Implementing A Deep Learning Framework for Short Term Traffic Flow Prediction

Hongsuk Yi*
Korea Institute of Science and Technology Information
hsyi@kisti.re.kr

Khac-Hoai Nam Bui
Korea Institute of Science and Technology Information
hoainam.bk2012@gmail.com

Heejin Jung
Korea Institute of Science and Technology Information
bigbear@kisti.re.kr

ABSTRACT

Traffic congestion in highway systems has become more serious. Recently, with the development of advanced technologies for intelligent transportation system, many studies focus on proposing new models and algorithms for time series analysis to predict short-term traffic flow. In this paper, we take a comprehensive study in implementing a deep learning framework for the short term traffic flow prediction in highway systems. Specifically, we apply long-short term memory for analyzing data in Gyeongbu Expressway of Korean transportation system. Experiments indicate promising results for predicting short term traffic flow in highway systems.

CCS CONCEPTS

• **Computing Methodologies** → *Deep Learning Method for Information Extraction and Knowledge Discovery*; • **Applied Computing** → *Traffic Flow Prediction.*

KEYWORDS

intelligent transportation system, traffic prediction, deep learning, long-short term memory, highway systems

ACM Reference Format:

Hongsuk Yi, Khac-Hoai Nam Bui, and Heejin Jung . 2019. Implementing A Deep Learning Framework for Short Term Traffic Flow Prediction. In *9th International Conference on Web Intelligence, Mining and Semantics (WIMS2019), June 26–28, 2019, Seoul, Republic of Korea.* ACM, New York, NY, USA, 8 pages. https://doi.org/10.1145/3326467.3326492

1 INTRODUCTION

In recent years, there has been a vast increase of available data with the advancement of smart cities [6]. Especially, in Intelligent Transportation Systems (ITS), an application of sensing, analysis, control and communications technologies, this issue can positively affect to traffic problems. Specifically, ITS produces a large amount of data in which big data will have profound impacts on the design and application of ITS, which makes the systems safer, more efficient, and profitable [19]. Recently, deep learning is introduced as a promising solution for big data analytics which address important issues such as extracting complex patterns, semantic indexing, and data tagging [11, 16]. Therefore, several domains in big data applications, such as computer vision [5] and speech recognition [7], have utilized deep learning to improve performances of the systems. In case of ITS, deep learning methods are applied for accurate and timely traffic flow information which is an important component for improving transportation systems. Recently, researchers and developers focus on deep neural network for traffic prediction, especially, in urban areas [9, 15, 18]. Particularly, comparing with conventional methods (e.g., statistical or machine learning models), deep learning enable models to use multi-layer architecture to discover intricate structures and complex patterns.

In this paper, we take a study in implementing Deep Long-Short Term Memory Recurrent Neural Network (LSTM-RNN) in highway systems. Based on the implementation, we are able to determine appropriate values of hyper-parameters such as dropout, learning rate, and neurons for optimizing short term traffic flow prediction. Specifically, our contributions include as follows:

- We collect and implement data of Highway system from Seoul to Busan, which is a very important path of public transportation in Korea, to predict the congestion in each particular region.
- Based on the implementation, we determine values of hyper-parameters (e.g., dropout) in the proposed model for predicting efficiently traffic congestion in highway systems.
- Results in our implementation indicate the promising values for the short-term traffic flow prediction in highway systems.

The rest of this paper is organized as follows: In Section 2, we take a brief background and description of deep learning models for traffic prediction such as RNN in deep learning and LSTM-RRN for time series prediction. In section 3, we propose a deep learning framework using LSTM-RRN architecture for short term traffic flow prediction. Section 4 presents the experimental results. Some discussions and future works are concluded in Section 5.

2 BACKGROUND

2.1 Deep Learning for Time Series Analysis

In real-world applications, such as speech recognition or sleep stage classification, data are captured over the course of time, constituting time-series [3]. Technically, the formal definition of time-series can be represented as a *vector* as follows [10]:

$$X = \{x^{(1)}, x^{(2)}, ..., x^{(n)}\} \tag{1}$$

where each element $x^{(t)} \in R^m$, pertaining to V is an array of m values (e.g., $\{a_1^{(t)}, a_2^{(t)}, ..., a_m^{(t)}\}$), and each element of set m corresponds to an input variable which is measured in the time-series.

*Coresponding Author.

WIMS2019, June 26–28, 2019, Seoul, Republic of Korea

ACM ISBN 978-1-4503-6190-3/19/06…$15.00
https://doi.org/10.1145/3326467.3326492

By applying deep learning for time series, several independent neural network layers work together to provide better results. Therefore, many applications have applied deep learning for time series forecasting. For instance, the automatic learning and handling of temporal dependence and structures, respectively [1, 12]. In particular, there have four typical deep learning methods that applied for time series analysis as follows:

- **MLPs** (Multilayer Perceptrons): a classical neural network architecture which is characterized by several layers of input nodes. Specifically, the nodes connect with each other as a directed graph between the input and output layers. MLPs use back-propagation for training the network.
- **CNNs** (Convolutional Neural Networks): a simple CNN model which includes multi-channel models, advanced multi-headed, and multi-output models.
- **LSTM** (Long Short-Term Memory): a type of RNN which includes Simple LSTM models, Stacked LSTMs, Bidirectional LSTMs and Encoder-Decoder models for sequence-to-sequence learning.
- **Hybrids**: the combination of MLP, CNN and LSTM models such as CNN-LSTMs and ConvLSTMs [14].

2.2 LSTM-Recurrent Neural Network for Traffic Prediction

LSTM model is a variant of RNN architecture which is able to learn information with long time spans and determine the optimal time lags in an automatic manner. Recent approaches take into account using LSTMs for predicting traffic flow in ITS [8, 13].

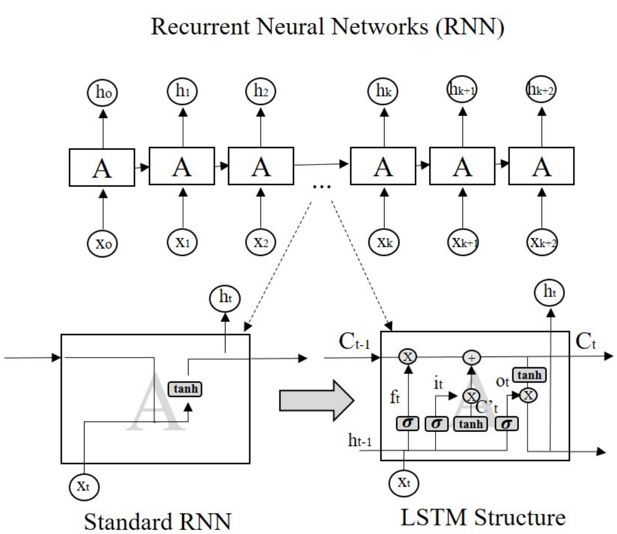


Figure 1: LSTMs Structure.

Fig. 1 depicts the difference between LSTM and standard RNN structures. Specifically, LSTM is an extension of RNN, which basically extends their memory. Therefore, it is able to learn from important experiences that have very long time lags. Supporting input of the historical traffic flow sequence is denoted as Eq. 1, the hidden state of memory block h and the output sequence y can be respectively calculated as follows:

$$h_t = H(W_{xh}x_t + W_{hh}h_{t-1} + b_h) \tag{2}$$
$$y_t = W_{hy}h_t + b_y \tag{3}$$

where W denotes weight matrices, b is bias vectors, and H is the hidden layer function. Supporting f, i, and C' are the forget gate, input gate, and candidate value, respectively (Fig. 2), the hidden layer function H can be sequentially calculated as follows:

$$f_t = \sigma(W_f \cdot [h_{t-1}, x_t] + b_f) \tag{4}$$
$$i_t = \sigma(W_i \cdot [h_{t-1}, x_t] + b_i) \tag{5}$$
$$C'_t = tanh(W_C \cdot [h_{t-1}, x_t] + b_C) \tag{6}$$

Where σ is the standard logistic sigmoid function which can be calculated as follows:

$$\sigma(x) = \frac{1}{1 + e^{-x}} \tag{7}$$

In this regard, the cell activation vectors C, the output gate o, and h can be respectively calculated as follows:

$$C_t = f_t * C_{t-1} + i_t * C'_t \tag{8}$$
$$o_t = \sigma(W_o \cdot [h_{t-1}, x_t] + b_o) \tag{9}$$
$$h_t = o_t * tanh(C_t) \tag{10}$$

In following section, we present our proposed framework for traffic prediction using LSTM-RNN model in which we focus on analyzing data from Gyeongbu Expressway for short-term traffic flow prediction with slightly modify LSTM structures.

3 DEEP LEARNING FRAMEWORK FOR SHORT TERM TRAFFIC PREDICTION USING LSTM NETWORK

3.1 System Architecture

Fig. 2 depicts the process of our proposed framework for short term traffic prediction using LSTM models. Specifically, traffic data from

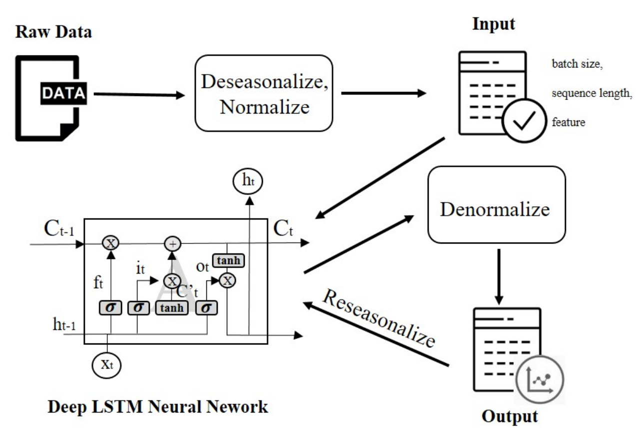


Figure 2: System Architecture.

highway systems will be collected and pre-processed for the input data. Then, a modify LSTM model is applied for extracting predicted values. Based on the implementations, parameters such as dropout, sequence length, and the number of neurons are determined to optimize the short-term traffic flow prediction.

3.2 Modify LSTM models

We modify the LSTM models based on the work in [4]. Specifically, the limitation of LSTM model is that there has no direction from Constant Error Carousels (CECs) to LSTM cell for supporting to control. In this regard, we add "peepholes connection" to all the gates which means the gate layers are able to look at the cell state (Fig. 2). Therefore, the input gate, forget gate, and output gate can be re-calculated as follows:

$$i_t = \sigma(W_i \cdot [h_{t-1}, x_t, C_{t-1}] + b_i) \tag{11}$$

$$f_t = \sigma(W_f \cdot [h_{t-1}, x_t, C_{t-1}] + b_f) \tag{12}$$

$$o_t = \sigma(W_o \cdot [h_{t-1}, x_t, C_t] + b_o) \tag{13}$$

3.3 Objective Function

The objective function of our approach is to minimize the square error e for predicting traffic flow, which is formulated as follows:

$$e = \sum_{t=1}^{n} (y_t - p_t)^2 \tag{14}$$

p is the predicted traffic flow value. Technically, if LSTM model includes more layers, the ability of the learning is stronger. However, if there has too many layers, overfitting problem can be occurred [17]. In this regard, in this study, we implement a model which includes two LSTM layers, a dropout layer, and a dense layer to learn the dynamic of traffic flow. The detail of our process for training data by using Deep LSTM-RNN is shown in Fig. 3. Particularly, $batch$, α, and β represents for the value of batch size, sequence length, and the number of neurons, respectively. Hence, by implementing the model with various of parameters, the objective of this study is to define the best values which can optimize the square error for short term traffic flow prediction.

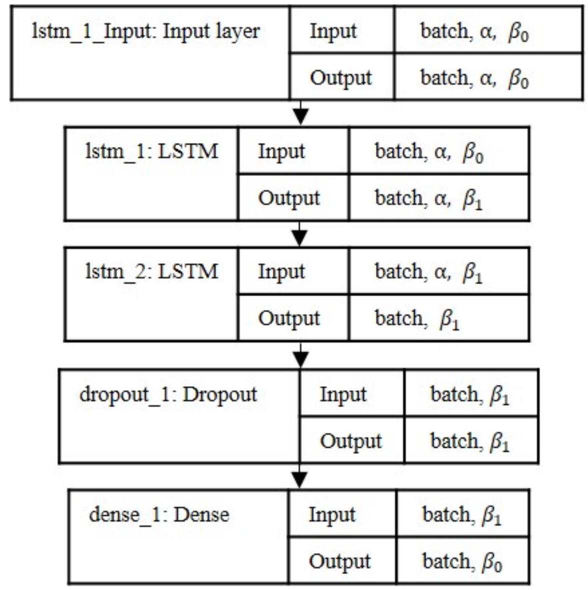


Figure 3: The architecture of proposed LSTM model.

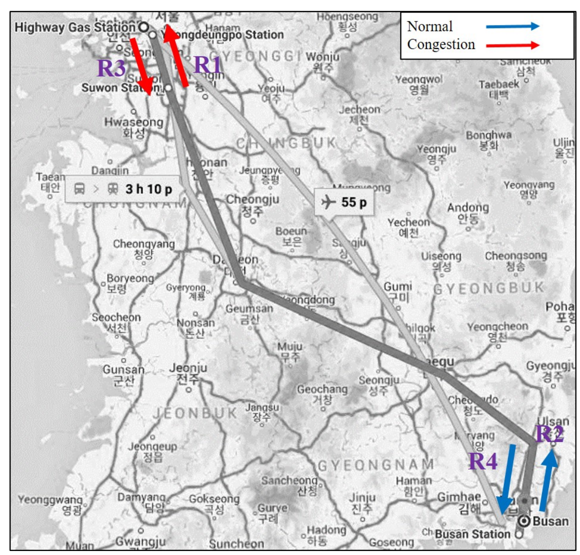


Figure 4: Gyeongbu Expressway from Seoul to Busan.

4 IMPLEMENTATION

4.1 Data Description and Experiment Design

In this study, we take a Korean highway, which names Gyeongbu Expressway [1] from Seoul to Busan (As shown in Fig. 4[2]) into account for implementing the short-term traffic flow prediction.

Specifically, real-time traffic data that collecting by a commercial navigation company is used for our implementation. Technically, each vehicle has a navigation system to store the electronic map of the road network that consists of enormous links and nodes. Each node represents an intersection with branch or merging points. On the other hand, each link represents a road between two nodes. On-Board Equipment (OBE) including Global Positioning System (GPS) of a vehicle traveling on the road network collects its' coordinates. The OBE measures how long the vehicle takes for entering and exiting the link, and computes the link-based average speed when the vehicle complete exiting. On the other hand, average travel time is computed based on the time for passing between two nodes. Furthermore, we aggregate the average link travel speed of vehicles passing through a specific link and compiles the average traffic speed every 5 minutes. Therefore, in our implementation, we use 5 minutes average traffic speed data collected from May to July, 2017 (3 months). Specifically, the usage data includes 23616 average traffic speed data. The number of datasets means that 288 average traffic speed data had aggregated every day, and stored them for 82 days excepting holidays. Moreover, since links have several posted limited speed from 80 KPH (Kilometers per hour) to 110 KPH, the 5

[1] https://en.wikipedia.org/wiki/Gyeongbu_Expressway
[2] https://www.google.com/maps/

minutes average speed were converted to the Traffic Performance Index (TPI), which is calculated as follows:

$$TPI_t = \frac{v_f - v_t}{v_f} \tag{15}$$

where v_f is the free-flow speed and v_t represents the observed 5 minutes average traffic speed. Particularly, the free-flow speed is defined as the maximum speed which a vehicle can make without interruption by any surrounding vehicles. Basically, the free-flow speed can be calculated as the following:

$$v_f = c * v_l \tag{16}$$

where v_l is the posted limited speed of a specific link and c is a constant number which can be varied from 1.2 to 1.5. By converting the 5 minutes average speed to TPI presents, the data is logically normalized which can be used as input data for traffic prediction.

Fig. 5 and 6 depicts the tracking of vehicles based on TPI. Specifically, for defining the congestion, we classify the roots of destinations (e.g., Seoul and Busan stations) into four regions such as $R1$, $R2$, $R3$, and $R4$ as shown in Fig. 4.

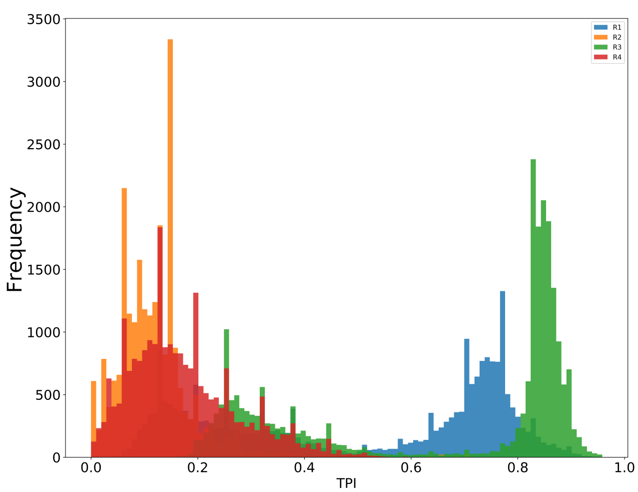


Figure 5: Frequency of Vehicles.

Technically, in continuous traffic flow such as freeways, the TPI less than 0.4 represents *non-congested condition*, the TPI from 0.4 to 0.8 represents *congested condition* and the TPI greater than 0.8 represents *serious congested condition* in general.

4.2 Experimental Results

4.2.1 TPI Prediction: Fig. 7 presents the predictions of traffic congestion in $R3$ and R_4. As shown in the figure, traffic congestion occurs in R_3 from 6:30 *AM* to 8:00 *PM* every day. Particularly, in the experimental results, we only extract two religions which are $R3$ and $R4$ since as our observation, the results of $R1$ and $R2$ are quite similar with $R3$ and $R4$, respectively. For more detail, Fig. 8 represent the TPI values in each region.

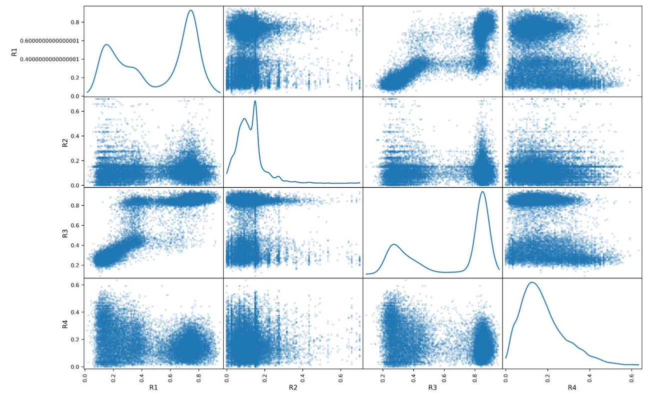


Figure 6: Heat map of tracking Vehicles.

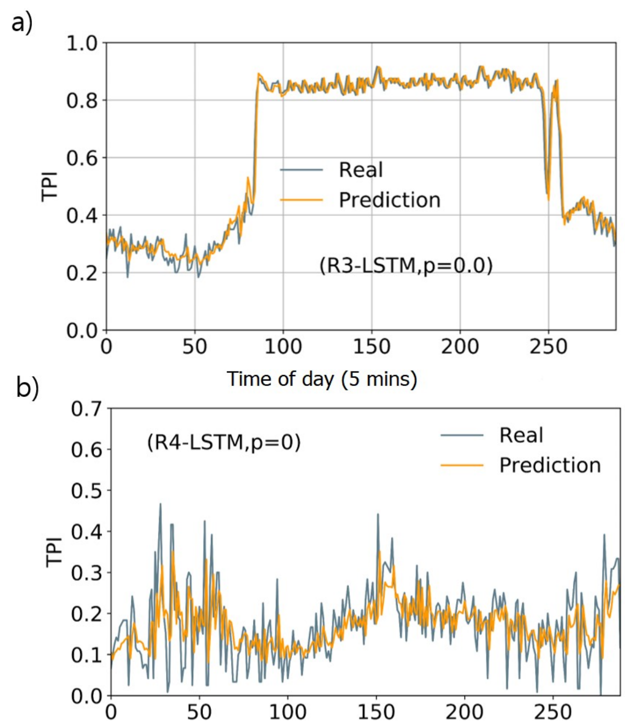
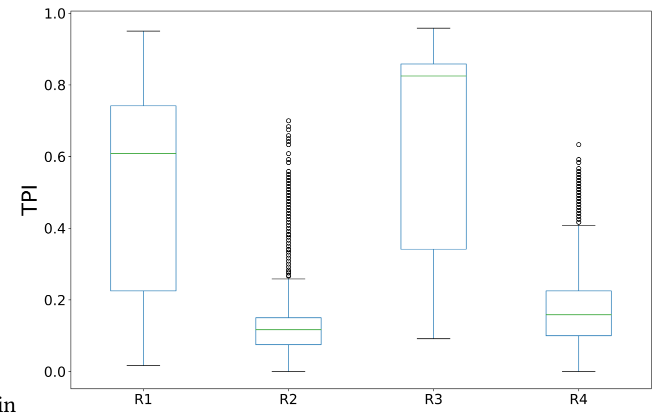


Figure 7: TPI Prediction.

4.2.2 Prediction Accuracy: we first calculate the Mean Squared Error (MSE) which are defined as follows:

$$MSE = \frac{1}{n}\sum_{i=1}^{n}(f_i - f_i')^2 \tag{17}$$

where f is the observation value of traffic flow, and f' is the prediction value. Fig. 9 and 10 depict the MSE with different values of dropout (0 and 0.7) of R_3 and R_4, respectively. As shown in figures, the overfitting occurs when we do not consider the value of dropout. In this regard, we have implemented with different values



inin

Figure 8: TPI values in each region.

of dropout to determine the best value. For more detail, Fig. 11 depicts the influence of dropout values to MSE.

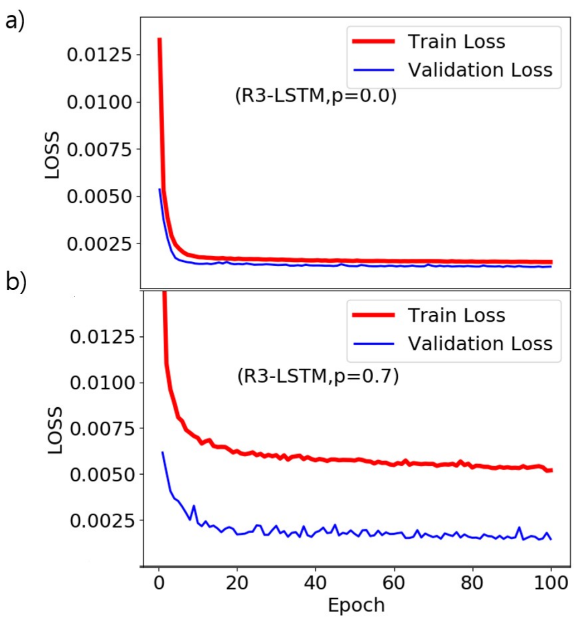


Figure 9: MSE results of $R3$ with different values of dropout.

Furthermore, Mean Absolute Percentage Error (MAPE) is taken into account which can be calculated as follows:

$$MAPE = \frac{1}{n} \sum_{i=1}^{n} \frac{| f_i - f_i' |}{f_i} \tag{18}$$

Fig. 12(a) and Fig. 12(b) show the results of $MAPE$ for $R3$ and $R4$, by applying different values of dropout which equals 0.1 and 0.5, respectively.

4.2.3 Sequence Length and Neurons: in this study, we also try to evaluate the influence of sequence length and neurons for training data. Fig. 13 depicts the results of MSE based on values of Sequence

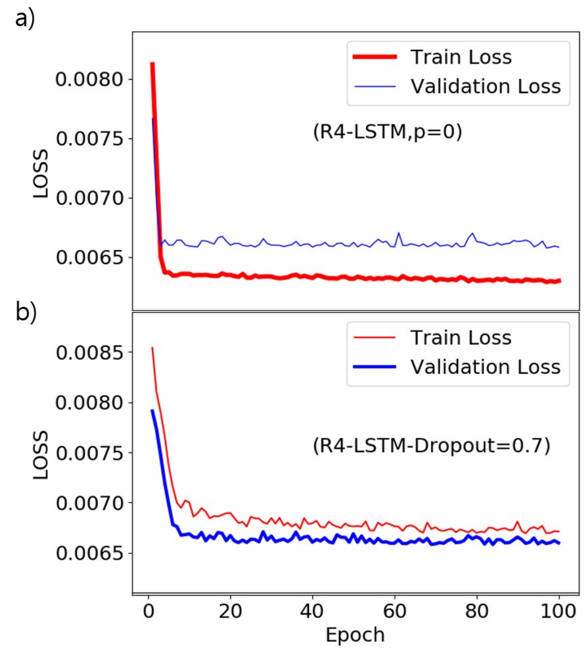


Figure 10: MSE results of $R4$ with different values of dropout.

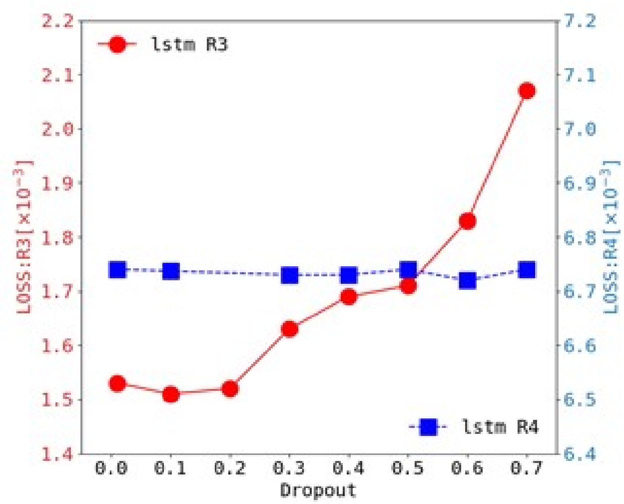
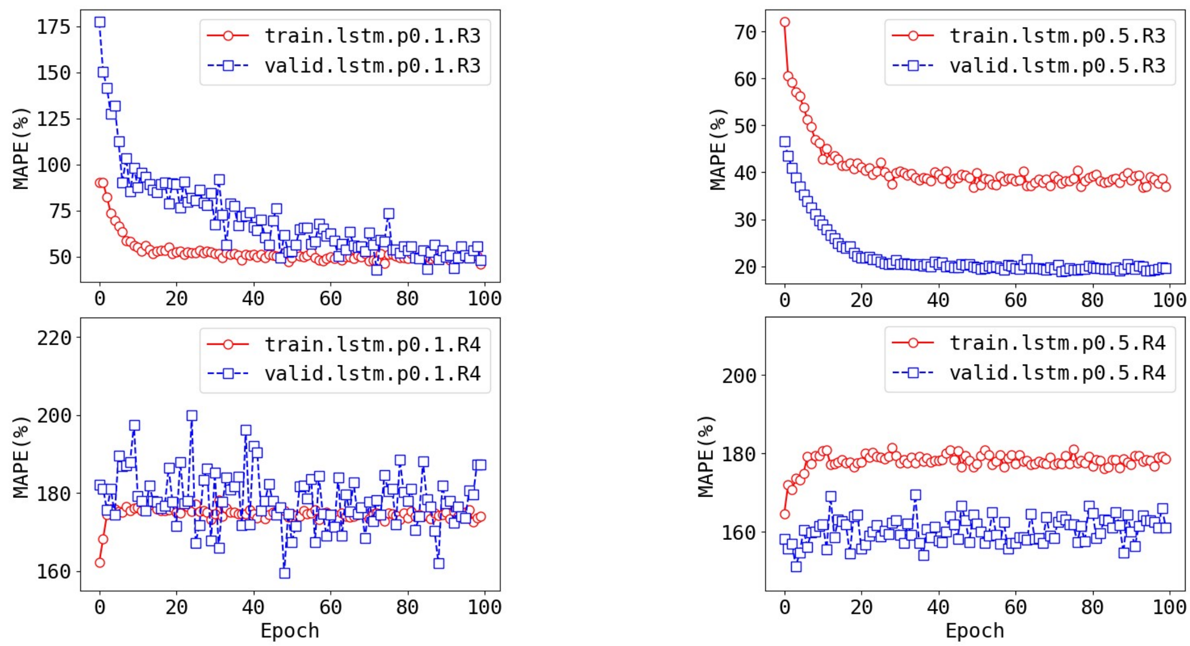


Figure 11: MSE results following variations of dropout.

Length. Specifically, we vary the value of sequence length from 10 to 60. As results, the best values of sequence length for short term prediction are around 20 to 30. Specifically, we do not need increasing the memory (sequence length) with the short term prediction. Moreover, the results of $MAPE$ following the values of sequence length are shown in Fig. 14.

In case of neurons, there is not much different when we vary the number of neurons. This means in case of short term prediction for highway systems, the number of neurons is not significant



(a) p = 0.1 (b) p =0.5

Figure 12: MAPE results with different values of dropout.

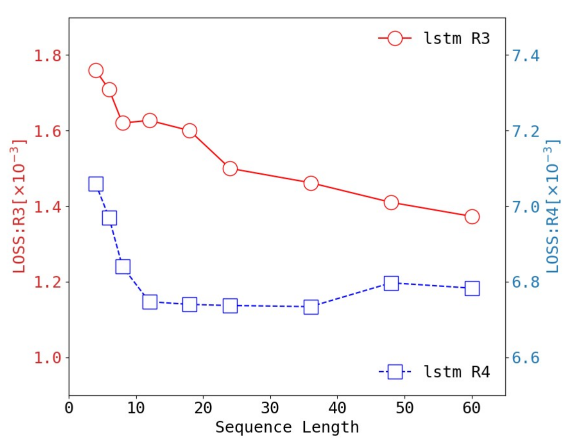


Figure 13: MSE results based on the values of Sequence Length.

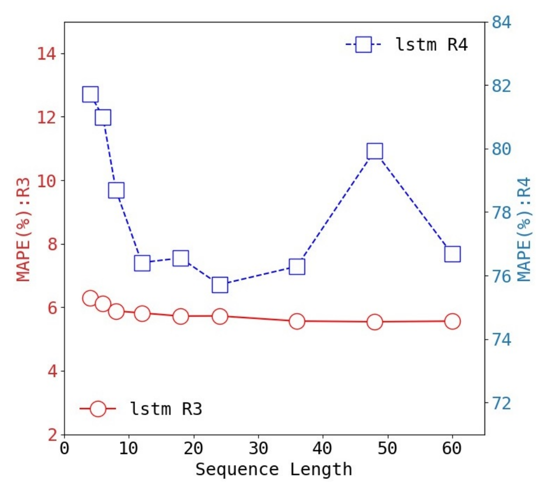


Figure 14: MAPE results based on the values of Sequence Length.

influence to the accuracy. Specifically, the appropriate value for the prediction is from 20 to 40 neurons.

4.2.4 Elapsed Times: Fig. 15 depicts the elapsed times of our implemented model based on the variation of sequence length. Particularly, we compare our model with Gated Recurrent Units (GRU),

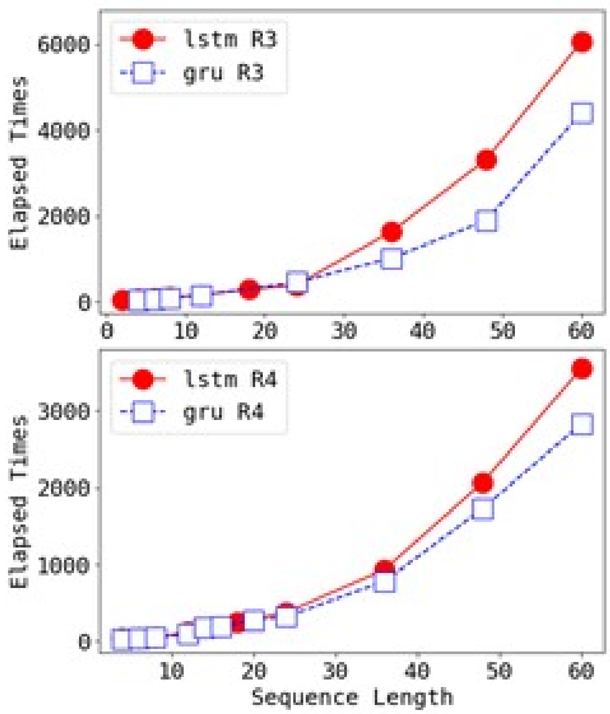


Figure 15: Elapsed times.

the model with the cost of less flexibility [2]. Results in figures indicate that with the small of training data, *GRU* is able to reduce the complexity time, however, with the increasing of dataset, LSTM is more appropriate for training data.

5 CONCLUSION

Time series forecasting is a challenging task, especially for the data with long sequences and noisy. Recently, with the advent of artificial intelligence and machine learning, deep learning methods provide promising results for time series analysis. In transportation system, deep learning based time series model for prediction has exploited the effectiveness of analyzing data sequence and extracting traffic flow features. In this regard, this paper focuses on implementing deep learning framework using LSTM-RNN models for predicting the congestion in highway system. The implementation indicates promising results that is worth continuing to work on.

For the future work of this study, we are going to exploit in more detail about the influences of parameters such as dropout, sequence length, and number of neurons for short-term prediction in highway systems. Moreover, in LSTM model, the weights of artificial neural networks are randomized for the initialization. Hence, determining average values from different results in each epoch of short-term prediction will be an important issue. There are interesting issues that we are taking into account regarding the research area of this study.

ACKNOWLEDGMENTS

This work was partly supported by Institute for Information & communications Technology Promotion(IITP) grant funded by the Korea government(MSIT) (No.2018-0-00494, Development of deep learning-based urban traffic congestion prediction and signal control solution system) and Korea Institute of Science and Technology Information(KISTI) grant funded by the Korea government(MSIT) (K-19-L02-C07-S01).

REFERENCES

[1] Wei Bao, Jun Yue, and Yulei Rao. 2017. A deep learning framework for financial time series using stacked autoencoders and long-short term memory. *PloS one* 12, 7 (July 2017), e0180944. https://doi.org/10.1371/journal.pone.0180944

[2] Rui Fu, Zuo Zhang, and Li Li. 2016. Using LSTM and GRU neural network methods for traffic flow prediction. In *Proceeding of 31st Youth Academic Annual Conference of Chinese Association of Automation (YAC '16)*. IEEE, Wuhan, China, 324–328. https://doi.org/10.1109/YAC.2016.7804912

[3] John Cristian Borges Gamboa. 2017. Deep Learning for Time-Series Analysis. *CoRR* abs/1701.01887 (Jan. 2017). http://arxiv.org/abs/1701.01887

[4] Felix A Gers and Jürgen Schmidhuber. 2000. Recurrent nets that time and count. In *Proceedings of the IEEE-INNS-ENNS International Joint Conference on Neural Networks (IJCNN 2000)*, Vol. 3. IEEE, Como, Italy, 189–194. https://doi.org/10.1109/IJCNN.2000.861302

[5] Kasthurirangan Gopalakrishnan, Siddhartha K Khaitan, Alok Choudhary, and Ankit Agrawal. 2017. Deep Convolutional Neural Networks with transfer learning for computer vision-based data-driven pavement distress detection. *Construction and Building Materials* 157 (Dec. 2017), 322–330. https://doi.org/10.1016/j.conbuildmat.2017.09.110

[6] Ibrahim Abaker Targio Hashem, Victor Chang, Nor Badrul Anuar, Kayode Adewole, Ibrar Yaqoob, Abdullah Gani, Ejaz Ahmed, and Haruna Chiroma. 2016. The role of big data in smart city. *Int J. Information Management* 36, 5 (Oct. 2016), 748–758. https://doi.org/10.1016/j.ijinfomgt.2016.05.002

[7] Geoffrey Hinton, Li Deng, Dong Yu, George Dahl, Abdel-rahman Mohamed, Navdeep Jaitly, Andrew Senior, Vincent Vanhoucke, Patrick Nguyen, Brian Kingsbury, et al. 2012. Deep Neural Networks for Acoustic Modeling in Speech Recognition: The Shared Views of Four Research Groups. *IEEE Signal Process. Mag.* 29, 6 (Nov. 2012), 82–97. https://doi.org/10.1109/MSP.2012.2205597

[8] Danqing Kang, Yisheng Lv, and Yuan-yuan Chen. 2017. Short-term traffic flow prediction with LSTM recurrent neural network. In *Proceeding of 20th IEEE International Conference on Intelligent Transportation Systems (ITSC '17)*. IEEE, Yokohama, Japan, 1–6. https://doi.org/10.1109/ITSC.2017.8317872

[9] Yisheng Lv, Yanjie Duan, Wenwen Kang, Zhengxi Li, and Fei-Yue Wang. 2015. Traffic Flow Prediction With Big Data: A Deep Learning Approach. *IEEE Trans. Intelligent Transportation Systems* 16, 2 (Apr. 2015), 865–873. https://doi.org/10.1109/TITS.2014.2345663

[10] Pankaj Malhotra, Lovekesh Vig, Gautam Shroff, and Puneet Agarwal. 2015. Long Short Term Memory Networks for Anomaly Detection in Time Series. In *Proceeding of 23rd European Symposium on Artificial Neural Networks (ESANN '15)*. Presses universitaires de Louvain, Bruges, Belgium, 89. https://www.elen.ucl.ac.be/esann/proceedings/papers.php?ann=2015

[11] Maryam M. Najafabadi, Flavio Villanustre, Taghi M. Khoshgoftaar, Naeem Seliya, Randall Wald, and Edin Muharemagic. 2015. Deep learning applications and challenges in big data analytics. *J. Big Data* 2 (Feb. 2015), 1. https://doi.org/10.1186/s40537-014-0007-7

[12] Heng Shi, Minghao Xu, and Ran Li. 2018. Deep Learning for Household Load Forecasting - A Novel Pooling Deep RNN. *IEEE Trans. Smart Grid* 9, 5 (Sep. 2018), 5271–5280. https://doi.org/10.1109/TSG.2017.2686012

[13] Yongxue Tian and Li Pan. 2015. Predicting Short-Term Traffic Flow by Long Short-Term Memory Recurrent Neural Network. In *Procceding of 2015 IEEE International Conference on Smart City (Smart City '15)*. IEEE Computer Society, Chengdu, China, 153–158. https://doi.org/10.1109/SmartCity.2015.63

[14] Di Zang, Jiawei Ling, Zhihua Wei, Keshuang Tang, and Jiujun Cheng. 2018. Long-Term Traffic Speed Prediction Based on Multiscale Spatio-Temporal Feature Learning Network. *IEEE Transactions on Intelligent Transportation Systems* (2018), 1 – 10. https://doi.org/10.1109/TITS.2018.2878068 (Early Access).

[15] Junbo Zhang, Yu Zheng, and Dekang Qi. 2017. Deep Spatio-Temporal Residual Networks for Citywide Crowd Flows Prediction. In *Proceedings of the 31st Conference on Artificial Intelligence (AAAI '17)*. AAAI Press, San Francisco„ USA., 1655–1661. http://aaai.org/ocs/index.php/AAAI/AAAI17/paper/view/14501

[16] Qingchen Zhang, Laurence T. Yang, Zhikui Chen, and Peng Li. 2018. A survey on deep learning for big data. *Information Fusion* 42 (July 2018), 146–157. https://doi.org/10.1016/j.inffus.2017.10.006

[17] Songyang Zhang, Xiaoming Liu, and Jun Xiao. 2017. On Geometric Features for Skeleton-Based Action Recognition Using Multilayer LSTM Networks. In

Proceeding of 2017 IEEE Winter Conference on Applications of Computer Vision (WACV '17). IEEE Computer Society, Santa Rosa, USA, 148–157. https://doi.org/10.1109/WACV.2017.24

[18] Zheng Zhao, Weihai Chen, Xingming Wu, Peter CY Chen, and Jingmeng Liu. 2017. LSTM network: a deep learning approach for short-term traffic forecast. *IET Intelligent Transport Systems* 11, 2 (Mar. 2017), 68–75. https://doi.org/10.1049/iet-its.2016.0208

[19] Li Zhu, Fei Richard Yu, Yige Wang, Bin Ning, and Tao Tang. 2019. Big Data Analytics in Intelligent Transportation Systems: A Survey. *IEEE Trans. Intelligent Transportation Systems* 20, 1 (Jan. 2019), 383–398. https://doi.org/10.1109/TITS.2018.2815678