Regular paper

Comparative Analysis of PM10 Prediction Performance between Neural Network Models

Yong-Jin Jung [ID] and Chang-Heon Oh[*] [ID], *Member*, *KIICE*

Department of Electrical, Electronics and Communication Engineering, Korea University of Technology and Education (KOREATECH), 31253 Korea

Abstract

Particulate matter has emerged as a serious global problem, necessitating highly reliable information on the matter. Therefore, various algorithms have been used in studies to predict particulate matter. In this study, we compared the prediction performance of neural network models that have been actively studied for particulate matter prediction. Among the neural network algorithms, a deep neural network (DNN), a recurrent neural network, and long short-term memory were used to design the optimal prediction model using a hyper-parameter search. In the comparative analysis of the prediction performance of each model, the DNN model showed a lower root mean square error (RMSE) than the other algorithms in the performance comparison using the RMSE and the level of accuracy as metrics for evaluation. The stability of the recurrent neural network was slightly lower than that of the other algorithms, although the accuracy was higher.

Index Terms: Neural network, Deep neural network, Recurrent neural network, Long short-term memory, Particulate matter

I. INTRODUCTION

Particulate matter (PM) is dust whose particles are invisible to the naked eye. According to some studies, PM has a variety of negative effects, including increasing the risk of development of cardiovascular, respiratory, and cerebrovascular diseases, as well as harming the body's defense system. The International Agency for Research on Cancer under the World Health Organization has designated PM as a group 1 carcinogen. PM has also been analyzed as a cause of declining economic activities within societies [1-7]. PM is formed as a result of air pollutants emitted from automobiles, factories, and cooking processes, among other sources. Industrial activities based on fossil fuels, such as coal and petroleum, have a significant impact on it. As a result, many people are increasingly showing interest in PM and requesting information to prepare for its occurrence in advance. In South Korea,

forecasts for PM have been implemented by combining various numerical model results with the prediction results of the community multi-scale air quality model, which is an air quality prediction model [8, 9]. People, however, are demanding higher forecast accuracy because the accuracy of existing models has not met expectations.

As a result, numerous studies have been conducted to increase the accuracy of PM prediction, and studies using neural networks are actively underway. In relation to PM predictions based on a deep neural network (DNN) among the available neural network algorithms, Dedovic et al. conducted a study using the weather variables and PM concentrations in Sarajevo for three years to achieve a prediction through an artificial neural network consisting of a single hidden layer. The study confirms that the performance in predicting the PM concentration can be improved using an extended input dataset containing PM concentration data from previous years

[10]. In a study related to recurrent neural network (RNN)-based PM prediction, Lim et al. proposed an RNN model that uses sequential data of previous PM concentrations to predict air pollutants. The proposed model was used to set the parameters that demonstrated the optimal prediction performance by changing the length of the input data, optimization function, and number of layers and nodes. It was demonstrated that the set parameters improved the PM prediction performance [11]. In a study related to long short-term memory (LSTM)-based PM prediction, Kang et al. conducted an LSTM-based PM prediction using weather data as the parameters. Data normalization was conducted to match the weather data range, and the PM concentration after 1 h was predicted using 3 h data, or the concentration after 12 h was predicted using 24 h data. The results of this study demonstrate that the data prediction performance can be improved using related weather factors [12].

In this study, we constructed PM_{10} models based on various neural network algorithms and trained them using the same data; we then conducted a comparative analysis of the performance of the algorithms. The aim was to investigate the impact of each algorithm characteristics on the prediction of PM; DNN, RNN, and LSTM algorithms were used to compare the performance of the three prediction models. We used the overall accuracy, detailed accuracy based on the air quality index (AQI), and the root mean square error (RMSE) to evaluate the performance of each model.

II. DATASET CONSTRUCTION AND PREPROCESSING

A. Dataset Construction

To design and test the prediction models, data were compiled based on the results of prior studies. Major data measured at 1 h intervals in Cheonan City, South Korea, during 2009-2018 were collected, as shown in Table 1. In the collected data, we observed that some data were not measured because of the environment and equipment, and for efficient

Table 1. Collected data

Category	Variable	Number of Data	Number of Missing Data
Meteorological elements	Temperature	87,622	26
	Wind Speed	87,612	36
	Wind Direction	87,601	47
Air pollutants	PM_{10}	249,268	13,676
	O_3	255,464	8,480
	CO	252,924	10,020
	NO_2	254,244	8,700
	SO_2	252,623	10,321

training, all data from the same time were removed when the data were constructed.

B. Data Preprocessing

The data used for training consisted of numerical and categorical data. The data scale may affect the effectiveness of the training, which may degrade the performance of the prediction models.
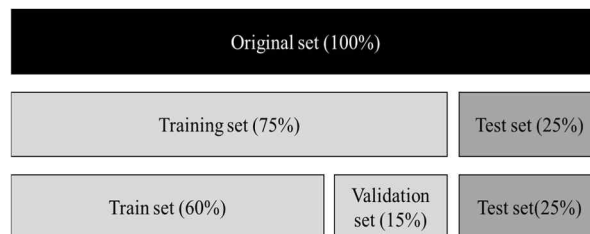

Fig. 1. Structure of dataset.

Therefore, a preprocessing process is required for the collected data to make them suitable for training. Because the wind direction had angles in 16 directions and the data were categorical, they were converted into 0s and 1s using one-hot encoding. The remaining data were numerical data that had different scales and were applied after conversion into values between zero and one using min-max scaling to unify them at the same scale. Hence, after preprocessing, all data could be represented as a value between zero and one.

III. Design of Prediction Models

A. Design of DNN-based Prediction Model

In the field of artificial intelligence, an artificial neural network (ANN) has an architecture that is modeled after the information-processing structure of the human brain and is used in machine learning and cognitive science. An ANN is defined through the connection pattern between neurons (nodes), that are the smallest unit of a neural network, the learning process for updating the weights of the connections, and finally the activation function, that converts the weighted inputs of neurons into activation level outputs. Sigmoid, hyperbolic tangent, and rectified linear unit (ReLU) functions are commonly used as activation functions; depending on the type, the input value is converted into an output value between zero and 1, or -1, which is then delivered to the next neuron.

The layers of a neural network primarily consist of an input, hidden, and output layers. If a single hidden layer is included, the neural network is classified as an ANN, and if

multiple hidden layers are included, it is classified as a DNN. The synapses connecting the neurons of each layer have their own weights. A DNN first initializes them with a certain value and then updates the weights in the direction of reducing the loss of the final output layer. To this end, the DNN uses a gradient descent method that updates the weights of all connections that constitute the DNN using the differential values for the loss function. The gradient descent method consists of two steps, that is, forward and back propagations, in which the final loss is calculated using the forward propagation and each weight is updated in the direction to reduce the loss using the back propagation [10, 13-16].

Therefore, the intensity of the training was adjusted using various parameters to train the prediction model. There are various types of parameters, and the training effect is sensitive to their values. Accordingly, we derived and applied the optimal values for the main parameters through a hyper-parameter search. Table 2 shows the hyper-parameter search results corresponding to the top three ranks; the parameter corresponding to the first rank was applied. ReLU was used as the activation function in the model, and Adam was used as the optimization function. In addition, 100 epochs were applied in the design of the prediction model.

B. Design of RNN-based Prediction Model

An RNN is a type of neural network specializing in time-series data, which has a sequential order, such as in natural language processing or speech recognition. An RNN facilitates the effective modeling of time-series information, such as text or speech, because it has a recursive structure, in which the output value of the hidden layer affects the output value of the next state [15, 17]. Unlike feedforward neural networks, such as a DNN, the RNN algorithm has a structure in which the neural network is recursive between the input and output through a cyclic loop in the internal nodes. Furthermore, it has a structure in which all the weights connecting the internal layers have the same value. A hyperbolic tangent or ReLU function is commonly used as the activa-

tion function of an RNN [18, 19].

In the case of an RNN, similar to that of a DNN, the intensity of the training is also adjusted using various parameters for training the prediction model. Similar to a DNN, an RNN has different types of parameters, and timesteps for recursive training are added and applied. A period of 1 d was used as the standard for applying the timesteps to the model, and accordingly, a fixed value of 24 was applied. In the case of other main parameters, the optimal values were checked using a hyper-parameter search. Table 2 shows the hyper-parameter search results corresponding to the top three ranks, and the parameter corresponding to the first rank was applied.

C. Design of LSTM-based Prediction Model

Proposed by Hochreiter and Schmidhuber in 1997, LSTM is a type of RNN that supplements the vanishing gradient problem of an RNN, which is difficult to train on long sequence data. The basic structure of LSTM is similar to that of an RNN; however, the structure of the memory cell that remembers the previous state has been modified using a method for solving the vanishing gradient problem of an RNN. Both RNN and LSTM have a structure in which the neural network is recursive between the input and output. However, RNN controls the information transfer and propagation through a recursion on a single layer, whereas LSTM controls the information transfer and propagation through four different fully connected layers inside the cell: three types of gates, that is, a forget gate, an input gate, and an output gate, and the cell state [20, 21].

Every gate of an LSTM outputs a value between zero and one through a sigmoid operation, and the weight is set individually for each gate. The forget gate remembers all previous information if the value is close to one in a sigmoid operation of the previous output and input values of the current stage and deletes the previous information if it is close to zero. The input gate determines the degree of reflection of the current information, and the data to be stored in the cell state are determined based on the product of the resulting values of the hyperbolic tangent function and the sigmoid function of the input value of the current stage and the previous output. Subsequently, the cell state of the current stage is updated using the information obtained from the forget and input gates. Finally, the final output is determined based on the product of the sigmoid operation result of the previous output, the input value of the current stage, and the result of the tanh function of the cell state [22–24].

Therefore, the intensity of training is adjusted using various parameters for the training of the model, and the training effect is sensitive to the corresponding values in the same manner as with the other algorithms. Accordingly, we derived and applied the optimal values for the main parameters using a hyper-parameter search. Herein, identical to the

Table 2. Hyper-parameter search result (DNN, RNN, LSTM)

	Rank	Layer	Hidden Node	L2	Dropout Rate	Batch Size
DNN	1	2	100	0.01	0.1	100
	2	3	140	0.001	0.5	60
	3	2	60	0.01	0.1	80
RNN	1	1	80	0	0	20
	2	1	60	0	0	20
	3	2	100	0	0	20
LSTM	1	2	80	0	0.3	60
	2	2	40	0	0.2	40
	3	2	40	0	0.1	20

case of the RNN model, the time steps were set to 24. ReLU and Adam were used as the activation and optimization functions, respectively, and 100 epochs were applied. Table 2 shows the hyper-parameter search results corresponding to the top three ranks, and the parameter corresponding to the first rank was applied to the model design.

IV. Performance Evaluation

The optimal values derived using the hyper-parameter search were applied to the design of each model, and the training set constructed as shown in Fig. 1 was used to train each model. Subsequently, the test set was used to evaluate the performance based on the predicted values, which were the training results of each model. The RMSE was used as a basis for the performance evaluation, and the accuracy of each AQI was used to determine the detailed prediction accuracy. Fig. 2 shows the prediction results of the sections that showed rapid changes in concentration among the test data for each model. Table 3 shows the RMSE and detailed prediction accuracy based on the prediction results of each model.

In the case of a DNN, as shown in Fig. 2, the difference between the actual PM concentrations and the predicted values was small; however, in a section of 150-200 $\mu g/m^2$, where the concentration changed rapidly, the PM concentrations were overestimated or underestimated compared to the actual concentrations. Furthermore, in the case of a PM concentration higher than 100 $\mu g/m^2$ in that section, the difference in concentration was relatively large compared to that of the other sections. In the case of an RNN, the errors were not large between the actual PM concentrations and the predicted values. However, the concentration was overestimated for a typically high-concentration PM higher than 100 $\mu g/m^2$. In contrast, in the case of a low-concentration PM, the actual PM concentration and predicted value were relatively close. In the case of LSTM, the concentration was overestimated for a typically high-concentration PM higher than 100 $\mu g/m^2$. In a section of 150-170 $\mu g/m^2$, particularly, where rapid changes in concentration were observed, the error was large between the actual PM concentration and the predicted value.

When the overall RMSE was compared, the RMSE of the

DNN was 8.3459, indicating that it was a better indicator than the other models. However, in the case of overall accuracy, the RNN had a value of 87.58%, showing a higher accuracy compared to the other models. For a "good" AQI, the RNN model showed the highest accuracy with a value of 84.28%, and in the case of a "moderate" AQI, the DNN model showed the highest accuracy with a value of 93.14%. In the case of a "bad" AQI, the accuracy of the LSTM model was 76.83%, which was higher than that of other models, and in the case of a "very bad" AQI, the accuracy of the RNN model was 72.79%, which was higher than that of other models. When the DNN and RNN models were compared based on the RMSE and overall accuracy, the prediction accuracy of the RNN model was higher than that of the DNN model. However, the margin of error for the samples that were not successfully predicted was narrower in the DNN model than in the RNN model. In this case, the evaluation results showed that the DNN model was more stable than the RNN model in terms of the margin of error.

V. CONCLUSIONS

The air pollution problem caused by increases in population, automobiles, and industrial activities has emerged as a major environmental problem worldwide. In particular, as the effect of PM on the human body has been revealed, people are using PM forecasts in daily life to respond preemptively to the air environment. Accordingly, people are demanding a high level of PM prediction accuracy, and prediction studies using a variety of methods are underway. Among them, studies are actively being conducted on the prediction of PM concentrations using neural networks.

In this study, we used neural network algorithms that have been applied in many studies to predict the PM concentration, and conducted a performance analysis and evaluation for each model to select an algorithm that is suitable for PM concentration prediction. To this end, we collected weather and air pollutant data monitored over a 10 year period in the Cheonan region of South Korea. The training, validation, and test sets were constructed from the collected data for use in the training of the prediction models. The data were then preprocessed to minimize the learning problem that occurs when the characteristics of the data differ from each other. Because the wind direction is categorical data expressed in terms of 16 directions, it was converted into a vector type of 0s and 1s using one-hot encoding. However, because the other data had different scales, min-max scaling was used to convert the range of numerical representations of the data into a range of values between zero and one. DNN, RNN, and LSTM were selected as the neural network algorithms used for the performance evaluation of the prediction models. For optimal prediction performance, the optimal values

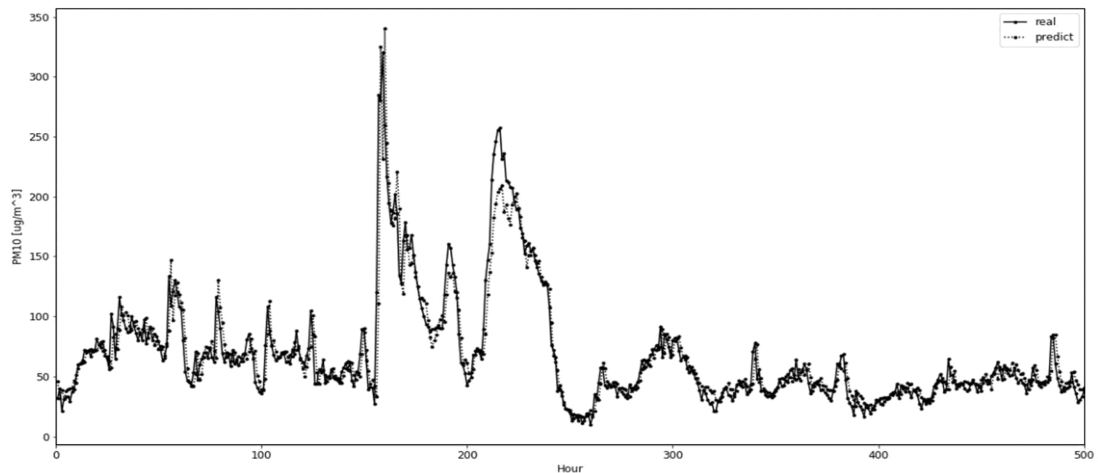**Table 3.** Comparison of prediction performance between models

Indicator	DNN	RNN	LSTM
RMSE	8.3459	8.3653	8.3964
Overall accuracy	87.38%	87.58%	87.1%
Accuracy for "good" AQI	79.9%	84.28%	79.2%
Accuracy for "moderate" AQI	93.14%	91.22%	92.8%
Accuracy for "bad" AQI	75.5%	75.24%	76.83%
Accuracy for "very bad" AQI	65.44%	72.79%	66.18%
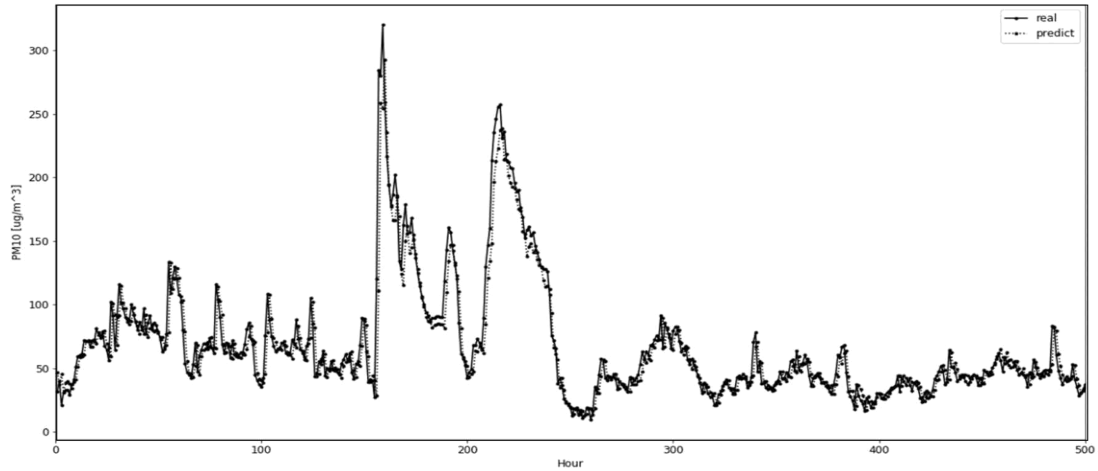
(a) DNN model
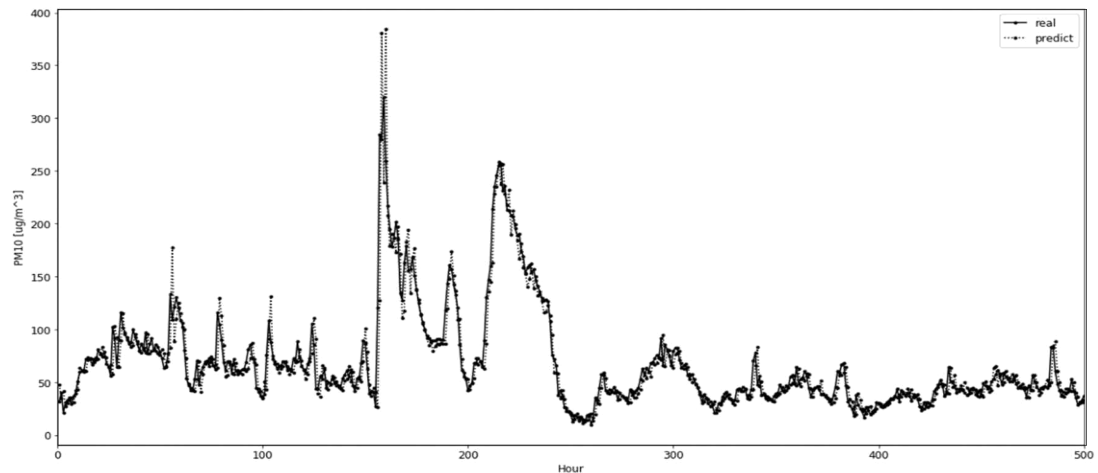
(b) RNN model

(c) LSTM model

Fig. 2. PM10 prediction result of each model.

of the parameters applied to each algorithm are required. A hyper-parameter search was conducted to obtain them, which were then applied to the model design. The same data were

used to train the designed models and evaluate their performances.

To compare the performance of the prediction models con-

structed using neural network algorithms, we checked the trend changes of the actual and predicted values. Furthermore, the detailed prediction accuracies were examined by classifying them based on RMSE and AQI. For the trend changes derived using each model, the differences between the models were minimal. However, in the performance comparison using the RMSE and the level of accuracy, the DNN model showed a lower RMSE than the other algorithms, showing stably predicted values. In the case of an RNN, the stability was slightly lower than that of the other algorithms, although the accuracy was higher.

For a predicted PM concentration, the accuracy is important in terms of consistency between the actual and predicted values; however, considering the wide concentration range, it was determined that a DNN-based prediction model with a small margin of error is suitable. In the future, we plan to use neural network algorithms to conduct a study to reduce the margin of error in PM concentration prediction. It is expected that the construction of a prediction model with better performance will increase the use of reliable prediction information.

ACKNOWLEDGEMENTS

REFERENCES

[1] C. A. Pope III and D. W. Dockery, "Health effects of fine particulate air pollution: line that connect," *Journal of the Air & Waste Management Association*, vol. 56, no. 6, pp. 709-742, 2006. DOI: 10.1080/10473289.2006.10464485.

[2] S. Fuzzi, U. Baltensperger, K. Carslaw, S. Decesari, H. D. Gon, M. C. Facchini, D. Fowler, I. Koren, B. Langford, U. Lohmann, E. Nemitz, S. Pandis, I. Riipinen, Y. Rudich, M. Schaap, J. G. Slowik, D. V. Spracklen, E. Vignati, M. Wild, M. Williams, and S. Gilardoni, "Particulate matter, air quality and climate: Lessons learned and future needs," *Atmospheric Chemistry and Physics*, vol. 15, no. 14, pp. 8217–8299, 2015. DOI: 10.5194/acp-15-8217-2015.

[3] A. Valavanidis, K. Fiotakis, and T. Vlachogianni, "Airborne particulate matter and human health: toxicological assessment and importance of size and composition of particles for oxidative damage and carcinogenic mechanisms," *Journal of Environmental Science and Health, Part C*, vol. 26, no. 4, pp. 339-362, 2008. DOI: 10.1080/10590500802494538.

[4] J. O. Anderson, J. G. Thundiyil, and A. Stolbach, "Clearing the air: A review of the effects of particulate matter air pollution on human health," *Journal of Medical Toxicology*, vol. 8, no. 2, pp. 166-175, 2012. DOI: 10.1007/s13181-011-0203-1.

[5] K. H. Kim, E. Kabir, and S. Kabir, "A review on the human health impact of airborne particulate matter," *Environment International*, vol. 74, pp. 136-143, 2015. DOI: 10.1016/j.envint.2014.10.005.

[6] N. J. Hime, G. B. Marks, and C. T. Cowie, "A comparison of the health effects of ambient particulate matter air pollution from five emission sources," *International Journal of Environmental Research and Public Health*, vol. 15, no. 6, 2018. DOI: 10.3390/ijerph15061206.

[7] World Health Organization (WHO), *Health Effects of Particulate Matter: Policy Implications for Countries in Eastern Europe, Caucasus and Central Asia, Regional Office for Europe*, 2013.

[8] National Institute of Environmental Research (NIER), "A study of data accuracy improvement for national air quality forecasting (III)," *National Institute of Environmental Research*, NIER-RP2016-248, 11-1480523-002809-01, 2016.

[9] Board of Audit and Inspection (BAI), "Weather forecast and earthquake notification system operation," *International THE Board of Audit and Inspection of KOREA*, 2017.

[10] M. M. Dedovic, S. Avdakovic, I. Turkovic, N. Dautbasic, and T. Konjic, "Forecasting PM10 concentrations using neural networks and system for improving air quality," *2016 XI International Symposium on Telecommunications(BIHTEL)*, pp. 1–6, 2016. DOI: 10.1109/BIHTEL.2016.7775721.

[11] Y. B. Lim, I. Aliyu, and C. G. Lim, "Air pollution matter prediction using recurrent neural networks with sequential data," in *Proceedings of the 2019 3rd International Conference on Intelligent Systems, Metaheuristics & Swarm Intelligence*, pp. 40-44, 2019. DOI: 10.1145/3325773.3325788.

[12] S. W. Kang, N. G. Kim, and B. D. Lee, "Fine dust forecast based on recurrent neural networks," in *2019 21st International Conference on Advanced Communication Technology (ICACT)*, pp. 456-459, 2019. DOI: 10.23919/ICACT.2019.8701978.

[13] J. B. Ahn and Y. M. Cha, "A comparison study of corrections using artificial neural network and multiple linear regression for dynamically downscaled winter temperature over South Korea," *Asia-Pacific Journal of Atmospheric Sciences*, vol. 41, no. 3, pp. 401-413, 2005.

[14] J. W. Oh, J. H. Song, K. H. Kim, and S. H. Jung, "Automatic composition using training capability of artificial neural networks and chord progression," *Journal of Korea Multimedia Society*, vol. 18, no. 11, pp. 1358-1366, 2015. DOI: 10.9717/kmms.2015.18.11.1358.

[15] V. Sze, Y. H. Chen, T. J. Yang, and J. S. Emer, "Efficient processing of deep neural networks: A tutorial and survey," in *Proceedings of the IEEE*, vol. 105, no. 12, pp. 2295-2329, 2017. DOI: 10.1109/JPROC.2017.2761740.

[16] N. D. Al-Shakarchy and I. H. Ali, "Detecting abnormal movement of driver's head based on spatial-temporal features of video using deep neural network DNN," *Indonesian Journal of Electrical Engineering and Computer Science*, vol. 19, no. 1, pp. 344-352, 2020. DOI: 10.11591/ijeecs.v19.i1.pp344-352.

[17] K. I. Funahashi and Y. Nakamura, "Approximation of dynamical systems by continuous time recurrent neural networks," *Neural Networks*, vol. 6, no. 6, pp. 801–806, 1993. DOI: 10.1016/S0893-6080(05)80125-X.

[18] Z. W. Yahaya, F. H. K. Zaman, and M. F. A. Latip, "Prediction of energy consumption using recurrent neural networks (RNN) and nonlinear autoregressive neural network with external input (NARX)," *Indonesian Journal of Electrical Engineering and Computer Science*, vol. 17, no. 3, pp. 1215-1223, 2020. DOI: 10.11591/ijeecs.v17.i3.pp1215-1223.

[19] S. Y. Yoo, J. C. Lee, J. H. Lee, H. J. Hwang, and S. S. Lee, "A study on time series data filtering of spar platform using recurrent neural network," *Journal of the Korean Society of Marine Engineering*, vol. 43, no. 1, pp. 8-17, 2019. DOI: DOI: 10.5916/jkosme.2019.43.1.8.

[20] X. Wang and H. C. Kim, "Text categorization with improved deep

learning methods," *Journal of Information and Communication Convergence Engineering*, vol. 16, no. 2, pp. 106-113, 2018. DOI: 10.6109/jicce.2018.16.2.106.

[21] C. H. Hwang, H. S. Kim, and H. K. Jung, "Detection and correction method of erroneous data using quantile pattern and LSTM," *Journal of Information and Communication Convergence Engineering*, vol. 16, no. 4, pp. 242-247, 2018. DOI: 10.6109/jicce.2018.16.4.242.

[22] Y. H. Kim, Y. K. Hwang, T. G. Kang, and K. M. Jung, "LSTM language model based Korean sentence generation," *The Journal of Korean Institute of Communications and Information Sciences*, vol. 41, no. 5, pp. 592-601, 2016. DOI: 10.7840/kics.2016.41.5.592.

[23] S. U. Kwon, D. H. Han, S. Y. Park, and J. H. Kim, "Long short term memory-based state-of-health prediction algorithm of a rechargeable lithium-ion battery for electric vehicle," *The Transactions of The Korean Institute of Electrical Engineers*, vol. 68, no. 10, pp. 1214–1221, 2019. DOI: 10.5370/KIEE.2019.68.10.1214.

[24] R. W. Kadhim and M. T. Gaata, "A hybrid of CNN and LSTM methods for securing web application against cross-site scripting attack," *Indonesian Journal of Electrical Engineering and Computer Science*, vol. 21, no. 2, pp. 1022–1029, 2021. DOI: 10.11591/ijeecs.v21.i2.pp1022-1029.

Yong-Jin Jung

received his B.S. in Electronics Engineering from Kongju National University, Cheonan, South Korea, in 2014, and his M.S. in Electrical, Electronics, and Communication Engineering from Korea University of Technology and Education (KOREATECH) in 2016. He is currently pursuing a Ph.D. in Electrical, Electronics, and Communication Engineering at the Korea University of Technology and Education (KOREATECH). His research interests include machine learning, data analysis, and deep learning.

Chang-Heon Oh

received his B.S. and M.S.E. degrees in telecommunication and information engineering from Korea Aerospace University, Kyunggi-Do, Korea, in 1988 and 1990, respectively. He received Ph.D. degree in avionics engineering from Korea Aerospace University, Kyunggi-Do, Korea, in 1996. From February 1990 to August 1993, he worked with Hanjin Electronics Co., where he was involved in the research and development of radio communication and monitoring systems. From October 1993 to February 1999, he worked with the CDMA R&D Center of Samsung Electronics Co., where he was involved in the design and development of CDMA cellular systems and CDMA PCS systems for successful commercial CDMA deployment in Korea. Since March 1999, he has been with the School of Electrical, Electronics and Communication Eng., Korea University of Technology and Education (KOREATECH), Cheonan, Korea, where he is currently a professor. His research interests are in the areas of wireless/mobile communication, wireless localization, IoT, and engineering education.