Emerging Technologies*, vol. 58, 2015.
- [174] X. Geng, Y. Li, L. Wang, L. Zhang, Q. Yang, J. Ye, and Y. Liu, “Spatiotemporal multi-graph convolution network for ride-hailing demand forecasting,” in *AAAI*, 2019.
- [175] E. Lee, H. Choi, D.-G. Kim *et al.*, “Pgdr: Prediction demand based on graph convolutional network for regional demand-responsive transport,” *JAT*, 2023.
- [176] T. Zhao, Z. Huang, W. Tu, F. Biljecki, and L. Chen, “Developing a multiview spatiotemporal model based on deep graph neural networks to predict the travel demand by bus,” *IJGIS*, 2023.
- [177] F. Huang, P. Yi, J. Wang, M. Li, J. Peng, and X. Xiong, “A dynamical spatial-temporal graph neural network for traffic demand prediction,” *Information Sciences*, vol. 594, 2022.
- [178] Y. Wen, Z. Li, X. Wang, and W. Xu, “Traffic demand prediction based on spatial-temporal guided multi graph sandwich-transformer,” *Information Sciences*, vol. 643, 2023.
- [179] J. Li, F. Lin, G. Han, Y. Wang, R. Yu, A. M. Oguni, and Z. Li, “Pag-tsn: Ridership demand forecasting model for shared travel services of smart transportation,” *TITS*, 2023.
- [180] M. Nazzal, A. Khreichah, J. Lee, and S. Angizi, “Semi-decentralized inference in heterogeneous graph neural networks for traffic demand forecasting: An edge-computing approach,” *arXiv preprint arXiv:2303.00524*, 2023.
- [181] L. Liao, B. Li, F. Zou, and D. Huang, “Mfgcn: A multimodal fusion graph convolutional network for online car-hailing demand prediction,” *IEEE Intelligent Systems*, 2023.
- [182] Y. Yang, X. Shao, Y. Zhu, E. Yao, D. Liu, F. Zhao *et al.*, “Short-term forecasting of dockless bike-sharing demand with the built environment and weather,” *JAT*, 2023.
- [183] G. Jin, Z. Xi, H. Sha, Y. Feng, and J. Huang, “Deep multi-view graph-based network for citywide ride-hailing demand prediction,” *Neurocomputing*, vol. 510, 2022.
- [184] Y. Rong, Z. Xu, R. Yan, and X. Ma, “Du-parking: Spatio-temporal big data tells you realtime parking availability,” in *KDD*, 2018.
- [185] N. Arora, J. Cook, R. Kumar, I. Kuznetsov, Y. Li, H.-J. Liang, A. Miller, A. Tomkins, I. Tsogtsuren, and Y. Wang, “Hard to park? estimating parking difficulty at scale,” in *KDD*, 2019.
- [186] W. Zhang, H. Liu, Y. Liu, J. Zhou, and H. Xiong, “Semi-supervised hierarchical recurrent graph neural network for city-wide parking availability prediction,” in *AAAI*, 2020.
- [187] W. Zhang, H. Liu, Y. Liu, J. Zhou, T. Xu, and H. Xiong, “Semi-supervised city-wide parking availability prediction via hierarchical recurrent graph neural network,” *TKDE*, vol. 34, no. 8, pp. 3984–3996, 2020.
- [188] Y. Feng, Y. Xu, Q. Hu, S. Krishnamoorthy, and Z. Tang, “Predicting vacant parking space availability zone-wisely: A hybrid deep learning approach,” *Complex & Intelligent Systems*, vol. 8, no. 5, 2022.
- [189] D. Zhao, C. Ju, G. Zhu, J. Ning, D. Luo, D. Zhang, and H. Ma, “Mepark: Using meters as sensors for citywide on-street parking availability prediction,” *TITS*, vol. 23, no. 7, 2021.
- [190] P. Fafoutellis and E. I. Vlahogianni, “Traffic demand prediction using a social multiplex networks representation on a multimodal and multisource dataset,” *IJTST*, 2023.
- [191] C. Wang, Y. Xia, and H.-L. Shen, “Routing and congestion in multimodal transportation networks,” *International Journal of Modern Physics C*, vol. 34, no. 03, 2023.
- [192] W. Shao, Z. Jin, S. Wang, Y. Kang, X. Xiao, H. Menouar, Z. Zhang, J. Zhang, and F. Salim, “Long-term spatio-temporal forecasting via dynamic multiple-graph attention,” *arXiv preprint arXiv:2204.11008*, 2022.
- [193] A. Sharma, A. Sharma, P. Nikashina, V. Gavrilenko, A. Tselykh, A. Bozhenyuk, M. Masud, and H. Meshref, “A graph neural network (gnn)-based approach for real-time estimation of traffic speed in sustainable smart cities,” *Sustainability*, vol. 15, no. 15, 2023.
- [194] C. Conlan, J. Oakley, G. V. Demirci, A. Sfyridis, and H. Ferhatosmanoglu, “Real-time spatio-temporal forecasting with dynamic urban event and vehicle-level flow information,” in *CEUR Workshop Proceedings*, vol. 3379. RWTH Aachen University, 2023.
- [195] J. Jiang, B. Wu, L. Chen, K. Zhang, and S. Kim, “Enhancing the robustness via adversarial learning and joint spatial-temporal embeddings in traffic forecasting,” in *Proceedings of the 32nd ACM International Conference on Information and Knowledge Management*, 2023, pp. 987–996.
- [196] J. Zhang, Z. Qu, C. Chen, H. Wang, Y. Zhan, B. Ye, and S. Guo, “Edge learning: The enabling technology for distributed big data analytics in the edge,” *CSUR*, vol. 54, no. 7, 2021.
- [197] Z. Mao, J. Li, N. Zheng, K. Tei, and S. Honiden, “Transfer learning method in reinforcement learning-based traffic signal control,” in *GCCE*, 2021.
- [198] T. Yin, X. Zhou, and P. Krähenbühl, “Multimodal virtual point 3d detection,” in *NeurIPS*, 2021.
- [199] S. Zhancheng, “Research on application of deep reinforcement learning in traffic signal control,” in *ICFSP*, 2021.
- [200] T. Shi, F.-X. Devailly, D. Larocque, and L. Charlin, “Improving the generalizability and robustness of large-scale traffic signal control,” *arXiv preprint arXiv:2306.01925*, 2023.