Original Article
ISSN (Online): 2582-7472

ShodhKosh: Journal of Visual and Performing Arts
June 2024 5(6), 595–609

ENHANCING LARGE SCALE TRAFFIC CONGESTION PREDICTION WITH ATTENTION-AUGMENTED BILSTM NEURAL NETWORK

Deepti Soni [1], Dr. Shraddha Masih [2]

[1] Research Scholar, SCSIT, DAVV, Indore, Madhya Pradesh, India
[2] Professor, SCSIT, DAVV, Indore, Madhya Pradesh, India

DOI
10.29121/shodhkosh.v5.i6.2024.1824

Funding: This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

ABSTRACT

One major issue impeding the sustainable growth of urban traffic is traffic congestion. To avoid traffic congestion, it is crucial to assess the current state of the traffic and project future traffic patterns. Identifying and creating prediction techniques for a city wide in urban setting is the key objective of this study. Predicting long-term levels of road congestion can help commuters avoid crowded regions and allow traffic agencies to take the necessary measures. In this paper, we present attention-augmented BiLSTM neural network approach for hourly based monthly traffic congestion prediction. Our experiments show that the proposed model outperforms the baselines in terms of accuracy and error.

An increasing number of studies are using real -time data-which is gathered using various devices including GPS, loop detectors and fixed location traffic sensors-to improve the prediction impact. Fixed location traffic sensors are more affordable, therefore, in our research we estimate the traffic congestion using real-time data obtained by these sensors.

Keywords: Neural Network, Traffic Congestion, Sustainable Growth

1. INTRODUCTION

A vital aspect of traffic management and control is the accurate and precise examination of the traffic condition on urban arterial. Until recently, estimates of traffic metrics including link occupancy, average speed and flow estimation have been made using a range of traffic data sources including radar, Bluetooth, magnetic loops and road tube counts. In the end, these variables are integrated to find the traffic condition by leveraging the basic correlation between flow, density and speed. Identification of different traffic conditions (congested, free flow etc.) is made possible by this connection. A significant portion of the time, traffic conditions rely on the parameters used to identify congestion. These variables are crucial for both detecting and predicting congestion. Considering the urban setting, many arterial road types exist, each with a unique maximum speed limit and lane configuration. There are roads leading to both city and state highways which are called primary roads. Istanbul features primary roads, secondary roads and local boulevards.

Certain elements are dependent only on the type of road and the location that the where the prediction is carried out, while some rules are generalized to apply on all road types. This paper is distinctive because it consists of all primary and secondary roads of Istanbul connecting city's core up to the state highways. Again, for the hourly prediction we choose two busiest sections of the city around- the-clock.

Thereby, in order to obtain better findings, it is crucial to select multiple zones inside a city and examine them according to various factors. The paper's primary contribution will be fostering city wide traffic congestion in a monthly and hourly fashion.

1.1. RELEVANT CONCEPT AND TERMS IN TRANSPORTATION

There are terms used in transportation that are essential to comprehending the discussion that is being convened. Here some basic ideas given by Hall (1996), Gerlough and Huber (1976, Chapter2) and Immers and Logghe (2002) is replicated to aid the better understanding of the terms. The factors that are most frequently employed to gauge traffic condition are density, flow and speed [1][2].

Traffic Density, is the number of vehicles per unit length of the road section. Due to the scarcity of the sensors to gauge vehicle presence, it has historically been challenging to determine traffic density for whole length of the road. Although the traffic cameras installed at places made it feasible [3].

Occupancy, it is the percentage of the time that a point in the road network is occupied by vehicles. Vehicle loop detectors (VLDs) are most often used to measure occupancy.

In a homogenous stream of traffic when all the vehicles are of the same length, occupancy is exactly proportional to the flow of traffic [6].

Traffic Speed, in transportation engineering speed is the average speed/segment speed of the traffic or the mean speed or space mean speed of the traffic at some time interval.

The fundamental relation between density and speed is given by, $q=k*u$, where k is the density, u is the speed and q is the traffic flow [5].

Traffic Flow, it is the total number of vehicles passing a reference spot per unit of time. Typically, these reference points are selected at the end or middle of the segment [4].

2. RELATED WORK

Congestion is a traffic state, which can be calculated from multiple traffic flow parameters like average speed, density or combination of these. Average speed estimation holds the key for prediction of congestion. It is done to analysis the speed profile of the selected segment, region or the selected corridor, in order to understand the pattern of the speed measured in some particular time limits, and this provides the signs of congestion.

There are number of papers till now focusing on the predicting traffic flow, giving the estimates of traffic flow by measuring the traffic speed [7, 8,9,10,11], the estimation of speed is for the next 5 min, next 15 mins or next 25 mins, a very lessor number of literature exist focusing on the hourly prediction or 24 hours prediction. The first one called as short-term traffic flow prediction and the whole day prediction is called long term prediction.

Despite papers exists, but lesser have been published on the traffic congestion detection and prediction. Although, there is a narrow gap between traffic flow and congestion prediction, there are differences between the two. While traffic flow determines how traffic will behave in a given length of time, congestion prediction quantifies likelihood of congestion and divides it into different levels. This, in turn estimates the total time that could be taken for a trip. Congestion prediction encompasses a variety of factors for its occurrence, here we focus on parameters that have the biggest effect. Considering an urban scenario, there are different types of roads and have different names in different countries. In Istanbul, there are primary boulevards & Avenues, for heavy traffic and has number of lanes for incoming and outgoing traffic, Secondary boulevards & Avenues, with fewer lanes than the first one, Local boulevards & Avenues are roads in the residential area with single lane in each direction, primary street roads and secondary street roads. This research focus on the secondary boulevard's roads and the primary street roads. Hongsuk et.al in their presents traffic prediction method using LSTM-RNN models, it introduces Hypernet framework for automated hyperparameter tuning using Bayesian optimization(M-SMBO). It used benchmark highway traffic performance index dataset to measure the congestion in a highway road section [12]. Saleem et.al proposed fusion- based intelligent traffic congestion control system (FITCCS-VN) to improve traffic flow and reduce congestion, with the help of this system drivers are able to view traffic flow and volume of traffic remotely and avoid that route. Although the system shows accuracy of 95% with a miss rate of 5%but the system was simulated and does not fit in the real road scenario [13]. Guanxiong Liu et.al. used a video frame captured from camera and consider a threshold value of speed=0 m/h or at least 3 consecutive vehicles for 10 second as a congestion. It used GMM and GFF methods for foreground detection in a video frame. The dataset includes a 30 minutes video of weather condition that are-sunny rainy and snowy but again it is a short traffic prediction for 30 mins [14]. In order to address the issue of congestion G.Kothai et.al devised a new BLSTME(hybrid boosted long short- term memory ensemble) and CNN model by increasing the traffic flow and compared their result with that of Autoencoder, ConvLSTM and PredNet in terms of recall and accuracy. This experiment, which used SUMO and OMNet++ in a simulated environment, reported 98% accuracy, nevertheless this was another artificial setting [15]. In the same proceedings, Chenyu et.al integrated CNN and attention mechanism for spatial-temporal dependencies to create STAWnet for traffic speed prediction. STAWnet uses self-learned node embedding for latent spatial relationships and does not require and prior knowledge of graph structure for prediction. It uses three state of the art dataset- METR-LA, PEMS-BAY and PEMS07 to prediction traffic and compare it from 11 baseline models- HA, ARIMA, FC-LSTM, T-GCN, DCRNN, STGCN, GaAN, Graph WaveNet, APTN and ST-GRAT and found it with a faster computation speed with lesser data processing, but it does not consider other speed aspects of traffic prediction [19].

These most recent works are largely route-specific and concentrate on short-term traffic congestion prediction over the next five, fifteen or thirty minutes. The majority of research took traffic flow into account when predicting congestion in the highway stretch. The necessity of long-term congestion prediction in an urban road scenario is acknowledged in this work. The major areas(districts) of a city are taken into account in this paper, which predicts congestion for the whole month and within a 24-hour window.

3. METHODOLOGY

3.1. MATHEMATICAL MODEL

Level of Service (LOS) and State definitions for urban corridors

LOS is a metric that quantifies the level of service quality (HCM,2010). For various kinds of roads-ways there are six distinct LOS states defined, with LOS A denoting the optimum operating condition and LOS F the worst. LOS for urban roads was described by HCM (2010) as the "decrease in travel speed as a percentage of the free-flow speed of the corridor". The official speed limits on urban arterial roadways in TURKEY is 50km/h, but local government have authorities to raise it to 70km/h or even 82 km/h (with a 10% tolerance gap) before imposing fines, which corresponds to 90 km/h in actuality. LOS A and LOS B states corelated to a speed interval of"90 km/h- 77 km/h" and "77 km/h- 60 km/h", assuming 90 km/h as the maximum permissible speed on the study region. The corresponding speed intervals for LOS C and LOS Dare- "45 km/h- 60 km/h" and "36 km/h- 45 km/h". The worst situation, which led to LOS F, was a segment speed of less than 27 km/h [6].

As previously stipulated, traffic patterns are created based on segment speed on urban arterial roads in HCM (2010), which is in accordance with the LOS definition. The speed information from the "Istanbul traffic density" dataset is transformed to a qualitative "state" characteristics, as shown in the table and this allows for the identification of important trends, such as areas with persistent traffic, bottleneck release points and many more. The speed data was terminated at 70 km/h since municipality has increased the speed limit in the research region to 82 km/h or (90 km/h for speeding penalty). Owing to this truncation a shared state (referred as state 1) of free flow conditions above 60 km/h was established for LOS A and LOS B. The state 2 represents a steady flow condition. In the present study, state 3 corresponds to LOS D which stands for those situations approaching unstable flow. State 4, represents unstable flow circumstances (LOS E and F) involving speed less than 36.

Table 1

Table 1 Speed Performance Index with Traffic State

LOS	Speed limit of 70km/hrs	Speed Performance Index (SPI)	Congestion Trends	Description
A	>85	[67,100]	0(normal)	High average speed, road traffic state good
B	67<=V<58			
C	50<=V<67	[50,67]	1(light)	Higher average speed, road traffic state better
D	40<=V<50	[40,50]	2(medium)	Lower average speed, road traffic state bit week
E	30<=V<40	[0,39]	3(heavy)	Low average speed, poor traffic state with frequent stop and go.
F	<30			

Referring to the above table, we calculated the congestion index using speed performance index (SPI).

$$SPI = \frac{V_{as}}{v_{max}} \quad (1)$$

Where V_as indicates average travel speed and V_max is the maximum permissible roadway speed.

3.2. BILSTM-ATTENTION MODEL
3.2.1. BILSTM

The bidirectional LSTM (BiLSTM) is an extension of the standard Long Short-term Memory (LSTM) network that processes the sequences in both forward and backward directions. This allows the model to capture context from past and future states, improving its performance on sequence-based tasks. Also, this helps in learning dependencies that may span across long sequences. This comprehensive view indirectly helps to maintain a more stable gradient flow and help to alleviate the issue of gradient vanishing [17][20].

The structure of BiLSTM is explained below:

This architecture can be interpreted as having two separate LSTM networks, one gets the sequence of tokens as it is while the other gets in the reverse order. Both of these LSTM network returns a probability vector as output and the final output is the combination of both of these probabilities.

X_iis the input token, Y_i is the output token, and A and A' are LSTM nodes. The final output of Y_i is the combination of A and A' LSTM nodes. The input to the model will be speed values denoted by x1, x2...... and output will be congestion levels [18].

Figure 1

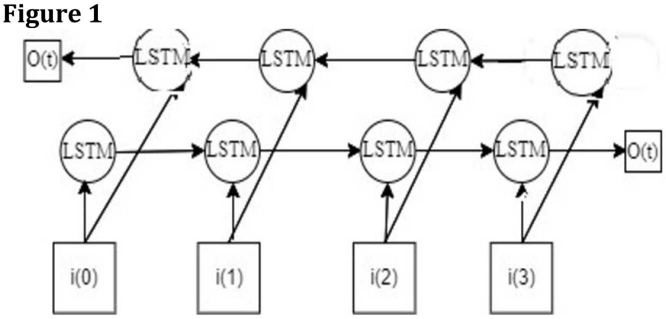

Figure 1 Structure of BiLSTM Model

The following equations explain the working of BiLSTM which is used here to calculate the predicted values

Forget gate:

$$(f_t^f) = \sigma_u(W_g X_t + R_g h_{t-1} + b_g) \quad (2)$$

$$f_t^b = \sigma_u(W_g' X_t + R_g' h_{t+1}^b + b_g') \quad (3)$$

Input gate

$$(i_t^f) = \sigma_u(W_j X_t + R_j h_{t-1} + b_j) \quad (4)$$

$$i_t^b = \sigma_u\left(W_j' X_t + R_j' h_{t+1}^b + b_j'\right) \quad (5)$$

Output gate:

$$(o_t^f) = \sigma_u(W_p X_t + R_p h_{t-1} + b_p) \tag{6}$$

$$o_t^b = \sigma_u(W_p' X_t + R_p' h_{t+1}^b + b_p') \tag{7}$$

Candidate cell update:

$$g_t^f = tanh(W_g X_t + R_p h_{t-1} + b_g) \tag{8}$$

$$g_t^b = tanh(W_g' X_t + R_p' h_{t+1}^b + b_g') \tag{9}$$

Cell State Update:

$$c_t^f = f_t^f \odot c_{t-1} + i_t^f \odot g_t^f \tag{10}$$

$$c_t^b = f_t^b \odot c_{t+1}^b + i_t^b \odot g_t^b \tag{11}$$

Hidden State Update:

$$h_t^f = o_t^f \odot \tanh c_t^f \tag{12}$$

$$h_t^b = o_t^b \odot \tanh c_t^b \tag{13}$$

Concatenation of Hidden States:

At each time step t, the hidden states from the forward and backward LSTMs are concatenated

$$h_t = [h_t^f \oplus h_t^b\ ] \tag{14}$$

In this way hidden states from both LSTMs are concatenated to form the final representation 'h'.

Table 2

Table 2 Notation for BiLSTM model

Notation	Definition
σ_u	Sigmoid Activation function
W_g, W_j, W_d, W_p	Input weight matrices for the forward LSTM gates and candidate cell state
R_g, R_j, R_d, R_p	Weight matrices for the hidden state in the forward LSTM gates and candidate cell state.
X_t	It is the input historical traffic flow
h_{t-1}	It is output at previous time (t-1)
b_*/b_*'	Bias vectors for forward and backward gates and candidate cell state.
$\oplus$	Concatenation operation
$\odot$	Element-wise multiplication

tanh	Hyperbolic tangent activation function
h_t^f, h_t^b	Forward LSTM process the sequence from start to end, producing hidden states h_f
h_t^b	Backward LSTM processes the sequence from end to start, producing hidden state h_b

Attention Mechanism

Neural networks' attention mechanism mimics how the human brain processes attention. Originally, it was the one suggested in the field of image recognition [14]. Whenever, individuals look at photos, they frequently selectively concentrate on a few key details. The attention mechanism has been widely used to natural language processing (NLP) in recent years [15,16], as traditional neural networks presume that each word in the input has the same weight. As a result, they are unable to discern the significance of various words. In order to determine the correlation between the fundamental model. Sequence data also describes traffic flow data, much like natural language does. Additionally, the importance of different traffic states in the flow data is not the same. However, as the issue has not been adequately resolved by the current approaches, we suggest BiLSTM attention-based model.

3.2.2. BILSTM BASED ATTENTION MODEL

Figure 2

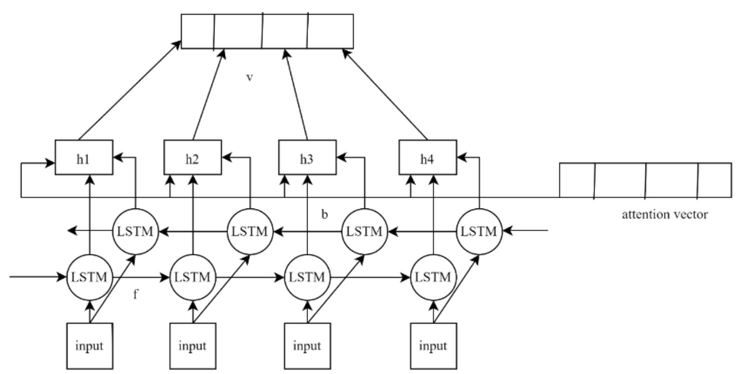

Figure 2 Structure of BilSTM Based Attention Model

Fig.2 show the structure of BiLSTM based attention model. It consists of two LSTM layers, the attention mechanism layer and the output layer. The attention mechanism layer [17] can take into account additional contextual association and emphasize the significance of a certain traffic condition to the overall traffic flow.

$$\overrightarrow{h_t} = LSTM_{forward}(vect) \quad (15)$$

$$\overleftarrow{h_t} = LSTM_{backward}(vect) \quad (16)$$

h_t = [(h_t) ⃗; (h_t) ⃖]
(17)

h_t is the concatenated hidden state at time step combing the forward and backward states.

u_t = tanh (Wh_t + b) where u_t is the intermediate representation at time step t.

Tanh is the hyperbolic tangent activation function, W is a weight matrix, b is a bias vector.

Attention weight, a_t $a_t = \frac{exp(u_t^T a)}{\sum_k \exp(u_k^T a)}$ (18)

a_t is the attention weight at time step t.

u_t^T is the transpose of vector u_t

a is another vector (part of attention mechanism)

denominator is the sum of the exponentiated scores over all time steps k, ensuring the attention weights sum to 1. All these set of equations describes the processing of input sequences using a BiLSTM with attention mechanism.

$v = \sum a_t * h_t$ (19)

After this, the output layer receives the vector v to carry out the last prediction.

3.3. TRAFFIC CONGESTION PREDICTION ARCHITECTURE

Fig 3 shows the architecture mainly consist of-embedding layer, forward and backward LSTM, attention mechanism layer and the prediction layer. The sequence of input {i (0), i (1),……, i (n-1)} represents the traffic flow data and each i(t) is a piece of data at a time interval encoded by one-hot representation. After the embedding layer, the data is mapped into the same dimensional vector space. Then, the LSTM network will process time -aware embedding vector and produce a hidden sequence {h (0), h (1),……,h()n-1)}.An attention mechanism is used to extract traffic embedding features through the output attention probability matrix that is produced by the process in Section 3.2. Then, the prediction layer extracts mean values of the sequence over time intervals and makes the features encoded into a classified vector. Then it is fed into the logistic regression layer at the top of prediction architecture.

4. PERFORMANCE ANALYSIS

In the following section, we first explain our dataset and experimental setup in order to assess the efficacy of our suggested methodology. Next, we display the results based on several metrics. Lastly, we depict the comparative outcomes using a few baselines.

4.1. DATASET AND EXPERIMENTS SETTINGS

1) Dataset Description

This study uses hourly average speed and traffic density data of Istanbul city collected from fixed sensors and cameras installed over the city. Each sensor collects the real -time traffic stream per hour. which contains attributes- date & time, latitude, longitude, minimum speed, maximum speed, average speed, geohash and number of vehicles attribute. Here, the geohash refers to the geolocation code which is obtained from the respected latitude/longitude values. The ‘number of vehicles’ attribute gives the hourly number of the vehicles passing a specific location i.e. the density of vehicles/vehicle count.

Figure 3

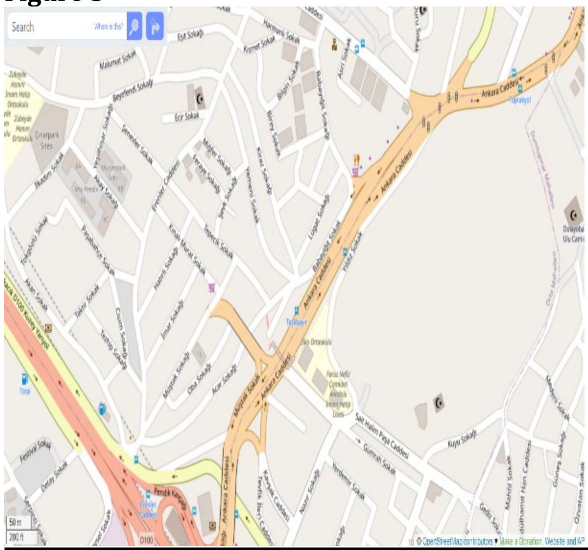

Figure 3 OpenStreetMap View of Roads in Istanbul

Figure 4

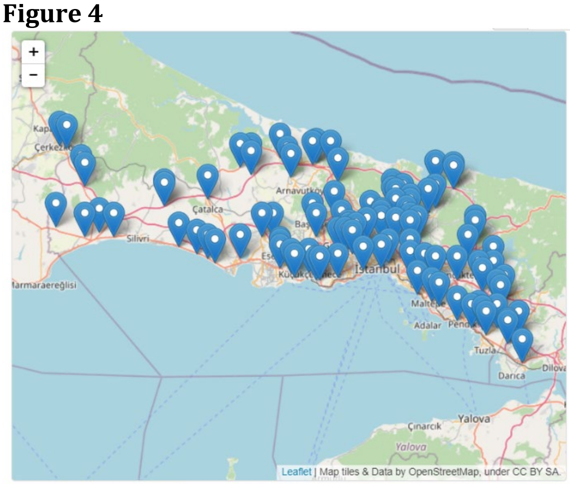

Figure 4 Locations for Experimentation

Figure 5

Figure 5 Sample of Roads

Figure 6

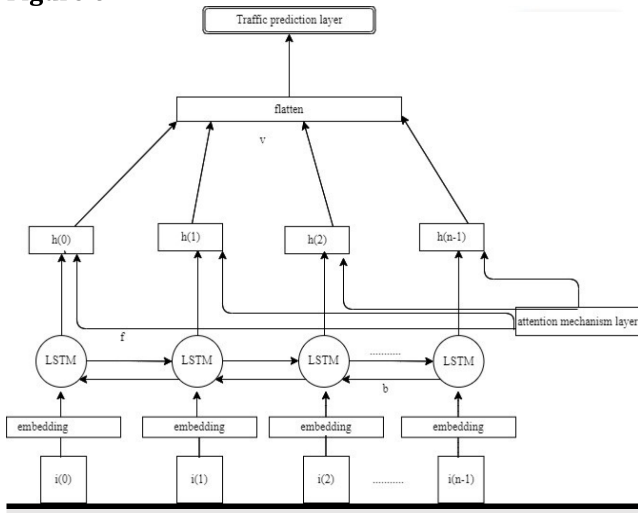

Figure 6 Traffic Prediction Architecture

Figure 7

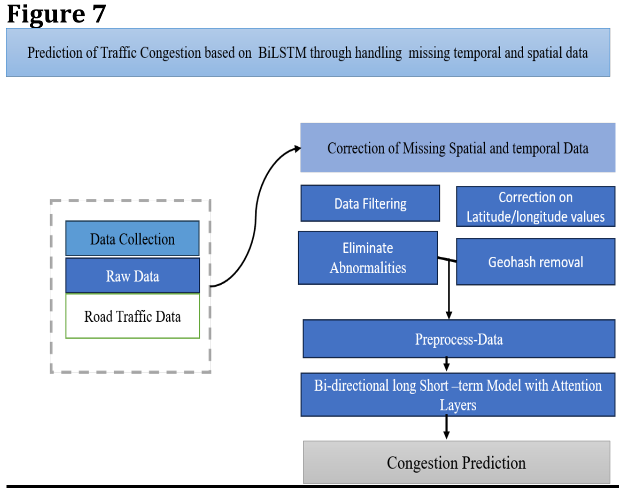

Figure 7 Steps Of Data Preprocessing

We have collected dataset that spans from January 1st 2023 to Dec 31st 2023.The datasets for the months of January through December is combined and we get 8 columns overall and 1+ GB of data after merging. There are 12,582900 entries of records for this year. The original size of the dataset was reduced after down casting the datatypes of the majority of the variables during merging and it came down to 0.5+GB. The latitude and longitude column names were adjusted in several instances because the provided lat.log values were referring to some location in Saudi Arabia. Following this, we preprocessed the dataset and discovered 2464 distinct geohash locations. We ensured all these geohash have single latitude and longitude value. The next issue we needed to explore was if all these geohashes contain data for every period of time and noticed that many of the geohashes contains single measurement of traffic density which did not seem to be very helpful. It is important to eliminate geohashes with inadequate data. There are 365 days in a year and 24 hours per day. So, 365 when multiplied by 24 gives the total count of hours per year, i.e.8,736 hours/year. Naturally, we expect locations to contain around 8736 hours of records as the dataset have hourly records starting from time 00:00 to 23:00. We removed other geohashes and took those with a total of 5000 hours of records for a year. The selection of geohashes with 5000 hours of records was chosen as there all entries were less than 5000 and this selection helps to handle missing hours. Subsequently, we discovered that several of the geohashes had missing hours of data. We ensure all the geohash must have datetime starting at 00:00. Now the complete dataset is transformed to ensure hourly frequency across all geohashes. As a result, some of the rows will have null values. Forward fill method these null values.

The plot depicts the daily time series for NUMBER_OF VEHICLES. Missing values were clearly present in this data frame which can be of few days or for few hours a day. The linear interpolate will capture the hourly and weekly pattern during interpolation. The plot seems to capture the seasonality because it populates some of the missing values by the mean each hour and day of the week. Linear interpolation is then used to better fill the remaining empty values because it performs well when adjacent values to the missing values are not missing. We now look for the dates with the highest total missing values across all geohashes and found 2 dates one in the month of July and other in May, and use interpolation to fill in the missing values. There was very slight change in the mean before imputation and after imputation. The mean value after imputation is 104.35381157520. The dataset sample is shown in the below figure (Fig: 8)

Figure 8

	A	B	C	D	E	F	G	H
1	DATE_TIME	LATITUDE	LONGITUD	GEOHASH	MAXIMUM	MINIMUM	AVERAGE_	NUMBER_OF_VEHICLES
2	6/1/2023 0:00	29.13025	40.9433	sxk8yv	169	5	65	242
3	6/1/2023 1:00	29.13025	40.9433	sxk8yv	161	8	66	108
4	6/1/2023 2:00	29.13025	40.9433	sxk8yv	152	10	71	103
5	6/1/2023 3:00	29.13025	40.9433	sxk8yv	152	13	68	69
6	6/1/2023 4:00	29.13025	40.9433	sxk8yv	162	6	61	94
7	6/1/2023 5:00	29.13025	40.9433	sxk8yv	160	6	64	277
8	6/1/2023 6:00	29.13025	40.9433	sxk8yv	127	3	53	832
9	6/1/2023 7:00	29.13025	40.9433	sxk8yv	108	1	42	601
10	6/1/2023 8:00	29.13025	40.9433	sxk8yv	71	1	22	444
11	6/1/2023 9:00	29.13025	40.9433	sxk8yv	82	2	24	460
12	6/1/2023 10:00	29.13025	40.9433	sxk8yv	104	2	37	362
13	6/1/2023 11:00	29.13025	40.9433	sxk8yv	98	1	35	330
14	6/1/2023 12:00	29.13025	40.9433	sxk8yv	89	2	39	392
15	6/1/2023 13:00	29.13025	40.9433	sxk8yv	107	2	39	343
16	6/1/2023 14:00	29.13025	40.9433	sxk8yv	93	2	29	352
17	6/1/2023 15:00	29.13025	40.9433	sxk8yv	98	2	29	339
18	6/1/2023 16:00	29.13025	40.9433	sxk8yv	89	1	28	333
19	6/1/2023 17:00	29.13025	40.9433	sxk8yv	71	2	25	487
20	6/1/2023 18:00	29.13025	40.9433	sxk8yv	66	1	21	599
21	6/1/2023 19:00	29.13025	40.9433	sxk8yv	105	1	36	574
22	6/1/2023 20:00	29.13025	40.9433	sxk8yv	125	2	46	447

Figure 8 Sample of Dataset

2) Experimental setting

Two time zones have been selected as the research region here: morning (6-11), which corresponds to traffic congestion check-in periods and afternoon (13-17), which corresponds to traffic congestion check-out times. For this, we utilized the Python Folium Library. After gazing into and analyzing the data, we saw that, the average speed is lower than usual in the morning and during afternoon and again slowly it is gaining the normal speed in the traffic. We divide this dataset in training from January to Oct, validation set November and test set December. The number of hidden units of LSTM is 64. We then use the ADAM (Adaptive Moment Estimation) to minimize the square errors in our congestion prediction. Moreover, the mini-batch size is set at 64. In order, to improve the generalization capability of our model and alleviate the overfitting problem [19], we adopt the dropout method proposed in [20,21], which randomly drops units from the network where the connection seems to be weakened and the dropout rate of the output layer is set at 0.6.

Figure 9

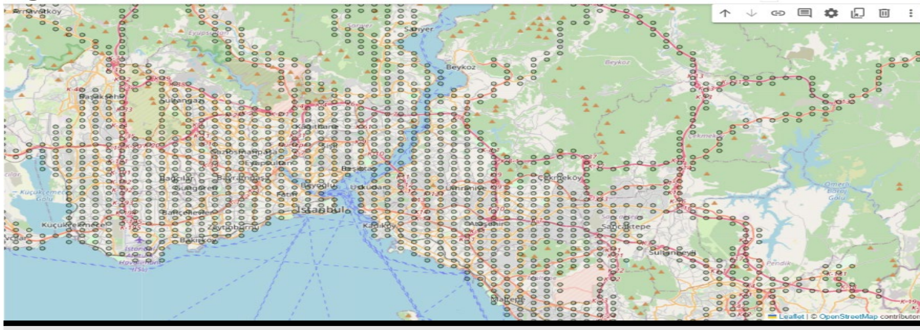

Figure 9 Folium Map

4.2. EXPERIMENTAL RESULTS, ANALYSIS AND DISCUSSION

We calculated the average monthly traffic congestion for the entire city of Istanbul for the month of December using data from 40 locations which corresponds to different districts of Istanbul city. Among them, we picked two locations for the mean hourly traffic congestion prediction forecast and selected one weekday in December for the hourly prediction. The two locations that were chosen are in the center of the city: Bagcilar (Loc 1) and Tulza (Loc-2). Fig:11 makes it evident that Loc 1 has more congestion than Loc 2, with the average speed profile for Loc 1 continuously decreasing throughout morning peak. It was discovered that, during peak hours, location1 had higher traffic than location 2.

We compare our proposed BiLSTM based attention model for traffic prediction with other methods in the prediction time series approaches ARIMA (autoregressive integrated moving average), KNN (K nearest neighbors) [23] LSTM ((long short-term memory network)), Attention- based LSTM. We use same dataset and measures to ensure a fair comparison. The algorithm's test result is shown in the below. Among the prediction models, the BiLSTM attention-based model has the lowest MAPE and RMSE.As a result, our proposed approach outperforms the baselines.

Figure 10

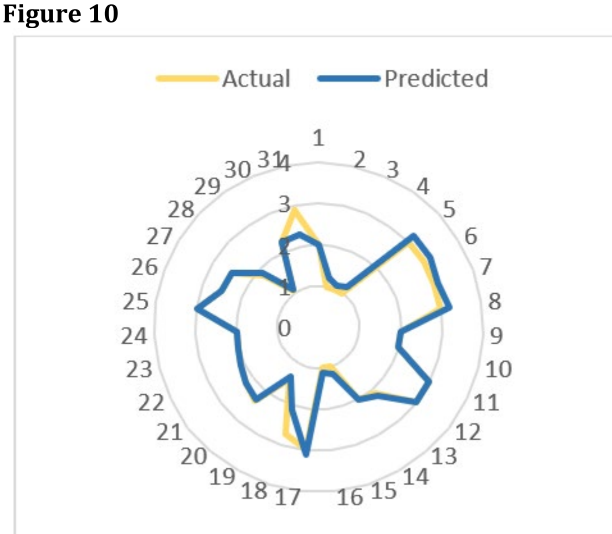

Figure 10 Mean Monthly Traffic Congestion Trends (December)

Figure 11

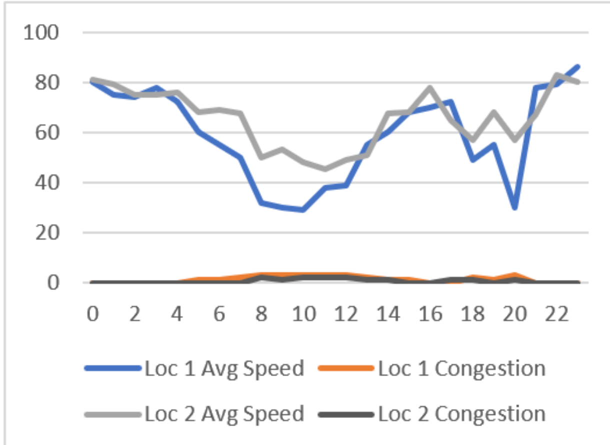

Figure 11 Mean Hourly Congestion Trends of Loc-1 and Loc-2

Measures

We utilize the two performance indicators- Root Mean Square Error (RMSE) and Mean Absolute Percentage Error (MAPE), Which are defined below to assess how proficiently the congestion forecast performed.

$$\text{RMSE}\ (\exists, \nexists) = \frac{1}{n}\sum_{i=1}^{n}\frac{|\exists_i - \nexists_i|}{\exists_i} \tag{20}$$

$$\text{RMSE}\ (\exists, \nexists) = \left[\frac{1}{n}\sum_{i=1}^{n}\left(\left|\exists i - \nexists i\right|\right)^2\right]^{\frac{1}{2}} \tag{21}$$

Table 3

Table 3 Prediction Performance of Attention Model

City	Model	RMSE	MAPE (%)
ISTANBUL	ARIMA	62.98	10.14
	KNN	59.82	9.98
	LSTM	50.55	8.72
	LSTM- based Attention Model	56.46	7.35
	BiLSTM- based attention model	46.21	6.68

Figure 12

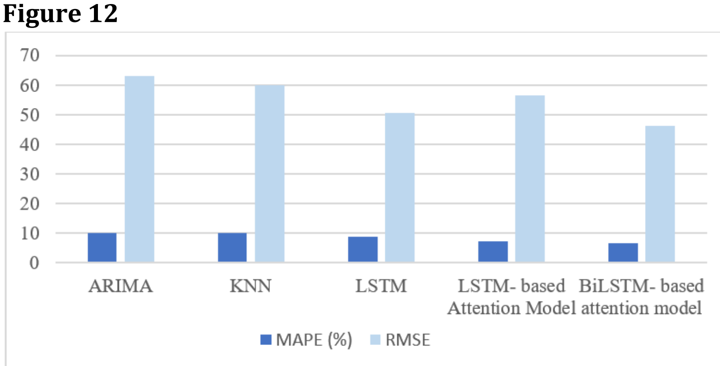

Figure 12 Experimental Result for Predicting Congestion

5. CONCLUSION

In this work, we present BiLSTM based Attention model to predict traffic congestion of thirty-one days. This model can fully use the time-aware traffic data by capturing several aspects at different times. When compared with ARIMA, KNN, LSTM and LSTM-based Attention Model in the experimental settings, our model has the lowest MAPE and RMSE. It is shown that the suggested approach performs noticeably better than baselines.

CONFLICT OF INTERESTS

None.

ACKNOWLEDGMENTS

None.

REFERENCES

Immers, L., Logghe, S., 2002. Traffic Flow Theory, Vol. 40 (21). Faculty of Engineering, Department of Civil Engineering, Section Traffic and Infrastructure, Kasteelpark Arenberg.

Gerlough, D.L., Huber, M.J., 1976. Traffic flow theory. In : Transportation Research Board (TRB) Special Report. Transportation Research Board.

Szegedy, C., Liu, W., Jia, Y., Sermanet, P., Reed, S., Anguelov, D., Erhan, D., Vanhoucke, V., Rabinovich, A., 2015. Going deeper with convolutions. In : Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 1–9.

Daganzo, C., Daganzo, C., 1997. Fundamentals of Transportation and Traffic Operations, Vol. 30. Pergamon Oxford

Gartner, N.H., Messer, C.J., Rathi, A., 2002. Traffic flow theory-a State-of-the-art Report : Revised Monograph on Traffic Flow Theory. In : Transportation Research Board (TRB) Special Report. Turner-Fairbank Highway Research Center.

De Fabritiis, C., Ragona, R., Valenti, G., 2008. Traffic Estimation and Prediction Based on Real Time Floating Car Data. In : 2008 11th International IEEE Conference on Intelligent Transportation Systems. IEEE, pp. 197–203.

Afrin, T., Yodo, N., 2020. A survey of Road Traffic Congestion Measures Towards a Sustainable and Resilient Transportation System. Sustainability 12 (11), 4660.

Fouladgar, M., Parchami, M., Elmasri, R., Ghaderi, A., 2017. Scalable deep Traffic Flow Neural Networks for Urban Traffic Congestion Prediction. In : 2017 International Joint Conference on Neural Networks, IJCNN, pp. 2251–2258.

Liu, J., Barreau, M., Cicic, M., Johansson, K.H., 2020. Learning-Based Traffic State Reconstruction Using Probe Vehicles. arXiv Preprint arXiv:2011.05031.

T. Zhang, J. Xu, S. Cong, C. Qu and W. Zhao, "A Hybrid Method of Traffic Congestion Prediction and Control," in IEEE Access, vol. 11, pp. 36471-36491, 2023, doi: 10.1109/ACCESS.2023.3266291

] Walid Fahs, Fadlallah Chbib, Abbas Rammal, Rida Khatoun, Ali El Attar, Issam Zaytoun, Joel Hachem, 2023.Traffic Congestion Prediction Based on Multivariate Modelling and Neural Networks Regressions, Procedia Computer Science, Volume 220, Pages 202-209, ISSN 1877-0509,

H. Yi and K. -H. N. Bui, "An Automated Hyperparameter Search-Based Deep Learning Model for Highway Traffic Prediction," in IEEE Transactions on Intelligent Transportation Systems, vol. 22, no. 9, pp. 5486-5495, Sept. 2021, doi: 10.1109/TITS.2020.2987614.

Saleem, Muhammad & Abbas, Sagheer & Ghazal, Taher & Khan, Muhammad & Sahawneh, Nizar & Ahmad, Munir. (2022). Smart cities : Fusion-based intelligent traffic congestion control system for vehicular networks using machine learning techniques. Egyptian Informatics Journal. In Press. 1-10. 10.1016/j.eij.2022.03.003.

G. Liu et al., "Smart Traffic Monitoring System Using Computer Vision and Edge Computing," in IEEE Transactions on Intelligent Transportation Systems, vol. 23, no. 8, pp. 12027-12038, Aug. 2022, doi: 10.1109/TITS.2021.3109481.

G. Kothai, E. Poovammal, Gaurav Dhiman, Kadiyala Ramana, Ashutosh Sharma, Mohammed A. AlZain, Gurjot Singh Gaba, Mehedi Masud, and Vimal Shanmuganathan. 2021. A New Hybrid Deep Learning Algorithm for Prediction of Wide Traffic Congestion in Smart Cities. Wirel. Commun. Mob. Comput. 2021 (2021).

HCM 2010 : highway capacity manual. (2010). Washington, D.C. : Transportation Research Board.

Schuster, M. and Paliwal, K.K. : Bidirectional Recurrent Neural Networks. IEEE Transactions on Signal Processing, 45(11), pp.2673-2681 (1997).

Brownlee, J. : Long Short-term Memory Networks with Python : Develop Sequence Prediction Models with Deep Learning. Machine Learning Mastery (2017).

Tian, C. and Chan, W. K. (2021). Spatial-Temporal Attention Wavenet: a Deep Learning Framework for Traffic Prediction Considering Spatial-Temporal Dependencies. IET Intelligent Transport Systems, 15(4), 549-561.

Zhai, Y., Wan, Y., & Wang, X. (2022). Optimization of Traffic Congestion Management in Smart Cities under Bidirectional Long and Short-Term Memory Model. Journal of Advanced Transportation, 2022(1), 3305400.