IEEE Open Journal of Intelligent Transportation Systems

Received XX Month, XXXX; revised XX Month, XXXX; accepted XX Month, XXXX; Date of publication XX Month, XXXX; date of current version November 18, 2026.

Digital Object Identifier 10.1109/OJITS.

Deep Learning–Driven Traffic Flow Prediction: A Systematic Review of Spatiotemporal and Generative Models

Aneri Kamal Patel, Ke Liu, Jing Ma (Member, IEEE)

[1]Department of Data Science & AI, Auckland University of Technology, Auckland, New Zealand

Corresponding Author: Jing Ma (e-mail:jing.ma@aut.ac.nz).

ABSTRACT Traffic flow prediction is a critical task in intelligent transportation systems (ITS), enabling applications such as traffic management, route planning, and autonomous driving. The rapid growth of large-scale spatiotemporal traffic data and advances in deep learning have driven significant improvements in forecasting accuracy and adaptability. This systematic review synthesizes deep learning–based traffic flow prediction research published between January 2020 and September 2025, focusing on traffic data characteristics, spatiotemporal representations, and modeling paradigms, including graph neural networks (GNNs), transformer architectures, hybrid approaches, and emerging generative models.

Particular emphasis is placed on the role of generative and uncertainty-aware models in capturing complex spatiotemporal dependencies and providing robust predictions under real-world variability. Persistent challenges, including heterogeneous data fusion, dynamic graph adaptation, asynchronous dependency propagation, and scalability in large-scale networks, are critically examined. By highlighting methodological innovations and gaps in prior surveys, this review provides a structured reference and identifies future research directions for developing robust, adaptive, and generalizable traffic prediction systems.

INDEX TERMS Traffic flow prediction, intelligent transportation systems, spatiotemporal forecasting, graph neural networks, transformer models, generative models, uncertainty-aware prediction.

I. Introduction

With the rapid pace of urbanization and the deployment of large-scale intelligent transportation systems (ITS), traffic flow prediction has become a cornerstone of modern urban mobility management. Reliable short-term and medium-term traffic forecasts support a wide range of applications, including adaptive signal control, dynamic route guidance, congestion mitigation, and autonomous driving systems [1].

Meeting these operational requirements has driven continuous advances in traffic flow prediction methodologies. Early approaches relied on statistical time-series models, such as ARIMA [2], and classical machine learning techniques [3], including support vector regression [4] and random forests [5]. While effective under simplified assumptions, these methods struggle to capture the highly nonlinear, non-stationary, and spatiotemporally coupled dynamics of real-world traffic systems. The increasing availability of high-resolution traffic data from heterogeneous sources, such as loop detectors, GPS trajectories, connected vehicles, and external contextual information, has further exposed the limitations of these approaches [6]–[8].

Since approximately 2020, advances in deep learning have fundamentally reshaped research on traffic flow prediction. In this context, graph neural networks (GNNs) have emerged as a dominant paradigm for modeling spatial dependencies by explicitly encoding road-network topology, overcoming the constraints of grid-based convolutional models on non-Euclidean structures [9], [10]. In parallel, transformer-based architectures and attention mechanisms have demonstrated strong capabilities in capturing long-range spatiotemporal dependencies and periodic traffic patterns spanning hours to weeks through global attention [11]. These developments have significantly improved predictive accuracy and modeling flexibility, particularly for large-scale and complex transportation networks [12].

Despite this progress, several fundamental challenges remain unresolved and continue to limit the robustness and real-

This work is licensed under a Creative Commons Attribution 4.0 License. For more information, see https://creativecommons.org/licenses/by/4.0/

world applicability of existing traffic flow prediction models. First, modern traffic systems generate heterogeneous and asynchronous data from multiple sources with varying spatiotemporal granularity, sampling frequencies, and reliability levels. Although deep learning frameworks can learn rich representations from individual data streams, many existing approaches process heterogeneous information in isolation or rely on simple fusion strategies. Therefore, effectively integrating multi-source traffic data while accounting for temporal asynchronicity, variable data availability, and differing reliability profiles remains an open challenge [6], [7], [13], [14].

Second, traffic networks are inherently dynamic and non-stationary. Spatial correlations between road segments vary over time due to fluctuating demand, traffic signal control, incidents, construction activities, and road closures [10], [15]. However, many state-of-the-art GNN-based models rely on static or slowly adapting adjacency matrices, which fail to capture rapid topological evolution and time-varying propagation patterns. This mismatch between model assumptions and real-world traffic dynamics remains a critical bottleneck for accurate and robust traffic forecasting [7], [16].

Third, most existing traffic flow prediction models are deterministic, producing single-point forecasts without quantifying predictive uncertainty. In practice, traffic systems exhibit significant stochasticity arising from unpredictable events, data sparsity on low-volume roads, and complex human driving behaviors [17]. The lack of uncertainty-aware forecasting limits the usefulness of predictions for risk-sensitive decision-making in traffic management, navigation systems, and autonomous vehicles. These limitations have motivated growing interest in probabilistic and generative modeling approaches capable of producing distributional forecasts and multiple plausible future traffic states [18]–[20]. Recent diffusion-based forecasting frameworks further extend this line of work by learning complex predictive distributions under uncertainty.

These challenges also reveal important gaps in the existing survey literature. While prior reviews have provided valuable overviews of traffic flow prediction methods, they have primarily emphasized deterministic modeling performance and static spatiotemporal representations. Systematic discussions of dynamic graph modeling, asynchronous dependency propagation, multi-source information fusion, and generative artificial intelligence (AI) for uncertainty quantification remain relatively limited. Moreover, the rapid emergence of hybrid GNN–transformer architectures has yet to be comprehensively synthesized from a challenge-driven perspective.

To address these gaps, this paper presents a systematic literature review of recent advances in traffic flow prediction, with a particular focus on deep learning-based spatiotemporal modeling. Following the PRISMA 2020 guidelines, we analyze and synthesize representative studies published between January 2020 and September 2025, covering data characteristics, modeling paradigms, architectural innovations, and evaluation practices. In contrast to prior surveys, this review emphasizes the underlying modeling challenges that drive methodological design choices and highlights emerging trends toward adaptive, uncertainty-aware, and generative forecasting frameworks, as summarized in Table 1.

The main contributions of this review are summarized as follows:

- We present a comprehensive taxonomy of traffic data types, spatiotemporal representations, and intrinsic characteristics relevant to traffic flow prediction.
- We review widely used public traffic datasets and benchmarking platforms, highlighting their roles in enabling reproducible evaluation and standardized comparison.
- We provide a structured and critical analysis of GNN- and transformer-based traffic flow prediction models, with particular emphasis on dynamic graph evolution, asynchronous propagation mechanisms, and multi-source information fusion.
- We present a critical comparison of representative traffic flow prediction architectures, analyzing their strengths, limitations, computational characteristics, and suitability for different traffic forecasting scenarios.
- We examine emerging generative AI and probabilistic modeling approaches and discuss their potential for uncertainty quantification and data sparsity mitigation.
- We identify persistent research challenges and outline future directions toward robust, scalable, and generalizable traffic forecasting systems.

The remainder of this paper is organized as follows. Section II describes the systematic review methodology, including data retrieval and screening procedures. Section III discusses traffic data characteristics and spatiotemporal representations. Section IV reviews deep learning-driven traffic flow prediction models, ranging from GNNs and transformers to hybrid and generative AI frameworks. Section V presents a critical comparison of representative traffic flow prediction architectures and analyzes their strengths, limitations, and application scenarios. Section VI discusses the remaining challenges and emerging research directions. Finally, Section VII concludes the paper.

II. Research Approach

This study follows a systematic literature review (SLR) methodology to provide a comprehensive and systematic synthesis of recent advances in traffic flow prediction, with particular emphasis on deep learning-based and generative AI approaches. The review process was designed to ensure transparency, reproducibility, and methodological rigor and

TABLE 1. Summary of Related Reviews on Traffic Flow Prediction

Source	Data Collection			Data Analysis			Model Architecture				Modeling Features		
	Sensor	Trajectory	External	Statistics	Big Data	AI-Based	GNN	Transformer	Hybrid	Gen-AI	Dynamic Graph	Periodic	Async
[1]	✓	✓	✓	✓		✓	✓	✓	✓		✓	✓	✓
[6]	✓	✓		✓		✓	✓						
[12]	✓	✓				✓	✓		✓			✓	
[17]	✓	✓	✓	✓		✓	✓	✓	✓		✓		
[13]	✓	✓	✓	✓	✓								
[7]	✓	✓		✓		✓	✓	✓	✓		✓	✓	
Our review	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓

was conducted in accordance with the PRISMA 2020 guidelines [21].

A. Review Protocol and Scope

The objective of this review is to analyze, categorize, and synthesize recent research on traffic flow prediction, with particular emphasis on how modern deep learning architectures address spatiotemporal dependencies, network dynamics, heterogeneous data sources, and predictive uncertainty. Given the rapid evolution of the field and the emergence of generative AI-based forecasting approaches, the review focuses on studies published between January 2020 and September 2025. This period encompasses the recent evolution from predominantly deterministic deep learning models toward probabilistic and generative approaches, including variational, diffusion-based, and adversarial frameworks.

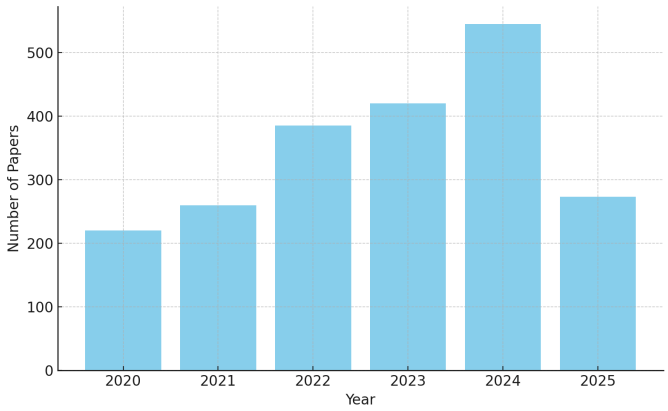

FIGURE 1. Annual distribution of traffic flow prediction studies published between January 2020 and September 2025. The 2025 count represents partial-year coverage up to September 2025.

B. Data Sources and Search Strategy

A comprehensive literature search was conducted across four major academic databases: Scopus, ACM Digital Library, IEEE Xplore, and ScienceDirect. The search strategy combined keywords related to traffic flow prediction, spatiotemporal modeling, and generative AI. Search keywords and their logical combinations were formulated using the following query:

> ("Traffic Flow Prediction" OR "Traffic Flow Forecasting" OR "Traffic Volume Prediction") AND ("Spatio-temporal" OR "Spatial" OR "Temporal") AND ("Convolutional Neural Networks" OR "Deep Learning" OR "Neural Networks") AND ("Transformer" OR "Attention Mechanism" OR "Generative Model" OR "Graph Neural Networks" OR "Recurrent Neural Networks" OR "GNN")

The search strategy was structured around three key dimensions: (i) the application domain (traffic flow prediction or forecasting), (ii) the spatiotemporal characteristics of traffic systems, and (iii) methodological approaches encompassing convolutional, recurrent, graph-based, transformer-based, and generative architectures. Synonymous terms and commonly used abbreviations were incorporated using logical OR operators to account for variations in terminology across studies. The search strategy was further validated through pilot searches to confirm the retrieval of key benchmark and highly cited studies.

To define the scope, studies based solely on traditional statistical methods (e.g., ARIMA), simulation-only analyses without predictive modeling, or non–data-driven approaches were excluded. The review focuses on short-term traffic flow prediction at the network or sensor level rather than individual vehicle trajectories.

The annual distribution of retrieved studies is illustrated in Fig. 1. A notable increase in publications after 2020 motivated our focus on recent literature, capturing advances in deep learning, graph-based modeling, and attention mechanisms for traffic flow prediction. It should be noted that the 2025 publication count reflects only studies available up to September 2025 and therefore represents partial-year coverage.

C. Screening and Quality Assessment

To manage and screen the retrieved records, a combination of Microsoft Excel, Zotero, Python-based scripts, and the Rayyan screening platform was employed. Zotero was used for reference management and duplicate detection, while Python scripts were used for metadata cleaning and

preliminary filtering. Rayyan facilitated title and abstract screening using semi-automated recommendations to support inclusion and exclusion decisions. The final database update was completed on September 30, 2025.

The study selection process consisted of three stages: duplicate removal, title and abstract screening, and full-text review. Duplicate records across databases were first eliminated. Titles and abstracts were then screened to remove clearly irrelevant studies. The remaining papers underwent full-text assessment to confirm compliance with the inclusion criteria and to evaluate methodological quality.

The initial search yielded 3,272 records, which were reduced to 2,506 after removing duplicate entries based on titles and metadata. Abstract-level screening excluded studies unrelated to urban traffic flow prediction, such as those focusing on non-road transportation modes or data security. Subsequently, studies retained for full-text assessment were restricted to publications appearing in leading journals and top-tier conferences. Journal quality was assessed using the Journal Citation Reports (JCR) quartile classification, and only journals ranked in the first quartile (Q1) within relevant categories (e.g., Transportation Science & Technology and Artificial Intelligence) were retained [22]. Conference quality was evaluated using the CORE Conference Ranking maintained by the Computing Research and Education Association of Australasia, and only conferences classified as A* or A were considered [23].

In addition to venue-based filtering, the methodological quality of selected studies was assessed during the full-text review stage. Specifically, studies were evaluated based on (i) clarity of the proposed methodology, (ii) presence of experimental validation, (iii) comparison with baseline or existing methods, and (iv) relevance to real-world traffic prediction scenarios.

These criteria were applied as structured eligibility checks rather than being combined into a numerical quality score. Methodological clarity required that the proposed approach be described in sufficient detail to understand its principal technical contribution; experimental validation required empirical evaluation using traffic data with clearly reported results; baseline comparison required evaluation against relevant existing or benchmark methods; and practical relevance required a direct connection to traffic flow prediction.

The full-text quality assessment was independently conducted by two authors using the predefined criteria. Any disagreement concerning study eligibility or methodological quality was discussed between the two assessors and resolved through consensus before the final inclusion decision was made. This procedure was adopted to improve consistency and reduce subjectivity during full-text screening.

Studies lacking sufficient methodological clarity, experimental evidence, meaningful baseline comparison, or practical relevance were excluded from the final analysis. Studies primarily focused on edge computing, lacking spatial or temporal analysis, or presenting purely methodological contributions without direct traffic applications were also excluded. The inclusion and exclusion criteria are summarized in Table 2. Following the full-text evaluation, 100 academic papers were included in the final analysis. The overall identification, screening, eligibility assessment, and inclusion process is illustrated in Fig. 2.

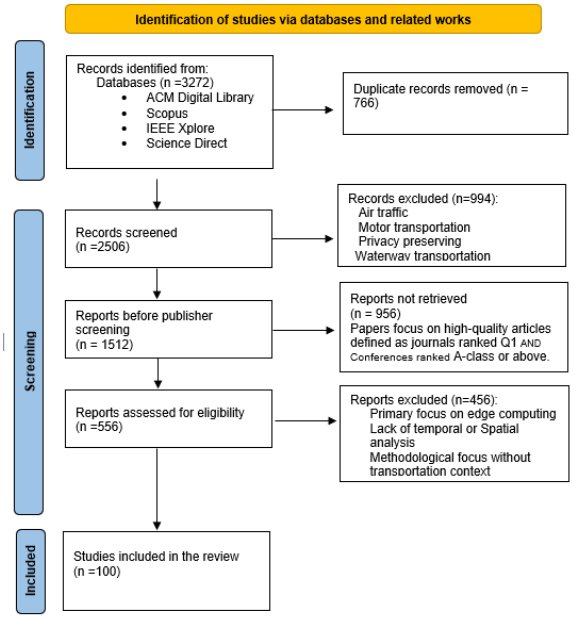

FIGURE 2. PRISMA-style flow diagram illustrating the study identification, screening, eligibility assessment, and inclusion process.

Although restricting the review to Q1 journals and A*/A-ranked conferences provides a consistent quality-oriented venue criterion, this restriction may introduce venue-selection bias by excluding relevant studies published in other reputable or emerging venues. Consequently, the final corpus should be interpreted as a quality-filtered representation of the literature rather than an exhaustive collection of all studies on traffic flow prediction. This limitation should be considered when interpreting the findings and research trends identified in this review.

D. Data Extraction and Analysis

To address the objectives of this review, the selected studies were analyzed according to the following research questions:

1) What traffic data sources are utilized, and how are they characterized in terms of spatial and temporal granularity?
2) How are spatiotemporal dependencies modeled, and which representation strategies are adopted?

TABLE 2. Inclusion and Exclusion Criteria

Criteria	Inclusion	Exclusion
Date	Publications between January 2020 and September 2025	Publications before 2020
Topic	Urban traffic flow prediction	Other transportation modes; data security
Quality	Q1 journals or A*/A-ranked conferences; studies meeting methodological quality criteria	Lower-ranked venues; studies not meeting methodological quality criteria
Scope	Studies incorporating spatial and/or temporal analysis	Edge computing; non-traffic-related studies

3) Which machine learning and deep learning paradigms are employed, and how have they evolved over time?
4) What evaluation metrics are commonly used, and which aspects of predictive performance do they emphasize?
5) How are unexpected traffic events, anomalies, and non-recurrent congestion addressed?
6) What are the emerging generative and uncertainty-aware deep learning approaches for traffic flow prediction?

These research questions guided the data extraction process and structured the comparative analysis presented in subsequent sections. In particular, studies were first examined with respect to traffic data characteristics and spatiotemporal representations, which form the foundation for understanding model design and performance.

III. Traffic Data Characteristics and Spatiotemporal Representations

Traffic flow prediction is fundamentally influenced by the characteristics of traffic data and the manner in which spatial and temporal dependencies are represented [7], [24]. Unlike conventional time-series forecasting, traffic flow prediction involves complex interactions among road networks, heterogeneous data sources, and dynamic operating conditions.

Accordingly, this section establishes a data-centric foundation by reviewing traffic data types and classifications, fundamental temporal and spatial characteristics, common spatiotemporal representation paradigms, and widely adopted evaluation metrics. This foundation provides the context for understanding the modeling choices and performance comparisons discussed in subsequent sections.

A. Data Types, Datasets, and Platforms

Traffic flow prediction relies on diverse data sources that differ in sensing mechanisms, spatial resolution, temporal granularity, and application context. This subsection reviews the classification of traffic data, widely used public benchmark datasets, and commonly adopted platforms and tools for traffic prediction research.

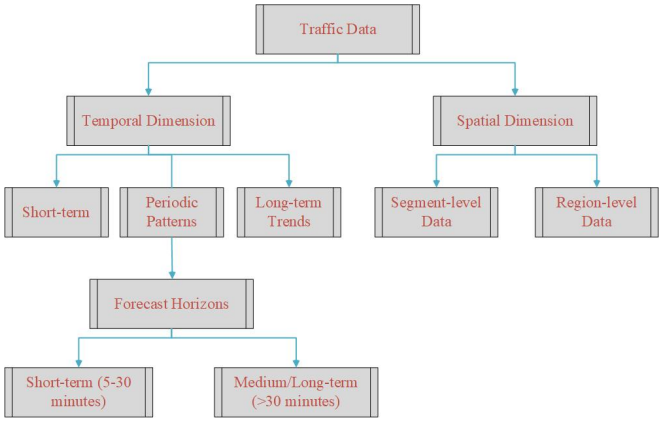

FIGURE 3. A unified classification of traffic data according to temporal horizons and spatial granularity.

1) Data Types and Classification

As illustrated in Fig.3, traffic data can be systematically structured along temporal and spatial dimensions. Commonly measured variables such as traffic flow, speed, and density evolve over time while simultaneously interacting across spatially connected road networks [24]. This temporal–spatial classification provides a structured basis for organizing traffic information according to prediction horizon and spatial granularity.

Classifying traffic data along temporal and spatial dimensions therefore provides a principled framework for organizing traffic information and guiding predictive modeling. From a temporal perspective, traffic data reflect short-term fluctuations, periodic patterns, and long-term trends. In contrast, from a spatial perspective, they capture complex dependencies induced by road network topology, traffic control mechanisms, and regional connectivity [24], [25]. This inherent spatiotemporal structure motivates the adoption of advanced learning models capable of jointly modeling temporal dynamics and spatial interactions [9], [10].

a: Temporal Dimension

From the temporal dimension, traffic data typically exhibit three distinct patterns: short-term fluctuations, periodic variations, and long-term trends [24]. Short-term fluctuations represent rapid changes in traffic conditions over short intervals, periodic variations correspond to recurring behaviors such as morning and evening peak hours, and long-term trends capture gradual changes driven by seasonal effects, infrastructure evolution, or demand growth. These temporal characteristics directly influence the selection of historical input windows and the design of traffic forecasting models [1], [12].

Building upon these temporal dimensions, traffic flow forecasting tasks are commonly categorized according to prediction horizons, as summarized in Fig.3. Based on the prediction horizon, traffic flow forecasting is generally categorized into short-term, medium-term, and long-term forecasting tasks. Short-term forecasting typically covers horizons of 5–30 minutes [9] and is widely used in real-time traffic management, including congestion monitoring, incident detection, and adaptive signal control. Owing to its high operational relevance, most existing studies focus on short-term prediction. In contrast, medium-term and long-term forecasting tasks, which typically extend beyond 30 minutes and may range from several hours to days, primarily support transportation planning, infrastructure development, and policy analysis.

b: Spatial Dimension

From the spatial dimension, traffic flow is governed by road network topology, connectivity between adjacent road segments, traffic signal control, and other infrastructural factors, resulting in complex spatial dependencies. For example, congestion occurring on a major road segment may propagate to neighboring links and surrounding areas. As illustrated in Fig.3, traffic data can therefore be categorized according to spatial granularity into segment-level and region-level representations [26].

Segment-level data describe traffic conditions on individual road segments, where each segment is defined as the connection between two nodes $S_i = (n_1, n_2)$. Traffic states on segment S_i are commonly represented using variables such as traffic flow $f_i(t)$, speed $v_i(t)$, and density $\rho_i(t)$, which are related through fundamental traffic flow relationships:

$$f_i(t) = v_i(t)\rho_i(t). \tag{1}$$

Such representations support fine-grained analysis of localized traffic behavior and are well suited for short-term prediction and real-time traffic control applications [27].

In contrast, region-level data aggregate traffic information across multiple road segments and intersections within a defined geographical area. A region R_j can be represented as a collection of segments $\{S_1, S_2, \ldots, S_k\}$, where k denotes the total number of segments in the region. The traffic state of R_j is commonly characterized using aggregated metrics such as average traffic flow $\bar{f}_j(t)$, average speed $\bar{v}_j(t)$, and average density $\bar{\rho}_j(t)$, defined as:

$$\begin{aligned} \bar{f}_j(t) &= \frac{1}{k}\sum_{i=1}^{k} f_i(t), \\ \bar{v}_j(t) &= \frac{1}{k}\sum_{i=1}^{k} v_i(t), \\ \bar{\rho}_j(t) &= \frac{1}{k}\sum_{i=1}^{k} \rho_i(t). \end{aligned} \tag{2}$$

Region-level data enable macro-scale traffic analysis and are widely used to investigate city-wide congestion patterns, regional mobility trends, and long-term transportation planning [13], [26].

While the preceding temporal-spatial classification defines how traffic data are organized and represented at different resolutions, it does not fully capture the intrinsic properties governing traffic evolution. Beyond data types, traffic data exhibit several fundamental characteristics (Section B) that directly determine the complexity of traffic dynamics and the design of forecasting models.

2) Publicly Available Datasets

Numerous public datasets support traffic prediction research, differing in sensing modality, spatial representation, temporal resolution, and prediction target [6], [13]. Table 3 summarizes widely adopted benchmarks covering diverse spatial scales, sensing technologies, and forecasting tasks.

Representative datasets can be broadly grouped according to their primary data characteristics. Sensor-based datasets, including PeMS variants [28], [29], METR-LA [30], TOPIS, WebTRIS, and LOOP [31], provide traffic flow, speed, and occupancy measurements from fixed detectors. Trajectory-based datasets, such as T-Drive [32], NYC Taxi [33], Porto Taxi [34], SZ-Taxi [35], and TaxiCD [36], contain vehicle movement records that support mobility analysis and travel-time prediction. Demand-oriented datasets, including TaxiBJ [37], [38] and DiDi datasets [39], capture passenger-demand and crowd-flow patterns, while metro, bike-share, and bus datasets extend prediction tasks to multimodal transportation systems [40].

Graph-oriented resources, such as GAIA [41] and Q-Traffic [42], incorporate road-network information, while OpenStreetMap [43] is commonly used to construct network topologies for graph-based modeling. Emerging multimodal datasets, such as MiTra [44], further combine multiple sources of mobility and network information to represent more complex traffic dynamics.

Collectively, these datasets support the evaluation of traffic forecasting methods across different sensing modalities, spatial representations, and urban environments [6], [13]. Their

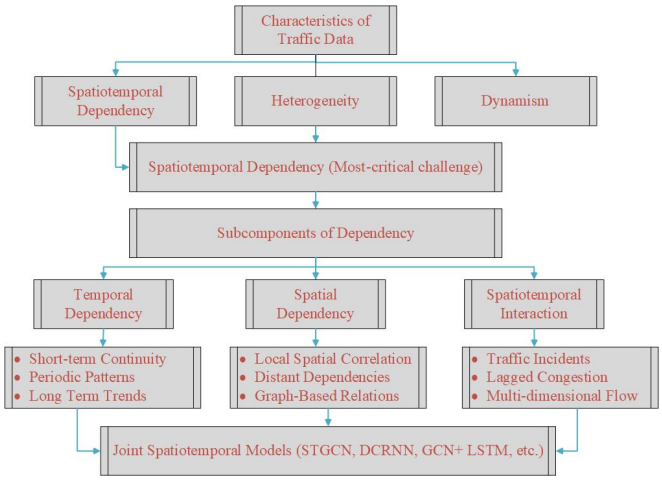

FIGURE 4. Conceptual overview of the fundamental characteristics of traffic data, emphasizing spatiotemporal dependency and its decomposition into temporal dependency, spatial dependency, and spatiotemporal interaction.

public availability also facilitates reproducible experimentation and standardized benchmarking [31].

3) Platforms and Tools

Traffic prediction research is supported by a range of mapping, simulation, benchmarking, and real-time data platforms [6], [13]. OpenStreetMap [43], [45] provides road-network information for graph construction and spatial modeling, while SUMO [6], [46], [47] and MATSim [48] provide simulation environments for traffic analysis and scenario-based experimentation. Benchmarking platforms such as TrafficNet [13], [49] and LibCity [31], [50] provide datasets, model implementations, and evaluation pipelines that support reproducible model comparison. Real-time services, such as the Google Traffic API [6], [51], additionally provide dynamic traffic information for deployment-oriented applications.

Together, these resources support network construction, simulation, benchmarking, and real-world data integration across different stages of traffic prediction research.

B. Fundamental Characteristics of Traffic Data

Traffic data are characterized by three fundamental properties: spatiotemporal dependency, heterogeneity, and dynamism, which jointly influence the design and performance of traffic flow forecasting models [7], [24]. As shown in Fig. 4, this subsection focuses on spatiotemporal dependency as a central characteristic of traffic data, encompassing both the temporal evolution of traffic states and their spatial interactions across road networks [25]. Accordingly, understanding temporal dependency, spatial dependency, and their interaction provides a foundation for selecting appropriate traffic-data representations and forecasting architectures, as discussed in the subsequent sections.

1) Temporal Dependency

As shown in the left branch of Fig. 4, temporal dependency characterizes how historical traffic conditions influence future traffic states. Core traffic variables, including flow, speed, and density, exhibit strong temporal correlations, evident in short-term continuity, periodic patterns, and long-term trends. Accurately capturing these temporal characteristics is fundamental to achieving reliable traffic forecasting [12].

Short-term continuity reflects the smooth evolution of traffic conditions over brief time intervals. For instance, recently observed traffic volume and speed often provide strong cues about near-future states. Periodic patterns arise from recurring human mobility behaviors, such as morning and evening peak hours, reduced late-night demand, and systematic differences between weekdays, weekends, and holidays. These regularities introduce strong cyclic structures into traffic time series. Long-term trends represent gradual shifts in traffic demand driven by factors such as urban expansion, infrastructure development, policy interventions, and seasonal variations.

Together, these short-term, periodic, and long-term characteristics require forecasting models to represent traffic evolution across multiple temporal scales, as conceptually summarized in Fig. 4.

2) Spatial Dependency

Corresponding to the middle branch in Fig. 4, spatial dependency describes the interdependence of traffic conditions across neighboring road segments and regions. These dependencies are shaped by road-network topology, traffic management mechanisms, and vehicle routing behavior [6], [13].

Local spatial dependencies arise when congestion on a major road segment generates spillback effects that influence adjacent links and surrounding areas. Distant dependencies extend beyond immediate neighbors; for example, congestion on a highway segment may affect downstream ramps or arterial roads located farther from the original bottleneck. In road-network-based settings, such relationships are naturally non-Euclidean because traffic locations are connected according to network topology rather than regular spatial grids. Consequently, road segments, sensors, or intersections can be represented as graph nodes, with edges encoding physical or functional dependencies.

These characteristics motivate graph-based representations capable of capturing both local and non-local traffic interactions; the corresponding graph-learning architectures are reviewed in Section IV.

3) Spatiotemporal Interaction

As illustrated in the right branch of Fig. 4, temporal and spatial dependencies in traffic systems are inherently inter-

TABLE 3. Overview of Public Traffic Datasets by Data Source and Available Traffic Attributes

Dataset		Data Source/Modality					Traffic Attributes					
		Sensor	Traj.	Weather	Ride-Hail	Road Net.	Flow	Speed	Occ.	Travel Time	Weather	Topology
PeMS	PeMS3	✓					✓	✓	✓			
	PeMS4	✓					✓	✓	✓			
	PeMS7	✓					✓	✓	✓			
	PeMS8	✓					✓	✓	✓			
	PeMS-BAY	✓					✓	✓	✓			
T-Drive			✓					✓		✓		
TaxiBJ				✓				✓			✓	
DiDi-Chengdu					✓		✓					
DiDi-Xi'an					✓		✓					
GAIA						✓						✓
Q-Traffic						✓	✓	✓				✓
TOPIS		✓					✓	✓				
WebTRIS		✓					✓	✓				
LOOP		✓					✓	✓	✓			
Porto			✓					✓				
SHSpeed		✓						✓				
SZ-Taxi			✓					✓		✓		
TaxiCD			✓							✓		
NYC Taxi			✓					✓		✓		
BikeDC			✓							✓		
NYC Bike			✓							✓		
NYC Bus			✓							✓		
CHI Bus			✓							✓		
SHMetro			✓			✓				✓		✓
HZMetro			✓			✓				✓		✓
METR-LA		✓					✓	✓				

connected. Their dynamic interaction governs traffic evolution and gives rise to complex, nonlinear spatiotemporal behaviors [7].

Traffic incidents, such as accidents or road closures, disrupt traffic conditions at specific locations and trigger cascading effects that propagate spatially across the road network and temporally over subsequent periods [52]. Lagged congestion provides another clear example of spatiotemporal coupling. For instance, traffic demand on major corridors may increase sharply during peak hours, while congestion on adjacent arterial or local roads emerges with a temporal delay due to flow redistribution, spillback, and route-switching behaviors [7].

These interactions demonstrate that traffic evolution cannot be fully characterized by temporal or spatial information alone. Effective forecasting therefore requires representations that jointly preserve temporal evolution and network-level spatial relationships. Building on these characteristics, the following subsection examines the principal data representations used in traffic forecasting models.

C. Representation and Structure

Building on the traffic characteristics discussed in Section B, this subsection focuses on how traffic observations are structurally organized as inputs to deep learning models. An appropriate data representation not only facilitates effective learning of traffic patterns but also directly influences model performance, generalization, and scalability [7], [24]. Based on how temporal and spatial information are encoded, existing studies can be broadly categorized into four main representation paradigms: temporal sequence representation, spatial grid representation, graph-based representation, and spatiotemporal hybrid representation [6]. These paradigms differ primarily in how they organize observations and encode relationships across time and network locations.

1) Temporal Sequence Representation

Temporal sequence representation focuses on modeling the temporal evolution of traffic states. In this paradigm, traffic observations are organized as ordered time series, enabling models to learn temporal dependencies from historical data. Traffic variables such as flow, speed, and occupancy are typically recorded at fixed time intervals, producing multivariate time series that form the basis of short- and medium-term traffic prediction tasks [12], [24], [27]. However, purely sequential representations do not explicitly encode relationships among spatially connected road segments, limiting their ability to characterize network-level traffic interactions.

2) Spatial Grid Representation

Spatial grid representation partitions the study area into discrete spatial units and maps traffic states onto grid cells to model the spatial distribution of traffic flow. This representation is widely used in large-scale urban traffic analysis, particularly when trajectory-based data or region-level aggregation is employed [13]. Two principal partitioning strategies are regular grids and region-based partitions.

Regular grid partitioning divides the urban area into fixed-size rectangular cells, with each cell storing aggregated traffic features such as average flow or speed. This approach is computationally efficient and facilitates the application of CNNs [26]. However, regular grids do not align well with road-network topology and may obscure fine-grained traffic interactions.

By contrast, region-based partitioning may incorporate administrative boundaries, functional zones, or other semantically meaningful spatial divisions, producing spatial units that better reflect urban structure. Although spatial grid representations enable effective extraction of spatial patterns, their performance is sensitive to spatial resolution, and they may struggle to represent non-Euclidean road connectivity. Consequently, these representations can be less suitable for fine-grained traffic prediction tasks that require explicit modeling of complex road-network connectivity.

3) Graph-Based Representation

Graph-based representation explicitly models the topological structure of road networks and has emerged as a prominent paradigm in recent traffic flow prediction research. In this formulation, the transportation network is represented as a graph, where nodes correspond to road segments, intersections, or sensors, and edges represent physical or functional connectivity between them [7], [13]. Node attributes may include traffic flow, speed, density, or historical statistics, while edges encode adjacency relationships, distances, or other interaction measures.

Compared with spatial grid-based modeling, graph-based representations provide a more topology-aware description of real-world traffic networks by preserving non-Euclidean road topology and fine-grained connectivity. A systematic comparison between grid-based and traffic-network representations, including their structural characteristics, data sources, and typical application scenarios, is summarized in Table 4. Nevertheless, the effectiveness of graph-based representations depends strongly on how the graph structure is constructed, while scalability becomes increasingly important as network size grows [53].

4) Spatiotemporal Hybrid Representation

Spatiotemporal hybrid representations combine temporal observations with explicit spatial structures, allowing traffic states to be organized jointly across time and network locations. A common formulation represents traffic observations as sequences of graph- or grid-structured states, thereby preserving both historical evolution and spatial organization.

Graph-based hybrid representations organize successive traffic states over fixed, learned, or adaptive network topologies, whereas grid-based hybrid representations combine spatially organized traffic maps with temporal sequences. The former preserve irregular road-network connectivity, while the latter provide regular spatial structures suitable for region-level traffic analysis.

Overall, the choice of representation determines the structural information available to the forecasting model. Temporal sequences emphasize historical evolution, spatial grids organize regional patterns, graphs preserve road-network connectivity, and hybrid representations jointly encode spatial and temporal information. Building on these representation paradigms, the following section reviews the deep learning architectures used to extract predictive features from these structures.

D. Evaluation Metrics for Deterministic and Probabilistic Traffic Flow Prediction

Traffic flow prediction performance is evaluated using metrics selected according to the prediction task, data representation, and application requirements. While deterministic forecasting primarily assesses point prediction accuracy, probabilistic and uncertainty-aware forecasting also requires evaluation of distributional quality, calibration, sharpness, and interval reliability [18]–[20].

For deterministic prediction, the most widely adopted metrics are Mean Absolute Error (MAE), Root Mean Squared Error (RMSE), and Mean Absolute Percentage Error (MAPE). MAE measures the average magnitude of prediction errors, whereas RMSE assigns greater weight to large deviations and is therefore more sensitive to large errors. MAPE expresses prediction errors relative to observed values, facilitating comparison across different scales; however, it can become unstable when observed values are zero or close to zero [7], [24], [30]. These metrics remain standard measures for evaluating point forecasting performance.

For traffic prediction tasks formulated over spatial grids or regions, additional task-specific metrics can be used to assess the structural consistency between predicted and observed spatial patterns. The Structural Similarity Index (SSIM) evaluates similarity in terms of luminance, contrast, and structural information and can therefore be applied to grid-based traffic representations [13], [54]. Intersection over Union (IoU) measures the overlap between predicted and observed regions and is applicable when the prediction task is formulated as spatial segmentation or region-level pattern estimation [26], [55].

For spatiotemporal forecasting, conventional error metrics such as RMSE and MAE can be aggregated across both

TABLE 4. Comparison of Grid-Based Modeling and Traffic Network (Graph-Based) Modeling

Feature	Grid-Based Modeling	Traffic Network (Graph-Based) Modeling
Node Definition	Regular spatial grid cells covering a geographic area	Road segments, intersections, or traffic sensors
Edge Definition	Implicit adjacency based on spatial proximity	Explicit edges defined by physical road connectivity
Spatial Structure	Euclidean and fixed	Non-Euclidean and topology-driven
Adjacency Modeling	Predefined grid neighborhood (e.g.4- or 8-neighbors)	Physical road links; static or dynamically learned
Spatial Granularity	Region-level (coarse, city-scale representation)	Segment-level (fine-grained modeling)
Modeling Flexibility	Limited ability to represent irregular road layouts	Naturally captures complex road network structures
Application Scenarios	Urban-wide traffic monitoring and trend analysis	Road-network-specific prediction and signal control
Data Sources	Trajectory data or region-aggregated traffic data	Sensor-based, loop detector, or station-level data
Typical Model Choices	CNN-based spatiotemporal models	GNN-based spatiotemporal models
Limitations	Loss of fine-grained road topology information	Higher computational and modeling complexity

spatial locations and temporal prediction steps to evaluate performance over the complete forecasting horizon [9], [10]. Such deterministic measures remain important because they provide interpretable assessments of point accuracy and support comparison with established forecasting methods.

However, deterministic metrics alone are insufficient for models that produce predictive distributions rather than single-point estimates [56]. Probabilistic forecasts should additionally be evaluated in terms of distributional quality, calibration, sharpness, and interval reliability [57]. A model may achieve competitive MAE or RMSE while producing poorly calibrated or excessively diffuse uncertainty estimates, limiting its usefulness in risk-sensitive applications [58].

Negative Log-Likelihood (NLL) evaluates the likelihood assigned by the predicted distribution to the observed outcomes, with lower values indicating better probabilistic fit [56], [58]. The Continuous Ranked Probability Score (CRPS) measures the discrepancy between the predicted cumulative distribution and the observed value, providing an assessment of the full predictive distribution rather than only a point estimate [56].

For interval-based forecasts, Prediction Interval Coverage Probability (PICP) measures the proportion of observations contained within the predicted intervals [59]. Because high coverage can be obtained using excessively wide intervals, PICP should be interpreted together with measures such as Mean Prediction Interval Width (MPIW), which evaluates interval sharpness [57], [59]. Calibration and sharpness are therefore complementary: calibration assesses the consistency between predicted uncertainty and observed outcomes, whereas sharpness reflects the concentration and informativeness of the predictive distribution [57].

Overall, metric selection should reflect the forecasting objective and output type. Deterministic models are typically evaluated using complementary point-error and, where appropriate, spatial or spatiotemporal metrics, whereas probabilistic models additionally require distributional and uncertainty-aware criteria. Reporting complementary deterministic and probabilistic measures provides a more comprehensive assessment of predictive accuracy, uncertainty calibration, and practical utility in intelligent transportation applications.

IV. Deep Learning–Driven Traffic Flow Prediction: From Deterministic Spatiotemporal Fusion to Generative AI

This section systematically reviews deep learning-based traffic flow prediction approaches from the perspective of major architectural paradigms, with particular emphasis on the methodological evolution from deterministic spatiotemporal modeling toward generative and uncertainty-aware approaches. Earlier studies primarily focused on capturing temporal dynamics and spatial dependencies using graph neural networks, transformer-based architectures, and hybrid spatiotemporal models, as illustrated in Fig. 5. Although these approaches have substantially improved point-forecasting accuracy, most remain deterministic and provide limited capability to characterize predictive uncertainty and multimodal traffic dynamics.

More recently, generative and probabilistic approaches have extended traffic forecasting beyond point prediction by enabling distributional forecasting and uncertainty quantification. Accordingly, this section categorizes existing deep learning-based methods into four representative classes: graph neural network-based models, transformer-based models, hybrid spatiotemporal architectures, and generative and uncertainty-aware frameworks. This taxonomy highlights the methodological progression from deterministic spatiotemporal representation learning toward probabilistic forecasting and provides a structured basis for examining the capabilities and limitations of different architectural paradigms.

Fig. 6 complements this taxonomy by providing a qualitative comparison of the four architectural paradigms across key modeling dimensions, including spatial dependency modeling, long-range temporal reasoning, nonlinearity handling, generative capability, uncertainty quantification, computational efficiency, and interpretability. The comparison highlights distinct strengths and trade-offs across the model

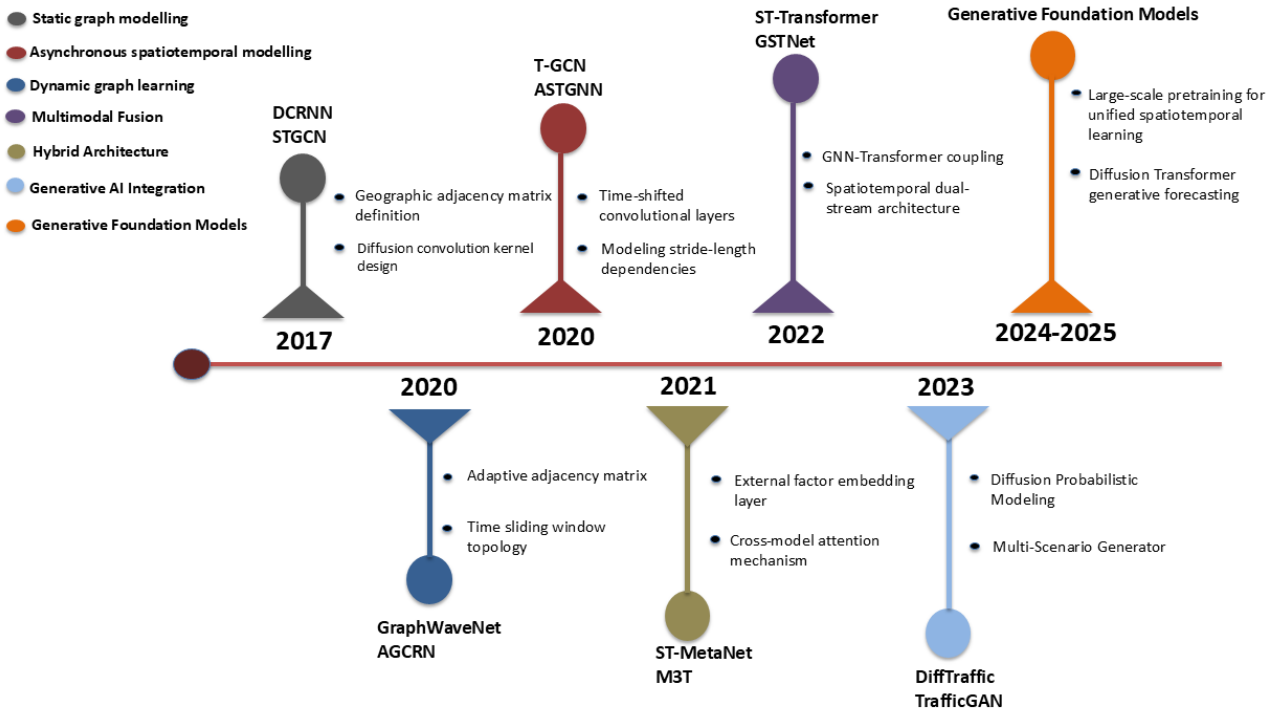

FIGURE 5. Evolution of spatiotemporal traffic forecasting models from static graph learning to generative and uncertainty-aware models (2017–2025).

families, indicating that no single paradigm consistently dominates across all dimensions. Accordingly, the figure provides a multidimensional perspective on architectural capabilities beyond point-forecasting performance alone.

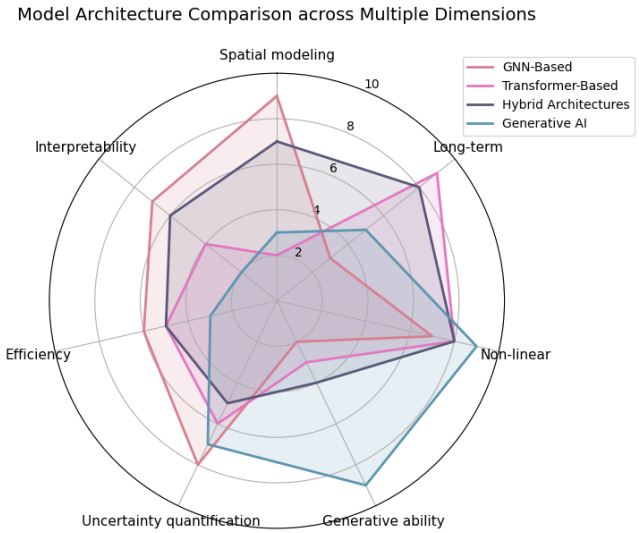

FIGURE 6. Conceptual comparison of prominent deep learning architectures for traffic flow prediction, highlighting their strengths and trade-offs across key modeling dimensions.

A. Graph Neural Networks: Topology-Driven Approaches

Graph Neural Networks (GNNs) have emerged as a prominent architectural paradigm for traffic flow prediction because of their ability to explicitly model road-network topology and capture non-Euclidean spatial dependencies that are difficult to represent using conventional grid-based convolutions. Traffic networks exhibit irregular connectivity, directional flow propagation, and heterogeneous interactions among road segments, making graph-based representations particularly suitable for spatial traffic modeling.

By representing road segments, intersections, or sensors as nodes and their relationships as edges, GNN-based models can learn spatial correlations and traffic propagation patterns across transportation networks. This capability enables them to capture both local and broader network-level dependencies, making GNNs particularly effective for traffic forecasting in complex and large-scale road networks.

A comprehensive overview of representative GNN-based traffic flow prediction models is presented in Table 5, which summarizes representative studies in chronological order. This taxonomy highlights the evolution of GNN architectures from early static-graph formulations to more recent dynamic, attention-enhanced, and multi-source learning frameworks, reflecting sustained research efforts to address the non-

TABLE 5. Chronological categorization of representative GNN-based models for traffic flow prediction.

Category	Model	Citation	Year
GNN-based	IGAGCN	[60]	2021
	GTA	[61]	2021
	ADGCN	[62]	2022
	AARGNN	[63]	2022
	Loc-GCLSTM	[64]	2022
	ASTFGCN	[65]	2022
	STDEN	[66]	2022
	DMGNN	[52]	2023
	ST-CGCN	[67]	2023
	GDGCN	[68]	2023
	STGDAE	[69]	2023
	STRGAT	[70]	2023
	STHGCN	[71]	2023
	STTGCN	[72]	2023
	DGCRN	[73]	2024
	TAGNN	[74]	2024
	ByGCN	[75]	2024
	LSSTF-GWN	[76]	2024
	GDSTGCN	[77]	2024
	PSPJSTGCN	[78]	2024
	ISGCN	[79]	2024
	MGC-DMF	[80]	2025
	DSTF	[81]	2025
	MVDGCN	[82]	2025
	DGSTA	[83]	2025
	DA-RMN	[84]	2025
	TFSG	[85]	2025
	MFE-GNN	[86]	2025
	CGNN	[87]	2025
	SCA-GCN	[88]	2025
	MuSrc	[89]	2025

stationary and heterogeneous characteristics of real-world traffic systems.

1) Evolution of Graph Modeling: From Static Topologies to Dynamic Representations

GNNs have emerged as a prominent paradigm for modeling spatial dependencies in traffic flow forecasting due to their ability to explicitly encode road network topology. Early GNN-based approaches typically assume that the underlying traffic network is static and known a priori, representing traffic systems using a predefined graph structure. In these methods, a fixed adjacency matrix is commonly constructed from geographical distance, road connectivity, or sensor proximity and remains unchanged throughout training and inference.

A representative example is the Diffusion Convolutional Recurrent Neural Network (DCRNN), which defines a diffusion process over the road network and integrates it within a recurrent neural network framework to capture spatiotemporal dependencies on a static graph [30]. Under this formulation, traffic information propagates along predefined network links, while the underlying spatial connectivity remains fixed throughout prediction.

However, real-world traffic systems are inherently non-stationary. Spatial relationships may evolve because of congestion, signal timing, incidents, roadworks, and fluctuating demand. For example, congestion may propagate along different routes under peak and off-peak conditions, while temporary disruptions can alter effective network connectivity. Fixed adjacency structures cannot explicitly adapt to such changes, limiting their ability to represent transient traffic propagation patterns [52], [73].

To address the limitations of fixed adjacency, a substantial body of work has focused on dynamic graph modeling,

TABLE 6. Graph Neural Network Approaches for Traffic Flow Prediction: Challenge-Based Categorization

Challenge Category	Model Name	Dynamic Graph	Async. Modeling	External Factors	Periodic / Long-term
Dynamic Graph Modeling	DMGNN [52]	✓		✓	
	DGCRN [73]	✓			✓
	TAGNN [74]	✓	✓	✓	✓
	ByGCN [75]	✓			✓
	LSSTF-GWN [76]	✓			✓
	GDGCN [68]	✓	✓		✓
	GDSTGCN [77]	✓			✓
	STGDAE [69]	✓			
	STPGCN [90]	✓			✓
	PSPJSTGCN [78]	✓			✓
	STTGCN [72]	✓			✓
	MVDGCN [82]	✓			
	DGSTA [83]	✓	✓		
Asynchronous Dependency	ADGCN [62]	✓	✓		✓
Multi-Source Fusion	AARGNN [63]			✓	
	ST-CGCN [67]			✓	✓
	STHGCN [71]	✓			✓
	ISGCN [79]	✓			✓
	MGC-DMF [80]			✓	✓
	DSTF [81]				
	DA-RMN [84]				
	MFE-GNN [86]			✓	
	CGNN [87]			✓	
	SCA-GCN [88]			✓	
	MuSrc [89]			✓	
Periodic Dependency	IGAGCN [60]				✓
	GTA [61]				✓
	ASTFGCN [65]	✓			✓
	Loc-GCLSTM [64]	✓			✓
	STRGAT [70]				✓
	TFSG [85]			✓	✓

Note: Models without checkmarks do not explicitly address the corresponding challenges and instead focus on baseline spatiotemporal representation learning.

as summarized in the first block of Table 6. Early dynamic approaches, such as DMGNN and DGCRN, move beyond manually specified graphs by learning time-varying adjacency structures or multi-graph ensembles directly from data [52], [73]. DMGNN constructs multiple graphs conditioned on traffic states and accident information to capture regime changes during disruptive events, whereas DGCRN, in contrast, generates a dynamic affinity matrix at each time step and fuses it with a static base graph, thereby balancing structural prior knowledge with data-driven adaptability. These approaches illustrate the transition from fixed spatial structures toward data-driven graph adaptation under changing traffic conditions.

Despite their effectiveness, these early dynamic models often rely on coarse global graph updates, where adjacency changes reflect overall traffic states rather than localized structural variations. As a result, such models may struggle to disentangle localized phenomena, such as spillback near individual intersections, from global network-wide fluctuations.

Subsequent models, including TAGNN, ByGCN, LSSTF-GWN, GDGCN, and GDSTGCN, refine this idea by reconstructing graph structures from hidden representations or history-driven interactions [68], [74]–[77]. TAGNN incorporates time-aware node embeddings and external contextual features, enabling the adjacency structure to vary across temporal regimes, such as peak and off-peak hours. ByGCN and LSSTF-GWN adopt multi-branch or multi-graph designs, where distinct graphs model different temporal scales or structural views of the network. GDGCN and GDSTGCN emphasize history-driven graph construction to capture non-local propagation patterns, particularly in freeway and highway networks.

While these designs substantially improve modeling flexibility, they also introduce increased architectural complexity

and raise important concerns regarding identifiability and interpretability. Multiple learned graphs may encode overlapping or redundant dependencies, making it difficult to associate learned structures with physical road characteristics or traffic mechanisms. Moreover, insufficient regularization may reduce the stability and generalizability of learned graph representations.

More recent developments, such as MVDGCN and DGSTA, further advance dynamic modeling through multiview and attention-enhanced formulations, in which adjacency weights are updated at each time step under the guidance of attention mechanisms or view-specific embeddings. These models demonstrate the potential of fine-grained, context-aware adaptation of spatial dependencies. However, increasingly parameterized dynamic and attention-based mechanisms can introduce additional computational cost and may increase sensitivity to data sparsity and distribution shifts [82], [83].

Complementing the challenge-based taxonomy, Table 7 provides a mechanism-level comparison of representative dynamic GNN models. Specifically, it summarizes the spatiotemporal modeling strategies, core architectural mechanisms, and targeted application scenarios associated with each approach. By explicitly linking architectural design choices to practical deployment contexts such as urban networks with strong periodicity, highways with rapidly evolving propagation patterns, or networks subject to unexpected disruptions, this table highlights how different spatiotemporal fusion strategies are tailored to distinct traffic environments.

Overall, the progression from static graph models such as DCRNN to dynamic and attention-driven architectures reflects a transition from predefined network topologies toward increasingly adaptive spatial representations. While these developments improve the ability to represent time-varying traffic interactions, they introduce additional challenges related to computational complexity, interpretability, and generalization. These broader challenges are discussed in Section VI.

2) Asynchronous Propagation: Temporal Decoupling Mechanisms

In real-world traffic systems, disturbances do not necessarily propagate instantaneously across connected road segments. Queue spillback, signal timing, and traffic redistribution can introduce temporal delays between changes at different network locations. However, many conventional spatiotemporal GNNs primarily model spatial relationships among nodes at the same time step, which may limit their ability to represent such lagged interactions.

To address this limitation, asynchronous modeling introduces cross-time dependencies into graph-based traffic representations. A representative example is the Asynchronous Dilated Graph Convolutional Network (ADGCN), which explicitly models lagged traffic interactions [62]. Its Asynchronous Spatiotemporal Graph Convolution (ASTGC) operator extends spatial relationships across multiple time lags, allowing the traffic state of a road segment at time t to influence connected segments at subsequent time steps $t+\Delta$. This formulation enables spatial propagation to be modeled without constraining interacting nodes to the same temporal instant.

In addition, ADGCN extends one-dimensional dilated causal convolution to graph-structured data, allowing the temporal receptive field to grow with network depth while preserving temporal causality [62]. By combining asynchronous graph convolution with dilated temporal convolutions, ADGCN demonstrates improved performance over conventional synchronous graph-based baselines on several benchmark traffic forecasting datasets, particularly for long-horizon prediction tasks where lagged propagation effects are more pronounced. These results suggest that explicitly modeling temporal decoupling can be beneficial for capturing longer-range traffic dynamics.

Beyond ADGCN, several dynamic-graph models summarized in Table 6, including TAGNN, GDGCN, and DGSTA, partially capture asynchronous behavior by allowing adjacency structures to evolve over time or by combining dynamic graphs with multi-scale temporal modeling [68], [74], [83]. In these methods, asynchronous effects typically arise indirectly as a consequence of dynamic topology learning or multi-step temporal aggregation, rather than being an explicit modeling objective. Moreover, limited analysis is usually provided on which specific time delays contribute most to performance improvements, making interpretation and robustness under irregular traffic conditions less clear.

Overall, asynchronous propagation has received relatively less attention compared to other research directions such as dynamic graph construction and multi-source data fusion. While ADGCN offers a clear example of explicitly modeling lagged spatial interactions, models such as TAGNN primarily focus on alleviating graph staticity through temporal embeddings and only indirectly capture delay effects. From a systematic review perspective, it is therefore important to distinguish models that explicitly model cross-time dependencies from those that capture asynchrony implicitly, in order to avoid conflating different modeling objectives within the spatiotemporal traffic forecasting literature.

3) Multisource Fusion: Integrating External and Contextual Factors

Traffic dynamics are influenced not only by historical traffic patterns and road network topology but also by a wide range of external factors, including weather conditions, holidays, public events, land-use patterns, and traffic incidents. Models that rely solely on traffic time series and spatial structure often struggle to capture anomalous behaviors or sudden disruptions induced by these exogenous factors. Consequently, an important research challenge in traffic forecasting is how

TABLE 7. Spatiotemporal Modeling Approaches, Core Mechanisms, and Application Scenarios

Model Name	Spatiotemporal Approach	Core Mechanism (Challenge Focus)	Application Scenario
DMGNN [52]	GAT + Bi-LSTM	Multi-perspective dynamic graph structure + dynamic adjacency matrix update	Multi-perspective traffic prediction under unexpected events
DGCRN [73]	Dynamic graph convolution + GRU + affinity matrix via linear constraints	Supernet-generated dynamic graph + static-affinity fusion for enhanced propagation	Multi-step prediction in frequently changing road networks
TAGNN [74]	Adaptive GCN + grouped TCN + external factor embedding	Time-related adjacency matrix + node embedding to enhance time-varying spatial dependency	Urban roads with high periodicity and dynamic dependencies
ByGCN [75]	Branch recognition module + multi-period modeling + spatial attention + GCN	Reconstructing graph structure with hidden branch identification + spatial-temporal decoupling	Complex networks with strong branch influence or incomplete topology
LSSTF-GWN [76]	Graph wavelet convolution + multi-graph fusion + temporal convolution + CBAM attention	Fusion of static, long-term interaction, and short-term dynamic graphs	Long-term prediction in highly volatile traffic regions
GDGCN [68]	Dynamic graph construction + spatial GCN block + temporal GCN block	Dynamic graph constructor for spatial graph + integrated temporal graph modeling	Urban arterials and highways with significant multi-scale variations
GDSTGCN [77]	Dynamic graph generation + diffusion convolution + multi-layer temporal convolution	History-driven dynamic graph generation + diffusion modeling of non-adjacent propagation	Dynamic structure evolution modeling on highways

to effectively integrate heterogeneous external data sources such as meteorological information, incident records, and points of interest (POIs) to improve prediction accuracy and robustness.

A representative example of multisource fusion is the Attentive Attributed Recurrent Graph Neural Network (AARGNN) [63]. In this framework, traffic sensors are modeled as attributed graph nodes enriched with both static and dynamic features. Static attributes encode spatial distances and semantic similarities between sensors, while dynamic attributes capture local traffic states and broader contextual information, such as regional congestion levels. These heterogeneous features are embedded as node and edge attributes and processed using a hybrid architecture that combines GNN layers with LSTM units. An attention mechanism is further employed to adaptively learn the relative importance of different information sources. Empirical results indicate that carefully designed cross-modal fusion can lead to substantial improvements over traffic-only baselines.

Beyond attribute-based integration, another line of research adopts dedicated fusion modules or multi-branch network architectures to handle heterogeneous inputs. The Spatiotemporal Composite Graph Convolutional Network (ST-CGCN), for instance, encodes spatial topology, temporal periodicity, and environmental context into a multi-relational adjacency matrix, thereby allowing the model to modulate edge strengths based on external conditions [67]. STHGCN and ISGCN follow a similar philosophy but introduce dynamic or interactive graph construction, in which external factors influence node embeddings and edge weights over time. These approaches enable external context to directly shape spatial dependency modeling rather than being treated as auxiliary features [71], [79].

More recent models, including MGC-DMF, MFE-GNN, CGNN, SCA-GCN, and MuSrc, further extend multisource fusion to richer and more diverse data modalities, such as weather variables, traffic incidents, POIs, regional functional zones, and multi-modal sensors. These methods often employ attention-based or metric-learning-based fusion mechanisms to dynamically weight heterogeneous inputs in a task-adaptive manner [80], [86]–[89]. In parallel, DSTF and DA-RMN focus on diversified spatiotemporal feature extraction and denoising-based multigraph fusion, respectively, and can be viewed as general-purpose fusion backbones capable of accommodating external inputs when available [81], [84].

Despite these advances, multisource fusion remains far from solved. External data sources are typically heterogeneous in spatial resolution, sampling frequency, and reliability. Some modalities, such as event logs, are sparse and noisy, while others, such as weather fields, are dense but only weakly correlated with localized traffic flows. Naïve fusion strategies, including direct feature concatenation or uniform attention across modalities, can increase model complexity, introduce spurious correlations, and exacerbate overfitting, particularly when rare events are underrepresented in the training data. Moreover, many existing studies evaluate their methods

under limited experimental settings and provide only shallow ablation analyses, making it difficult to disentangle the true contribution of individual external factors from performance gains attributable to increased model capacity.

From a systematic review perspective, the multisource fusion category of Table 6 demonstrates that integrating exogenous information can enhance model robustness, particularly under non-recurrent and event-driven conditions. At the same time, the literature highlights open challenges related to modality selection, alignment, and regularization. Future research should move beyond simply adding more data toward more principled fusion strategies that quantify the marginal utility of each source and maintain generalization when external signals are missing, noisy, or distributionally shifted.

4) Long-Term Periodicity: Multi-Scale Temporal Modeling

Traffic flow exhibits pronounced temporal regularities at multiple scales, most notably daily and weekly periodic patterns such as recurrent morning and evening peak hours on weekdays and distinct travel behaviors during weekends. Effectively capturing these periodic signals is critical for accurate long-term traffic forecasting. At the same time, long-horizon prediction (e.g., beyond 60 minutes) remains challenging due to error accumulation in autoregressive forecasting and increasing uncertainty about distant future states. Consequently, spatiotemporal traffic models must jointly capture long-range temporal dependencies and periodic structure while mitigating the propagation of prediction errors.

The models grouped under the periodic dependency category in Table 6 explicitly encode such regularities within GNN-based architectures. IGAGCN combines information-geometry-based graph convolutions with attention mechanisms to enhance sensitivity to long-range temporal dependencies and periodic fluctuations in traffic series [60]. ASTFGCN further advances this direction by decomposing historical observations into multiple temporal components, such as recent, daily, and weekly patterns, using adaptive temporal filters [65]. Separate time-aware graphs are constructed for each component and subsequently fused, enabling explicit separation of short-term dynamics from longer-term periodic signals. This structured decomposition has been shown to improve forecasting accuracy on datasets with strong and stable periodicity.

Other approaches capture periodic behavior through hybrid spatiotemporal designs rather than explicit decomposition. Loc-GCLSTM combines graph convolution with LSTM units and localized temporal windows to jointly model short-term continuity and recurring temporal trends [64]. STRGAT employs graph attention for spatial dependency modeling alongside temporal self-attention to selectively emphasize historically informative time steps, allowing the model to focus on periodic intervals relevant to the prediction task [70]. TFSG further introduces a self-supervised guidance signal into graph learning, encouraging better alignment between periodic components and auxiliary temporal factors such as time-of-day or day-of-week embeddings [85]. Similarly, although PSPJSTGCN does not explicitly decompose temporal patterns, its joint use of predefined and self-learned graphs over extended temporal horizons enables it to capture long-term temporal correlations and periodic traffic behaviors [78].

Beyond explicit periodic modeling, substantial progress in capturing long-range temporal dependencies has been achieved through dilated convolutional architectures inspired by WaveNet. Models such as Graph WaveNet and ADGCN employ stacked dilated causal convolutions to exponentially expand the temporal receptive field while preserving causality, enabling direct modeling of relationships between distant time points from relatively short input sequences [10], [62]. These studies demonstrate that dilated temporal convolutions can effectively capture long-range temporal dependencies and achieve competitive performance in long-horizon traffic forecasting tasks, including 60-minute-ahead prediction [10], [62].

Attention-based encoder-decoder architectures provide another effective mechanism for modeling long-term dependencies. GMAN, for instance, leverages multi-head self-attention to directly connect all time steps in the input and output sequences, allowing the entire prediction horizon to be generated in a single forward pass [91]. This non-autoregressive design substantially reduces the cumulative error commonly observed in autoregressive forecasting frameworks. Periodicity-aware variants of attention-based models often incorporate temporal positional encodings, such as hour-of-day or day-of-week embeddings, enabling the attention mechanism to distinguish between different temporal contexts.

To balance short-term responsiveness and long-term stability, several recent works adopt hybrid temporal modeling strategies. In these architectures, short-term fluctuations are captured using fine-grained recurrent or convolutional modules, while long-term trends and periodic patterns are modeled through explicit decomposition, trend estimation, or periodicity-aware attention components. Periodicity-focused GNNs such as IGAGCN, ASTFGCN, Loc-GCLSTM, STRGAT, and TFSG can be viewed as instances of this broader paradigm, each making different trade-offs between architectural complexity, interpretability, and forecasting horizon.

Overall, the integration of periodic priors with expanded temporal receptive fields has proven effective for enhancing long-range traffic forecasting performance. The models in the periodic dependency block of Table 6 illustrate how explicit temporal decomposition, dilated convolutions, and attention mechanisms can be combined to exploit daily and weekly

regularities while still accommodating irregular disturbances in real-world traffic systems.

B. Transformer-Based Paradigms: Attention Beyond Grids

Transformer-based architectures have emerged as a powerful alternative to recurrent and convolutional models for traffic flow prediction by leveraging self-attention mechanisms to directly model long-range dependencies [11]. Unlike RNNs and CNNs, which rely on sequential processing or localized receptive fields, transformers enable global interactions across distant time steps and spatial locations, alleviating gradient vanishing and limited temporal context. This capability is particularly advantageous for traffic forecasting, where congestion propagation and demand patterns often exhibit long-range temporal correlations and non-local spatial effects [7], [92].

However, standard transformer formulations are agnostic to the physical structure of the road network and spatial constraints. As a result, recent transformer-based traffic forecasting models differ substantially in how attention is structured and how spatial topology is encoded. To provide a systematic overview, Table 8 summarizes representative transformer-based models in chronological order. This chronological ordering highlights the rapid evolution of transformer-based traffic forecasting models since 2022 and reflects a growing focus on richer spatiotemporal interaction modeling.

Building on this overview, we analyze transformer-based traffic flow prediction models along two orthogonal design dimensions: (i) the formulation of attention mechanisms and (ii) the strategy used to encode spatial topology. These dimensions jointly define the core design space of transformer-based approaches, governing how long-range temporal dependencies are captured and how road network structure is incorporated into attention-based reasoning.

This design-space evolution is illustrated in Fig.7, which provides a conceptual timeline of attention mechanisms and spatial encoding strategies in transformer-based traffic forecasting. As shown in the figure, early models adopt decoupled and structure-agnostic attention with static adjacency matrices, while subsequent approaches introduce graph-guided attention, joint spatiotemporal cross-attention, and multi-scale hierarchical designs. More recent models further extend this paradigm through global similarity-based attention and dynamic graph representations, enabling globally expressive spatiotemporal interactions.

Importantly, the 2025 stage in Fig.7 does not correspond to a single dominant architecture but instead represents an emerging foundation-level transformer paradigm. Recent models proposed during this period increasingly emphasize robustness, uncertainty awareness, and transferability across heterogeneous traffic networks through adaptive attention mechanisms and uncertainty-aware graph encoding strategies. This shift reflects a transition from task-specific optimisation to more generalisable, deployment-ready transformer-based traffic forecasting models.

1) Heterogeneous Attention: Spatiotemporal Dependency Modeling

Transformer-based models have gained popularity in traffic flow prediction largely because self-attention mechanisms can directly capture long-range temporal dependencies without relying on step-by-step recurrence. By allowing direct interactions between distant time steps, attention mechanisms reduce issues such as gradient vanishing and limited temporal context that commonly affect recurrent and convolutional models. An early example is GMAN, which introduces a graph-based multi-attention encoder–decoder framework that separately learns spatial and temporal representations and combines them using attention to generate multi-step forecasts [91]. This work demonstrates that attention mechanisms can flexibly model both temporal evolution and spatial influence patterns within traffic networks.

Subsequent studies extended this idea by introducing heterogeneous attention mechanisms, in which different attention components are designed to capture complementary aspects of spatiotemporal dependencies. As summarized in Table 9, one group of models focuses mainly on temporal attention. These methods emphasize long-range temporal correlations and multi-scale trends, while spatial relationships are handled using relatively simple static or adaptive graph structures. Examples include TE-TCN, GSTA, STFGCN, MC-STTM, and DDGformer. Such designs are computationally efficient and scalable but may struggle to represent complex spatial interactions in highly irregular road networks [93], [94], [96]–[98].

In contrast, another group of models jointly models spatial and temporal dependencies within a unified attention framework. In these approaches, attention is computed over both nodes and time steps, allowing the model to learn how traffic conditions at one location and time influence conditions at other locations and times. Models such as STGAGRTN and LOGO use spatiotemporal gated attention to capture both local and global propagation patterns [95], [101]. STGAFormer further strengthens this integration by tightly coupling spatial and temporal attention blocks via gate-fusion mechanisms [99]. More recent models, including STPGCN, iTransMamba, MSGFormer, and HLFAN, extend this idea by introducing scale-aware or frequency-aware attention heads, where different heads specialize in short-term fluctuations, long-term trends, or global contextual patterns [90], [100], [102], [103].

From a critical perspective, heterogeneous attention improves modeling flexibility and has led to strong performance on benchmark datasets such as PeMS [28], [29] and METR-LA [30]. However, this flexibility also increases model com-

TABLE 8. Model Categorization Overview (Part II: Transformer-Based)

Category	Model	Citation	Time
Transformer-Based	GSTA	[93]	2022
	STPGCN	[90]	2022
	TE-TCN	[94]	2023
	STGAGRTN	[95]	2023
	DDGformer	[96]	2024
	STFGCN	[97]	2024
	MC-STTM	[98]	2024
	STGAFormer	[99]	2024
	MSGFormer	[100]	2024
	LOGO	[101]	2025
	iTransMamba	[102]	2025
	HLFAN	[103]	2025

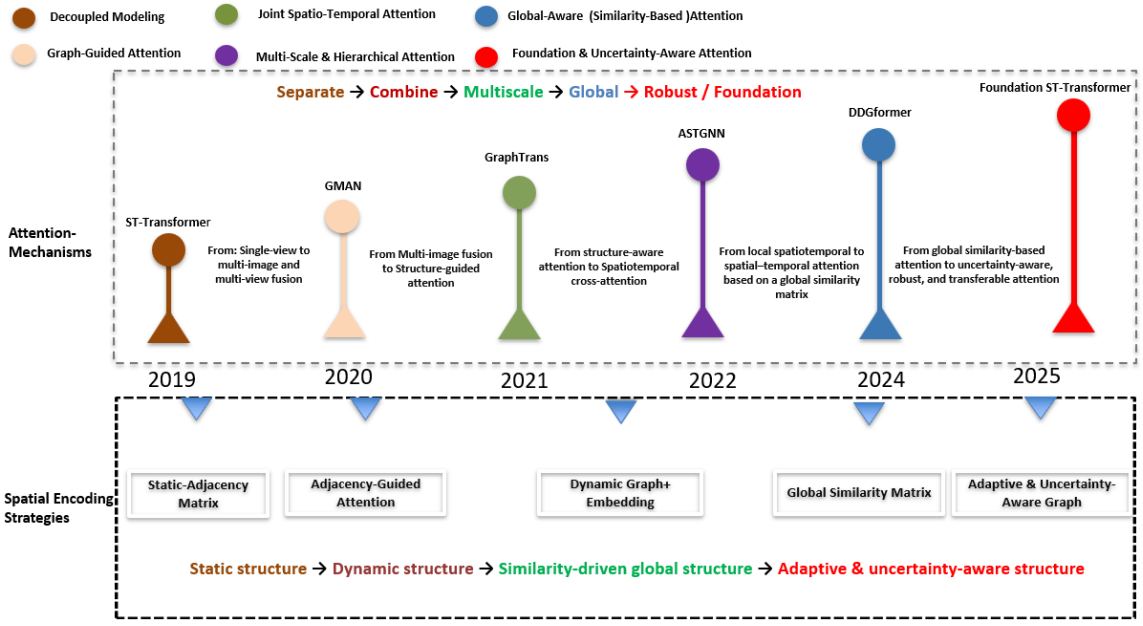

FIGURE 7. Progression of attention mechanisms and spatial encoding strategies in transformer-based traffic forecasting models, highlighting the shift from static and decoupled designs to adaptive, multiscale, and globally expressive formulations.

plexity and reduces interpretability. Many studies introduce multiple specialized attention modules but provide limited analysis of how each component contributes to performance or how robust the model is under distribution shifts, such as traffic incidents or structural network changes.

Overall, heterogeneous attention provides a useful framework for understanding transformer-based traffic models. Temporal-attention-focused methods prioritize sequence modeling, joint spatiotemporal attention aims to directly capture traffic propagation, and scale-specific attention mechanisms attempt to separate different types of variability. The attention classification presented in Table 9 clarifies these design choices and highlights the trade-offs between modeling expressiveness, interpretability, and computational cost.

2) Adaptive Topology Encoding: Dynamic Graph Representation

While attention mechanisms control how information is aggregated across time steps and nodes, spatial encoding strategies determine how road network structure is incorporated into the model. Early attention-based traffic models typically used either fully connected attention across all sensors or fixed adjacency matrices inherited from graph neural networks. Both approaches have limitations: fully connected attention is computationally expensive and may learn spurious correlations, while fixed adjacency matrices cannot adapt to changing traffic conditions.

The models in Table 9 illustrate three main strategies for topology encoding in transformer-based traffic forecasting:

static adjacency, adaptive adjacency, and distance-aware or globally constrained attention. Models such as TE-TCN and the temporal component of DDGformer rely on a fixed adjacency matrix derived from geographical distance or predefined sensor connections, which they use as a mask or bias in the attention mechanism [94], [96]. STGAFormer similarly combines static physical adjacency with learned attention weights, which preserves physical interpretability but limits adaptability under dynamic traffic patterns [99].

Adaptive adjacency approaches, used in models such as GSTA, STFGCN, MC-STTM, DDGformer, and more recent architectures like iTransMamba, MSGFormer, and HLFAN, learn additional graph structure directly from data [93], [96]–[98], [100], [102], [103]. In these models, node embeddings or latent relation matrices are optimized jointly with attention, allowing spatial relationships to evolve over time and capture functional similarities between road segments. While this improves flexibility and can reveal non-obvious interactions, it also raises concerns about interpretability and robustness, since learned connections are not explicitly constrained to reflect physical road topology.

A third class of approaches introduces distance-aware or globally constrained attention, as in STGAGRTN, LOGO, and STPGCN [90], [95], [101]. These models incorporate pairwise distances, travel times, or reachability information as continuous edge features that influence attention scores. This design allows global attention while still enforcing spatial decay principles, for example, by reducing the influence of very distant nodes unless supported by strong evidence. Distance-aware attention, therefore, offers a compromise between rigid fixed graphs and unconstrained global attention.

Despite these advances, adaptive topology encoding remains an open challenge. Most studies rely on qualitative visualizations of attention patterns and do not systematically compare learned graphs with known physical or functional connectivity. In addition, few works evaluate the stability of learned spatial structures under sensor failures, network changes, or deployment in new cities. Finally, adaptive and distance-aware encoding strategies introduce additional parameters, increasing the risk of overfitting when training data are limited.

In summary, heterogeneous attention mechanisms and adaptive topology encoding strategies jointly define modern transformer-based traffic forecasting models. Attention determines how information flows across time and space, while topology encoding constrains which interactions are plausible. Insights from Table 9 and Fig.7 suggest that future research should focus on more principled topology learning approaches that incorporate physical constraints, uncertainty modeling, and cross-city generalization, rather than relying solely on increasingly complex attention mechanisms.

C. Hybrid Frameworks: Integrating GNNs and Transformers

The integration of graph-based and temporal modeling components has become increasingly prominent in traffic forecasting, motivated by their complementary abilities to represent spatial interactions and temporal dynamics. As summarized in Table 10, existing hybrid approaches can be broadly categorized into three fusion paradigms: cascaded, parallel, and iterative/integrated architectures. These paradigms differ primarily in how spatial and temporal representations are processed and combined within the forecasting framework. Early approaches combined graph convolutions with sequence models in a pipeline (cascade fusion), while more recent methods exploit parallel fusion of spatial and temporal branches or fully integrated architectures that interleave GNN and transformer components. Although this evolution has generally improved predictive accuracy on long-horizon benchmarks, it has also increased model complexity and reduced interpretability.

1) Cascaded Fusion: Sequential Feature Abstraction

Cascaded fusion architectures model spatial and temporal dependencies as two sequential processing stages within a unified prediction pipeline. In a typical design, a graph-based module (e.g., GCN or GAT) is first applied at each time step to encode spatial correlations over the road network, and the resulting node-level representations are then fed into a temporal model such as a transformer, GRU, BiLSTM, or temporal convolutional network. The reverse ordering is also adopted in some studies, where temporal patterns are extracted first and subsequently refined through spatial graph modeling.

The core rationale of cascaded fusion is to decouple spatial and temporal learning objectives: the spatial module focuses on capturing network topology and inter-sensor dependencies, while the temporal module specializes in modeling sequential dynamics and long-range temporal patterns. Conceptually, this paradigm extends earlier GNN–RNN frameworks by combining graph-based spatial encoding with increasingly expressive temporal modeling components, including recurrent, attention-based, and dilated convolutional mechanisms.

Representative examples of cascaded fusion architectures, as summarized in Table 11, include STSSN, PDG2Seq, PKET-GCN, CNN-BiLSTM, SASTDGCN, MDTP, and MTDLTFP [105], [106], [109]–[111], [114], [122]. These models typically exhibit clear architectural modularity and relatively stable training behavior, which facilitates implementation and interpretation in terms of spatial versus temporal feature contributions. However, the strict stage-wise separation inherent in cascaded fusion may limit the exploitation of fine-grained interactions between evolving spatial patterns and temporal context, motivating the development of more integrated fusion mechanisms.

TABLE 9. Classification of attention mechanisms and spatial encoding strategies in transformer-based traffic forecasting models.

Model	Attention Type		Spatial Encoding Strategy		
	Temporal	Spatio-Temporal	Static	Adaptive	Distance/Global
GSTA	✓			✓	
STPGCN	✓		✓		✓
TE-TCN		✓	✓		
STGAGRTN		✓			✓
DDGformer	✓		✓	✓	
STFGCN	✓			✓	
MC-STTM	✓			✓	
STGAFormer		✓	✓	✓	
LOGO		✓			✓
iTransMamba	✓	✓		✓	
MSGFormer	✓	✓		✓	
HLFAN	✓	✓		✓	

2) Parallel Processing: Joint Spatiotemporal Learning

Parallel fusion architectures are designed to address the limitations of strictly stage-wise pipelines by modeling spatial and temporal dependencies simultaneously. Instead of sequentially feeding one module into another, these approaches process spatial and temporal information in parallel branches, followed by an explicit fusion stage. Typically, a spatial sub-network (e.g., a GNN or graph filtering module) and a temporal sub-network (e.g., a transformer, temporal CNN, or recurrent model) operate on shared inputs, and their latent representations are subsequently merged. This design enables spatial and temporal representations to be learned simultaneously and combined through an explicit fusion mechanism.

Formally, let $Z^{(S)} = f_{\text{spatial}}(X, A)$ denote spatial representations produced by a graph-based encoder given traffic observations X and adjacency matrix A, and let $Z^{(T)} = f_{\text{temporal}}(X)$ denote temporal representations learned by a sequence model. A fusion function $g(\cdot)$ then combines these representations to produce the final prediction $\hat{Y} = g(Z^{(S)}, Z^{(T)})$. In practice, the fusion operator ranges from simple feature concatenation followed by fully connected layers to more advanced cross-attention or gating mechanisms, in which one branch adaptively modulates the other.

Representative parallel fusion models, summarized in Table 11, include ST-InNet, Gboot, MFGCN, POI-MetaBlock, ST-SSL, STA-GCNT, GE-STT, LEISN, MIFGCN, and several recent hierarchical or multi-branch architectures [26], [107], [112], [116], [118]–[121], [124]. Many of these architectures incorporate hierarchical or multi-scale mechanisms to combine features across different temporal resolutions and spatial granularities. Such designs provide greater flexibility in modeling coupled spatiotemporal dependencies, although they typically introduce additional architectural and tuning complexity.

3) Iterative and Integrated Architectures: Interweaving GNN and Transformer

Iterative and integrated architectures provide a more tightly coupled alternative to cascaded and parallel fusion strategies. Rather than relying on a single-pass pipeline or loosely coupled branches, these models interleave spatial and temporal processing across multiple layers or iterations, enabling continuous mutual refinement. A common design pattern involves stacking spatiotemporal blocks, each consisting of a spatial sub-layer (e.g., GNN, graph filter, or spatial attention) followed by a temporal sub-layer (e.g., transformer block, GRU, or temporal convolution). This structure is conceptually analogous to standard transformer architectures, where self-attention and feed-forward layers are repeatedly stacked, with graph-based operations replacing or augmenting conventional feed-forward components.

Representative examples include AutoMSNet, HSTGNN, ODENet-GNN, MGODENN, AGCN-DP, MSTFF, D-STANet, BSFN, LSTM-GAT, ATSDHNN, MSTAFN, MSTAT-GRU, FTL-CCTP, STU, ALTCN, and related adaptive or ODE-inspired models, which are categorized as "Adaptive/Iterative" in Table 11 [108], [125]–[138]. These models typically perform multiple rounds of interaction, in which updated temporal states influence subsequent spatial aggregation, and vice versa, often mediated by attention or gating mechanisms. Such iterative fusion enables spatial and temporal representations to be progressively refined across successive processing stages, providing a more tightly coupled mechanism for modeling evolving spatiotemporal interactions.

Compared with cascaded and parallel strategies, integrated architectures provide tighter interaction between spatial and temporal representations but generally involve greater architectural complexity. This progression from sequential to parallel and iterative fusion illustrates how hybrid traffic

TABLE 10. Model Categorization Overview (Part III: Hybrid Architectures)

Category	Model	Citation	Time
Hybrid Architectures	STGODE	[104]	2021
	MDTP	[105]	2021
	STSSN	[106]	2022
	ST-InNet	[107]	2022
	LSTM_BiLSTM	[27]	2022
	AutoMSNet	[108]	2022
	PKET-GCN	[109]	2023
	CNN-BiLSTM	[110]	2023
	SASTDGCN	[111]	2023
	POI-MetaBlock	[26]	2023
	SDGDN	[72]	2023
	ST-SSL	[112]	2023
	GMHANN	[113]	2023
	MTDLTFP	[114]	2023
	RTM-GCN	[115]	2024
	Gboot	[116]	2024
	GECRAN	[117]	2024
	MFGCN	[118]	2024
	STA-GCNT	[119]	2024
	LEISN	[120]	2024
	MIFGCN	[121]	2024
	PDG2Seq	[122]	2025
	DSTGNDE	[123]	2025
	GE-STT	[124]	2025
	HSTGNN	[125]	2025
	ODENet-GNN	[126]	2025
	MGODENN	[127]	2025
	AGCN-DP	[128]	2025
	MSTFF	[129]	2025
	D-STANet	[130]	2025
	BSFN	[131]	2025
	LSTM-GAT	[132]	2025
	ATSDHNN	[133]	2025
	MSTAFN	[134]	2025
	MSTAT-GRU	[135]	2025
	FTL-CCTP	[136]	2025
	STU	[137]	2025
	ALTCN	[138]	2025

forecasting models have increasingly moved toward more tightly coupled spatiotemporal representation learning.

D. Generative and Uncertainty-Aware Models

While deep learning has substantially advanced traffic flow prediction, existing spatiotemporal models still face fundamental limitations that constrain their applicability in complex, dynamic transportation environments. Specifically, three interrelated challenges remain: (i) non-Euclidean spatial dependencies inherent in real-world road networks; (ii) long-range temporal dependencies caused by delayed traffic propagation and multiscale periodic patterns; and (iii) inherent uncertainty due to stochastic disruptions, sensor noise, and open-world anomalies [53], [145]–[147].

Traditional deterministic models typically produce single-point forecasts, failing to capture the full distributional nature of traffic dynamics. This shortcoming is particularly problematic under conditions of data sparsity, sensor malfunction, or rare events, where robustness and reliability are critical. To overcome these limitations, a new generation of gen-

TABLE 11. Model Classification Based on Fusion Architecture

Model	Spatio-Temp. Division	Encoder–Decoder Fusion	Hier./Parallel	Adaptive/Iterative	Ref.
STGODE	✓				[104]
MDTP		✓			[105]
STSSN		✓			[106]
ST-InNet			✓		[107]
LSTM_BiLSTM		✓			[27]
AutoMSNet				✓	[108]
PKET-GCN		✓			[109]
CNN-BiLSTM		✓			[110]
SASTDGCN		✓			[111]
POI-MetaBlock			✓		[26]
SDGDN	✓				[72]
ST-SSL			✓		[112]
RTM-GCN	✓				[115]
Gboot			✓		[116]
GECRAN	✓				[117]
GMHANN	✓				[113]
MFGCN			✓		[118]
MTDLTFP		✓			[114]
STA-GCNT			✓		[119]
LEISN	✓				[120]
MIFGCN			✓		[121]
PDG2Seq		✓			[122]
DSTGNDE	✓				[123]
GE-STT	✓				[124]
HSTGNN			✓	✓	[125]
ODENet-GNN			✓	✓	[126]
MGODENN			✓	✓	[127]
AGCN-DP				✓	[128]
MSTFF		✓	✓	✓	[129]
D-STANet			✓	✓	[130]
BSFN		✓	✓		[131]
LSTM-GAT	✓		✓		[132]
ATSDHNN			✓	✓	[133]
MSTAFN		✓	✓	✓	[134]
MSTAT-GRU			✓	✓	[135]
FTL-CCTP			✓	✓	[136]
STU				✓	[137]
ALTCN				✓	[138]

erative, uncertainty-aware models has emerged, employing probabilistic reasoning and stochastic generation to better represent spatiotemporal evolution under uncertainty.

As summarized in Table 12, recent approaches integrate deep generative architectures, such as variational autoencoders, Bayesian recurrent networks, normalizing flows, and diffusion-based frameworks to capture multimodal distributions over future traffic states. These approaches represent a shift from deterministic pattern fitting toward probabilistic distribution modeling, offering improved flexibility, robustness, and interpretability for diverse traffic forecasting scenarios.

Recent studies have further advanced this line of research by introducing diffusion-based and probabilistic deep learning frameworks for traffic forecasting. Graph latent diffusion models align spatial and temporal dependencies in

TABLE 12. Model Categorization Overview (Part IV: Generative and Uncertainty Models)

Category	Model	Citation	Time
Generative and Uncertainty Models	MRN	[139]	2023
	STGNF	[18]	2024
	TADGCN	[20]	2024
	BEDMA	[19]	2024
	Graph Latent Diffusion	[140]	2024
	Disentangled Diffusion	[141]	2025
	MP-DGGRT	[142]	2025
	UQGNN	[143]	2025
	GenSTNet	[144]	2025

latent graph spaces, enabling the generation of multiple plausible traffic dynamics under complex network interactions [140]. Disentangled diffusion models further decompose spatiotemporal traffic patterns to capture multi-scale uncertainty in probabilistic forecasting [141]. Similarly, multi-period diffusion generative graph recurrent transformer networks combine diffusion-based generation with recurrent and transformer modules to model long-horizon uncertainty propagation in vehicular traffic flow prediction [142]. Beyond diffusion models, uncertainty-quantified GNN frameworks explicitly estimate predictive uncertainty and provide distribution-aware outputs for multivariate spatiotemporal forecasting [143]. These recent works indicate that traffic forecasting is gradually shifting from single-point estimation toward distributional, scenario-based, and risk-aware prediction.

Beyond this high-level categorization, we also provide a detailed mapping of each model to the specific challenges it addresses. Table 13 evaluates models in terms of spatial uncertainty, temporal uncertainty, and robustness in open-world conditions, highlighting their respective design focuses and deployment strengths.

1) Non-Euclidean Spatial Dependencies: Modeling Road Network Topology and Dynamic Interactions

Urban traffic networks are intrinsically non-Euclidean, characterized by irregular, asymmetric, and dynamically evolving topologies. Nodes representing sensors, intersections, or road segments are connected via complex spatial relationships that cannot be captured by conventional grid-based convolutional operations. While GNNs offer a significant improvement by modeling structural dependencies through adjacency matrices, most existing GNN-based traffic prediction models assume static or manually defined graph structures, limiting their adaptability to time-varying traffic conditions, network perturbations, and unobserved spatial interactions.

To address these limitations, recent generative modeling approaches introduce probabilistic representations of road network topology, enabling dynamic and context-aware spatial reasoning. Variational Graph Autoencoders (VGAEs) [148] and Graph Diffusion Models (GDMs) learn latent stochastic embeddings of node connectivity, allowing the generation of plausible topological configurations under disruptions such as traffic accidents, road closures, or sensor failures. A representative example is the Spatio-Temporal Graph Normalizing Flow (STGNF) [18], which leverages invertible transformations to map observed traffic states into a continuous latent space. The adjacency matrix is treated as a stochastic variable, facilitating uncertainty-aware learning over both explicit and latent spatial dependencies.

Additionally, diffusion-based generative frameworks such as graph latent diffusion models [140], extend this paradigm by iteratively denoising perturbed traffic representations to recover the underlying flow distributions. These methods offer increased robustness to sensor noise, missing data, and topology shifts, making them well-suited for real-world deployments. Collectively, these generative models mark a shift from structure-constrained spatial encoding toward distributional graph learning, laying the groundwork for dynamic, flexible, and uncertainty-aware traffic forecasting across diverse urban scenarios.

2) Long-Range Temporal Dependencies: Cross-Period Flow Propagation and Event-Delay Effects

Urban traffic systems exhibit intricate multi-scale temporal dynamics, encompassing short-term fluctuations, medium-term cyclic variations, and long-term evolutionary trends. Effectively modeling these long-range temporal dependencies is critical for accurate multi-horizon forecasting, particularly under non-stationary and high-variance traffic conditions. However, this remains a persistent challenge due to cumulative prediction errors, asynchronous propagation of events, and nonlinear cross-period interactions across spatially distributed regions.

Traditional sequence modeling approaches including RNNs and transformer-based architectures have shown limited effectiveness in capturing long-distance dependencies within highly stochastic traffic time series. Recurrent models suffer from vanishing gradients and limited memory capacity, while standard self-attention mechanisms entail high computational

TABLE 13. Challenge-Based Categorization of Generative Models

Model	Spatial Uncertainty Modeling	Temporal Uncertainty Modeling	Open-World Robustness
STGNF [18]	✓		✓
TADGCN [20]	✓		✓
MRN [139]		✓	
BEDMA [19]		✓	✓
GenSTNet [144]	✓	✓	✓
Graph Latent Diffusion [140]	✓	✓	
Disentangled Diffusion [141]		✓	✓
MP-DGGRT [142]	✓	✓	✓
UQGNN [143]	✓	✓	✓

complexity and limited scalability. As a result, these deterministic architectures often struggle to generalize under dynamic conditions with delayed causal effects and cross-temporal noise.

Recent developments in generative temporal modeling introduce stochastic latent variables to explicitly capture temporal uncertainty and multi-modal future trajectories. Temporal Variational Autoencoders (TVAEs), for instance, model future evolution as a distribution over latent sequences conditioned on historical context, thereby accounting for both epistemic and aleatoric uncertainty. Similarly, Diffusion Probabilistic Models (DPMs) learn a stochastic generative process that iteratively denoises random temporal states, enabling the simulation of diverse future flow realizations and mitigating long-horizon error accumulation.

A representative example is the Bayesian Encoder–Decoder Model with Attention (BEDMA) [19], which incorporates a variational inference mechanism into a deep spatiotemporal forecasting framework. By learning a posterior distribution over latent states conditioned on past traffic observations and exogenous variables such as weather, holidays, or social events BEDMA enables multi-modal prediction with calibrated uncertainty estimates. In parallel, diffusion-based generative models trained via score-based learning have demonstrated state-of-the-art performance in long-term forecasting tasks by modeling the uncertainty dynamics as a continuous stochastic process.

Beyond predictive accuracy, these generative frameworks enable distribution-aware decision support by producing full predictive distributions rather than single-point estimates. Such probabilistic outputs allow traffic management systems to quantify forecast confidence and implement risk-sensitive strategies for adaptive signal control, congestion mitigation, and autonomous decision-making. These advances collectively mark a paradigm shift from deterministic sequence modeling toward probabilistic temporal reasoning, bridging the gap between accuracy and reliability in real-world intelligent transportation scenarios.

3) Uncertainty in Open-World Scenarios: Extreme Events, Data Sparsity, and Noise Robustness

Real-world intelligent transportation systems operate under open-world conditions, where traffic environments are inherently unpredictable, partially observable, and continuously evolving. Unlike controlled experimental settings, operational road networks must handle abrupt disruptions such as accidents, construction, extreme weather, and public events that introduce complex spatiotemporal anomalies and distributional shifts. These irregularities give rise to both epistemic uncertainty, resulting from incomplete knowledge of unobserved variables, and aleatoric uncertainty, arising from intrinsic stochasticity and sensor-level noise.

Traditional deep learning models, typically trained under closed-world assumptions, are prone to overfitting and generalize poorly under out-of-distribution scenarios. Their performance deteriorates significantly when facing missing trajectories, sensor failures, or unanticipated external signals. To address this challenge, generative modeling frameworks explicitly learn the joint distribution of traffic states and uncertainty factors, thereby enhancing robustness and adaptive capacity.

Bayesian deep networks and variational inference-based architectures introduce stochastic latent variables to model posterior uncertainty, enabling calibrated confidence estimation for downstream predictions. Likewise, diffusion-based generative models learn denoising transitions from corrupted traffic states, providing resilient reconstructions under sparse or noisy input conditions.

An illustrative example is the Traffic Anomaly-aware Dynamic Graph Convolutional Network (TADGCN) [20], which integrates probabilistic spatial representations with adaptive attention mechanisms. TADGCN dynamically reweights graph edges to emphasize reliable connections while suppressing anomalous signals, enhancing robustness under perturbations. Similarly, Variational Bayesian Encoder–Decoder models [19] generate diverse future trajectories conditioned on uncertainty-aware inputs, improving

generalization in the presence of extreme events or data scarcity.

Beyond prediction performance, explicit uncertainty quantification supports decision reliability in real-time transportation operations. Generative models producing full predictive distributions enable downstream modules such as traffic control centers and autonomous systems to reason about confidence intervals, propagate risk estimates, and implement safety-critical strategies. For instance, probabilistic forecasts can support robust route planning, adaptive signal control under uncertainty, and risk-sensitive traffic regulation.

Overall, addressing uncertainty in open-world traffic scenarios requires a synergistic integration of probabilistic reasoning, stochastic generation, and robust learning. These components collectively form the foundation for the next generation of resilient, trustworthy, and uncertainty-aware traffic forecasting systems operating under real-world constraints. To synthesize the technical foundations and practical roles of generative models, Table 14 summarizes the representative architectures, highlighting their core generative mechanisms, targeted challenges, and relevant application scenarios.

In summary, generative and uncertainty-aware models represent a fundamental departure from traditional point forecasting paradigms in traffic flow prediction. By modeling predictive distributions rather than single estimates, these approaches provide richer information for downstream traffic management and control, particularly under rare events, non-recurrent congestion, and distribution shifts. This emerging direction is expected to play a central role in next-generation intelligent transportation systems, motivating further research into scalable, interpretable, and physically consistent generative traffic models.

V. Critical Comparison of Model Architectures for Traffic Flow Prediction

Building on the architecture-specific analysis presented in the preceding sections, this section compares the major model families from an operational perspective. Rather than reiterating their individual architectural characteristics, the comparison focuses on five practical criteria: computational complexity, scalability, robustness to missing data, interpretability, and recommended use cases, as summarized in Table 15.

Computational Complexity and Scalability. GNN-based models generally exhibit moderate computational cost, with message-passing complexity depending on graph size and connectivity, making them suitable for large-scale sensor networks [9], [30]. Transformer-based models typically incur quadratic complexity with respect to sequence length under standard self-attention, which can limit scalability unless efficient attention mechanisms are employed [11], [94], [98]. Hybrid models combine multiple spatial and temporal components and therefore often introduce additional computational and memory overhead [99], [125]. Generative models may further increase inference cost because of distributional estimation or iterative sampling procedures [18], [141], [142].

Robustness to Missing Data. GNN-based and hybrid models can be sensitive to missing or noisy observations, particularly when spatial dependencies are learned from incomplete sensor inputs [9], [52]. Transformer-based models can accommodate missing observations through masking mechanisms, although their robustness depends on the training strategy and data characteristics [92]. Generative models provide additional capabilities for modeling incomplete data by learning underlying data distributions and, in some formulations, reconstructing missing traffic states [19].

Recent work has also explored lightweight recurrent architectures specifically designed for traffic forecasting under incomplete sensor observations, combining enhanced recurrent units with adaptive imputation strategies to improve robustness and computational efficiency [149].

Interpretability. GNN-based models provide structural interpretability through graph representations aligned with road-network topology, although their learned representations may remain difficult to explain [61]. Transformer-based models provide some insight through attention patterns; however, attention weights should not necessarily be interpreted as causal explanations [92]. Hybrid and generative architectures often introduce additional model components and latent representations, which can further complicate interpretation [18], [95].

As summarized in Table 15, no single architectural paradigm consistently dominates across all evaluation dimensions. GNN-based models provide effective network-level spatial modeling but can be sensitive to incomplete observations; transformer-based models capture long-range dependencies at increased computational cost; hybrid architectures integrate complementary spatial and temporal mechanisms at the expense of additional model complexity; and generative approaches provide flexible probabilistic modeling capabilities but often incur greater computational demands [6], [7], [18], [19], [141]. These trade-offs indicate that architecture selection should be guided by application-specific data, forecasting objectives, and deployment requirements. Building on this comparative assessment, Section VI examines the unresolved challenges and emerging research directions for next-generation traffic forecasting.

VI. Toward Next-Generation Traffic Forecasting: Open Challenges and Emerging Paradigms

Building on the limitations and trade-offs identified in the preceding analysis, this section synthesizes the key open challenges and emerging research directions in deep learning-based traffic flow prediction. The discussion focuses on heterogeneous data integration, adaptive spatiotemporal

TABLE 14. Representative Generative Models: Mechanism, Challenge Addressed, and Application

Model	Generative Mechanism	Challenge Addressed	Application Scenario
STGNF [18]	Normalizing Flow + Stochastic Adjacency	Spatial uncertainty modeling	Robust flow prediction under topology shifts
TADGCN [20]	Dynamic Probabilistic Graph + Attention	Anomaly-aware generalization	Traffic event robustness and fault tolerance
MRN [139]	Temporal Variational Encoder	Long-range temporal uncertainty	Multi-horizon trend prediction
BEDMA [19]	Variational Attention Encoder–Decoder	Temporal multi-modality + open-world exogenous inputs	Forecasting with weather/events/holiday priors
GenSTNet [144]	Diffusion + VAE Hybrid Generator	Full-spectrum uncertainty modeling	Distribution-aware real-time forecasting
Graph Latent Diffusion [140]	Latent graph diffusion	Joint spatial-temporal dependency uncertainty	Scenario-based traffic dynamics prediction
Disentangled Diffusion [141]	Disentangled denoising diffusion	Multi-scale probabilistic temporal uncertainty	Probabilistic spatiotemporal traffic forecasting
MP-DGGRT [142]	Diffusion + Graph Recurrent Transformer	Long-horizon uncertainty propagation	Vehicular traffic flow prediction
UQGNN [143]	Probabilistic GNN uncertainty quantification	Calibrated predictive uncertainty	Multivariate spatiotemporal forecasting under uncertainty

TABLE 15. Structured Comparison of Deep Learning Model Families for Traffic Flow Prediction Across Key Deployment Dimensions

Model Family	Complexity	Scalability	Robustness*	Interpretability	Typical Use Case	Examples
GNN	$\mathcal{O}(E)$ per graph-propagation layer; moderate	High; suitable for large road networks	Low–Medium	Medium	Real-time network-wide prediction	DCRNN, STGCN, GWN
Transformer	$\mathcal{O}(T^2)$ for standard temporal attention; high	Medium; constrained by sequence length and memory	Medium	Medium	Long-range and long-horizon forecasting	GSTA, DDGformer
Hybrid	Architecture-dependent; typically very high	Medium; computationally resource-intensive	Medium–High	Low	High-accuracy spatiotemporal forecasting	STGAFormer, HSTGNN
Generative	Architecture-dependent; high–very high	Low–Medium	High	Low	Probabilistic and uncertainty-aware forecasting	STGNF, BEDMA, GenSTNet

**Robustness* denotes the qualitative ability of a model family to maintain prediction performance under noisy or incomplete observations, traffic-pattern variations, and distribution changes. The Low–High ratings are comparative assessments synthesized from the reviewed literature rather than standardized quantitative scores.

modeling, uncertainty-aware forecasting, scalability, practical deployment, and reproducible evaluation [6]–[8], [12], [17].

A. Unified Modeling of Heterogeneous Traffic Data

The reviewed literature reveals that heterogeneous traffic data sources are often handled in a fragmented manner [6], [13]. Many existing models process fixed sensors, probe vehicle trajectories, and external contextual information separately, or combine them through straightforward feature concatenation [80], [108]. Such approaches largely overlook fundamental differences in temporal granularity, spatial coverage, data reliability, and missingness across modalities [121].

A key research direction is therefore the development of unified multimodal learning frameworks that explicitly account for these heterogeneities during data integration. Attention-guided fusion, reliability-aware weighting, and uncertainty-aware imputation provide promising mechanisms for dynam-

ically emphasizing informative sources while reducing the influence of noisy, incomplete, or asynchronous observations [63], [115]. Such integration strategies are particularly important for real-world traffic forecasting, where data availability and quality may vary substantially across sensing modalities and operating conditions [13].

B. Dynamic and Adaptive Graph Representation Learning

Building on the graph-based forecasting approaches reviewed earlier, a key unresolved challenge is to represent spatial relationships that evolve with traffic conditions rather than relying on fixed or weakly adaptive connectivity. Real-world interactions between road segments may change substantially under incidents, demand fluctuations, congestion propagation, and infrastructure modifications [52], [150].

Recent research increasingly explores dynamic and adaptive graph representations that allow spatial dependencies to evolve over time [74], [123]. Promising directions include event-driven and time-varying graph construction, joint learning of multiple connectivity views (e.g., physical, functional, and semantic), and explicit modeling of asynchronous spatiotemporal interactions [72], [82], [122].

Beyond improving adaptivity, future graph-learning methods should ensure that dynamically inferred relationships remain physically meaningful, stable under distribution shifts, interpretable, and transferable across different urban networks [109]. Addressing these requirements is essential for moving dynamic graph learning from dataset-specific optimization toward robust network-level traffic forecasting.

C. Adaptive Spatiotemporal Representation Learning

Beyond adapting spatial connectivity, a complementary challenge is to learn unified representations of traffic dynamics operating across heterogeneous temporal and spatial scales. Traffic states evolve through a combination of short-term fluctuations, recurrent periodic patterns, and delayed network-wide propagation, requiring models to capture dependencies with different durations and spatial extents [1], [8].

Traffic time series exhibit both long-term periodic patterns (e.g., daily cycles) and short-term irregular fluctuations driven by stochastic events [20], [115], [122]. A fixed representation may therefore be insufficient across different traffic regimes, since the relative importance of recent observations, periodic histories, and spatial context can vary with congestion level, time of day, and unexpected disturbances.

Existing temporal architectures offer complementary capabilities but also distinct trade-offs. Recurrent models such as LSTMs capture sequential dependencies but struggle with long horizons [27], [110]. Attention-based models improve long-range modeling but introduce scalability challenges due to quadratic complexity [11], [94]. Temporal convolutional networks offer efficient multi-scale modeling, though fixed receptive fields may limit flexibility in highly dynamic scenarios [62], [107]. Complementary to purely data-driven representation learning, physics-regularized approaches integrate established traffic-flow dynamics with learned models to improve the physical consistency of traffic state representations. For example, recent work combines Gaussian Processes with traffic-flow dynamics to regularize data-driven traffic modeling, illustrating the potential of incorporating domain knowledge into adaptive forecasting frameworks [151].

Future research should therefore move toward adaptive spatiotemporal representations that dynamically determine which temporal scales and spatial contexts are most informative for the prevailing traffic state. Such frameworks should jointly balance multi-scale dependency modeling, computational efficiency, interpretability, and robustness across heterogeneous traffic regimes rather than relying on a single fixed spatial or temporal representation.

D. Uncertainty-Aware and Probabilistic Traffic Forecasting

A major limitation of current research is the predominance of deterministic forecasting paradigms [7]. While effective in reducing average error, such models fail to quantify predictive uncertainty, which is critical for safety-critical applications [14], [17].

Future work should emphasize probabilistic and uncertainty-aware forecasting frameworks that model full predictive distributions. Generative models, normalizing-flow frameworks, Bayesian learning, and diffusion-based approaches offer promising capabilities for capturing variability, multimodality, and rare events [18], [19], [140]–[143]. These methods support confidence estimation, scenario generation, and robust decision-making under uncertainty [69].

A remaining challenge is the rigorous evaluation of predictive uncertainty. Future studies should complement conventional point-error metrics with measures of calibration, sharpness, predictive intervals, and distributional quality to determine whether estimated uncertainty is reliable for operational decision-making [6], [8].

E. Scalability and Real-Time Deployment

Despite strong performance on benchmark datasets, many advanced models incur high computational and memory costs [95], [98], [99]. These constraints hinder real-time deployment in large-scale traffic systems [94], [102].

Future research should prioritize computationally efficient and scalable architectures through strategies such as graph sparsification, efficient attention, and model compression [68], [79], [138]. Beyond theoretical complexity, evalu-

ation on realistic hardware and network scales is necessary to establish whether forecasting gains can be maintained under practical latency and resource constraints.

F. Integration with Traffic Control and Decision-Making Systems

Most traffic prediction models are evaluated independently using error-based metrics, with limited consideration of their downstream impact on traffic management systems [6], [17]. This disconnect diminishes the practical relevance of improvements in forecasting.

Bridging this gap requires closer integration between forecasting and downstream traffic-management tasks, including signal control, routing, and congestion management. Evaluation should therefore extend beyond prediction error to system-level outcomes such as delay, safety, and environmental impact. Joint optimization of prediction and control represents a particularly important direction for translating forecasting improvements into measurable operational benefits [1], [152].

G. Standardized Benchmarks and Reproducibility

The lack of standardized benchmarks and evaluation protocols remains a major challenge for reproducible research [6]–[8]. Differences in data preprocessing, input features, experimental settings, and evaluation procedures can hinder reliable comparison across studies.

Future research should promote open datasets, shared experimental pipelines, and unified evaluation protocols that assess not only predictive accuracy but also robustness and generalization. Such benchmarks should increasingly accommodate probabilistic and generative forecasting to support consistent evaluation under uncertainty and distribution shifts [18].

VII. Conclusion

This paper presented a systematic literature review of traffic flow prediction research, with particular emphasis on deep learning-driven approaches published between 2020 and September 2025. Following the PRISMA 2020 guidelines, the review examined traffic data characteristics, spatiotemporal representations, and major modeling paradigms, including graph neural networks, transformer-based architectures, hybrid spatiotemporal models, and emerging generative approaches.

The reviewed literature reveals a clear progression from conventional sequence-based forecasting toward increasingly integrated spatiotemporal architectures that exploit graph representation learning, attention mechanisms, adaptive feature fusion, and probabilistic modeling. The comparative analysis further indicates that no single model family consistently dominates across predictive accuracy, scalability, adaptability, interpretability, robustness, and deployment feasibility.

By synthesizing these developments within a structured taxonomy and critically examining their methodological and operational trade-offs, this review provides a consolidated reference for researchers and practitioners in intelligent transportation systems. The open challenges identified in Section VI provide a foundation for advancing traffic forecasting toward more reliable, generalizable, and practically deployable systems.

Acknowledgment

The authors would like to thank former PhD candidate, Haidi Fan, for her initial contributions to this research and for drafting an early version of the manuscript.

REFERENCES

[1] Y. He, P. Huang, W. Hong, Q. Luo, L. Li, and K.-L. Tsui, "In-depth insights into the application of recurrent neural networks (rnns) in traffic prediction: A comprehensive review," *Algorithms*, vol. 17, no. 9, p. 398, 2024.

[2] B. M. Williams and L. A. Hoel, "Modeling and forecasting vehicular traffic flow as a seasonal arima process," *Transportation Research Record*, vol. 1644, pp. 1–10, 2003.

[3] E. I. Vlahogianni, J. C. Golias, and M. G. Karlaftis, "Optimized and meta-optimized neural networks for short-term traffic flow prediction," *Transportation Research Part C*, vol. 13, no. 3, pp. 211–234, 2005.

[4] C.-H. Wu, J.-M. Ho, and D. T. Lee, "Travel-time prediction with support vector regression," *IEEE Transactions on Intelligent Transportation Systems*, vol. 5, no. 4, pp. 276–281, 2004.

[5] G. A. Davis and N. L. Nihan, "Nonparametric regression and short-term freeway traffic forecasting," *Journal of Transportation Engineering*, vol. 117, no. 2, pp. 178–188, 1991.

[6] B. Medina-Salgado, E. Sánchez-DelaCruz, P. Pozos-Parra, and J. E. Sierra, "Urban traffic flow prediction techniques: A review," *Sustainable Computing: Informatics and Systems*, vol. 35, p. 100739, 2022.

[7] H. Yuan and G. Li, "A survey of traffic prediction: from spatio-temporal data to intelligent transportation," *Data Science and Engineering*, vol. 6, no. 1, pp. 63–85, 2021.

[8] A. Carianni and A. Gemma, "Overview of traffic flow forecasting techniques," *IEEE Open Journal of Intelligent Transportation Systems*, vol. 6, pp. 848–882, 2025.

[9] B. Yu, H. Yin, and Z. Zhu, "Spatio-temporal graph convolutional networks: A deep learning framework for traffic forecasting," in *Proceedings of the 27th International Joint Conference on Artificial Intelligence (IJCAI)*, 2018, pp. 3634–3640.

[10] Z. Wu, S. Pan, G. Long, J. Jiang, and C. Zhang, "Graph wavenet for deep spatial-temporal graph modeling," in *Proceedings of the Twenty-Eighth International Joint Conference on Artificial Intelligence (IJCAI-19)*. AAAI Press, 2019, pp. 1907–1913.

[11] A. Vaswani, N. Shazeer, N. Parmar, J. Uszkoreit, L. Jones, A. N. Gomez, L. Kaiser, and I. Polosukhin, "Attention is all you need," vol. 30, 2017, pp. 5998–6008.

[12] 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, pp. 54 739–54 756, 2021.

[13] W. Jiang and J. Luo, "Big data for traffic estimation and prediction: a survey of data and tools," *Applied System Innovation*, vol. 5, no. 1, p. 23, 2022.

[14] H. Tian, Y. Feng, M. Quddus, Y. Demiris, and P. Angeloudis, "Multimodal learning for traffic risk prediction: Combining aerial imagery with contextual data," *IEEE Open Journal of Intelligent Transportation Systems*, vol. 6, pp. 1–14, 2025.

[15] B. Li, Q. Yang, J. Chen, D. Yu, D. Wang, and F. Wan, "A dynamic spatio-temporal deep learning model for lane-level traffic prediction," *Journal of Advanced Transportation*, vol. 2023, no. 1, p. 3208535, 2023.

[16] L. Bai, L. Yao, C. Li, X. Wang, and C. Wang, "Adaptive graph convolutional recurrent network for traffic prediction," in *Advances in Neural Information Processing Systems*, vol. 33. Curran Associates Inc., 2020, pp. 16 566–16 576.

[17] N. Kumar and M. Raubal, "Applications of deep learning in congestion detection, prediction and alleviation: A survey," *Transportation Research Part C: Emerging Technologies*, vol. 133, p. 103432, 2021.

[18] Y. An, Z. Li, W. Liu, H. Sun, M. Chen, W. Lu, and Y. Gong, "Spatio-temporal graph normalizing flow for probabilistic traffic prediction," in *Proceedings of the 33rd ACM International Conference on Information and Knowledge Management*, 2024, pp. 45–55.

[19] J. Kong, X. Fan, X. Jin, S. Lin, and M. Zuo, "A variational bayesian inference-based en-decoder framework for traffic flow prediction," *IEEE Transactions on Intelligent Transportation Systems*, vol. 25, no. 3, pp. 2966–2975, 2023.

[20] C. Wang, K. Zuo, S. Zhang, C. Liu, H. Peng, W. Li, Z. Shen, P. Hu, R. Wang, and B. Jie, "Tadgcn: A time-aware dynamic graph convolution network for long-term traffic flow prediction," *Expert Systems with Applications*, vol. 258, p. 125134, 2024.

[21] M. J. Page, J. E. McKenzie, P. M. Bossuyt, I. Boutron, T. C. Hoffmann, C. D. Mulrow, L. Shamseer, J. M. Tetzlaff, E. A. Akl, S. E. Brennan *et al.*, "The prisma 2020 statement: An updated guideline for reporting systematic reviews," *BMJ*, vol. 372, p. n71, 2021.

[22] Clarivate, "Journal citation reports," 2025, accessed: 2026-02-02. [Online]. Available: https://clarivate.com/webofsciencegroup/solutions/journal-citation-reports/

[23] CORE, "Core conference ranking portal," 2025, accessed: 2026-02-02. [Online]. Available: https://portal.core.edu.au/conf-ranks/

[24] 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*, vol. 43, pp. 3–19, 2014.

[25] Y. Zheng, L. Capra, O. Wolfson, and H. Yang, "Urban computing: concepts, methodologies, and applications," *ACM Transactions on Intelligent Systems and Technology (TIST)*, vol. 5, no. 3, pp. 1–55, 2014.

[26] K. Wang, L. Liu, Y. Liu, G. Li, F. Zhou, and L. Lin, "Urban regional function guided traffic flow prediction," *Information Sciences*, vol. 634, pp. 308–320, 2023.

[27] C. Ma, G. Dai, and J. Zhou, "Short-term traffic flow prediction for urban road sections based on time series analysis and lstm_bilstm method," *IEEE Transactions on Intelligent Transportation Systems*, vol. 23, no. 6, pp. 5615–5624, 2021.

[28] P. Varaiya, "Freeway performance measurement system (pems) version 4," 2004. [Online]. Available: https://escholarship.org/uc/item/4bt6h98r

[29] X. Liu, Y. Xia, Y. Liang, J. Hu, Y. Wang, L. Bai, C. Huang, Z. Liu, B. Hooi, and R. Zimmermann, "Largest: A benchmark dataset for large-scale traffic forecasting," *Advances in Neural Information Processing Systems*, vol. 36, pp. 75 354–75 371, 2023.

[30] Y. Li, R. Yu, C. Shahabi, and Y. Liu, "Diffusion convolutional recurrent neural network: Data-driven traffic forecasting," *International Conference on Learning Representations*, 2018.

[31] "Bigscity-libcity documentation: Raw data," 2022, documentation of METR_LA, PEMSD4, PEMSD7, PEMSD8, LOOP, and other datasets. [Online]. Available: https://bigscity-libcity-docs.readthedocs.io/en/latest/user_guide/data/raw_data.html

[32] J. Yuan, Y. Zheng, C. Zhang, W. Xie, X. Xie, G. Sun, and Y. Huang, "Driving directions based on taxi trajectories," *Proceedings of the 18th SIGSPATIAL International Conference on Advances in GIS*, pp. 405–414, 2010.

[33] N. Y. C. Taxi and L. Commission, "Tlc trip record data," 2024, official NYC taxi data; yellow and green taxis; pickup/dropoff times and locations. [Online]. Available: https://www.nyc.gov/site/tlc/about/tlc-trip-record-data.page

[34] M. Tenzer, Z. Rasheed, K. Shafique, and N. Vasconcelos, "Meta-learning over time for destination prediction tasks," in *Proceedings of the 30th International Conference on Advances in Geographic Information Systems*, 2022, pp. 1–10.

[35] X. Tian, C. Zou, Y. Zhang, L. Du, and S. Wu, "Na-dgru: A dual-gru traffic speed prediction model based on neighborhood aggregation and attention mechanism," *Sustainability*, vol. 15, no. 4, p. 2927, 2023.

[36] A. Kumar, D. Garg, and G. Sharma, "Three-tier survey of deep learning based traffic prediction schemes," in *2024 11th International Conference on Reliability, Infocom Technologies and Optimization (Trends and Future Directions)(ICRITO)*. IEEE, 2024, pp. 1–6.

[37] J. Zhang, Y. Zheng, and D. Qi, "Deep spatio-temporal residual networks for citywide crowd flows prediction," *AAAI Conference on Artificial Intelligence*, vol. 31, no. 1, 2017, original TaxiBJ dataset; 32x32 grid, Beijing, 2013-2016.

[38] W. Jiang *et al.*, "Taxibj21: An open crowd flow dataset based on beijing taxi gps traces," *International Journal of Geo-Information*, vol. 11, no. 10, p. 524, 2022, updated TaxiBJ; 1,000 drivers, 3 months, same grid format.

[39] G. Jin, Y. Cui, L. Zeng, H. Tang, Y. Feng, and J. Huang, "Urban ride-hailing demand prediction with multiple spatio-temporal information fusion network," *Transportation Research Part C: Emerging Technologies*, vol. 117, p. 102665, 2020.

[40] L. Bai, L. Yao, S. S. Kanhere, X. Wang, W. Liu, and Z. Yang, "Spatio-temporal graph convolutional and recurrent networks for citywide passenger demand prediction," in *Proceedings of the 28th ACM international conference on information and knowledge management*, 2019, pp. 2293–2296.

[41] DiDi Chuxing, "GAIA Open Dataset Initiative," [Online]. Available: https://outreach.didichuxing.com/research/opendata/, 2020, accessed: 2026-02-02. Trajectory data of DiDi Express and Premier drivers in Chengdu and Xi'an, China (Oct–Nov 2016).

[42] B. Liao, J. Zhang, C. Wu, D. McIlwraith, T. Chen, S. Yang, Y. Guo, and F. Wu, "Deep sequence learning with auxiliary information for traffic prediction," in *Proceedings of the 24th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD)*, 2018, pp. 537–546.

[43] J. Xue, R. Tan, J. Ma, and S. V. Ukkusuri, "Data science in transportation networks with graph neural networks: a review and outlook," *Data Science for Transportation*, vol. 7, no. 2, p. 10, 2025.

[44] A. A. Chaudhari, M. Treiber, and O. Okhrin, "Mitra: A drone-based trajectory data for an all-traffic-state inclusive freeway with ramps," *Scientific Data*, vol. 12, p. 1174, 2025. [Online]. Available: https://doi.org/10.1038/s41597-025-05472-0

[45] OpenStreetMap Foundation, "Openstreetmap," https://www.openstreetmap.org, 2024, accessed: 2024-12-01.

[46] P. A. Lopez *et al.*, "Microscopic traffic simulation using sumo," *IEEE Intelligent Transportation Systems Conference (ITSC)*, pp. 2575–2582, 2018, official SUMO reference; microscopic traffic simulator for transportation research.

[47] G. A. C. (DLR), "Sumo traffic simulation," 2024, open-source microscopic traffic simulator; widely used for dataset generation and validation. [Online]. Available: https://www.eclipse.org/sumo/

[48] A. Horni, K. Nagel, and K. W. Axhausen, *The Multi-Agent Transport Simulation MATSim*. Ubiquity Press, 2016, comprehensive reference for MATSim agent-based simulation framework. [Online]. Available: https://www.ubiquitypress.com/site/books/10.5334/bax/

[49] "Trafficnet: Traffic flow benchmark dataset," 2020, benchmark dataset aggregating traffic flow data from multiple sensors and cities. [Online]. Available: https://trafficnet.org/

[50] J. Jiang, C. Han, W. Jiang, W. X. Zhao, and J. Wang, "Libcity: A unified library towards efficient and comprehensive urban spatial-temporal prediction," *arXiv preprint arXiv:2304.14343*, 2023.

[51] G. LLC, "Google maps platform - traffic api," 2024, web-based service providing real-time traffic data, congestion levels, and travel times. [Online]. Available: https://developers.google.com/maps/documentation/roads

[52] Y. Ye, Y. Xiao, Y. Zhou, S. Li, Y. Zang, and Y. Zhang, "Dynamic multi-graph neural network for traffic flow prediction incorporating traffic accidents," *Expert Systems with Applications*, vol. 234, p. 121101, 2023.

[53] P. Tam, I. Song, S. Kang, S. Ros, and S. Kim, "Graph neural networks for intelligent modelling in network management and orchestration: a survey on communications," *Electronics*, vol. 11, no. 20, p. 3371, 2022.

[54] Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli, "Image quality assessment: From error visibility to structural similarity," *IEEE Transactions on Image Processing*, vol. 13, no. 4, pp. 600–612, 2004.

[55] M. Everingham, L. Van Gool, C. K. I. Williams, J. Winn, and A. Zisserman, "The pascal visual object classes (voc) challenge," *International Journal of Computer Vision*, vol. 88, no. 2, pp. 303–338, 2010.

[56] T. Gneiting and A. E. Raftery, "Strictly proper scoring rules, prediction, and estimation," *Journal of the American Statistical Association*, vol. 102, no. 477, pp. 359–378, 2007.

[57] T. Gneiting, F. Balabdaoui, and A. E. Raftery, "Probabilistic forecasts, calibration and sharpness," *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, vol. 69, no. 2, pp. 243–268, 2007.

[58] B. Lakshminarayanan, A. Pritzel, and C. Blundell, "Simple and scalable predictive uncertainty estimation using deep ensembles," in *Advances in Neural Information Processing Systems*, vol. 30, 2017, pp. 6402–6413.

[59] A. Khosravi, S. Nahavandi, D. Creighton, and A. F. Atiya, "Comprehensive review of neural network-based prediction intervals and new advances," *IEEE Transactions on Neural Networks*, vol. 22, no. 9, pp. 1341–1356, 2011.

[60] J. An, L. Guo, W. Liu, Z. Fu, P. Ren, X. Liu, and T. Li, "Igagcn: Information geometry and attention-based spatiotemporal graph convolutional networks for traffic flow prediction," *Neural Networks*, vol. 143, pp. 355–367, 2021.

[61] S. Zhang, Y. Guo, P. Zhao, C. Zheng, and X. Chen, "A graph-based temporal attention framework for multi-sensor traffic flow forecasting," *IEEE Transactions on Intelligent Transportation Systems*, vol. 23, no. 7, pp. 7743–7758, 2021.

[62] T. Qi, G. Li, L. Chen, and Y. Xue, "Adgcn: An asynchronous dilation graph convolutional network for traffic flow prediction," *IEEE Internet of Things Journal*, vol. 9, no. 5, pp. 4001–4014, 2021.

[63] L. Chen, W. Shao, M. Lv, W. Chen, Y. Zhang, and C. Yang, "Aargnn: An attentive attributed recurrent graph neural network for traffic flow prediction considering multiple dynamic factors," *IEEE Transactions on Intelligent Transportation Systems*, vol. 23, no. 10, pp. 17 201–17 211, 2022.

[64] Z. Chen, Z. Lu, Q. Chen, H. Zhong, Y. Zhang, J. Xue, and C. Wu, "Spatial–temporal short-term traffic flow prediction model based on dynamical-learning graph convolution mechanism," *Information Sciences*, vol. 611, pp. 522–539, 2022.

[65] S. Li, L. Ge, Y. Lin, and B. Zeng, "Adaptive spatial-temporal fusion graph convolutional networks for traffic flow forecasting," in *2022 International Joint Conference on Neural Networks (IJCNN)*. IEEE, 2022, pp. 1–8.

[66] J. Ji, J. Wang, Z. Jiang, J. Jiang, and H. Zhang, "Stden: Towards physics-guided neural networks for traffic flow prediction," in *Proceedings of the AAAI Conference on Artificial Intelligence*, vol. 36, no. 4, 2022, pp. 4048–4056.

[67] Y. Bao, J. Huang, Q. Shen, Y. Cao, W. Ding, Z. Shi, and Q. Shi, "Spatial–temporal complex graph convolution network for traffic flow prediction," *Engineering Applications of Artificial Intelligence*, vol. 121, p. 106044, 2023.

[68] Y. Xu, L. Han, T. Zhu, L. Sun, B. Du, and W. Lv, "Generic dynamic graph convolutional network for traffic flow forecasting," *Information Fusion*, vol. 100, p. 101946, 2023.

[69] J. Peng, T. Guan, and J. Liang, "Spatial-temporal graph discriminant autoencoder for traffic congestion forecasting," in *2023 IEEE 26th International Conference on Intelligent Transportation Systems (ITSC)*. IEEE, 2023, pp. 23–28.

[70] Q. Zhang, C. Li, F. Su, and Y. Li, "Spatiotemporal residual graph attention network for traffic flow forecasting," *IEEE Internet of Things Journal*, vol. 10, no. 13, pp. 11 518–11 532, 2023.

[71] C. Zhang, Q. Zheng, and Y. Zhang, "Spatial-temporal flow holistic interaction graph convolution network for bidirectional traffic flow forecasting," in *2023 IEEE International Conference on Big Data (BigData)*. IEEE, 2023, pp. 1262–1268.

[72] Q. Li, X. Yang, Y. Wang, Y. Wu, and D. He, "Spatial–temporal traffic modeling with a fusion graph reconstructed by tensor decomposition," *IEEE transactions on intelligent transportation systems*, vol. 25, no. 2, pp. 1749–1760, 2023.

[73] H. Xiao, Z. Zhao, and T. Yang, "A traffic flow prediction method based on constrained dynamic graph convolutional recurrent networks," *Engineering Applications of Artificial Intelligence*, vol. 133, p. 108486, 2024.

[74] F. Chen, X. Sun, Y. Wang, Z. Xu, and W. Ma, "Adaptive graph neural network for traffic flow prediction considering time variation," *Expert Systems with Applications*, vol. 255, p. 124430, 2024.

[75] T. Dan, X. Pan, B. Zheng, and X. Meng, "Bygcn: Spatial temporal byroad-aware graph convolution network for traffic flow prediction in road networks," in *Proceedings of the 33rd ACM International Conference on Information and Knowledge Management*, 2024, pp. 415–424.

[76] W. Xu, J. Liu, J. Yan, J. Yang, H. Liu, and T. Zhou, "Dynamic spatiotemporal graph wavelet network for traffic flow prediction," *IEEE Internet of Things Journal*, vol. 11, no. 5, pp. 8019–8029, 2023.

[77] R. Gan, B. An, L. Li, X. Qu, and B. Ran, "A freeway traffic flow prediction model based on a generalized dynamic spatio-temporal graph convolutional network," *IEEE Transactions on Intelligent Transportation Systems*, vol. 25, no. 10, pp. 13 682–13 693, 2024.

[78] X. Yang, Q. Li, P. Xu, D. He, and H. Tan, "Parallel self-learned and predefined joint spatio-temporal graph convolutional networks for traffic flow prediction," *IEEE Internet of Things Journal*, vol. 12, no. 9, pp. 11 698–11 707, 2025.

[79] Q. Li, P. Xu, X. Yang, Y. Wu, H. He, and D. He, "Spatial-temporal traffic prediction with an interactive spatial-enhanced graph convolutional network model," *IEEE Transactions on Intelligent Transportation Systems*, vol. 25, no. 12, pp. 20 767–20 778, 2024.

[80] J. Feng, J. Huang, W. Ding, C. Guo, and Z. Shi, "Mgc-dmf: A traffic flow forecasting method based on multi-graph spatio-temporal convolution and dynamic metric fusion with multi-source basic information," *Applied Soft Computing*, vol. 177, p. 113178, 2025.

[81] X. Wang, X. Wang, F. Huang, F. Zou, L. Liao, and R. Zeng, "Dstf: A diversified spatio-temporal feature extraction model for traffic flow prediction," *Neurocomputing*, vol. 621, p. 129280, 2025.

[82] L. Zhong, B. Wang, Z. Tian, T. K. Rodrigues, W. Liu, and W. She, "Multiview spatiotemporal dynamic graph convolution network for traffic flow prediction," *IEEE Internet of Things Journal*, vol. 12, no. 18, pp. 37 062–37 076, 2025.

[83] Z. Liu, Z. Qin, T. Huang, and G. Ding, "Dynamic graph convolution and spatiotemporal self-attention network for traffic flow prediction," *IEEE Internet of Things Journal*, vol. 12, no. 14, pp. 27 379–27 392, 2025.

[84] Y. Bao, Q. Shen, Y. Cao, and Q. Shi, "Da-rmn: Denoising attention enhanced recurrent multigraph convolutional network for traffic flow prediction," *IEEE Internet of Things Journal*, vol. 12, no. 5, pp. 4820–4833, 2025.

[85] Y. Zheng, C. Luo, and R. Shao, "Tfsg: A self-supervised guided graph learning method for traffic flow prediction," *IEEE Internet of Things Journal*, vol. 12, no. 13, pp. 24 345–24 357, 2025.

[86] R. Zhong, B. Hu, F. Wang, Y. Feng, Z. Li, X. Song, Y. Wang, S. Lou, and J. Tan, "Multi-factor embedding gnn-based traffic flow prediction considering intersection similarity," *Neurocomputing*, vol. 620, p. 129193, 2025.

[87] M. Patil, Q. Ahmed, and S. Midlam-Mohler, "Travel time and weather-aware traffic forecasting in a conformal graph neural network framework," *IEEE Transactions on Intelligent Transportation Systems*, vol. 26, no. 12, pp. 21 734–21 744, 2025.

[88] L. Chen, Q. Zhao, G. Li, M. Zhou, C. Dai, Y. Feng, X. Liu, and J. Li, "A sparse cross attention-based graph convolution network with auxiliary information awareness for traffic flow prediction," *IEEE Transactions on Intelligent Transportation Systems*, vol. 26, no. 3, pp. 3210–3222, 2025.

[89] W. Xiong, R. Fonod, A. Alahi, and N. Geroliminis, "Multi-source urban traffic flow forecasting with drone and loop detector data," *IEEE Transactions on Intelligent Transportation Systems*, vol. 26, no. 11, pp. 18 637–18 652, 2025.

[90] Y. Zhao, Y. Lin, H. Wen, T. Wei, X. Jin, and H. Wan, "Spatial-temporal position-aware graph convolution networks for traffic flow forecasting," *IEEE Transactions on Intelligent Transportation Systems*, vol. 24, no. 8, pp. 8650–8666, 2022.

[91] C. Zheng, X. Fan, C. Wang, and J. Qi, "Gman: A graph multi-attention network for traffic prediction," in *Proceedings of the AAAI Conference on Artificial Intelligence*, vol. 34, no. 01, 2020, pp. 1234–1241.

[92] S. Reza, M. Mirza, M. Iftikhar, and M. B. Zafar, "A multi-head attention-based transformer model for traffic flow forecasting," *Expert Systems with Applications*, vol. 202, p. 117348, 2022.

[93] Y. Feng, F. Han, and S. Zhao, "A graph convolutional stacked temporal attention neural network for traffic flow forecasting," in *2022 International Joint Conference on Neural Networks (IJCNN)*. IEEE, 2022, pp. 1–7.

[94] Q. Ren, Y. Li, and Y. Liu, "Transformer-enhanced periodic temporal convolution network for long short-term traffic flow forecasting," *Expert Systems with Applications*, vol. 227, p. 120203, 2023.

[95] D. Wu, K. Peng, S. Wang, and V. C. Leung, "Spatial–temporal graph attention gated recurrent transformer network for traffic flow forecasting," *IEEE Internet of Things Journal*, vol. 11, no. 8, pp. 14 267–14 281, 2023.

[96] Y. Li, H. Xu, T. Zhang, X. Li, G. Li, and W. Tian, "Ddgformer: Direction-and distance-aware graph transformer for traffic flow prediction," *Knowledge-Based Systems*, vol. 302, p. 112381, 2024.

[97] Y. Ma, H. Lou, M. Yan, F. Sun, and G. Li, "Spatio-temporal fusion graph convolutional network for traffic flow forecasting," *Information Fusion*, vol. 104, p. 102196, 2024.

[98] J. Xiao and B. Long, "A multi-channel spatial-temporal transformer model for traffic flow forecasting," *Information Sciences*, vol. 671, p. 120648, 2024.

[99] Z. Geng, J. Xu, R. Wu, C. Zhao, J. Wang, Y. Li, and C. Zhang, "Stgaformer: Spatial–temporal gated attention transformer based graph neural network for traffic flow forecasting," *Information Fusion*, vol. 105, p. 102228, 2024.

[100] W. Li, J. Chen, Y. Zhang, R. Sun, S. Xia, Z. Pan, and J. Luo, "Msgformer: Revolutionizing traffic flow prediction with multiscale and gated transformer architecture," *IEEE Internet of Things Journal*, vol. 12, no. 2, pp. 2014–2025, 2024.

[101] H. Chi, Y. Lu, C. Xie, W. Ke, and B. Chen, "Spatio-temporal attention based collaborative local–global learning for traffic flow prediction," *Engineering Applications of Artificial Intelligence*, vol. 139, p. 109575, 2025.

[102] J. Cao, X. Sheng, J. Zhang, and Z. Duan, "itransmamba: A lightweight spatio-temporal network based on long-term traffic flow forecasting," *Knowledge-Based Systems*, vol. 317, p. 113416, 2025.

[103] J. Hu and S. Baldi, "High and low frequency attention network for long-term traffic flow prediction," *IEEE Transactions on Intelligent Transportation Systems*, vol. 26, no. 9, pp. 13 811–13 821, 2025.

[104] Z. Fang, Q. Long, G. Song, and K. Xie, "Spatial-temporal graph ode networks for traffic flow forecasting," in *Proceedings of the 27th ACM SIGKDD conference on knowledge discovery & data mining*, 2021, pp. 364–373.

[105] Z. Fang, L. Pan, L. Chen, Y. Du, and Y. Gao, "Mdtp: A multi-source deep traffic prediction framework over spatio-temporal trajectory data," *Proceedings of the VLDB Endowment*, vol. 14, no. 8, pp. 1289–1297, 2021.

[106] S. Cao, L. Wu, J. Wu, D. Wu, and Q. Li, "A spatio-temporal sequence-to-sequence network for traffic flow prediction," *Information Sciences*, vol. 610, pp. 185–203, 2022.

[107] F. Dai, P. Huang, Q. Mo, X. Xu, M. Bilal, and H. Song, "St-innet: Deep spatio-temporal inception networks for traffic flow prediction in smart cities," *IEEE Transactions on Intelligent Transportation Systems*, vol. 23, no. 10, pp. 19 782–19 794, 2022.

[108] S. Fang, C. Zhang, S. Xiang, and C. Pan, "Automsnet: Multi-source spatio-temporal network via automatic neural architecture search for traffic flow prediction," *IEEE Transactions on Intelligent Transportation Systems*, vol. 24, no. 3, pp. 2827–2841, 2022.

[109] Y. Bao, J. Liu, Q. Shen, Y. Cao, W. Ding, and Q. Shi, "Pket-gcn: prior knowledge enhanced time-varying graph convolution network for traffic flow prediction," *Information Sciences*, vol. 634, pp. 359–381, 2023.

[110] M. Méndez, M. G. Merayo, and M. Núñez, "Long-term traffic flow forecasting using a hybrid cnn-bilstm model," *Engineering Applications of Artificial Intelligence*, vol. 121, p. 106041, 2023.

[111] J. Wan and Y. Wu, "Sastdgcn: Self-attention based spatial-temporal double graph convolutional networks for traffic flow forecasting," in *2023 IEEE International Conference on Big Data (BigData)*. IEEE, 2023, pp. 1544–1549.

[112] J. Ji, J. Wang, C. Huang, J. Wu, B. Xu, Z. Wu, J. Zhang, and Y. Zheng, "Spatio-temporal self-supervised learning for traffic flow prediction," in *Proceedings of the AAAI conference on artificial intelligence*, vol. 37, no. 4, 2023, pp. 4356–4364.

[113] Q. Wang, W. Liu, X. Wang, X. Chen, G. Chen, and Q. Wu, "Gmhann: A novel traffic flow prediction method for transportation management based on spatial-temporal graph modeling," *IEEE Transactions on Intelligent Transportation Systems*, vol. 25, no. 1, pp. 386–401, 2023.

[114] H.-X. Hu, Q. Hu, G. Tan, Y. Zhang, and Z.-Z. Lin, "A multi-layer model based on transformer and deep learning for traffic flow prediction," *IEEE Transactions on Intelligent Transportation Systems*, vol. 25, no. 1, pp. 443–451, 2023.

[115] Y. Bao, Q. Shen, Y. Cao, W. Ding, and Q. Shi, "Residual attention enhanced time-varying multi-factor graph convolutional network for traffic flow prediction," *Engineering Applications of Artificial Intelligence*, vol. 133, p. 108135, 2024.

[116] Q. Gao, Z. Wang, L. Huang, G. Trajcevski, K. Zhang, and X. Chen, "Enhancing dependency dynamics in traffic flow forecasting via graph risk bootstrap," in *Proceedings of the 32nd ACM International Conference on Advances in Geographic Information Systems*, 2024, pp. 147–159.

[117] J. Yan, L. Zhang, Y. Gao, and B. Qu, "Gecran: Graph embedding based convolutional recurrent attention network for traffic flow prediction," *Expert Systems with Applications*, vol. 256, p. 125001, 2024.

[118] J. Tian, L. Han, M. Chen, Y. Xu, Z. Chen, T. Zhu, L. Sun, and W. Lv, "Mfgcn: Multi-faceted spatial and temporal specific graph convolutional network for traffic-flow forecasting," *Knowledge-Based Systems*, vol. 306, p. 112671, 2024.

[119] D. Lu, "Enhancing spatial-temporal awareness via graph convolutional networks and transformers for traffic flow forecasting," in *2024 International Joint Conference on Neural Networks (IJCNN)*. IEEE, 2024, pp. 1–8.

[120] Q. Lai and P. Chen, "Leisn: A long explicit–implicit spatio-temporal network for traffic flow forecasting," *Expert Systems with Applications*, vol. 245, p. 123139, 2024.

[121] Q. Li, P. Xu, D. He, Y. Wu, H. Tan, and X. Yang, "Multi-source information fusion graph convolution network for traffic flow prediction," *Expert Systems with Applications*, vol. 252, p. 124288, 2024.

[122] J. Fan, W. Weng, Q. Chen, H. Wu, and J. Wu, "Pdg2seq: Periodic dynamic graph to sequence model for traffic flow prediction," *Neural Networks*, vol. 183, p. 106941, 2025.

[123] H. Shi and W. Zhang, "Learning continuous spatial-temporal evolution via dynamic graph neural differential equations for traffic flow forecasting," *Information Sciences*, vol. 700, p. 121845, 2025.

[124] W. Kong, Y. Ju, S. Zhang, J. Wang, L. Huang, and H. Qu, "Graph enhanced spatial–temporal transformer for traffic flow forecasting," *Applied Soft Computing*, vol. 170, p. 112698, 2025.

[125] L. Wang, S. Hong, H. Chi, C. Xie, Y. Zhu, and H. Mao, "Hybrid spatio-temporal graph neural network with attention fusion for traffic flow prediction," *Knowledge-Based Systems*, vol. 324, p. 113813, 2025.

[126] Z. Chu, W. Ma, M. Li, and H. Chen, "Multi-time scale augmented neural odes graph neural for traffic flow prediction with elastic channel variation," *Applied Soft Computing*, vol. 182, p. 113513, 2025.

[127] Y.-T. Chen, C. Li, and S. Li, "Attention based spatial-temporal multi-graph ordinary differential equation network for traffic flow prediction," *Engineering Applications of Artificial Intelligence*, vol. 149, p. 110526, 2025.

[128] Z. Yingran, L. Chao, and S. Rui, "Enhancing traffic flow forecasting with delay propagation: Adaptive graph convolution networks for spatio-temporal data," *IEEE Transactions on Intelligent Transportation Systems*, vol. 26, no. 1, pp. 650–660, 2025.

[129] H. Chen, J. Huang, Y. Lu, and J. Huang, "Urban traffic flow prediction based on multi-spatio-temporal feature fusion," *Neurocomputing*, vol. 638, p. 130117, 2025.

[130] J. Tang, J. Yang, and Y. Zhang, "D-stanet: A delay-enhanced spatio-temporal attention network for traffic prediction," *Knowledge-Based Systems*, vol. 320, p. 113533, 2025.

[131] J. Chen, S. Pan, W. Peng, and W. Xu, "Bilinear spatiotemporal fusion network: An efficient approach for traffic flow prediction," *Neural Networks*, vol. 187, p. 107382, 2025.

[132] J. Wang, K. Liu, H. Li, Q. Gao, X. Wang, and Y. Gong, "Multiscale temporal features-based hybrid lstm-gat for traffic flow prediction," *IEEE Internet of Things Journal*, vol. 12, no. 21, pp. 44 909–44 926, 2025.

[133] A. Ali, I. Ullah, S. Ahmad, Z. Wu, J. Li, and X. Bai, "An attention-driven spatio-temporal deep hybrid neural networks for traffic flow prediction," *IEEE Transactions on Intelligent Transportation Systems*, vol. 26, no. 9, pp. 14 154–14 168, 2025.

[134] Q. Shen, Z. Ying, Z. Liu, and J. Lou, "Multi-form spatiotemporal feature fusion enhancement network for traffic flow prediction," *IEEE Transactions on Intelligent Transportation Systems*, vol. 26, no. 11, pp. 20 912–20 929, 2025.

[135] Y. Hu, S. Li, D. Xia, W. Zhang, P. Yuan, F. Wu, and H. Li, "A multiview spatial-temporal adaptive transformer-gru framework for traffic flow prediction," *IEEE Internet of Things Journal*, vol. 12, no. 6, pp. 7114–7132, 2025.

[136] X. Yuan, Z. Luo, N. Zhang, G. Guo, L. Wang, C. Li, and D. Niyato, "Federated transfer learning for privacy-preserved cross-city traffic flow prediction," *IEEE Transactions on Intelligent Transportation Systems*, vol. 26, no. 4, pp. 4418–4431, 2025.

[137] W. Ruan, W. Wang, S. Zhong, W. Chen, L. Liu, and Y. Liang, "Cross space and time: A spatio-temporal unitized model for traffic flow forecasting," *IEEE Transactions on Intelligent Transportation Systems*, vol. 26, no. 11, pp. 21 296–21 308, 2025.

[138] S. Zhang, X. Yin, W. Zhang, J. Xu, and X. Jing, "Adaptive lightweight temporal convolutional network with context-aware downsampling strategy for traffic flow prediction," *Engineering Applications of Artificial Intelligence*, vol. 156, p. 111208, 2025.

[139] D. Ma, J. Zhu, X. B. Song, and X. Wang, "Traffic flow and speed forecasting through a bayesian deep multi-linear relationship network," *Expert Systems with Applications*, vol. 213, p. 119161, 2023.

[140] Y. Liu, Y. Zhang, X. Zhang, Y. Yang, Y. Xie, S. G. Machiani, Y. Li, and J. Luo, "Align along time and space: A graph latent diffusion model for traffic dynamics prediction," in *2024 IEEE International Conference on Data Mining (ICDM)*, 2024. [Online]. Available: https://ieeexplore.ieee.org/document/10884127

[141] W. Zhang and K. Zheng, "Disentangled diffusion models for probabilistic spatio-temporal traffic forecasting," *Complex & Intelligent Systems*, vol. 11, no. 7, p. 322, 2025.

[142] Y. Bao, Q. Shen, Y. Cao, Y. Hou, W. Lu, and Q. Shi, "Multi-period diffusion generative graph recurrent transformer network for traffic flow prediction in vehicular networks," *IEEE Transactions on Intelligent Transportation Systems*, 2025. [Online]. Available: https://ieeexplore.ieee.org/document/11033208

[143] D. Yu, D. Zhuang, L. Jiang, R. Xu, X. Ye, Y. Bu, S. Wang, and G. Wang, "Uqgnn: Uncertainty quantification of graph neural networks for multivariate spatiotemporal prediction," in *Proceedings of the 33rd ACM International Conference on Advances in Geographic Information Systems*, ser. SIGSPATIAL '25. New York, NY, USA: Association for Computing Machinery, 2025, p. 52–65. [Online]. Available: https://doi.org/10.1145/3748636.3762709

[144] H. Byeon *et al.*, "Spatiotemporal network for accurate traffic flow prediction in intelligent transportation systems using generative ai," *IEEE Transactions on Intelligent Transportation Systems*, vol. 26, 2025.

[145] X. Zheng, Y. Liu, Z. Bao, M. Fang, X. Hu, A. W.-C. Liew, and S. Pan, "Towards data-centric graph machine learning: Review and outlook," *arXiv preprint arXiv:2309.10979*, 2023.

[146] Z. Chen, G. Cheng, D. Niu, X. Qiu, Y. Zhao, and Y. Zhou, "Wff-egnn: Encrypted traffic classification based on weaved flow fragment via ensemble graph neural networks," *IEEE Transactions on Machine Learning in Communications and Networking*, vol. 1, pp. 389–411, 2023.

[147] A. Younesi, M. Ansari, M. Fazli, A. Ejlali, M. Shafique, and J. Henkel, "A comprehensive survey of convolutions in deep learning: Applications, challenges, and future trends," *IEEE Access*, vol. 12, pp. 41 180–41 218, 2024.

[148] T. N. Kipf and M. Welling, "Variational graph auto-encoders," in *NIPS Workshop on Bayesian Deep Learning*, 2016.

[149] C.-I. Huang, L. Biechy, J.-S. Chang, C.-H. Tseng, J.-H. Wu, W.-Y. Chang, and Y.-T. Lin, "A lightweight recurrent architecture for robust urban traffic forecasting with missing data," *IEEE Open Journal of Intelligent Transportation Systems*, vol. 6, pp. 1235–1244, 2025.

[150] F. Li, J. Feng, H. Yan, G. Jin, F. Yang, F. Sun, D. Jin, and Y. Li, "Dynamic graph convolutional recurrent network for traffic prediction: Benchmark and solution," *ACM Transactions on Knowledge Discovery from Data*, vol. 17, no. 1, 2023.

[151] K. Binjaku, C. Pasquale, E. K. Meçe, and S. Sacone, "Freeway traffic modeling by physics-regularized gaussian processes," *IEEE Open Journal of Intelligent Transportation Systems*, vol. 6, pp. 116–130, 2025.

[152] G. Luo, H. Zhang, Q. Yuan, J. Li, W. Wang, and F.-Y. Wang, "One size fits all: A unified traffic predictor for capturing the essential spatial–temporal dependency," *IEEE transactions on neural networks and learning systems*, 2023.



Aneri Kamal Patel received the M.S. degree in Computer and Information Sciences from Gujarat Technological University, India. She is currently pursuing the Ph.D. degree at Auckland University of Technology, Auckland, New Zealand. Her research interests include intelligent transportation systems, traffic signal control, and reinforcement learning.



Ke Liu received his Masters degree in Computer and Information Sciences from The University of Sydney, Sydney, Australia in 2022. He is now a PhD student at the Auckland University of Technology, Auckland, New Zealand. His research focuses on the design and validation of algorithms for autonomous vehicles.



Jing Ma (Member, IEEE) received the Ph.D. degree in Computer Sciences from Auckland University of Technology, New Zealand, in 2019. She is currently a senior lecturer at the department of Data Science and AI at Auckland University of Technology, Auckland, New Zealand. Her research interests include intelligent transportation systems, artificial intelligence for autonomous vehicles, and robotics.