


SURVEY PAPER

Open Access



Gap, techniques and evaluation: traffic flow prediction using machine learning and deep learning

Noor Afiza Mat Razali^{1*} , Nuraini Shamsaimon¹, Khairul Khalil Ishak², Suzaimah Ramli¹, Mohd Fahmi Mohamad Amran¹ and Sazali Sukardi³

*Correspondence:

noorafiza@upnm.edu.my

¹ National Defence University of Malaysia, Kuala Lumpur, Malaysia

Full list of author information is available at the end of the article

Abstract

The development of the Internet of Things (IoT) has produced new innovative solutions, such as smart cities, which enable humans to have a more efficient, convenient and smarter way of life. The Intelligent Transportation System (ITS) is part of several smart city applications where it enhances the processes of transportation and commutation. ITS aims to solve traffic problems, mainly traffic congestion. In recent years, new models and frameworks for predicting traffic flow have been rapidly developed to enhance the performance of traffic flow prediction, alongside the implementation of Artificial Intelligence (AI) methods such as machine learning (ML). To better understand how ML implementations can enhance traffic flow prediction, it is important to inclusively know the current research that has been conducted. The objective of this paper is to present a comprehensive and systematic review of the literature involving 39 articles published from 2016 onwards and extracted from four main databases: Scopus, ScienceDirect, SpringerLink and Taylor & Francis. The extracted information includes the gaps, approaches, evaluation methods, variables, datasets and results of each reviewed study based on the methodology and algorithms used for the purpose of predicting traffic flow. Based on our findings, the common and frequent machine learning techniques that have been applied for traffic flow prediction are Convolutional Neural Network and Long-Short Term Memory. The performance of their proposed techniques was compared with existing baseline models to determine their effectiveness. This paper is limited to certain literature pertaining to common databases. Through this limitation, the discussion is more focused on (and limited to) the techniques found on the list of reviewed articles. The aim of this paper is to provide a comprehensive understanding of the application of ML and DL techniques for improving traffic flow prediction, contributing to the betterment of ITS in smart cities. For future endeavours, experimental studies that apply the most used techniques in the articles reviewed in this study (such as CNN, LSTM or a combination of both techniques) can be accomplished to enhance traffic flow prediction. The results can be compared with baseline studies to determine the accuracy of these techniques.

Keywords: Traffic flow, Congestion prediction, Machine learning, Deep learning, Connected vehicles, Hybrid machine learning

Introduction

With the growth of IoT technology in recent years, innovations are more focused on making the world ‘smarter’ through the implementation of smart technology, such as smart cities, smart industries and smart transportation. The Intelligent Transportation System (ITS) is a revolutionary technology for smart transportation, a component of IoT. As mentioned in [1], ITS is part of the implementation of smart cities in IoT which aims to solve urbanisation issues that involve human mobility through the application of machine and deep learning (as an example, see Fig. 1). ITS mainly aims to offer creative and enhanced services for traffic management as well as for various transportation modes so travellers can commute in an organised, safe and ‘smart’ manner [2, 3]. Undoubtedly, technological developments have transformed systems and innovations to be more organised and structured. ITS is also intended to minimise traffic problems, mainly congestion, to enhance traffic efficiency, reduce travel time for commuters and improve the comfort and safety levels of travellers. Other than ensuring the efficient use of infrastructure and road safety, ITS provides other benefits such as transport networks for accessibility and mobility of travellers in an area. It also makes the best use of infrastructure and can offer a range of affordable and convenient transportation options. ITS further oversees incidents and congestions to provide road users and travellers the best level of service [4, 5].

Connected or automated vehicles (CV) are part of the technology developed under ITS. CV technology enables various vehicle categories (such as cars, buses, trucks, etc.) to communicate with each other, either through built-in or separate devices that are wirelessly connected to share valuable and important mobility and safety information [6]. The CV environment utilises data obtained through short-range communication broadcasts as well as peer-to-peer (P2P) exchanges within an acceptable parameter to detect activities being performed by other travellers and to identify potential warning hazards. CV allows travellers to safely arrive at their destination, quickly and cost-efficiently [7]. Apart from communicating with other vehicles, CV technology can also enable communication with road-side infrastructures to provide valuable notification and information.

Various well-known vehicle companies and manufacturers (such as BMW, Audi, Tesla and Volvo) have been involved in the development of CV technology. Each brand has introduced various innovations for their CV. BMW offers driver-focused technology,

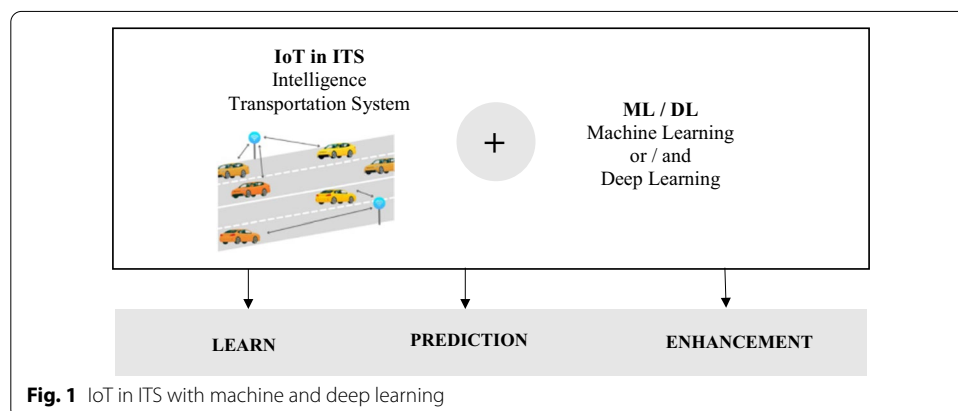


Fig. 1 IoT in ITS with machine and deep learning

digital services and a driver assistance system to assist drivers during their commute [8]. Audi designed a system that focuses on safety by reducing distractions for drivers [9]. Tesla, another growing manufacturer, has produced a vehicle with an advanced infotainment system, allowing drivers to fully utilise the technological functionality of their car [10]. Volvo has also focused on improving its infotainment system through a scheme called Sensus Connect [11].

With the advancement and development of CV, one of the main problems that researchers are still trying to solve is congestion [12]. Numerous works and research have been conducted to develop a prediction system for traffic flow and congestion. Distribution of congestion information is essential. With the development of CV, such information can be much easily distributed. However, before congestion information can be distributed, this study strives to determine how traffic information is generated through predictions and classifications. Since CV is part of IoT, the implementation of Artificial Intelligence (AI), such as machine learning (ML), is an innovative method that can provide a more reliable approach for producing and generating traffic flow predictions [13]. Figure 2 displays the relationship between the concepts of AI, ML and deep learning (DL) techniques. According to [14], DL is a subcategory of ML. They are closely related to each other in AI. Hence, this study presents a systematic review of traffic flow predictions using ML for CV. The introduction of CV promoted the use of ML methods to further enhance the capability of this technology. ML generates massive quantities of data, necessitating a better and faster data analysis which conventional methods cannot handle [15]. Examples of the real-life usage of ML applications, particularly for transportation, are seen through various applications such as Google Maps, Waze and HERE WeGo. They are useful platforms for users to plan and manage their travels. These applications collect big data through observations and users. The data is analysed using ML to provide real-time traffic data, as well as predictions of future events and traffic conditions, which are necessary information for users to efficiently plan their journeys. The authors in [16] proposed the use of a combination of big data and machine learning techniques to improve vehicle braking, thereby promoting eco-driving.

This systematic review focuses on the traffic flow prediction model using machine learning techniques in connected vehicles. Several research questions (RQ) have been established to develop ideas and guidelines for this study. The research questions are listed as follows:

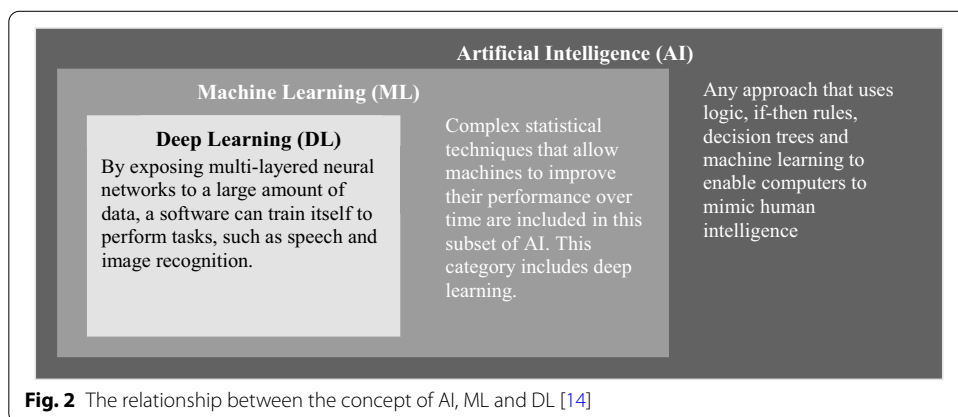


Fig. 2 The relationship between the concept of AI, ML and DL [14]

-
- RQ1 In studies on traffic flow prediction that implemented machine learning algorithms and techniques, what are the gaps found in each study which can be further improved in terms of methods, challenges, evaluation approach and results?
- RQ2 What are the main ML techniques and parameters used in the highlighted research, particularly for traffic flow prediction?
- RQ3 How is the effectiveness of the proposed methodology evaluated and how did the proposed solution influence traffic management in urban cities?
-

Numerous studies have proposed innovative techniques to establish congestion prediction models in CV that apply ML techniques, resulting in various outcomes and results. However, the recent literature still lacks systematic and structured research on the use of ML techniques in traffic flow prediction of CV. The aim of this structured review is to provide a comprehensive and dependable understanding of traffic flow predictions using ML techniques for CV, as well as to present a summary of relevant existing literature. For instance, a study by [17] had proposed the ANN based short term traffic volume prediction in India, describing the use of the ANN technique. The results and limitations revealed a research gap that can be further explored by other researchers. Another example would be the authors in [18] who proposed a Faster R-CNN detector to address the problem of traffic estimation using data from traffic videos.

Therefore, this study presents the different purposes, methodologies and outcomes from the list of available literature. It further highlights the frequently used ML techniques for traffic flow prediction as well as the variables and parameters involved. Lastly, this study determines the effectiveness of the methods that apply ML techniques and how they were evaluated.

“[Search methodology](#)” section presents the search methodology where the literature search strategy is explained. “[Gaps, techniques, evaluation and opportunities](#)” section discusses the results extracted from the literature review on traffic congestion prediction involving ML techniques. “[Conclusion, limitations and future work](#)” section displays the obtained results. Final section presents the conclusion, limitations of this current research and recommendations for future works.

Search methodology

Figure 3 displays the process of conducting a structural literature review (SLR), starting with problem recognition. During this phase, the problem statement is evaluated to determine the focus of this study and to drive the systematic review. This phase is important since the problem statement of each study should be known so as to ascertain its importance for discussion and how it can help enhance future works [19]. The next process is to create and generate the research questions once the problem has been determined. The research questions are the focus of this study in the path to learn answers and discuss key findings [20]. Literature search and analysis can be a challenging phase since researchers must study and review the existing algorithms, techniques and models of the focus research area. This



current study employs four databases to pinpoint the articles and journals that highlight the machine learning-based (ML) techniques widely used for prediction and analysis in CV, especially for traffic flow prediction. The results are subsequently explained based on the information gathered from the literature review. Lastly, the conclusion is constructed for this study.

Literature inclusion and exclusion criteria

Several criteria must be fulfilled for studies to be selected and considered in the systematic review:

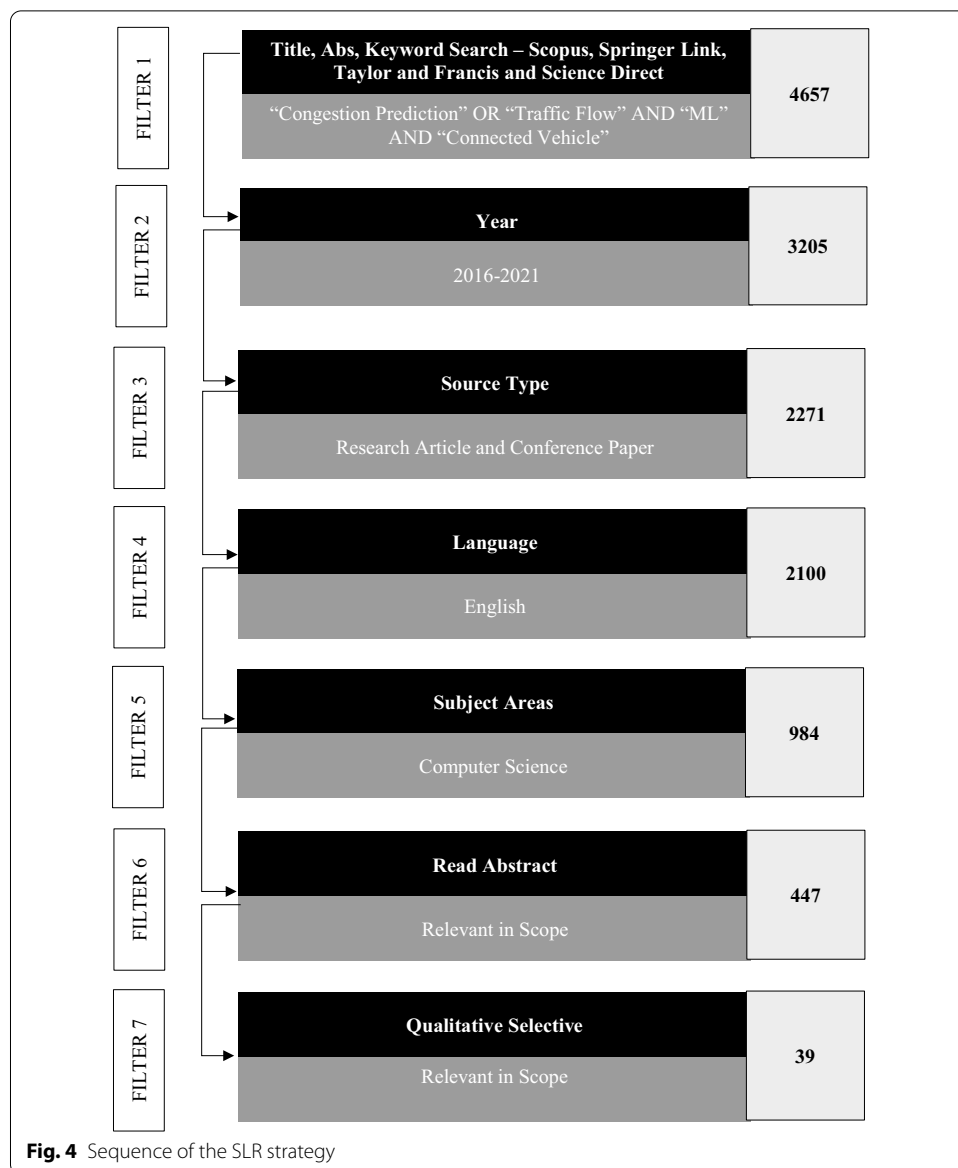
- Articles are indexed and published as journal articles and proceedings. Unpublished literatures are excluded.
- Articles are focused on 'traffic flow prediction using machine learning' within the domain and context of transportation and computer science. Articles on traffic flow prediction unrelated to transportation and computer science are excluded.
- Articles are easily accessible by the researcher, either through public/open access or institution access. Access that requires payment is excluded.
- Articles must contain basic sections of literature, namely, the abstract, introduction, methodology, results, discussion and conclusion. Articles that lack the required sections are excluded.
- Articles must be written in English. Articles written in languages other than English are excluded.

Study selection

Figure 4 presents the process of selecting the relevant studies for this research. This study initially conducted a search of keywords in the selected database, which produced 4657 search results. Due to the large number of results, the initial focus was on the first 3205 articles. A thorough preliminary screening from 2016 to 2021 produced 2271 articles based on source type filters, namely, research articles and conference papers. This study then focused on articles and papers written in English, producing 2100 results. Next, the computer science field was chosen which yielded 984 articles. 447 articles were then selected based on their title and relevance to the study, as assessed by their abstract. Subsequently, 245 articles were selected to be read in full. After thoroughly reviewing the full text, the results led to the final selection of 39 articles.

Gaps, techniques, evaluation and opportunities

A total of 39 articles have been extracted from four databases: Scopus, SpringerLink, ScienceDirect and Taylor & Francis. This was according to their relevance with the title and field of this study. The chosen studies are the result of various filters and the application of the inclusion and exclusion criteria, as previously mentioned in "Search methodology" section. The extensive review of these articles has led to the extraction of the main attributes; the gaps each study was trying to solve; the methodology, model or approach applied by the authors; the dataset used in the study; how the authors conducted the evaluation of their methodology; the parameters involved in the research;



the evaluation metric employed by the authors; and the results of the study. Table 1 displays the tabular analysis of these articles. The findings of this review are categorised into three groups: machine learning approach, the hybrid or ensemble machine learning approach and the machine learning approach with the inclusion of social media (which is further explained). Discussions are presented based on the research questions determined for this study.

Gaps, methods, challenges and evaluation methodology for traffic flow prediction

The aim of every research or literature is to discuss a certain topic or introduce and propose a solution to a problem, commonly known as gaps. Gaps are areas that can be further explored or improved to create a better understanding of the discussion. To answer the RQ1 formulated above, this study discusses the gaps found in each study based on

Table 1 (continued)

No	Authors	Model name	Machine learning and deep learning																									
			A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W			
28	Bartlett et al. (2019)	ML method comparative study					X							X													X	
29	Xu et al. (2018a)	Treating network status as a video for prediction of congestion level											X															
30	Shirazi and Morris (2018)	Feature collection system	X																									
31	Tampubolon and Hsiung (2018)	Supervised Deep Learning Based Traffic Flow Prediction (SDLTFP)										X																
32	Jin et al. (2018)	Spatio Temporal Recurrent Convolutional Network (STRCN)	X																									
33	Duan et al. (2018)	Deep hybrid neural network	X																									
34	Chen et al. (2018)	Fuzzy Deep convolutional Network (FDCN)	X																								X	
35	Kong et al. (2018)	Intelligent Traffic Recommendation System																										
36	Tian et al. (2018)	Traffic flow forecasting	X																									
37	Khan et al. (2017)	Framework that integrates CV with AI																								X		
38	Lawe and Wang (2016)	Deep-learning neural network for traffic flow optimization																										X
39	Wang et al. (2016)	Traffic Condition Estimation Integrated with GPS and tweet																								X		
Total			15	18	1	2	2	2	5	2	1	2	3	3	1	3	3	1	2	1	2	1	1	1	1	1	1	1

(A) Convolutional Neural Network [CNN]; (B) Long-Short Term Memory [LSTM]; (C) Random Forest [RF]; (D) XGBoost; (E) Artificial Neural Network [ANN]; (F) Graph Convolution Network [GCN]; (G) Sparse Autoencoder [SAE]; (H) Radial Basis Function [RBF]; (I) Hidden Markov Model; (J) Deep Neural Network [DNN]; (K) Recurrent Neural Network [RNN]; (L) C4.5; (M) K-Nearest Neighbour [KNN]; (N) Support Vector Machine [SVM]; (O) Fuzzy C-Means; (P) Random Subspace; (Q) Transpose CNN; (R) Support Vector Regression [SVR]; (S) Gated Recurrent Unit [GRU]; (T) Gradient Boost; (U) Fuzzy Network; (V) Case-Based Reasoning [CBR]; (W) Multi-Task Learning

the SLR results for traffic flow predictions which applied machine learning techniques. The purpose of each study, methods, challenges, evaluation approach and results are also presented in the following sections.

Gaps in SLR for Traffic flow prediction using machine learning

The research gap is an area that is either lacking or unexplored by previous authors, as listed in the second column of Table 1. Usually, authors would either try to develop a novel method/framework or enhance existing methods by proposing alternative techniques. Gaps are usually the motivation behind research and the inspiration for research expansion [21].

Each of the 39 articles selected in this literature review had reported a gap that they attempted to fill or solve, either by introducing new methodologies or by enhancing existing methods proposed in previous research. Ref. [22] stated that the current techniques for traffic prediction have failed to perform well due to the complex association features between road sections. Ref. [23] found that existing studies generally use similar traffic flow data of a city for the purpose of model training and testing. Ref. [24] discovered that the current models ignore and disregard the influence of weather conditions when making predictions on traffic flow. Ref. [25] claimed that a lack of literature is present regarding the prediction of congestion patterns in traffic zones or busy areas. Ref. [26] noted that existing LSTM models do not address the issue of enormous traffic flow data parallel to computing and distributed storage. Ref. [27] suggested that studies on traffic prediction should change focus towards the data-intensive era, which is currently lacking. Ref. [28] stated that current traffic classification methods are not solid against illumination conditions. Ref. [29] learned that studies which focus on time series for IoT traffic prediction are lacking. Ref. [30] found that existing studies have not effectively modelled or designed dynamic traffic patterns for irregular regions. Ref. [31] followed the current trend of growth on the use of LSTM techniques for traffic prediction. Ref. [32] stated that current approaches generally employ adjacency or distance matrix for the correlation between adjacent segments of the road, which have failed to dynamically capture spatiotemporal dependencies. Ref. [33] reported that previous studies have mainly focused on a single ML technique, causing the degeneration of traffic prediction accuracy to a certain extent. Ref. [34] found that most methodologies for prediction had focused on accuracy instead of immediacy. Ref. [35] concluded that traffic prediction study is unpopular because there is a lack of computationally efficient methods and algorithms, including good quality data. Based on the implementations of previous studies, [36] claimed that the performance of CNN for traffic prediction has been relatively unimpressive. Ref. [37] noted that there is a lack of research that considers road types as input for traffic prediction. Ref. [38] discovered that existing methodologies have only focused on a single road segment for making traffic predictions. Ref. [39] tracked the trend and growth of studies that have focused on traffic flow prediction using Application Programming Interface (API), such as Google Maps, for real-time traffic. Ref. [40] stated that previous studies have only considered the utility of a single factor, thus, there is currently a lack of research which considers multiple factors. Ref. [41] focused on studies that had only used analysed tweets for model training. Ref. [42] reported that existing methodologies with commonly known ML techniques do not consider the complexity

and heterogeneity of a traffic network system. According to [43], existing studies that addressed traffic time estimation issues are limited due to short-term or real-time predictions. Ref. [44] realised that it is a challenge to define and explain an appropriate traffic flow prediction, which heavily depends on data generation and collection. Ref. [45] perceived that former data mining algorithms typically address the issue using traditional mathematical and statistical theories, which made modelling spatial and temporal relationships impossible. Ref. [46] found that some researchers did not consider spatial and temporal relationships in deep learning methods. Ref. [47] discovered that existing models often suffer from a complex shift in factors and do not fully consider self-features and external factors of traffic flow. According to [48], existing studies that combined CNN and RNN techniques for traffic prediction were unable to collect the globality and connectivity of traffic networks. Ref. [49] lamented the current lack of traffic flow prediction studies on urbanised arterial roads. Ref. [50] reported the lack of studies that considered the overall road network status. Ref. [51] uncovered the current lack of studies that classify congestion using traffic camera images. Ref. [52] stated that existing studies which did not implement deep learning had eventually encountered difficulty when training a deep network due to complexity and time constraints. Ref. [53] noticed that existing studies have focused more on city regions instead of road sections for traffic flow prediction. Ref. [54] found that certain neural network techniques, such as CNN and LSTM, are only able to capture spatial or temporal data, respectively. Ref. [55] found that recent studies with DL models have failed to shed light on data uncertainty. Ref. [56] stated that there is concern over convergence and fitting due to the production of large data when discussing traffic flow prediction. Ref. [57] discovered that existing studies have only used valid data and excluded missing data when training network models, resulting in a relatively small training set size. Ref. [58] found several research works and studies that considered the use of the integrated CVT-AI method for traffic density estimation. Ref. [59] stated that existing studies on urban traffic optimisation are insufficient. Ref. [60] noticed that existing studies have generally focused on determining the application of social media as one of the major sources for traffic monitoring.

Techniques in SLR for traffic flow prediction using machine learning

The following section discusses the techniques for traffic flow prediction using machine learning based on SLR.

Machine learning approach

Most articles reviewed in this study had reported the implementation of a single machine learning approach through various strategies and methods. Ref. [61] proposed a model known as SG-CNN where the process of data training is optimised by an algorithm that groups road segments. By utilising the CNN algorithm, the proposed method revealed positive results compared to other baseline models. Ref. [22] proposed an approach based on multi-task learning for both spatial and temporal dependency extraction in multiple cities. Their model produced good results and showed superiority over other baseline models. In the study by [23], a model for congestion pattern prediction based on the Hidden Markov Model was proposed and compared with ARIMA, another well-known baseline model for traffic prediction. Overall, the proposed model performed

better than ARIMA. A distributed LSTM combined with a time window and normal distribution based on the MapReduce framework was proposed by [25], which also produced good results compared to existing LSTM models. Ref. [26] offered a machine learning-based framework for traffic speed prediction that compared three algorithms: CNN, LSTM-NN and XGBoost. The outcome revealed that XGBoost outperformed the other two algorithms. Ref. [62] developed an efficient automated system for congestion classification based on deep residual networks (CNN) and compact image representation. Ref. [63] proposed an LSTM-Deep Learning model for predicting IoT traffic in time series, in which the authors stated that the increased number of hidden layers provided the best accuracy. Ref. [64] proposed a Multitask Deep Learning Model (MTGCN) which used the graph convolution network (GCN) algorithm. A comparison between MTGCN and other baseline models have shown that the proposed method had the best performance. A traffic flow prediction model based on the LSTM network, known as LSTM_SPLSTM, was suggested by [30]. Ref. [65] proposed LGSTN, a local and global spatial-temporal network for traffic flow forecast on a road segment basis instead of regions. They evaluated LGSTN using two different datasets, with the proposed method performing well in both. Ref. [32] also used LSTM to propose a prediction model based on missing temporal and spatial values. This model was assessed in both urban and suburban areas. The proposed method produced better predictions in suburban areas compared to urban areas. Ref. [34] proposed a model that used a neural network, known as an Intelligent Traffic Congestion Prediction System, using Floating Car Data (FCD). This model was evaluated using same-region and cross-region data, and it was accurate in both. Ref. [37] developed a method called SSGRU which focused on multiple road segments for traffic prediction. The proposed method performed better in both urban and suburban areas compared to the Gated Recurrent Unit (GRU) and LSTM models. Ref. [38] conducted a case study to compare different machine learning algorithms for traffic congestion predictions and assessments using the Estimated Time of Arrival (ETA) based on the congestion index. Their results showed that Random Forest and XGBoost algorithms obtained the best results compared to other methods. Ref. [39] utilised the Graph Neural Network to develop TKGNN, which improved the accuracy more than the baseline models in unstable and fluctuating conditions. Ref. [42] utilised LSTM in their deep neural network model based on data analysis from a Vehicle Detection System. The results showed that in the case of high traffic density, the proposed model exhibited greater accuracy than in low traffic density, especially during weekdays. A model named C-LSTM was proposed by [44] as an end-to-end neural network model. Ref. [45] introduced a model based on a deep learning method called the varying spatiotemporal graph-based convolution model (VSTGC), which produced outstanding outcomes in different scenarios and timeframes. Ref. [46] introduced a CNN-based model (MF-CNN) with spatiotemporal features that is capable of performing network-scale traffic flow predictions. In a study by [47], a framework called MRes-RGNN was proposed based on multiple residual recurrent graph neural networks. It outperformed other existing baseline models. Ref. [48] conducted a comparative analysis experiment on machine learning methods for traffic flow prediction using three techniques (namely, k-NN, SVR and ANN) in which the latter outperformed the other techniques. Ref. [49] implemented the k-NN and C4.5 algorithms to predict traffic congestion by treating

network status as a video. A feature collection system was proposed by [66] to collect appearance-based features of traffic images where two machine learning techniques (DNN and CNN) were selected for evaluation. Based on the results, CNN performed with the highest accuracy. Ref. [51] proposed a supervised deep learning-based traffic flow prediction (SDLTFP), a fully connected deep neural network, with the application of Bayesian Network and Dropout for faster data training. Ref. [52] proposed an LSTM-based traffic recommendation and forecasting method to conduct two experiments. The proposed method outperformed all other methods. Refs. [56] and [57] introduced a framework that can increase the accuracy of real-time traffic condition assessment by integrating CV technology with AI methods, such as SVM and CBR. Ref. [58] developed a deep-learning neural network for optimising traffic flow.

Hybrid/ensemble machine learning approach

Several reviewed studies have implemented a combination of machine learning techniques in their methodology, producing a hybrid or ensemble learning approach. Ref. [59] introduced a framework that combines radial basis function (RBF) with stacked autoencoder (SAE) for traffic flow prediction, with the inclusion of weather influence. Their framework produced a higher prediction accuracy compared to other baseline models. Ref. [24] proposed an improved traffic state identification model based on selective ensemble learning of SVM, fuzzy C-means and random subspace. This model achieved better accuracy compared to other models tested in the study. Ref. [33] proposed a more efficient data collection strategy and a hybrid neural network model by combining LSTM, CNN and Transpose CNN to extract spatial and temporal information data. Ref. [35] developed a combination of GCN and LSTM-deep learning networks, which produced favourable results compared to other models. A deep ensemble-stacked LSTM (DE-LSTM) model was proposed by [36]. Although the proposed method failed to outperform other models during non-peak traffic testing data, it did perform better than LSTM. Ref. [43] suggested an urban traffic flow prediction based on spatial-temporal features using a combination of CNN and LSTM, which produced good results compared to other models. Ref. [46] introduced STRNC, a combination of CNN and LSTM techniques, to simultaneously capture temporal dependencies during the prediction of traffic flow. The STRNC was tested with different datasets, producing satisfactory results. Ref. [53] offered a deep hybrid neural network which was improved through the implementation of a greedy algorithm for tracing GPS in taxis. A combination of LSTM and CNN was also used for spatial and temporal dependencies. The results showed that the proposed model with the greedy algorithm performed better than other models. Ref. [54] introduced FDCN, a deep residual network model built with fuzzy theory that proved to be more powerful than existing models.

Machine learning approach with the inclusion of social media

Several studies have also developed machine learning approaches with the inclusion of social media data, such as tweets. Ref. [55] introduced the Extended Coupled Hidden Markov Model for the efficient integration of GPS probe readings and traffic-related tweets. The outcome of their study indicates that the proposed model performs better than other existing models without social media data. Ref. [60] proposed a deep learning

urban model combined with tweet information, adopting the SAE architecture and LSTM technique.

Evaluation from SLR for traffic flow prediction using machine and deep learning

Commonly used ML and DL techniques

From the extensive review of selected literature, this section addresses RQ2, which aims to list the commonly used ML and DL techniques. Based on the findings from the literature review, various ML techniques and algorithms have been employed for traffic flow prediction, such as SVM, CNN, LSTM, XGBoost, etc. The most common ML techniques used for predicting traffic congestion from 2016 to 2021 are **CNN** [41] and **LSTM** [22, 27, 35, 46, 47, 51, 53, 55, 63]. The combination or hybrid usage of these algorithms has also been reported in which CNN and LSTM were applied to extract spatial and temporal information, respectively [26, 34–36, 41, 43–46, 53, 54, 56, 57, 62, 64, 65]. The outcomes illustrate that CNN and LSTM are from the main ML algorithms that can be considered in future research, however, other ML techniques that perform equally well should not be excluded such as **Random Forest** [35, 46, 53, 54], **XGBoost** [39], **ANN** [27, 39] and **GCN** [49].

Parameters

Since the focus of SLR is on traffic flow prediction studies, the traffic datasets varied among different countries. Ref. [67] discussed several real-time traffic parameters such as average vehicle speed, object detection, distance calculation and vehicle counting that can be collected through multiple practises using data from traffic cameras. Table 2 presents the list of parameters used by each reviewed literature. Most of these datasets have produced common variables that can be applied to analyse and calculate traffic flow classifications, such as **traffic speed or velocity** [30, 36]; **traffic flow or number of vehicles that passed a certain point** [22, 25–28, 33, 34, 36, 37, 44, 45, 48, 50, 52, 54, 57, 59]; **traffic volume or the number of vehicles that crossed a road section per unit of time** [23, 27, 28, 31–33, 41, 46, 56, 68]; **time, time intervals, passage of time or time span** [22, 25, 26, 38, 44, 49, 52, 59]; and **traffic density or the average number of vehicles per kilometre of road space or area** [23, 25, 32, 37, 39, 40, 44, 46, 47, 50, 53, 55, 57, 58, 60]. Other variables, including **weather**, were also included in previous methodologies and considered as external factors [35, 41, 58]. These factors are capable of influencing the performance and outcome of the applied procedures.

Comparison with baseline model

Each reviewed study had compared their proposed model and ideas with other existing baseline models. To answer RQ3, which determines the effectiveness of the proposed models, it was found that most of the reviewed research had reported significant outcomes from their proposed methodology. This was accomplished through evaluations during their experimental setup when comparing with existing baseline models. These models have been used by numerous research works that reported satisfactory results, which were then later examined and enhanced by others. Examples of commonly compared baseline models are as follows [24, 39–41, 43, 45, 59]:

- Historical Average (HA): A time-series model that forecasts demands using values of previous demands at a given location during a similar time interval.
- Autoregressive Integrated Moving Average (ARIMA): A time-series model composed of a combination of autoregressive components and moving average in time-series modelling for traffic flow prediction.
- Support Vector Regression (SVR): A model that is based on linear SVM in performing regression tasks.

These baseline models are often compared with the proposed methodology to determine whether the suggested method is capable of performing at a similar or higher level. The other models used for comparison include ML algorithms.

Evaluation metric in SLR for traffic flow prediction using machine learning

In the surveyed articles, the evaluation metric is commonly applied to compare the proposed methodology with other baseline models. Table 3 presents a list of metrics used by the reviewed literature. Commonly used metrics include the Forecasting Value (FV), Root Mean Square Error (RMSE), Mean Absolute Error (MAE) and Mean Absolute Percentage Error (MAPE). The formulas are as follows:

Forecasting value (FV)

$$F_t = \frac{\sum_{t=1}^n (A_t + A_t)}{n}$$

Mean Absolute Error (MAE)

$$\text{MAE} = \frac{1}{n} \sum_{i=1}^n |A_t - F_t|$$

Root Mean Square Error (RMSE)

$$\text{RMSE} = \left[\frac{1}{n} \sum_{i=1}^n (A_t - F_t)^2 \right]^{\frac{1}{2}}$$

Mean Absolute Percentage Error (MAPE)

$$\text{MAPE} = \frac{\sum \frac{|A_t - F_t|}{A_t} \times 100}{n}$$

These metrics are used to compare between traffic flow prediction models. The lowest error value is the best value to determine the performance of each model in different scenarios.

Conclusion, limitations and future work

This work contributes to the expansion of research in the field of traffic flow prediction using ML and DL techniques. It adds to the literature and future studies by serving as a reference for other academicians, practitioners and researchers. This study provides a systematic and comprehensive review on recent research that implemented ML and DL

Table 3 (continued)

No	Authors	Evaluation metrics																
		A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q
27	Chen et al. (2019)	X	X	X														
28	Bartlett et al. (2019)		X															
29	Xu et al. (2018a)							X										
30	Shirazi and Morris (2018)							X										
31	Tampubolon and Hsiung (2018)	X	X	X														
32	Jin et al. (2018)		X															
33	Duan et al. (2018)		X															
34	Chen et al. (2018)	X	X														X	
35	Kong et al. (2018)		X														X	
36	Tian et al. (2018)	X	X														X	
37	Khan et al. (2017)																X	
38	Lawe and Wang (2016)		X															
39	Wang et al. (2016)		X				X											
Total		18	29	16	2	2	3	1	7	1	1	1	1	1	1	1	3	1

(A) Mean Absolute Error [MAE]; (B) Root Mean Square Error [RMSE]; (C) Mean Absolute Percentage Error [MAPE]; (D) Symmetric Mean Absolute Percentage Error [SMAPE]; (E) Mean Square Error [MSE]; (F) Precision; (G) L1; (H) Accuracy; (I) Correlation; (J) Coefficient; (K) Disagreement; (L) Double Fault Measure; (M) Recall; (N) Congestion Index [CI]; (O) R-Square; (P) Mean Root Error [MRE]; (Q) Mean Error [ME]

techniques as well as algorithms for traffic flow prediction. After a thorough selection process, a total of 39 articles were selected and comprehensively reviewed. The traffic flow prediction gap addressed in these articles include lack of computationally efficient methods and algorithms. Moreover, good quality data for data training are limited. Since similar traffic flow data of a city were used, this led to the utilisation of incomprehensive contents of data when training the network models. These factors have been determined to restrain the improvement of traffic flow prediction using ML and DL techniques. The gap is added by the failure of applying dynamically captured spatial and temporal relationships in deep learning due to complex association features between road sections and congestion patterns in traffic zones or busy areas. There is also a lack of computation and distributed storage, resulting in the limitation of time series for IoT traffic prediction. This issue must be addressed in future research.

Nevertheless, various methods, models and strategies have been introduced and proposed by researchers to determine the most accurate technique for traffic flow prediction. The application of machine learning techniques to predict traffic flow, especially CNN, LSTM and others, has shown positive results in enhancing the performance of the proposed methodologies such as C-LSTM, STRCN, MTN, etc. Comparisons with existing baseline models, such as ARIMA and HA, have further strengthened the idea behind the proposed models since the comparison results exhibited positive and enhanced outcomes.

This current study faced several constraints, such as being limited to only four main known databases (Scopus, ScienceDirect, SpringerLink and Taylor & Francis). Future research should include other literary sources and studies listed in more databases. This study was also limited to the techniques and algorithms found in the list of reviewed studies. There is a possibility that other techniques are present which are not listed in this study. For future endeavours, experiments that implement commonly used DL techniques (CNN and LSTM), extensively discussed in the reviewed articles, can be accomplished by utilising traffic data captured in various local urban regions to provide wider data patterns for model training. This will subsequently improve the prediction of traffic in local urban areas and further enhance the accuracy of ML and DL techniques for traffic flow prediction. However, cooperation to contribute the essential large data volume by local urban authority will be the main hurdle faced by researchers. Rules and regulations for traffic data sharing in local urban authority will also become another challenge. Connected IoT environments resulting from the process of setting up devices to capture traffic data for ML and DL training may contribute to cybersecurity issues as well. A framework must be established to address cybersecurity related matters for ITS in smart local urban cities. This will provide a broad area for future research. In our future efforts, we will extend this work to develop the IoT Monitoring Framework for Smart City Cybersecurity. The focus will be on ITS deployment for local urban cities. This will contribute towards the advancement of ITS technology in local urban smart cities. The evolution of ITS will establishes a more structured and effective management of transportation flow in smart cities.

Acknowledgements

This research is fully supported by National Defence University of Malaysia (UPNM) and the Ministry of Higher Education Malaysia (MOHE) under FRGS/1/2021/ICT07/UPNM/02/1. The authors fully acknowledge UPNM and the MOHE for the approved fund, which made this research viable and effective.

Authors' contributions

NAMR, suggested and conceived the core conception of this research work. NS presented the case study, findings and discussion. SR, MFMA and SS defined the overall organization of the manuscript. NAMR and NS has carried out thorough oversight of this work. The final manuscript was read and accepted by all contributors. All authors read and approved the final manuscript.

Funding

National Defence University of Malaysia (UPNM) and the Ministry of Higher Education Malaysia (MOHE) under FRGS/1/2021/ICT07/UPNM/02/1.

Availability of data and materials

All papers studied in this systematic review are available in SCOPUS, IEEE Xplore, ACM Digital Library, SPRINGERLINK and ScienceDirect. Please see the references below.

Declarations**Ethics approval and consent to participate**

Not applicable.

Consent for publication

Not applicable.

Competing interests

The authors declare that they have no competing interests.

Author details

¹National Defence University of Malaysia, Kuala Lumpur, Malaysia. ²Management and Science University, Shah Alam, Selangor, Malaysia. ³CyberSecurity Malaysia, Cyberjaya, Selangor, Malaysia.

Received: 18 July 2021 Accepted: 16 November 2021

Published online: 04 December 2021

References

- Kumar S, Tiwari P, Zymbler M. Internet of Things is a revolutionary approach for future technology enhancement: a review. *J Big Data*. 2019. <https://doi.org/10.1186/s40537-019-0268-2>.
- United Nations ESCAP. Intelligent transportation systems for sustainable development in Asia and the Pacific. 2021. pp. 1–38. <http://www.unescap.org/sites/default/files/ITS.pdf>.
- Gautam V. Analysis and application of vehicular ad hoc network as intelligent transportation system. *Lect Notes Netw Syst*. 2021;140:1–17. https://doi.org/10.1007/978-981-15-7130-5_1.
- Hassn HAH, Ismail A, Borhan MN, Syamsunur D. The impact of intelligent transport system quality: drivers' acceptance perspective. *Int J Technol*. 2016;7(4):553–61. <https://doi.org/10.14716/ijtech.v7i4.2578>.
- M. of works Malaysia. Malaysian its blueprint (2019–2023). 2013;53(9).
- Chang X, Li H, Rong J, Huang Z, Chen X, Zhang Y. Effects of on-board unit on driving behavior in connected vehicle traffic flow. *J Adv Transp*. 2019. <https://doi.org/10.1155/2019/8591623>.
- US Department of Transportation. U.S. Department of Transportation. 2020. <https://www.transportation.gov/research-and-technology/how-connected-vehicles-work>. Accessed 12 May 2021.
- BMW. "BMW," Digital Services, BMW. 2021. <https://www.bmw.com.my/en/topics/bmw-owners/bmw-connected-drive.html>. Accessed 13 May 2021.
- Audi. "Audi," Audi connect. 2021. <https://www.audi.com/en/experience-audi/models-and-technology/digital-services/audi-connect.html>. Accessed 13 May 2021.
- Tesla. Tesla. 2021. <https://www.tesla.com/support/connectivity>. Accessed 13 May 2021.
- Volvo. "Volvo," Volvo. 2021. <https://www.volvocars.com/us/own/connected-car>. Accessed 13 May 2021.
- Migliani A, Kumar N. Deep learning models for traffic flow prediction in autonomous vehicles: a review, solutions, and challenges. *Veh Commun*. 2019;20:100184. <https://doi.org/10.1016/j.vehcom.2019.100184>.
- Kamble SJ, Kounte MR. Machine learning approach on traffic congestion monitoring system in internet of vehicles. *Procedia Comput Sci*. 2020;171(2019):2235–41. <https://doi.org/10.1016/j.procs.2020.04.241>.
- Garbade DMJ. Clearing the confusion: AI vs machine learning vs deep learning differences. *Towards Data Science*. 2018. <https://towardsdatascience.com/clearing-the-confusion-ai-vs-machine-learning-vs-deep-learning-differences-fce69b21d5eb>. Accessed 17 May 2021.
- Bhavsar P, Saifro I, Bouaynaya N, Polikar R, Dera D. Machine Learning in transportation data analytics. In: *Data Anal. Intell. Transp. Syst.*, no. December, pp. 283–307. 2017. <https://doi.org/10.1016/B978-0-12-809715-1.00012-2>.
- Delnevo G, Di Lena P, Mirri S, Prandi C, Salomoni P. On combining Big Data and machine learning to support eco-driving behaviours. *J Big Data*. 2019;6(1):1–15. <https://doi.org/10.1186/s40537-019-0226-z>.
- Sharma B, Kumar S, Tiwari P, Yadav P, Nezhurina MI. ANN based short-term traffic flow forecasting in undivided two lane highway. *J Big Data*. 2018;5(1):48. <https://doi.org/10.1186/s40537-018-0157-0>.
- Fedorov A, Nikolskaia K, Ivanov S, Shepelev V, Minbaleev A. Traffic flow estimation with data from a video surveillance camera. *J Big Data*. 2019;6(1):73. <https://doi.org/10.1186/s40537-019-0234-z>.
- Nasution MKM, Onrizal, Aulia I. Design of the research problem statement. *J Phys Conf Ser*. 2019;1235(1):012115. <https://doi.org/10.1088/1742-6596/1235/1/012115>.

20. Bjørkvold T, Blikstad-Balas M. Students as researchers: what and why seventh-grade students choose to write when investigating their own research question. *Sci Educ.* 2018;102(2):304–41. <https://doi.org/10.1002/sce.21324>.
21. Khan KS, Kunz R, Kleijnen J, Antes G. Five steps to conducting a systematic review. *J R Soc Med.* 2003;96(3):118–21. <https://doi.org/10.1258/jrsm.96.3.118>.
22. Tu Y, Lin S, Qiao J, Liu B. Deep traffic congestion prediction model based on road segment grouping. *Appl Intell.* 2021. <https://doi.org/10.1007/s10489-020-02152-x>.
23. Zhang Y, Yang Y, Zhou W, Wang H, Ouyang X. Multi-city traffic flow forecasting via multi-task learning. *Appl Intell.* 2021. <https://doi.org/10.1007/s10489-020-02074-8>.
24. Hou Y, Deng Z, Cui H. Short-term traffic flow prediction with weather conditions: based on deep learning algorithms and data fusion. *Complexity.* 2021. <https://doi.org/10.1155/2021/6662959>.
25. Sun T, Huang Z, Zhu H, Huang Y, Zheng P. Congestion pattern prediction for a busy traffic zone based on the Hidden Markov Model. *IEEE Access.* 2021;9:2390–400. <https://doi.org/10.1109/ACCESS.2020.3047394>.
26. Xia D, et al. A distributed WND-LSTM model on MapReduce for short-term traffic flow prediction. *Neural Comput Appl.* 2021;33(7):2393–410. <https://doi.org/10.1007/s00521-020-05076-2>.
27. Romo L, Zhang J, Eastin K, Xue C. Short-term traffic speed prediction via machine learning. *Commun Comput Inf Sci.* 2020;1311:31–42. https://doi.org/10.1007/978-981-33-4532-4_3.
28. Abdelwahab MA, Abdel-Nasser M, Ichiro Taniguchi R. Efficient and fast traffic congestion classification based on video dynamics and deep residual network. *Commun Comput Inf Sci.* 2020;1212:3–17. https://doi.org/10.1007/978-981-15-4818-5_1.
29. Abdellah AR, Koucheryavy A. Deep learning with long short-term memory for iot traffic prediction. In: *Lect. Notes Comput. Sci. (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics)*, vol. 12525 LNCS, pp. 267–80. 2020. https://doi.org/10.1007/978-3-030-65726-0_24.
30. Wang F, Xu J, Liu C, Zhou R, Zhao P. MTGCN: a multitask deep learning model for traffic flow prediction. In: *Lect. Notes Comput. Sci. (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics)*, vol. 12112 LNCS, pp. 435–51. 2020. https://doi.org/10.1007/978-3-030-59410-7_30.
31. Lin Y, Wang R, Zhu R, Li T, Wang Z, Chen M. The short-term exit traffic prediction of a toll station based on LSTM. In: *Lect. Notes Comput. Sci. (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics)*, vol. 12275 LNAI, pp. 462–71. 2020. https://doi.org/10.1007/978-3-030-55393-7_41.
32. Qu Y, Zhu Y, Zang T, Xu Y, Yu J. Modeling local and global flow aggregation for traffic flow forecasting. 2020.
33. Wang Z, Chu R, Zhang M, Wang X, Luan S. An improved selective ensemble learning method for highway traffic flow state identification. *IEEE Access.* 2020;8:212623–34. <https://doi.org/10.1109/ACCESS.2020.3038801>.
34. Shin DH, Chung K, Park RC. Prediction of traffic congestion based on LSTM through correction of missing temporal and spatial data. *IEEE Access.* 2020;8:150784–96. <https://doi.org/10.1109/ACCESS.2020.3016469>.
35. Ranjan N, Bhandari S, Zhao HP, Kim H, Khan P. City-wide traffic congestion prediction based on CNN, LSTM and transpose CNN. *IEEE Access.* 2020;8:81606–20. <https://doi.org/10.1109/ACCESS.2020.2991462>.
36. Liu D, Hui S, Li L, Liu Z, Zhang Z. A method for short-term traffic flow forecasting based on GCN-LSTM. In: *Proc. - 2020 Int. Conf. Comput. Vision, Image Deep Learn. CVIDL 2020*, no. Cvidl, pp. 364–8. 2020. <https://doi.org/10.1109/CVIDL51233.2020.00-70>.
37. Elleuch W, Wali A, Alimi AM. Neural congestion prediction system for trip modelling in heterogeneous spatio-temporal patterns. *Int J Syst Sci.* 2020;51(8):1373–91. <https://doi.org/10.1080/00207721.2020.1760957>.
38. Sun P, Boukerche A, Tao Y. SSGRU: A novel hybrid stacked GRU-based traffic volume prediction approach in a road network. *Comput Commun.* 2020;160(April):502–11. <https://doi.org/10.1016/j.comcom.2020.06.028>.
39. Zafar N, Haq IU. Traffic congestion prediction based on estimated time of arrival. *PLoS ONE.* 2020;15(12 December):1–19. <https://doi.org/10.1371/journal.pone.0238200>.
40. Wang J, Chen Q. A traffic prediction model based on multiple factors. *J Supercomput.* 2020;77(3):2928–60. <https://doi.org/10.1007/s11227-020-03373-0>.
41. Essien A. A deep-learning model for urban traffic flow prediction with traffic events mined from twitter. 2020.
42. Ren Y, Xie K. Transfer knowledge between sub-regions for traffic prediction using deep learning method. In: *Lect. Notes Comput. Sci. (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics)*, vol. 11871 LNCS, pp. 208–19. 2019. https://doi.org/10.1007/978-3-030-33607-3_23.
43. Chou CH, Huang Y, Huang CY, Tseng VS. Long-term traffic time prediction using deep learning with integration of weather effect, vol. 11440 LNAI. Berlin: Springer International Publishing; 2019.
44. Yi H, Bui KHN. VDS data-based deep learning approach for traffic forecasting using LSTM network. In: *Lect. Notes Comput. Sci. (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics)*, vol. 11804 LNAI, pp. 547–58. 2019. https://doi.org/10.1007/978-3-030-30241-2_46.
45. Xu J, Zhang Y, Jia Y, Xing C. An efficient traffic prediction model using deep spatial-temporal network, vol. 2, pp. 386–99. 2019. <https://doi.org/10.1007/978-3-030-12981-1>.
46. Wang J, Cao Y, Du Y, Li L. DST: a deep urban traffic flow prediction framework based on spatial-temporal features. In: *Lect. Notes Comput. Sci. (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics)*, vol. 11775 LNAI, pp. 417–27. 2019. https://doi.org/10.1007/978-3-030-29551-6_37.
47. Yang D, Li S, Peng Z, Wang P, Wang J, Yang H. MF-CNN: traffic flow prediction using convolutional neural network and multi-features fusion. *IEICE Trans Inf Syst.* 2019;E102D(8):1526–36. <https://doi.org/10.1587/transinf.2018E DP7330>.
48. Chen C, et al. Gated residual recurrent graph neural networks for traffic prediction. In: *33rd AAAI Conf. Artif. Intell. AAAI 2019, 31st Innov. Appl. Artif. Intell. Conf. IAAI 2019 9th AAAI Symp. Educ. Adv. Artif. Intell. EAAI 2019*, pp. 485–92. 2019. <https://doi.org/10.1609/aaai.v33i01.3301485>.
49. Bartlett Z, Han L, Nguyen TT, Johnson P. A machine learning based approach for the prediction of road traffic flow on urbanised arterial roads. In: *Proc. - 20th Int. Conf. High Perform. Comput. Commun. 16th Int. Conf. Smart City 4th Int. Conf. Data Sci. Syst. HPCC/SmartCity/DSS 2018*, pp. 1285–92. 2019. <https://doi.org/10.1109/HPCC/SmartCity/DSS.2018.00215>.

50. Xu W, Yang G, Li F, Yang Y. Traffic congestion level prediction based on video processing technology. In: Lect. Notes Comput. Sci. (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics), vol. 10736 LNCS, pp. 970–80. 2018. https://doi.org/10.1007/978-3-319-77383-4_95.
51. Shirazi SM, Morris B. Traffic flow classification using traffic cameras, vol. 12112 LNCS. 2018.
52. Tampubolon H, Hsiung PA. Supervised deep learning based for traffic flow prediction. In: 2018 Int. Conf. Smart Green Technol. Electr. Inf. Syst. Smart Green Technol. Sustain. Living, ICSGTEIS 2018 - Proceeding, no. December 2019, pp. 95–100. 2018. <https://doi.org/10.1109/ICSGTEIS.2018.8709102>.
53. Jin W, Lin Y, Wu Z, Wan H. Spatio-temporal recurrent convolutional networks for citywide short-term crowd flows prediction. In: ACM Int. Conf. Proceeding Ser., pp. 28–35. 2018. <https://doi.org/10.1145/3193077.3193082>.
54. Duan Z, Yang Y, Zhang K, Ni Y, Bajgain S. Improved deep hybrid networks for urban traffic flow prediction using trajectory data. IEEE Access. 2018;6:31820–7. <https://doi.org/10.1109/ACCESS.2018.2845863>.
55. Chen W, et al. A novel fuzzy deep-learning approach to traffic flow prediction with uncertain spatial-temporal data features. Futur Gener Comput Syst. 2018;89(June):78–88. <https://doi.org/10.1016/j.future.2018.06.021>.
56. Kong F, Li J, Lv Z. Construction of intelligent traffic information recommendation system based on long short-term memory. J Comput Sci. 2018;26:78–86. <https://doi.org/10.1016/j.jocs.2018.03.010>.
57. Tian Y, Zhang K, Li J, Lin X, Yang B. LSTM-based traffic flow prediction with missing data. Neurocomputing. 2018;318:297–305. <https://doi.org/10.1016/j.neucom.2018.08.067>.
58. Khan SM, Dey KC, Chowdhury M. Real-time traffic state estimation with connected vehicles. IEEE Trans Intell Transp Syst. 2017;18(7):1687–99. <https://doi.org/10.1109/TITS.2017.2658664>.
59. Lawe S, Wang R. Optimization of traffic signals using deep learning neural networks. In: Lect. Notes Comput. Sci. (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics), vol. 9992 LNAI, pp. 403–15. 2016. https://doi.org/10.1007/978-3-319-50127-7_35.
60. Wang S, Li F, Stenneth L, Yu PS. Enhancing traffic congestion estimation with social media by coupled hidden Markov model. In: Lect. Notes Comput. Sci. (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics), vol. 9852 LNAI, pp. 247–64. 2016. https://doi.org/10.1007/978-3-319-46227-1_16.
61. Canales D, et al. Connected urban growth: public-private collaborations for transforming urban mobility. pp. 1–40. 2017. <https://www.mckinsey.com/business-functions/sustainability-and-resource-productivity/our-insights/public-private-collaborations-for-transforming-urban-mobility?cid=soc-web>.
62. Romo L, Zhang J, Eastin K, Xue C. Short-term traffic speed prediction via machine learning. 2020.
63. Abdelwahab MA, Abdel-Nasser M, Taniguchi R. Efficient and fast traffic congestion classification based on video dynamics and deep residual network. 2020.
64. Abdellah AR, Koucheryavy A. Deep learning with long short-term memory for IoT traffic prediction, vol. 9247. 2015.
65. Lin Y, et al. The short-term exit traffic prediction of a toll station based on LSTM, vol. 12275 LNAI. 2020.
66. Xu W, Yang G, Li F, Yang Y. Traffic congestion level prediction based on video processing technology. 2018; 1(92).
67. Khazukov K, et al. Real-time monitoring of traffic parameters. J Big Data. 2020. <https://doi.org/10.1186/s40537-020-00358-x>.
68. Crivellari A, Beinat E. LSTM-based deep learning model for predicting individual mobility traces of short-term foreign tourists. Sustain. 2020;12(1):349. <https://doi.org/10.3390/su12010349>.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Submit your manuscript to a SpringerOpen[®] journal and benefit from:

- Convenient online submission
- Rigorous peer review
- Open access: articles freely available online
- High visibility within the field
- Retaining the copyright to your article

Submit your next manuscript at ► [springeropen.com](https://www.springeropen.com)
