A Review of Recent Methods for Traffic Prediction with Few Data

Zhaowei Huang

Mathematics Department, Santa Monica College, California, USA
huang_zhaowei01@student.smc.edu

Abstract. Predicting traffic accurately is very important for building smart cities. But the deep learning models that achieve state-of-the-art performance today—especially graph neural networks (GNNs)—require massive amounts of past data. This is a problem in new urban areas or on roads that have recently been built—where the data doesn't yet exist. It's for this reason that Few-Shot Learning (FSL) and Zero-Shot Learning (ZSL) have played such a crucial role in recent traffic prediction research. In this survey, we examine four recent strategies that are practical and promising: (1) meta-learning, which helps models adapt quickly to new situations; (2) better GNNs, which we build by designing them with how traffic actually behaves in mind; (3) large foundation models trained on transportation data, and (4) the surprising use of large language models (LLMs) for forecasting—how each works, what they're good at, and where they fail. Our review makes clear where we're at and where we're going, and is intended for anyone who wants to build traffic predictors that work, even when data is hard to come by.

Keywords: Traffic Prediction, Few-Shot Learning, Zero-Shot Learning, Meta-Learning, Graph Neural Networks.

1 Introduction

Cities are more crowded than ever, and their transport systems are struggling. Congestion is now a daily reality in almost every city around the world, and safety concerns are rising as roads get more complicated. Regarding the future of transportation, Intelligent Transportation Systems are the backbone, and a huge part of those systems is being able to predict the way things are going to move.

In Intelligent Transportation Systems (ITS), models that predict accidents, speed, and volume are very important [1].

This review focuses on predicting traffic flow—that is, the speed, volume, and density on a network. These three measurements are very important because they help traffic managers [2], navigation systems, and urban planners make decisions in real time.

Over the past few years, deep learning has made much better traffic prediction possible [3]. Encoder-decoder models with attention mechanisms have achieved the

K. Subramanian (ed.), *Proceedings of the International Workshop on Advances in Deep Learning for Image Analysis and Computer Vision (IWADIC 2025)*, Advances in Computer Science Research 128,
https://doi.org/10.2991/978-94-6239-648-7_80

best results so far [4]. Graph Neural Networks (GNNs) are a key example, as they are very effective at modeling the complex relationships between roads across space and time [5].

But here's the rub: those highly-performing models require lots of labeled historical data to run. This is a problem. When a city deploys new sensors or builds new roads there is no historical data—a "cold-start"—that can lead to delays in deployment of months to several years. Even well-trained models have a hard time when real-world conditions change since they may encounter situations that were not well represented in the training data [6]. In familiar, low-data situations, conventional models overfit and their accuracy suffers.

Getting around these roadblocks is where approaches that learn from very little data or none at all come in. This is where Few-Shot Learning (FSL) and Zero-Shot Learning (ZSL) come in. The goal of FSL is to train models that can generalize from a small number of examples. The goal of ZSL is to make accurate predictions in a scenario where the model has never seen that particular situation before—no training examples whatsoever. This survey unites the latest work applying FSL and ZSL to traffic prediction.

This paper is organized as follows. Section 2 describes four main approaches: meta-learning, GNNs with built-in traffic knowledge, transportation-specific foundation models, and the use of LLMs. Section 3 compares these approaches. Section 4 steps back to look at bigger challenges and future opportunities and Section 5 concludes.

2 Key Approaches in Limited Data Traffic Prediction

2.1 Meta-Learning: Learn to Learn Quickly

The intuition behind meta-learning is simple: instead of training a model to perform well on one task, we train a model to quickly learn new tasks. In other words, "learn how to learn." The model gets exposure to experience across many similar tasks, so when presented with a new situation, it can "get up to speed" with a few examples.

One such approach is the STDA-Meta framework which is built on top of the Model-Agnostic Meta-Learning (MAML) algorithm [7]. The authors treated each traffic prediction in a city as a separate task. When training, the model would see data from several cities with rich information. But the model wasn't trying to master any single city, instead, they were trying to learn a good initialization that a new city could be quickly tuned for a new city with just a small amount of local data. To help models adapt to cities with different layouts and traffic, they also included a Spatio-Temporal Domain Adaptation (STDA) module.

2.2 Building Better GNNs with Traffic Knowledge

Instead of building models that adapt, another line of work takes the opposite approach: can we design models that by design work well with limited data. This line of research involves building "inductive biases" into the model architecture, sort of baking in our knowledge of how traffic behaves.

The LocaleGN model is one [8]. They begin with a simple observation: most traffic is local. What happens at an intersection tomorrow is mostly dependent on what happened at nearby intersections today. They hardcode this fact by letting only nearby nodes share info. The result is a simpler, more efficient model that's less likely to overfit when there's little data (giving it an automatic advantage in few-shot settings).

2.3 Transportation Foundation Models

Large pre-trained models like GPT and BERT have had so much success in NLP. Can we do the same in transportation. Train one single foundation model on massive amounts of traffic data from many cities hoping that it will learn generalizable traffic information and enable zero-shot transfer to new cities.

OpenCity is this [9]. They pre-trained a large model with Transformer and GNN components on diverse traffic data from multiple sources. After this broad training, they could immediately get reasonable zero-shot predictions for unseen cities.

2.4 Leveraging Language Models for Traffic Prediction

This might be the most surprising. Using large language models (LLMs) for traffic prediction. The issue is clear: large language models are made for language (text), but traffic data is numerical and structured.

The TPLLM study found a way out [10]. They designed an embedding module that feeds traffic speed information into a CNN to process it and road network info into a GCN to understand it. Then they transform the processed features into information that an LLM can understand (sequences of tokens). This enables the LLM to apply its sequence modeling ability to traffic prediction. When there's limited data available, they use parameter efficient fine tuning methods like LoRA to enable quick adaptation.

3 Comparative analysis of methods

Table 1. Comparison of Key Methods for Limited-Data Traffic Prediction

Approach	Example Model	Main Idea	Best For
Meta-Learning	STDA-Meta	Train models for fast adaptation	Few-shot: learning new tasks with minimal data
GNN with Inductive Bias	LocaleGN	Embed traffic knowledge into architecture	Few-shot: avoiding overfitting with scarce data
Foundation Model	OpenCity	Pre-train general model on diverse data	Zero-shot: prediction in new environments
LLM Adaptation	TPLLM	Convert traffic data for LLM processing	Few-shot: leveraging pre-trained knowledge

Making Sense of the Different Approaches To help make sense of these different strategies, we've summarized some key characteristics of each approach in Table 1.

The results so far are promising. STDA-Meta shows that models can adapt quickly to new cities when data is limited [7]. These studies showed some very interesting things. For example, LocaleGN found that a simple model can actually be better than a complex one when there is not much data [8]. OpenCity proved that we can predict traffic for a new city without training on its data at all [9]. And TPLLM did something completely new by using a large language model, the same kind of AI that powers chatbots, to predict traffic [10].

4 Challenges and Prospects

The field of traffic prediction is doing more than just making small improvements. Researchers are now rethinking the basic ways to predict traffic. Looking at the progress so far, we can see several challenges and also opportunities for what to do next.

4.1 Main Challenges Computational Costs.

Large models like foundations models and language models require substantial computational resources to train and run. This may limit their applicability to cities with limited budgets.

Current models have two main challenges.

One is that they have trouble using different kinds of data. Real traffic is affected by many things. Not only just old data, like weather, events, and accidents, but also the models find it hard to put all of that information together.

The other problem is that as they get more complex, they become harder to understand. They act like "black boxes," which makes them difficult to trust. Because of this, traffic engineers need better ways to explain why a model gives a certain answer and to see where it might be going wrong.

4.2 Future Prospects

In the future, there are three main things we need to focus on.

First, people need to make large models more practical. We can also use techniques like knowledge distillation to make them smaller and faster.

Second, people can build systems that could use different types of data, like weather, events, and social media. This will help them understand how these things affect traffic.

Finally, our AI systems need to know their own limits. They should be able to tell when they are in a situation they haven't seen before. When that happens, they should either adapt. Or at least give us a sign to show that they are uncertain.

5 Conclusion

There are four promising directions for traffic prediction in limited-data regimes such as meta-learning, GNNs with embedded traffic knowledge, transportation foundation models, and LLM adaptation. Each approach has one thing in common—quick adaptation, data efficiency, broad generalization, or powerful reasoning—but also each has one thing that is lacking. The right choice depends on your specific situation—what data you have, what resources you do, and what you need—but we are optimistic that the next generation of traffic prediction tools will be more versatile, more efficient, and more reliable. They will be better suited for the messy reality of urban transportation.

References

1. Alanazi, F., Umar, I. K., Yosri, A. M., & Okail, M. A.: Comparative evaluation of deep learning and traditional models for predicting traffic accident severity in Saudi Arabia. Scientific Reports, 15, 32568. (2025)
2. Kumar, P. M., Basheer, S., Rawal, B. S., Afghah, F., Babu, G. C., & Arunmozhi, M.: Traffic scheduling, network slicing and virtualization based on deep reinforcement learning. Computers and Electrical Engineering, 100, 107987. (2022)
3. Lv, Y., Duan, Y., Kang, W., Li, Z., & Wang, F. Y.: Traffic flow prediction with big data: a deep learning approach. IEEE Transactions on intelligent transportation systems, 16(2), 865-873. (2015)
4. Zhang, J., Pi, Y., Sun, C., Wei, Y., Yu, F., & Yao, W.: Spatial-temporal encoder-decoder model for traffic flow prediction. Journal of National University of Defense Technology, 47(3), 173-182. (2025)
5. Zhao, L., et al.: T-gcn: A temporal graph convolutional network for traffic prediction. IEEE transactions on intelligent transportation systems, 21(9), 3848-3858. (2019)
6. Guo, T., & Su, L.: Dynamic Aware: Adaptive Multi-Mode Out-of-Distribution Detection for Trajectory Prediction in Autonomous Vehicles. arXiv preprint arXiv:2509.13577. (2025)
7. Sun, M., Ding, W., Zhang, T., Liu, Z., & Xing, M.: STDA-Meta: A Meta-Learning Framework for Few-Shot Traffic Prediction. arXiv preprint arXiv:2310.20223. (2023)
8. Li, M., Tang, Y., & Ma, W.: Few-Sample Traffic Prediction with Graph Networks using Locale as Relational Inductive Biases. In Proceedings of the 31st ACM International Conference on Information & Knowledge Management (CIKM '22). (2022)
9. Li, Z., Xia, L., Shi, L., Xu, Y., Yin, D., & Huang, C.: OpenCity: Open Spatio-Temporal Foundation Models for Traffic Prediction. arXiv preprint arXiv:2408.10269. (2024)
10. Ren, Y., Chen, Y., Liu, S., Wang, B., Yu, H., & Cui, Z.: TPLLM: A Traffic Prediction Framework Based on Pretrained Large Language Models. arXiv preprint arXiv:2403.02221. (2024)