Turkish Journal of Electrical Engineering & Computer Sciences

http://journals.tubitak.gov.tr/elektrik/

Research Article

Turk J Elec Eng & Comp Sci
(2022) 30: 562 – 578
© TÜBİTAK
doi:10.3906/elk-2104-28

Long-term traffic flow estimation: a hybrid approach using location-based traffic characteristic

Tuğberk AYAR, Ferhat ATLİNAR, M. Amaç GÜVENSAN, H. İrem TÜRKMEN*
Department of Computer Engineering, Yıldız Technical University, İstanbul, Turkey

Received: 05.04.2021 • **Accepted/Published Online:** 27.09.2021 • **Final Version:** 21.03.2022

Abstract: Traffic speed estimation plays a key role in various situations, ranging from individual's trip planning to urban traffic management. Despite many studies on short-term prediction, there is only a limited number of studies focusing on long-term prediction and only a couple of them does go beyond 24 h. On the contrary, this study presents a novel hybrid architecture using location-based traffic characteristic for traffic speed estimation up to 7 days. In this architecture, the introduced mean filtering estimation (MFE) model and long short-term memory (LSTM) neural network are jointly utilized for minimizing the error for traffic flow estimation. Both MFE and LSTM utilizes the speed data, collected from roadside sensors in İstanbul, of previous weeks that have the same weekday and the same time with target time to be predicted. Results in this study indicate that the use of MFE gives lower error rates for locations with low traffic complexity while LSTM outperforms MFE model for locations with high traffic complexity. Thanks to the introduced MFE and the proposed hybrid architecture, we are able to predict the speed data of a given location with an error of lower than +/- 10 km/h.

Key words: Traffic flow estimation, long term traffic speed estimation, long short-term memory, mean estimation, standard deviation

1. Introduction

The correct estimation of traffic flow has been drawn attention for a long time due to its importance in our daily life. Traffic speed predictions bring many opportunities such as for individuals to better plan their routes or for traffic management agencies to prevent congestion resulting in saving time. Even though short-term traffic flow estimation problem is examined frequently, there are a limited number of studies that put their focus on long-term prediction have been reported so far. In this study, we propose a novel hybrid system that performs long term traffic flow prediction by combining a well known deep leaning architecture LSTM with a new statistical model, so called MFE.

The proposed architecture in this study is based on the assumption that the traffic flow characteristics of a location for a given day and time, can be predicted using the traffic flow data of the same day and time of the previous weeks. We first introduce a simple MFE model, which calculates the mean of previous weeks' data, for long-term traffic flow estimation. This model reveals an outstanding performance despite its simplicity. However, our experimental results demonstrate that MFE suffers at locations with complex traffic characteristic. Thus, models were trained for each individual location and forecasts were made up to 7 days ahead using the introduced MFE model and the state-of-the-art solution LSTM neural network [1, 2] both. The

*Correspondence: irem@yildiz.edu.tr

This work is licensed under a Creative Commons Attribution 4.0 International License.

results demonstrate that MFE gives better estimations for the locations with relatively less traffic complexity which means traffic characteristic of these locations do not fluctuate too much during time. On the other hand, LSTM gives considerably better results than MFE for the locations with high traffic complexity, i.e uncertain traffic characteristic. To smooth out the previous week's data in order to reduce the effect of the sudden random increases and decreases, time windows are used for both methods with different experimentally determined sizes. In addition to window size parameter, experiments were done with different numbers of previous weeks for MFE, while it is chosen as 2 for LSTM test scenarios based on experimental results. As a result, a novel hybrid architecture is proposed in this study which takes advantage of both algorithms MFE and LSTM. It uses a threshold value to select the proper algorithm based on the traffic flow complexity respect to the speed data at a given location. To the best of our knowledge, in the literature, no study was able to achieve less than +/- 10 km/h error rate for the traffic flow prediction of upcoming week in five minutes resolution. Experimental results show that due its simplicity and robustness the proposed architecture is a promising approach for long-term traffic flow estimation up to 7 days. The key points of the proposed system are given as follows:

- The MFE method having a very low computational complexity is presented for long-term traffic flow estimation.
- A successful approach which is able to make prediction for the one of the longest horizon (7 days ahead) and the highest resolution (5 min) in the literature is introduced.
- For the first time in the literature, a long-term traffic flow estimation system could perform with a +/- 10 km/h error using the combination of simple and complex methods based on the local characteristics of traffic.
- An objective measurement metric that can distinguish locations with more predictable speed characteristic from the complex ones has been proposed.

In Section 2, we first discuss the available solutions on traffic flow forecasting. Then, we introduce the details of our novel hybrid long-term traffic flow estimation algorithm in Section 3. In Section 4, we present our detailed test results. Finally, we conclude our paper in Section 5.

2. Related work

Traffic flow estimation is one of the real-life problems especially for metropolitan cities that has been studied frequently in recent years. These studies can be basically divided as short-term and long-term estimation from the time perspective whereas they also can be categorized into two groups solutions exploiting statistical information or neural network-based solutions from the methodology perspective.

Literature review shows us that research on short-term estimations have dominated the field since the challenge is slightly high for accurate long-term traffic flow estimation. Autoregressive integrated moving-average (ARIMA), various regression models [3, 4] and artificial neural networks [5–7] are most frequently used techniques for short-term prediction. The success of deep learning in modeling complex and nonlinear systems has led to the studies using deep learning approaches for traffic flow prediction in recent years [8–10]. Yufang et al. proposed a prediction system that combines back propagation and LSTM based on the road type (suburb, freeway, city) [11]. Even though they did not mentioned an exact prediction horizon, it is stated that they could predict the entire route before driving.

There is a few number of studies in the literature that focused on medium or long-term traffic flow estimation [12]. Zhao et al. proposed LSTM Networks for 15, 30, 45 and 60 min prediction horizons [13]. In [9], researchers proposed a hybrid model which uses recurrent neural networks (RNN) and convolutional neural networks (CNN) to predict traffic flow speed up to 45 min ahead. On the other hand, Wang et al. exploited a bidirectional LSTM (BDLSTM) model to forecast traffic speed up to 30 min [1]. Lu et al. proposed a novel graph LSTM framework in order to make predictions by modelling spatial-temporal dependencies in a one hour horizon [14]. They achieved a mean absolute percentage error (MAPE) of 29.5% and a MAPE of 66.5% in Xi'an and Beijing datasets, respectively. In [15], traffic flow estimations were made for up to 90 min using statistical approach with enhanced k-nearest neighbor (Enhanced k-NN) algorithm. Li et al. employed a type-2 fuzzy LSTM model for long-term traffic volume prediction [16].

In a recent article, a graph based CNN-LSTM model is trained using global positioning system (GPS) trajectory data for long-term traffic forecasting up to 4 h [17]. Chen et al. estimated the traffic flow rate for up to 6 h exploiting fuzzy deep convolutional neural networks model on the GPS data [18]. [19, 20] are two studies giving predictions up to 24 h using neural models where both studies exploit a combination of CNN and RNN algorithms. Another study that predicts next day's traffic flow is presented by Li et al. [21]. The researchers used a hybrid method that combines wavelet decomposition with CNN and LSTM. They stated that decomposing the original traffic data improves prediction accuracy while employing combination of CNN and LSTM enables a better performance on capturing and learning the long-term temporal features. They compared the performance of the proposed approach with stand alone LSTM and CNN methods and got better results. He et al. proposed spatio-temporal convolutional neural network (STCNN) which can capture general spatio-temporal traffic dependencies and the periodic traffic pattern exploiting Skip-ConvLSTM model [22]. The Skip-ConvLSTM extracts the periodic characteristics from skipped history traffic data which is very important for long-term traffic predictions. They conducted their experiments on TaxiBJ and BikeNYC datasets. They achieved a mean absolute error (MAE) value of 0.92 and a MAE value of 3.18 for BikeNYC and TaxiBJ datasets, respectively.

In addition to the studies using speed data and GPS for traffic flow speed forecasting, studies that also evaluate the environmental factors which may affect the traffic flow have been published recently. Peng et al. performed traffic flow forecasting up to 24 h for scenarios with and without rainfall data using seasonal ARIMA, exponential smoothing and feedforward neural networks methods. The seasonal self-connected integrated walking average method achieved the most successful result with a mean absolute percentage error of approximately 17% in the tests performed to predict the traffic flow after 24 h [23].

Belhadi et al. used RNN to predict the long-term traffic flow from multiple data sources [24]. In addition to traffic flow data, they also exploited weather condition and contextual information such as being weekend day and event day. They aimed to predict the number of the vehicles passing through a location during a given time interval. Predictive rate, which is defined as the number of long-term traffic flows that are correctly predicted over the tested ones is used as evaluation metric. A predictive rate of up to 80% is achieved.

In a recent study, Simunek et al. present an ensemble long-term traffic prediction model that combines parametric and nonparametric approaches including linear regression and case-based reasoning (CBR) [25]. They predicted the traffic flow within a week period exploiting weather information, calendar data and the features of road segment such as length of the segment, number of public transportation stops etc. Average MAE of all individual predictions was calculated as 4.67. Guo et al. proposed a double graph convolutional

neural network to predict traffic flow rate in peak hours by exploiting external factors such as working days, accident sings [26]. In a recent work researchers used fully BDLSTM for traffic volume prediction up to 2 h. They combined BDLSTM with an attention mechanism to capture the temporal shifting in the traffic volume and also exploited external features, including weather conditions and events [27].

Studies which perform long term traffic flow prediction (from 6 h to 7 days) are given in Table 1. Examining Table 1, it could be seen that there are only two studies that make long-term forecasts up to 1 week [22, 25]. The study of Simunek et al. is quite promising by their low MAE rate of 4.67 [25]. However, in order to obtain high success rates, very detailed information about road and weather conditions should be provided. These requirements complicate adapting the system to a new city or a new database. Another drawback of this study is working on low resolution data. Researchers sampled the traffic flow rate down to 1 hour resolution. In our proposed method, the system automatically determines the appropriate model according to the traffic flow rate characteristic without any need of additional data sources and works on high resolution (data frequency of 5 min). This facilitates the practical and more accurate use of the proposed system on different datasets. In the other study that make predictions up to 1 week, the researchers explore the useful periodic traffic patterns by employing ConvLSTM over skipped spatio-temporal traffic data [22]. Although it is not possible to make a fair comparison since it has been studied with different datasets, achieved success rates show a proper efficiency of the system in estimating long-term traffic flow rate. However, the time interval of the two datasets that used in experiments, TaxiBJ and BikeNYC, are 30 min and 1 hour, respectively. The low resolution of the results reduces the applicability of the system for a very dynamic city like Istanbul.

Table 1. State of the art studies about long term traffic flow rate prediction.

Study	Prediction horizon	Database	Data type	Method	Success rate
[18]	Up to 6 h	TaxiBJ	GPS	Fuzzy Deep CNN	MAE of 0.02, RMSE of 0.33
[23]	Up to 24 h	Georgia Department of Transportation	Traffic Flow Rate	Seasonal Self-Connected Integrated Walking Average	MAPE of 17%
[21]	Up to 24 h	Highways Agency Network Traffic Flow Data	Traffic Volume and Temporal Information	W-CNN-LSTM	MAE of 268.359, RMSE of 420
[24]	Up to 24 h	Odense Kommune (Denmark)	Traffic Flow, Weather Condition and Contextual Information	RNN	Predictive rate of up to 80%
[25]	Up to 7 days	Road and Motorway Directorate of the Czech Republic	Traffic Flow Speed, Weather Conditions, Road type etc.	Combination of CBR and Linear Regression	MAE of 4.67
[22]	Up to 7 days	TaxiBJ and BikeNYC	GPS, Bike rent	STCNN	MAE of 3.18 and 0.92, RMSE of 4.08 and 1.36

Apart from other studies on traffic flow speed estimation, our proposed architecture exploits traffic characteristic of a given location for long-term traffic prediction. This approach can also be applied to other systems that make traffic flow prediction using different methods. It is capable of employing the most effective algorithms that will work in different locations and thus obtaining the highest possible efficiency from the system.

3. Methodology

In this study, we introduce a hybrid model for long-term traffic speed estimation exploiting the traffic pattern complexity of a given location. The proposed model is constructed over the assumption that traffic flow characteristic of a location follows a similar trend for the same day and day part of past consecutive weeks. In order to capture this similarity, a simple statistical model, namely MFE model and a nonlinear predictor, LSTM network, are combined in an architecture, as shown in Figure 1, to decrease the error rate of long-term traffic speed estimation.

3.1. Preprocessing and model selection

The available raw traffic data needs to be preprocessed for further use in any of the proposed algorithms. The first step is to eliminate outliers and combine the velocities that are measured every minute into an average for a five-minute interval. Having done that, the standard deviation of all data points over a certain time frame can be computed. Standard deviation is a key parameter for indicating the traffic characteristic of a specific location and it tells us about the level of traffic complexity on that location. Higher standard deviations refer to a more complex traffic characteristic, whereas lower standard deviations mostly stand for predictable and certain traffic conditions. Thus, we chose the standard deviation as the determining factor whether the statistical model or the neural network approach should be used to make a prediction.

3.2. Mean filtering estimation algorithm

MFE has a simple working methodology which calculates the mean of the input data as output. Despite its simple formula, MFE is a powerful algorithm as it smooths the input data which works fine with time dependent problems. Using a time window helps to smooth out the data which prevents a high impact of sudden random changes in the data. The formula of MFE is given in Equation 1. In this equation, D_t^w denotes the speed data, w week before the forecast day, and t time steps away from the forecast time. k represents how many previous weeks are included to the prediction, i represents how many time steps before and after the prediction are included.

$$Prediction = \frac{1}{k(2i+1)} \sum_{w=1}^{k} \sum_{t=-i}^{i} D_t^w \tag{1}$$

Nevertheless, it lacks capturing an increasing or decreasing trend for time series along with not being able to figure out complex characteristics. Since the prediction error of MFE increases for locations with a complex traffic characteristic, we exploit long short-term memory network for such locations in order to benefit from its power on nonlinear problems.

3.3. Long short-term memory network

LSTM, as a member of RNN family, is widely used for time series problems and yields successful results [28]. Having the capabilities of an RNN algorithm, LSTM additionally has an advantage which gives the algorithm it's name: long short-term memory. This additional feature takes this algorithm one step ahead of classic RNN algorithms by solving the vanishing gradient problem. Vanishing gradient problem can be simply explained as gradients becoming extremely small during the back propagation. This reduces the learning of network by a

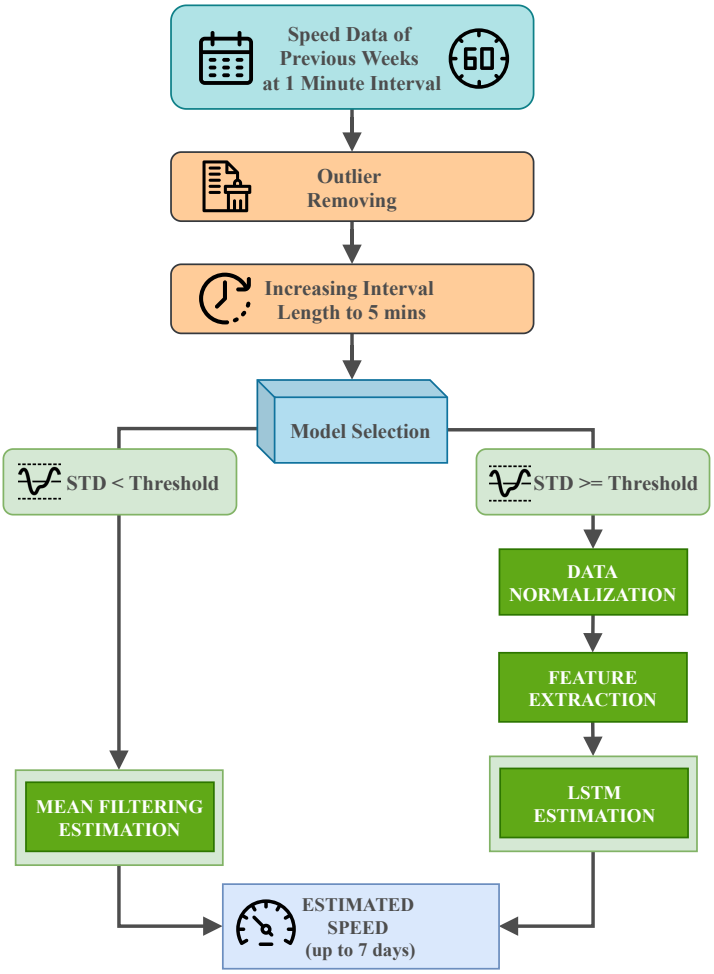


Figure 1. The proposed hybrid architecture consists of mean filtering estimation (MFE) model and long short-term memory(LSTM) network, where standard deviation of the speed data is a great indicator to designate the appropriate model in order to minimize the mean absolute percentage error.

great deal. Having a short term memory which includes gates that make algorithm capable of deciding which previous input to forget or to keep as it is illustrated at the left side of Figure 2.

$$LSTM\ Input\ Matrix: \begin{bmatrix} D^1_{-i} & D^2_{-i} & \dots & D^k_{-i} & \overrightarrow{dv_{-i}} \\ D^1_{1-i} & D^2_{1-i} & \dots & D^k_{1-i} & \overrightarrow{dv_{1-i}} \\ \vdots & \vdots & & \vdots & \vdots \\ D^1_{i-1} & D^2_{i-1} & \dots & D^k_{i-1} & \overrightarrow{dv_{i-1}} \\ D^1_{i} & D^2_{i} & \dots & D^k_{i} & \overrightarrow{dv_{i}} \end{bmatrix} \tag{2}$$

A simple illustration of used LSTM network is given in the right side of Figure 2. The network consists of 3 layers which include varying number of LSTM units. Even though MFE and LSTM consist of many common steps for the data preparation, LSTM additionally includes min-max normalization. Furthermore, the "partOftheday" feature, which indicates the exact four-hour interval among 6 units in a particular day, is added to the normalized data. This feature represents which part of the day, including the following hours (2-6,

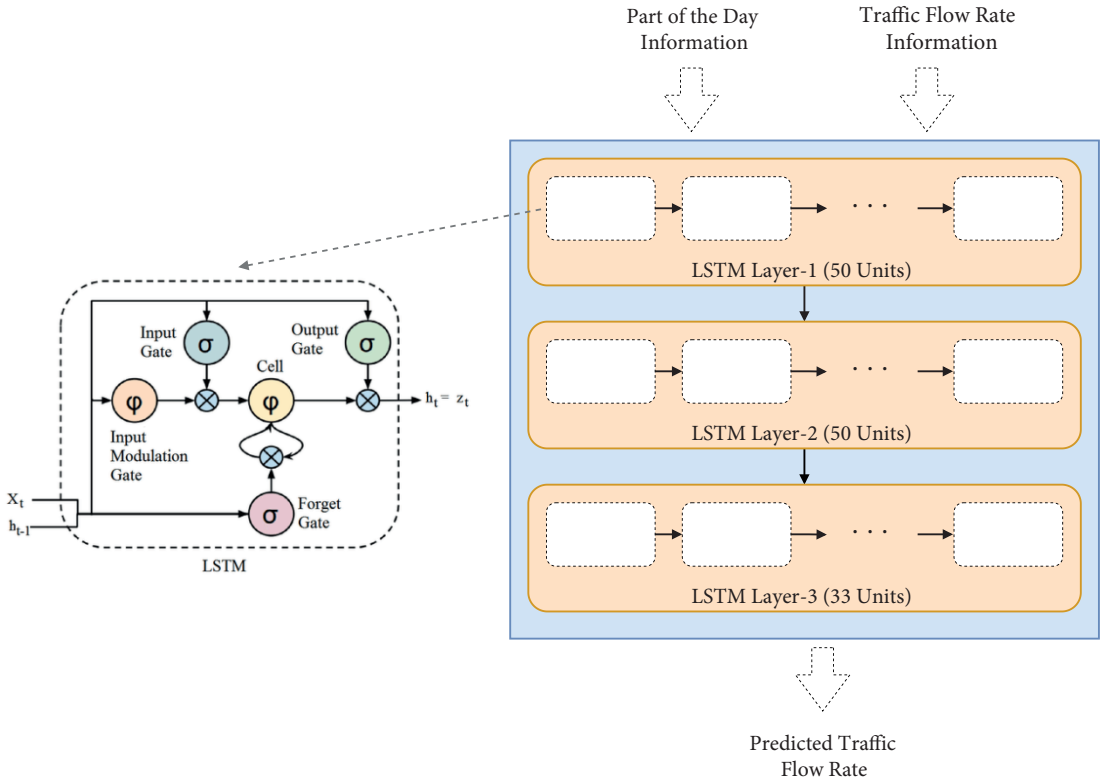


Figure 2. LSTM network for traffic flow rate prediction.

6-10, 10-14, 14-18, 18-22, 22-2), the data point belongs to. Afterwards, LSTM input data is formed using the traffic flow data of the previous weeks in combination with the daytime one hot encoding vector. The matrix in Equation 2 represents the data given to the LSTM model. In this matrix, D_i^k stands for the traffic speed data of the k weeks before and i time steps further than the target time to be forecasted. Finally, $\overrightarrow{dv_i}$ is a one hot code vector that represents the "partOftheday" feature of the corresponding time step.

4. Experimental results

In this section, we first introduce our dataset, the test environment and train/test parameters. Afterwards, we demonstrate the performance of the MFE and LSTM models, comparatively. We then present the relationship between MAPE values and standard deviation for both MFE and LSTM models. Finally, we compare the performance of the proposed hybrid approach against the well-known approaches including support vector regression and polynomial regression.

4.1. Experimental setup

The data set used in this study is provided by İstanbul Metropolitan Municipality and is gathered from road side sensors with a measurement frequency of one minute. The data consists of four features id of the sensor, the flow direction, recording time and flow speed. Although the original data set including data both from 2016 and 2017, we exploit data in 2017 since data records in 2016 are incomplete. The raw data requires a few preprocessing steps, including outlier detection and downsampling operations for better estimations. For

the outlier detection, a window with a size of 20 min is convoluted over the data and the data records farther than twice the standard deviation from the mean are removed. Following the outlier detection, data frequency is decreased to 5 min with the purpose of having a more complete data. Finally, having less than 10 percent missing data, there are only 187 left from 7152 sensor locations. The details of the introduced ready-to-use data set are given in Table 2.

Table 2. The details of our ready-to-use data set prepared by using the data provided by İstanbul Metropolitan Municipality.

Parameter	**Value**
Number of segments	187
Data resolution before preprocessing	Every minute
Data resolution after preprocessing	Every 5 min
Location	The main arterial roads in İstanbul
Time interval	from January to December in 2017
Features	SegmentID, direction, speed, time

In this study, we ran several tests to demonstrate the effectiveness of our hybrid approach consisting of the proposed statistical MFE model and an LSTM neural network. The error metric for these experiments is chosen as mean absolute percentage error (MAPE) whose formula is illustrated in Equation 3. In the equation, n represents the sample size while y_i and $\bar{y}_i$ represents the actual and predicted values respectively for each of the data point to be predicted. MAPE is calculated as the average of the absolute percentage errors. On the other hand, all the details about our train/test parameters both for MFE algorithm and LSTM network are given in Table 3.

$$MAPE = \frac{1}{n}\sum_{i=1}^{N}\frac{|y_i - \bar{y}_i|}{y_i} \tag{3}$$

4.2. Performance of MFE algorithm

For the mean filtering estimation model, various experiments are conducted with different hyperparameters. The results in Figure 3 show us how the number of previous weeks and the window size effect the error rate. In these experiments, the number of previous weeks are selected from 1 to 4, whereas for each week, the window size starts with 0 min and ends with 120 min with a 10-min increase at each iteration. The MAPE for MFE is calculated by taking the average error rates of all 187 sensor locations. The best results are obtained when the number of previous weeks and the window size are chosen as 3 and 40, respectively. In general, Figure 3 demonstrates that the mean absolute percentage error is decreasing between the window size 0 and 60, and starts to increase after 60. Including only the data of last week is not giving satisfactory results. On the other hand, analyzing only the window size parameter shows us that ”60” is the best window size independent from the number of previous weeks which indicates that smoothing the data is beneficial. The results are the average error of 187 different locations for the whole year of 2017 and the minimum MAPE value is calculated as 0.175 for the MFE model.

Table 3. The train/test parameters and the selected hyperparameters of the proposed LSTM and MFE model.

LSTM hyperparameters	
Parameter	**Value**
Train period	2 months prior than the test month
Test period	1 month
Tested months	10 months from March 2017 to December 2017
Size of train data	17280 instances (60 days × 24 h × 12 instances per hour)
Size of test data	8640 instances (30 days × 24 h × 12 instances per hour)
Total number of models	1870 (187 segments × 10 months)
Batch size	2048
Number of epochs	50
Number of layers	3
Number of units	50, 50, 33
Dropout rate	0.2
Window size	120 min (+/- 60 min)
MFE hyperparameters	
Parameter	**Value**
Train period	Unnecessary
Test period	10 months from March 2017 to December 2017
Window size	40 min
Number of previous weeks	3
Total number of instances for estimation	27 instances (3 weeks x 9 instances per week)

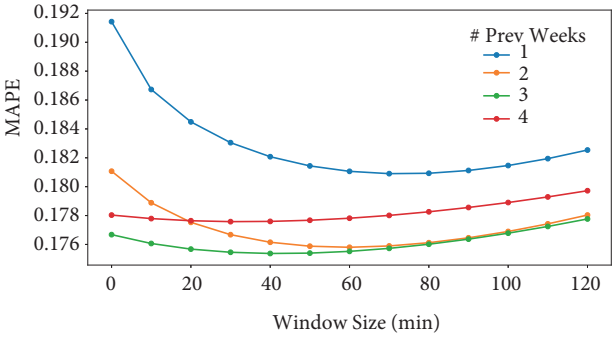


Figure 3. The MAPE values of mean filtering estimation model for both the number of different window sizes and the number of previous weeks.

4.3. Performance of LSTM network

In addition to the general data preprocessing steps, min-max normalization is applied before the LSTM training and the data which is fed into LSTM is refined by categorizing it into different parts of the day and the corresponding part of the day of a particular time step is represented with a one hot encoding vector. The daytime represents which part of the day the data belongs to. Each day is divided into six equal parts of 4 h length. The window sizes used for the LSTM models are 10, 30, 60, 90, 120, and 240, respectively.

LSTM models are trained with the data of 2 consecutive months and tested over the data of third consecutive month with a single month wide sliding window for the whole 2017. When the error rate is calculated for LSTM, all of the estimations of each location are concatenated and the MAPE value is calculated using these estimations. The used LSTM model consists of 3 hidden layers with 50 units for the first two layers and 33 for the third one. The dropout value is chosen as 0.2. In Figure 4, it is shown that an increase in window size decreases the error rate. Even though the error decreases with increasing window size, the computation time rises by great amounts as well. Therefore, based on the elbow technique, a window size of 120 is chosen which yields a MAPE value of 0.186.

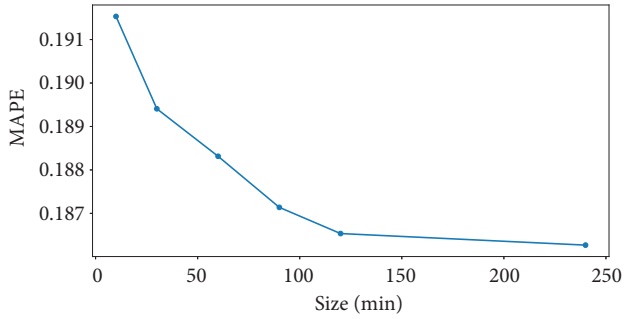


Figure 4. The MAPE values of LSTM model for the number of different window sizes.

We also analyzed the relationship between train/validation performance of our proposed LSTM network and the traffic pattern complexity of a given location. Figure 5 demonstrates that locations with higher complexity traffic characteristic present a bigger gap between train and validation loss values on behalf of train loss values. This outcome strengthens our hypothesis that standard deviation is a good metric to choose a proper traffic flow estimation algorithm.

4.4. Comparison of MFE and LSTM

In overall, MFE has a lower minimum MAPE value compared to LSTM. However, this does not mean MFE is better than LSTM at each sensor location. Traffic forecasts are especially important for locations where the traffic flow speed data is hard to be predicted. The more complex the traffic characteristic of a location is, the harder it is to forecast traffic speed accurately. In this study, we recommend to exploit the standard deviation of traffic speed for indicating the traffic complexity of a specific location. Figure 6 demonstrates the error rates related to standard deviation for MFE and LSTM, respectively. In this figure, each point belongs to a different location. It is important to note that the error rates are higher at locations with higher standard deviation values. Despite the fact that both methods have higher error rates for locations with higher standard deviation values, LSTM errors are lower at those points compared to MFE.

Figure 7 illustrates the error rates for each location. The MAPE values for MFE and LSTM approaches are represented at x axis and y axis, respectively. This figure demonstrates that at locations with lower error rates, MFE is usually more successful than LSTM.

Figure 7 shows that a model selection should be based on the standard deviation at a given location. Thus, it emerges the need of a threshold. The proposed architecture exploits MFE for locations that have a lower standard deviation than the threshold and LSTM for the locations with a higher values. The overall error rate is calculated from the average of all location error rates. The final threshold is decided after calculating

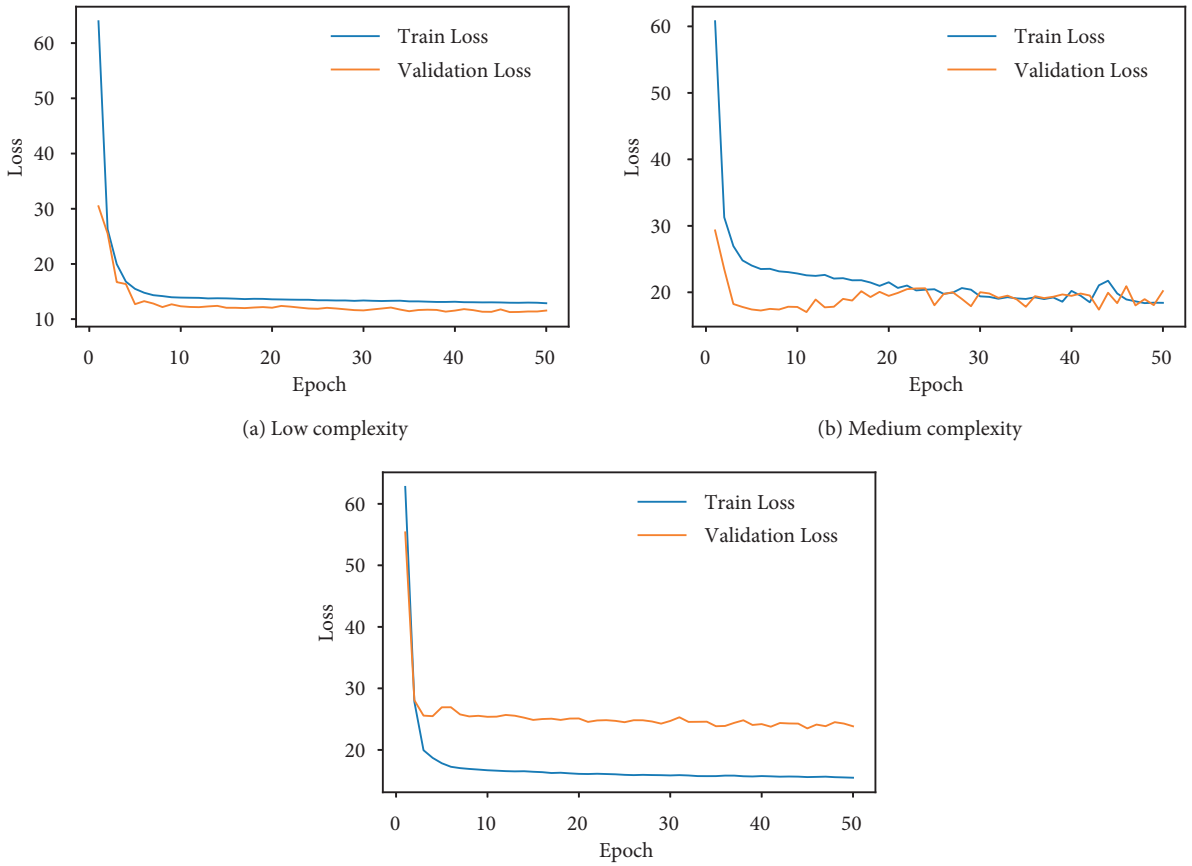


Figure 5. Train and validation losses for locations with low, medium and high complexity of traffic characteristic.

the architecture error rates by using each unique standard deviation of location points as threshold value of the architecture and is found to be 24. Figure 8 shows the error rates of the proposed architecture for different standard deviations.

Figure 9 demonstrates weekly estimations for 3 different locations comparatively, and Table 4 reveals the error rates and standard deviation of corresponding week's speed data for those locations. Both the Table 4 and subfigures illustrate even though error rates of both models increase when standard deviation is higher, LSTM's performance does not get affected by this change as much as MFE method, therefore it catches up and performs relatively better if the traffic characteristics are more complex. We believe that the obtained error rates are very promising to integrate the proposed approach into widely-used real-life applications.

Table 4. Standard deviations and errors for the subfigures in Figure 9.

Figure	Standard deviation	MFE MAPE	LSTM MAPE
(a)	16.43	0.121	0.138
(b)	18.37	0.125	0.137
(c)	27.43	0.344	0.303

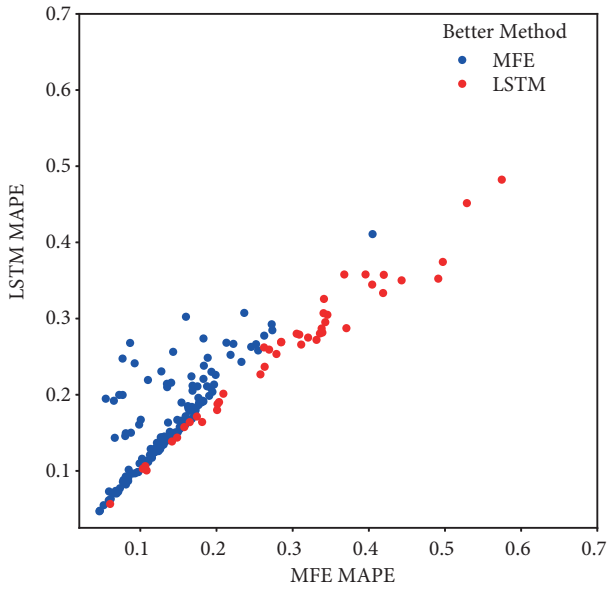


Figure 6. The relationship between MAPE values and standard deviation both for MFE and LSTM models.

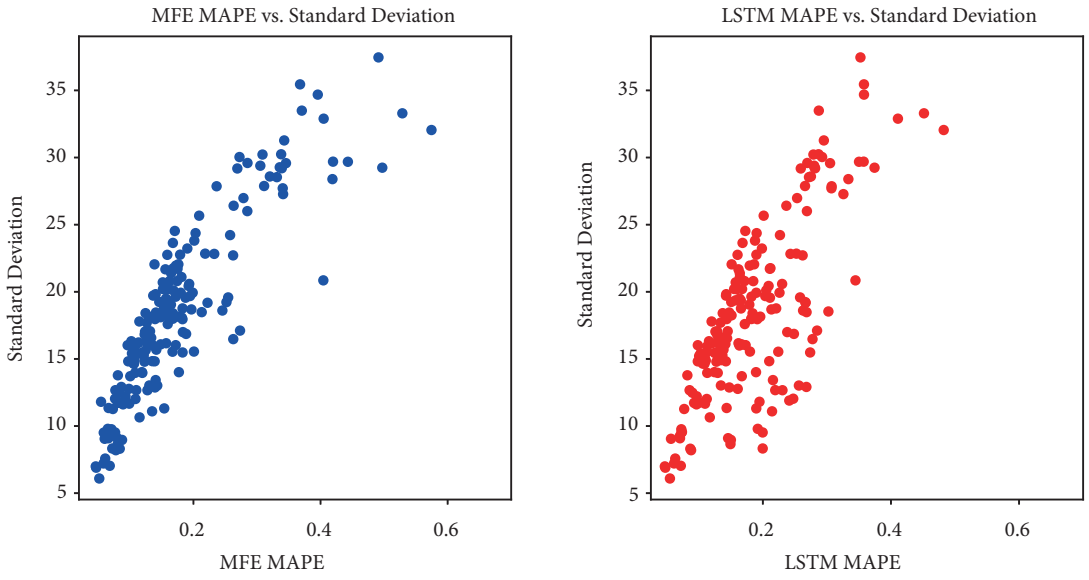


Figure 7. The correspondence of MAPE values of MFE and LSTM models. The blue and red points represent the preferred estimation model by the proposed hybrid architecture, MFE and LSTM, respectively for 187 sensor locations.

Finally, we analyzed the training and testing durations of the proposed models on a laptop with 2.2 GHz Quad Core Intel Core i7 CPU, 16 GB RAM and Intel Iris 1536 MB GPU. Test results showed that LSTM training, LSTM testing and MFE testing tasks lasts for 232.27, 0.55 and 0.42 seconds, respectively, whereas MFE does not require any training operation.

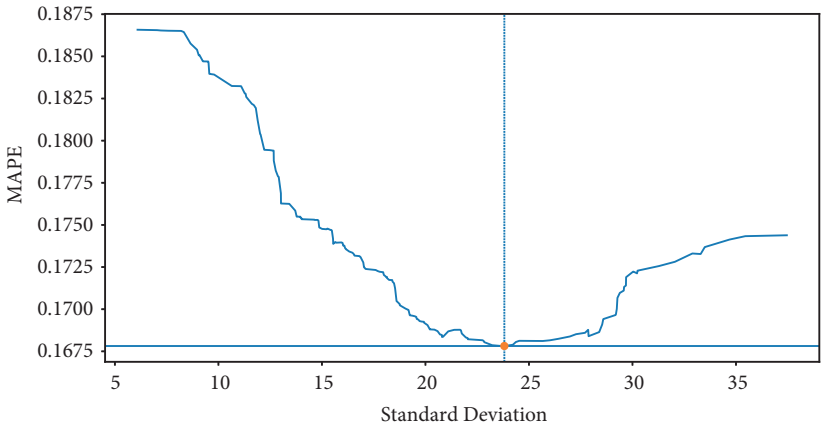


Figure 8. The optimum value for the standard deviation for the proposed hybrid architecture.

4.5. Performance of the proposed hybrid approach

In order to constitute a fair comparison of the proposed architecture with the well known methods in the literature, three different approaches polynomial regression (PR), ARIMA and support vector regression (SVR) are employed. In the conducted tests, it was observed that the ARIMA method's error rate is increasing directly proportional to the prediction horizon and therefore, it is not a suitable method for long-term forecasting. Table 5 gives the MAPE values of the candidate approaches for long-term traffic flow prediction.

The SVR is trained to find the nonlinear relationship between x_t and y_t, where x is the actual traffic flow rate, while y corresponds the predicted one for time variable t [29]. The prediction function g(x) is represented in Equation 4 where α_j and α_j^* are the Lagrange multipliers, k(x,x_t) is kernel function and b is the bias value.

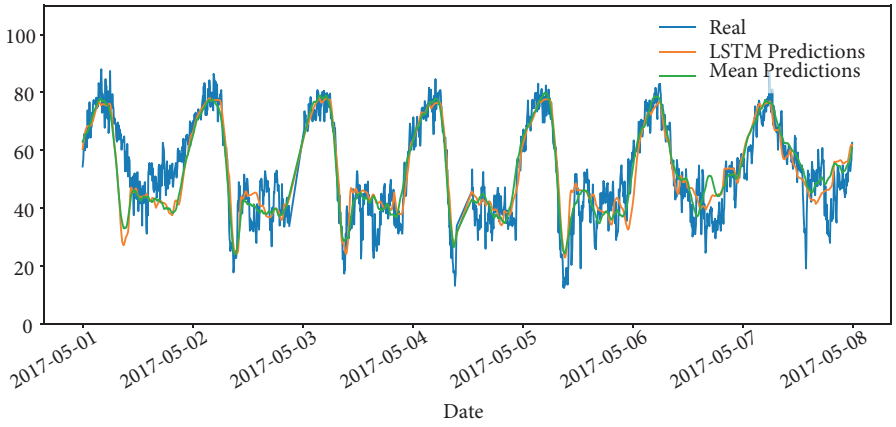
$$y_t = g(x_t) = \sum_{j=1}^{r} (\alpha_j - \alpha_j^*) k(x, x_t) + b \tag{4}$$

$$y_t = g(x_t) = \beta_0 + \beta_1 x_{t,1} + \beta_2 x_{t,2}^2 + ... + \beta_n x_{t,n}^n + \epsilon_t \tag{5}$$

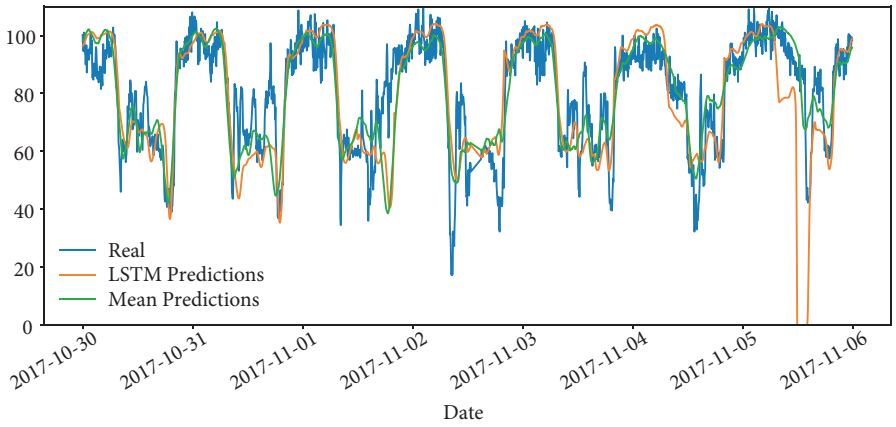
In order to apply SVR, traffic flow data corresponding to the last two months are used to predict traffic for seven days ahead. We use radial basis function kernel with the gamma value of 0.01 and C value of 10 for training.

Regression analysis is a powerful method that enables examining the relationship between two variables [30]. g(x_t), which is a PR model of order n, is defined by Equation 5 where y_t is the predicted value at time variable t, $x_{t,1}$, $x_{t,2}$,...,$x_{t,n}$ are observation group of t and finally β and ϵ are polynomial coefficients and error coefficients respectively. In the conducted experiments, a regression model of order 7 is calculated since it gives the best results.

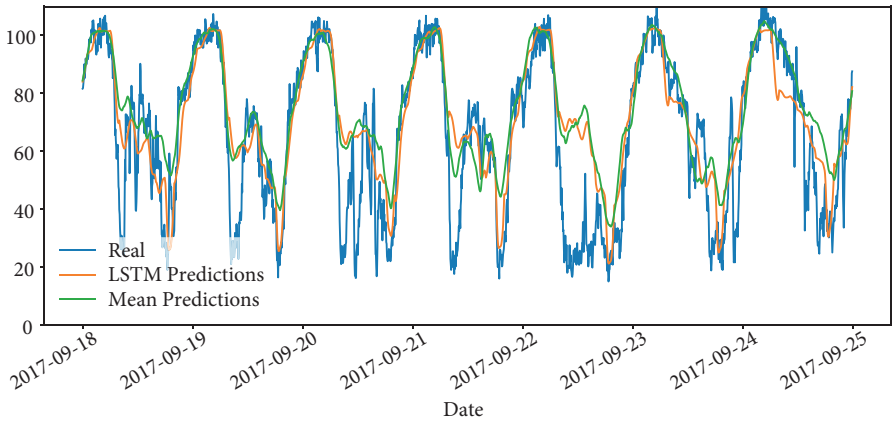
The results confirm that the proposed hybrid architecture gives a lower error rate than plain LSTM and MFE methods as well as PR and SVR algorithms. Obtained success rates of the presented method are comparable with the results of the studies in the literature that make predictions for 7 days later. The proposed approach that allows choosing the most efficient model dynamically according to the traffic characteristics improves the success rates of stand alone LSTM by %10. Since LSTM is a powerful network for modelling time series, variants of it are frequently employed by the systems that perform middle and long term traffic



(a) Low complexity



(b) Medium complexity



(c) High complexity

Figure 9. Weekly estimations for 3 different sensor points (a) a sample sensor where MFE has a better performance. (b) a sample sensor where both models have close performance. (c) a sample model where LSTM has a better performance.

prediction [1, 13, 14, 16, 17, 21, 22, 27]. The proposed hybrid approach can also be adapted to those systems in order to improve overall system performance by combining different networks which are trained depending on traffic characteristics of a given road segment.

Table 5. The performance comparison of the proposed architecture.

Method	**MAPE**
Long short-term memory (LSTM) network	0.187
Mean filtering estimation (MFE)	0.175
Support vector regression (SVR)	0.203
Polynomial regression (PR)	0.192
Proposed architecture (MFE/LSTM)	0.168

5. Conclusion

Accurate long-term estimation of traffic flow is hard, especially when the flow data has complex characteristics. The introduced MFE can easily figure out the certain patterns and it outperforms LSTM algorithm in terms of both estimation error and computational complexity. Therefore, when the flow is easier to predict, it might not be feasible to run a complex algorithm such as LSTM, while simply taking the averages from previous weeks performs better. On the other hand, MFE starts to fail when the patterns in the data start to be harder to detect, and thus, LSTM, a powerful algorithm, can deal with detecting more intricate patterns. Following these outcomes, we build a hybrid architecture which takes advantage of the strength of both models and it outperforms the LSTM and MFE solutions by 11% and 4%, respectively. Experimental results show that a combination of the aforementioned models is a promising approach for the long-term traffic flow prediction problem with a error of lower than +/- 10 km/h. We also demonstrate that the well-known algorithms such as ARIMA, SVR and polynomial regressor used for traffic flow estimation perform far behind our proposed solutions.

This study has proven that modelling the locations with different traffic characteristics by using different methods significantly increases the overall prediction success. In this context, exploiting the standard deviation of flow velocities in order to detect locations with different characteristics appears to be an effective solution. The experimental results show that, effective systems can be developed not only by combining MFE and LSTM methods, but also by employing multiple deep learning methods or regression-based algorithms together.

As a future work, we aim at deploying other deep learning approaches, such as convolutional neural networks, into our hybrid approach. We also will investigate the effect of their hyperparameters thoroughly. We then hope to feed these deep neural networks with meteorological data, accident statistics, road construction information, city events, in order to build more robust models by using the most influential features against changing environmental conditions.

Acknowledgement

This work was supported by the Turkish Scientific and Technological Research Council of Turkey (TÜBİTAK) under the grant number TÜBİTAK1001-120E357.

References

[1] Wang J, Chen R, He Z. Traffic speed prediction for urban transportation network: A path based deep learning approach. Transportation Research Part C: Emerging Technologies 2019; 100: 372–385. doi: 10.1016/j.trc.2019.02.002

[2] Wang Z, Su X, Ding Z. Long-term traffic prediction based on lstm encoder-decoder architecture. IEEE Transactions on Intelligent Transportation Systems 2020;1–11. doi: 10.1109/TITS.2020.2995546

[3] Jeong YS, Byon YJ, Castro-Neto MM, Easa SM. Supervised weighting-online learning algorithm for short-term traffic flow prediction. IEEE Transactions on Intelligent Transportation Systems 2013; 14 (4): 1700-1707. doi: 10.1109/TITS.2013.2267735

[4] Chang H, Lee Y, Yoon B, Baek S. Dynamic near-term traffic flow prediction: system-oriented approach based on past experiences. IET intelligent transport systems 2012; 6(3): 292-305. doi: 10.1049/iet-its.2011.0123

[5] Van Lint JWC, Hoogendoorn SP, van Zuylen HJ. Accurate freeway travel time prediction with state-space neural networks under missing data. Transportation Research Part C: Emerging Technologies 2005; 13 (5-6): 347-369. doi: doi.org/10.1016/j.trc.2005.03.001

[6] Zhong M, Sharma S, Lingras P. Short-term traffic prediction on different types of roads with genetically designed regression and time delay neural network models. Journal of computing in civil engineering 2005; 19 (1): 94-103.

[7] Kumar K, Parida M, Katiyar V. K. Short term traffic flow prediction for a non urban highway using artificial neural network. Procedia-Social and Behavioral Sciences 2013; 104, 755-764. doi: 10.1016/j.sbspro.2013.11.170

[8] Lv Y, Duan Y, Kang W, Li Z, Wang, FY. Traffic flow prediction with big data: a deep learning approach. IEEE Transactions on Intelligent Transportation Systems 2014; 16 (2): 865-873. doi: 10.1109/TITS.2014.2345663

[9] Wu Y, Tan H, Qin L, Ran B, Jiang Z. A hybrid deep learning based traffic flow prediction method and its understanding. Transportation Research Part C: Emerging Technologies 2018; 90:166–180. doi: 10.1016/j.t<c.2018.03.001

[10] Zhao J, Gao Y, Bai Z, Wang H, Lu S. Traffic speed prediction under non-recurrent congestion: Based on LSTM method and BeiDou navigation satellite system data. IEEE Intelligent Transportation Systems Magazine 2019; 11 (2): 70-81. doi: 10.1109/MITS.2019.2903431

[11] Yufang L, Mingnuo C, Wanzhong Z. Investigating long-term vehicle speed prediction based on BP-LSTM algorithms. IET Intelligent Transport Systems 2019; 13 (8): 1281-1290. doi: 10.1049/iet-its.2018.5593

[12] Bui KHN, Cho J, Yi H. Spatial-temporal graph neural network for traffic forecasting: An overview and open research issues. Applied Intelligence 2021; 1-12. doi: 10.1007/s10489-021-02587-w

[13] Zhao Z, Chen W, Wu X, Chen PC, Liu J. LSTM network: a deep learning approach for short-term traffic forecast. IET Intelligent Transport Systems 2017; 11 (2): 68-75. doi: doi.org/10.1049/iet-its.2016.0208

[14] Lu Z, Lv W, Cao Y, Xie Z, Peng H et al. LSTM variants meet graph neural networks for road speed prediction. Neurocomputing 2020; 400: 34-45. doi: 10.1016/j.neucom.2020.03.031

[15] Habtemichael FG, Cetin M. Short-term traffic flow rate forecasting based on identifying similar traffic patterns. Transportation research Part C: emerging technologies 2016; 66: 61-78. doi: 10.1016/j.trc.2015.08.017

[16] Li R, Hu Y, Liang Q. T2F-LSTM Method for Long-Term Traffic Volume Prediction. IEEE Transactions on Fuzzy Systems 2020; 28 (12): 3256-3264. doi: 10.1109/TFUZZ.2020.2986995

[17] Bogaerts T, Masegosa AD , Angarita-Zapata JS, Onieva E, Hellinckx P. A graph cnn-lstm neural network for short and long-term traffic forecasting based on trajectory data. Transportation Research PartC: Emerging Technologies 2020; 112:62–77. doi: 10.1016/j.trc.2020.01.010

[18] Chen W, An J, Li R, Fu L, Xie G et al. A novel fuzzy deep-learning approach to traffic flow prediction with uncertain spatial–temporal data features. Future generation computer systems 2018; 89: 78-88. doi: 10.1016/j.future.2018.06.021

[19] Zang D, Ling J, Wei Z, Tang K, Cheng J. Long-term traffic speed prediction based on multiscale spatio-temporal feature learning network. IEEE Transactions on Intelligent Transportation Systems 2018; 20 (10):3700–3709. doi: 10.1109/TITS.2018.2878068

[20] Qu L, Li W, Li W, Ma D, Wang Y. Daily long-term traffic flow forecasting based on a deep neural network. Expert systems with applications 2019; 121:304–312 doi: 10.1016/j.eswa.2018.12.031

[21] Li Y, Chai S, Ma Z, Wang G. A Hybrid Deep Learning Framework for Long-Term Traffic Flow Prediction. IEEE Access 2021; 9: 11264-11271. doi: 10.1109/ACCESS.2021.3050836

[22] He Z, Chow CY, Zhang JD. STCNN: A spatio-temporal convolutional neural network for long-term traffic prediction. In: 2019 20th IEEE International Conference on Mobile Data Management (MDM); 2019 pp. 226-233.

[23] Peng H, Bobade SU, Cotterell ME, Miller JA. Forecasting traffic flow: Short term, long term, and when it rains. In: Springer, Cham International Conference on Big Data 2018; pp. 57-71.

[24] Belhadi A, Djenouri Y, Djenouri D, Lin DJCW. A recurrent neural network for urban long-term traffic flow forecasting. APPLIED INTELLIGENCE 2020; 50, 3252-3265. doi: 10.1007/s10489-020-01716-1

[25] Simunek M, Smutny Z. Traffic information enrichment: creating long-term traffic speed prediction ensemble model for better navigation through waypoints. Applied Sciences 2021; 11 (1): 315. doi: 10.3390/app11010315

[26] Guo YR, Wang XM, Zhang H, Jim GJ. Traffic flow prediction method of diversion area in peak hours based on double flow graph convolution network. Advances in Transportation Studies 2021;

[27] Wang J, Zhu W, Sun Y, Tian C. An effective dynamic spatiotemporal framework with external features information for traffic prediction. Applied Intelligence 2021; 51 (6): 3159-3173. doi: 10.1007/s10489-020-02043-1

[28] Ünlü R. Cost-oriented LSTM methods for possible expansion of control charting signals. Computers & Industrial Engineering 2021; 154, 107163. doi: 10.1016/j.cie.2021.107163

[29] Smola AJ, Schölkopf B. A tutorial on support vector regression. Statistics and computing 2004; 14 (3): 199-222. doi: 10.1023/B:STCO.0000035301.49549.88

[30] Ahn J, Ko E, Kim EY. Highway traffic flow prediction using support vector regression and Bayesian classifier. In: IEEE International conference on big data and smart computing (BigComp); 2016. pp. 239-244. doi: 10.1109/BIGCOMP.2016.7425919