Hybrid Model of Convolutional LSTM and CNN to Predict Particulate Matter

Seonggu Lee and Jitae Shin

Abstract—**Particulate matter (PM) can harm human health by causing lung cancer, pneumonia, or cardiovascular disease. There is a growing awareness of dangerous PM among people and governments. In order to prepare for the risk, the prediction performance of PM is important. Therefore, many kinds of research are developing various prediction models. Among the models, LSTM based models show the best result and it uses various auxiliary data, including spatial features to improve performance. However, spatial features can be depreciated because all input data has to be unfolded to 1D vector. In this paper, we apply Convolutional LSTM to our model to take advantage of the spatiotemporal relation of the wind and PM forecasting problem. Also, we add CNN to extract temporal features of the dataset on our model in parallel. Finally, we combine both Convolutional LSTM and CNN to predict more accurate PM concentration. In the experiment, we compared this model with LSTM and CNN-LSTM models in previous studies. At the result, the hybrid model showed the best performance.**

Index Terms—**Deep learning, convolutional long short-term memory (ConvLSTM), CNN, particulate matter prediction.**

I. INTRODUCTION

Particulate matter (PM) is the term for a very small mixture of solid particles and liquid droplets found in the air [1]. This term includes PM10 whose diameters are generally 10 micrometers and smaller; and PM2.5 whose diameters are generally 2.5 micrometers and smaller [1]. PM is very small and light, floating in the atmosphere. People can breathe these PM which causes serious health problems including decreased lung function, obesity, asthma and exacerbation of chronic respiratory, and cardiovascular diseases [2]. PM also has a direct negative impact on other human life areas. It is the main problem of reduced visibility (e.g., haze) in many countries, which can increase the possibility of a traffic accident. Additionally, it contributes to acid rain effects, and damages farm crops [1]. Due to these bad effects on human life, many people are interested in an accurate prediction method of PM concentration. They can avoid contact with PM and prevent various problems that these can cause if they know the temporal change of PM in advance. To predict PM concentrations, many kinds of research are underway; deterministic method and statistical method are generally used.

A deterministic method establishes a numerical model which is based on the aerodynamic theory and physicochemical processes tracking the generation, dispersion, and transmission process of pollutants [3]; WRF/Chem model and CMAQ model [4] are commonly used. This method is helpful to understand the PM diffusion process. However, it needs too much prior complex knowledge [5].

A Statistical method is not sophisticated but simply applies statistics-based models to predict air quality [3]. The models established by the statistical method can be divided into a non-deep-learning based model and deep-learning based model. The non-deep-learning based model includes the autoregressive integrated moving average (ARIMA) [6], multiple linear regressions (MLR) and the support vector regression (SVR). The deep-learning based model includes the multilayer perceptron (MLP), the deep recurrent neural network (DRNN) [7], the long short-term memory neural network (LSTM NN) [3] and the hybrid model based on convolutional neural network and LSTM (CNN-LSTM) [8].

In the recent study, the deep-learning-based model showed better prediction accuracy than non-deep-learning-based method [3], [8]. Among these models, LSTM-based models, remembering information for long periods of time showed better performance. Especially, the CNN-LSTM model showed the highest accuracy. This model uses CNN, which extracts temporal feature to improve performance [8] and then uses LSTM to analyze the extracted features to estimate the PM concentration of the next point of time.

A used dataset is also important to predict PM concentration more accurately. Several studies selected nearby PM measurement stations as a source of the dataset based on the spatial correlation between these stations and a target station to apply the spatial information on the model, and used auxiliary data such as meteorological data or other air pollutants [3], [7]. Using this additional data, these experiments showed a better predictive performance.

However, LSTM model or CNN-LSTM model cannot handle both spatial data and temporal data at the same time. Hence, the additional data can be depreciated. Because all input data has to be unfolded to 1D vector before temporal data processing on the models, spatial information can be lost during this process [9]. Thus, these models have a limitation in predicting PM concentration which changes continuously depending to the spatiotemporal axis. For example, if a wind is strong, PM from a nearby area can flow in the target area where we want to measure. In this case, the PM concentration changes depending on spatiotemporal data, including wind intensity, wind direction, and surrounding PM concentration. However, the recent studied model cannot process spatiotemporal data simultaneously as mentioned above. This fact is the reason for deteriorating the predictive performance of the model.

In this paper, we use Convolutional LSTM (ConvLSTM), which combines convolution operation on LSTM and include 2D spatial information, thus it can analyze spatial and temporal information at the same time [9]. By doing so, we can take advantage of the spatiotemporal relation of the wind and PM forecasting problem and inflow of PM can be simulated on the forecasting model. Also, we add 3D CNN to extract the temporal feature of the dataset in parallel and finally combine both ConvLSTM and 3D CNN to predict PM concentration more accurately.

II. PROPOSED METHOD

A. Data Description

The PM dataset shows collected the PM concentration of Korea where it was relatively high in the world. In detail, PM10 data were measured every hour in 53 places in Seoul, Gyeonggi, and Incheon. We could access the data on https://www.airkorea.or.kr/. Further, since weather data has a correlation with the prediction of PM concentration, we also used weather data from three meteorological stations located in the area where the PM measurement stations had been installed. We could access the data on https://data.kma.go.kr/data/grnd/. The measuring stations are represented in the Google map as shown in Fig. 1. The measurement period of the data is from January 1, 2014 to December 31, 2016. We used 2014 and 2015 as training data and 2016 as test data. All raw data are listed in the Table I in detail.
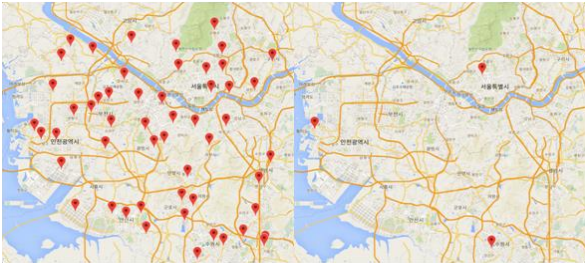

Fig. 1. PM station(left) and meteorological station(right).

TABLE I: INPUT PARAMETER

Region		
Gyeonggi, Incheon, Seoul		
Measurement Period		
2014.1.1 1:00 – 2016.12.31 24:00		
Feature	**Unit**	**Source**
1h PM10 concentration	ug/m³	53 of PM10 stations
1h Temperature	°C	3 of meteorological stations
1h Humid	%	3 of meteorological stations
1h Vapor Pressure	hPa	3 of meteorological stations
1hWind speed	m/s	3 of meteorological stations
1h Wind direction	°	3 of meteorological stations

B. Data Pre-processing

We need a pre-processing to train the model correctly. The ConvLSTM model used in this paper has to contain 2D spatial information of the data. Therefore, as shown in Fig. 2, the locations of the measuring stations were grid-mapped to each cell of the 2D matrix based on the Transverse Mercator (TM) coordinate. The TM coordinate is a conformal mapping of the earth ellipsoid where a central meridian is mapped into a straight line at constant scale and is used for the grid systems of several countries in geography [10]. The grid-mapping formula of i-th stations is in (1). X_{TM} and Y_{TM} denote the TM coordinate given by integer and min is the function that returns minimum value. According to the equation (1), the size of the matrix was set to 9x9.

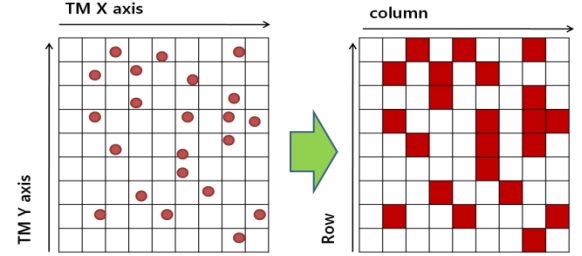$$(row, col) = (X_{TM}(i) - \min(X_{TM}), Y_{TM}(i) - \min(Y_{TM})) \quad (1)$$

Fig. 2. Grid mapping of station distribution.

Then, the data of measuring stations were grid-mapped to each cell in which the location had been mapped. The meteorological data and the PM data were combined and represented as three-dimensional tensor in every time slice. All missing values were interpolated through the inverted distance weight (IDW) used for spatial data interpolation [11]. Additionally, we need to compare other models which do not include 2D spatial information. Thus, we also unfolded the conjugated data to make the input of the models. All process is shown in Fig. 3.
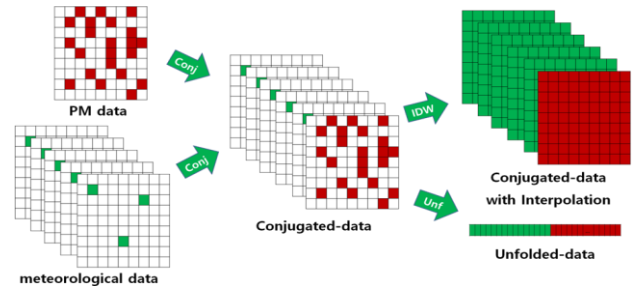

Fig. 3. Conjugation, interpolation and unfolding.

C. Modeling

The proposed model is shown in Fig. 4, which is a hybrid model of ConvLSTM and 3D CNN. The input data are four-dimension including past 24-hour time, row, column and features. The kernel size of ConvLSTM is two-dimension including row and column. The kernel size of 3D CNN is three-dimension including time, row and column.

The ConvLSTM model uses three kinds of features as inputs including PM10 concentration, wind intensity, and wind direction data. The ConvLSTM model structure applies encoder-decoder for future predictor [12]. Two 2D ConvLSTM layers represent encoder and decoder respectively. The encoder creates hidden representation from past 24-hour data. This is copied by four times and used to inputs of the decoder. Therefore, spatiotemporal features are extracted sequentially to predict the next 4-hour PM concentrations.

The 3D CNN Model uses all feature data except the direction of the wind as input. Three 3D CNN layer extracts

only accumulated temporal features from past 24-hours in parallel. The features extracted from the two models are merged, and three 2D CNN layers with a kernel size of 1x1 analyze the merged feature. Then, next 4-hour PM concentrations of all stations are extracted from the output of last 2D CNN layer.

In our model, we set the filter size to 128 on 3D CNN layers and 2D ConvLSTM layers and to 1024 on 2D CNN layers. Specifically, we make multiple models with the different kernel size, [k,k] of the 2D ConvLSTM layers, which was subdivided into [1,1], [3,3], [5,5], which is for we need to compare the result depending on the kernel size. Predicted value can be affected by the more distant measuring stations in every prediction time if the kernel size is larger [9]. Also, we set the padding to (k-1) / 2 to keep the spatial information.
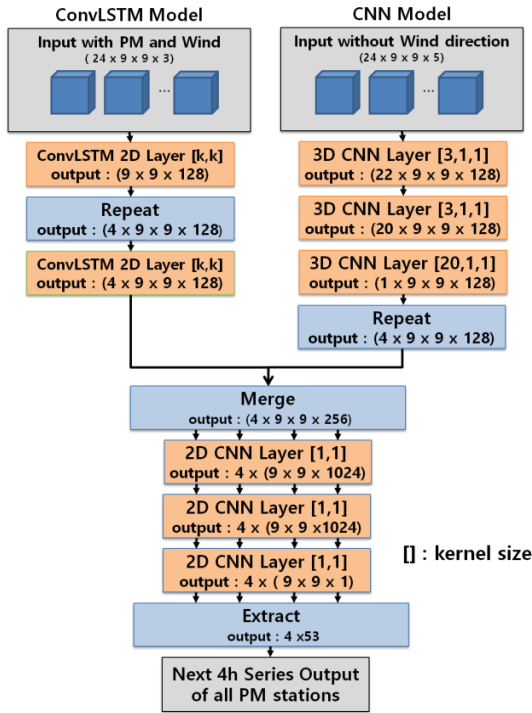

Fig. 4. Hybrid Model of ConvLSTM and CNN.

To verify the performance of the proposed models, various comparable models are needed. Those models maintain the encoder-decoder for future predictor structure to predict the next 4-hour PM concentration at same time.

1) ConvLSTM : The CNN model part is excluded from the Hybrid model and the ConvLSTM model part is directly connected to 2D CNN. Merging is also excluded. This model uses all features.

2) CNN-LSTM : This model is similar to the CNN-LSTM model [8] proposed in the previous study, but two LSTM layers are used to apply the encoder-decoder structure. Two 1D CNN layer and two LSTM layer are connected sequentially. Because this model does not need 2D spatial data, we use unfolded 1D vector from 3D tensors. Thus, 2D spatial information is excluded.

3) LSTM : This model is made from the CNN-LSTM model, but two 1D CNN layers are excluded. This model is the simplest model, thus this is used as the baseline for this paper.

III. EXPERIMENT

To evaluate the performance of the proposed model, three indicators are used including the root mean square error (RMSE), the mean absolute error (MAE), and the index of agreement (IOA), which are mainly used to show model accuracy in weather forecasting research. The formulas for each indicator are in (2), (3) and (4); N denotes the number of test samples, y_i denotes predicted value and y_i^* denotes true value. The IOA denotes how closely the predicted value matches the actual value. If the predicted value is close to the true value, the IOA value is close to 1.

$$RMSE = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(y_i - y_i^*)^2} \qquad (2)$$

$$MAE = \sqrt{\frac{1}{N}\sum_{i=1}^{N}|y_i - y_i^*|} \qquad (3)$$

$$IOA = 1 - \frac{\sum_{i=1}^{N}(|y_i - y_i^*|)^2}{\sum_{i=1}^{N}(|y_i - <y_i^*>| + |y_i^* - <y_i^*>|)^2} \qquad (4)$$

The average of RMSE, MAE, and IOA of PM concentration from all PM measuring stations for the next 4 hours are presented in Table II, Table III and Table IV, respectively. In the tables, Hybrid denotes the proposed model in this paper, and [1,1], [3,3] and [5,5] denotes the kernel size of Convolution LSTM.

TABLE II: COMPARISON OF RMSE OF THE DIFFERENT MODEL

Model	Next 1-hour	Next 2-hour	Next 3-hour	Next 4-hour	Average
Hybrid [1,1]	10.49	14.36	17.04	19.13	15.25
Hybrid [3,3]	**9.76**	**13.36**	**16.38**	**18.86**	**14.59**
Hybrid [5,5]	10.44	13.72	16.59	18.99	14.93
ConvLSTM [3,3]	9.85	13.61	16.69	19.19	14.84
CNN-LSTM	13.91	15.41	17.33	19.26	16.48
LSTM	13.90	16.00	17.99	19.67	16.89

TABLE III: COMPARISON OF MAE OF THE DIFFERENT MODEL

Model	Next 1-hour	Next 2-hour	Next 3-hour	Next 4-hour	Average
Hybrid [1,1]	7.35	9.77	11.47	12.80	10.35
Hybrid [3,3]	**6.89**	**9.11**	**11.03**	**12.62**	**9.91**
Hybrid [5,5]	7.41	9.46	11.30	12.86	10.26
ConvLSTM [3,3]	6.95	9.27	11.23	12.84	10.07
CNN-LSTM	9.80	10.63	11.75	12.89	11.27
LSTM	9.89	11.24	12.48	13.49	11.78

TABLE IV: COMPARISON OF IOA OF THE DIFFERENT MODEL

Model	Next 1-hour	Next 2-hour	Next 3-hour	Next 4-hour	Average
Hybrid [1,1]	0.9654	0.9309	0.8959	0.8624	0.9136
Hybrid [3,3]	**0.9703**	**0.9415**	**0.9089**	0.8768	**0.9244**
Hybrid [5,5]	0.9647	0.9361	0.9038	0.8717	0.9191
ConvLSTM [3,3]	0.9695	0.9378	0.9028	0.8686	0.9197
CNN-LSTM	0.9385	0.9240	0.9026	**0.8776**	0.9107
LSTM	0.9353	0.9109	0.8840	0.8579	0.8971

Among the proposed Hybrid models with different kernel size, the Hybrid model with [3,3] kernel size shows the best performance in all indicators (RMSE: 14.59, MAE : 9.91, IOA : 0.9244) and the Hybrid model with [5,5] kernel size (RMSE : 14.93, MAE : 10.26, IOA : 0.9191) shows the better performance than the Hybrid model with [1,1] kernel size (RMSE : 15.25, MAE : 10.35, IOA : 0.9136.) If the kernel size is [1,1], the ConvLSTM layers in the Hybrid model cannot extract additional spatial features from the near measuring stations. Therefore, the result shows that extracting additional spatial features makes the performance better. It also represents that considering the inflow of PM by wind effect is important. The best kernel size for effectively extracting spatial information from the given dataset is not [5,5] but [3, 3]. This means that the model with large kernel size can cause performance degradation because more distant measuring stations used to every time slices input is not relevant to predict target measuring stations.

Compared with all models, Hybrid models and ConvLSTM model which keep two-dimensional spatial information show better results than CNN-LSTM (RMSE : 16.48, MAE : 11.27, IOA : 0.9107) and LSTM (RMSE : 16.89, MAE : 11.78, IOA : 0.8971) which do not keep the spatial information through unfold process of input data. This result indicates that it is reasonable to keep the spatial information in order to estimate the PM concentration of many measuring stations from many measuring stations.

Also, the Hybrid model shows better performance than ConvLSTM model having same kernel size on ConvLSTM layers. This shows that the combination of 3D CNN layers extracting temporal features in parallel can help to improve the performance.

IV. CONCLUSION

In this paper, we proposed a Hybrid model, which combines the ConvLSTM model processing spatiotemporal information at the same time sequentially and the CNN model extracting temporal features in parallel. To evaluate the performance of the proposed model, we compared this model with other models including LSTM model or CNN-LSTM model which maintains temporal information for long periods of time, but those are not suitable to process spatial data. We used three indicators including RMSE, MAE and IOA indicators which are generally used in weather forecasting research. Some useful findings can be listed as follows:

1) The Hybrid model and ConvLSTM model which can keep and process spatial information exhibited better performance than the LSTM-based models which do not keep spatial information.

2) By choosing the appropriate kernel size and using the proper number of near measuring stations to every time slices input, the model performance could be improved.

3) Comparing the Hybrid model and ConvLSTM model, we demonstrated that combining with CNN can make performance better because CNN extracted temporal features in parallel.

For further work, we will extend the coverage of measuring stations and use more additional feature data to predict longer-term PM concentration. Also, we will improve the internal structure of the ConvLSTM layer to simulate inflow of PM from near areas by the wind effect better.

REFERENCES

[1] United States Environmental Protection Agency. Particulate Matter (PM) Pollution. [Online]. Available: https://www.epa.gov/pm-pollution/particulate-matter-pm-basics

[2] K. H. Kim, E. Kabir, and S. Kabir, "A review on the human health impact of airborne particulate matter," *Environment International*, vol. 74, p. 138, Jan. 2015.

[3] X. Li, L. Peng, X. Yao, S. Cui, Y. Hu, C. You, and T. Chi, "Long short-term memory neural network for air pollutant concentration predictions: Method development and evaluation," *Environmental Pollution*, vol. 231, pp. 997-1004, Sep. 2017.

[4] J. Chen, J. Lu, J. C. Avise, J. A. Damassa, M. J. Kleeman, and A. P. Kaduwela, "Seasonal modeling of PM2.5 in California's San Joaquin valley," *Atmospheric Environment*, vol. 92, pp. 182-190, Aug. 2014.

[5] R. Stern, P. Builtjes, M. Schaap, R. Timmermans, R. Vautard, A. Hodzic, and A. Kerschbaumer, "A model intercomparison study focussing on episodes with elevated PM10 concentrations," *Atmospheric Environment*, vol. 42, no. 19, pp. 4567-4588, Jun. 2008.

[6] P. Goyal, A. T. Chan, and N. Jaiswal, "Statistical models for the prediction of respirable suspended particulate matter in urban cities," *Atmospheric Environment*, vol. 40, no. 11, pp. 2068-2077, Apr. 2006.

[7] J. Fan, Q. Li, J. Hou, X. Feng, H. Karimian, and S. Lin, "A spatiotemporal prediction framework for air pollution based on deep RNN," *Remote Sensing and Spatial Information Sciences*, vol. IV-4(4W2), pp. 15-22, Aug, 2017.

[8] C. J. Huang and P. H. Kuo, "A deep cnn-lstm model for particulate matter (Pm2.5) forecasting in smart cities," *Sensors(Basel)*, vol. 18, no. 7, Jun. 2018.

[9] X. Shi, Z. Chen, H. Wang, D. Y. Yeung, W. Wong, and W. Woo, "Convolutional LSTM network: A machine learning approach for precipitation nowcasting," in *Proc. 2015 NIPS Conf.*, 2015, pp. 802-810.

[10] F. Charles and F. Karney, "Transverse Mercator with an accuracy of a few nanometers," *Journal of Geodesy*, vol. 85, no. 8, pp. 475-485, Aug. 2011.

[11] P. So-Woo, K. Joo-wook, and S. Doo-sam, "A proposal of an interpolation method of missing wind velocity data in writing a typical weather data," *Journal of the Korean Solar Energy Society*, vol. 37, no. 6, pp. 79-91, Dep 2017.

[12] N. Srivastava, E. Mansimov, and R. Salakhutdinov, "Unsupervised learning of video representations using LSTMs," in *Proc. 2015 ICML Conf.*, 2015, pp. 843-852.

Seonggu Lee was born in Daejeon, South Korea, in 1994. In 2017, he received his degree from KyungHee University in the department of information display. From the year 2017 until today, he is currently working toward the M.S. degree at Sungkyunkwan University. His current research interests include an embedded system and deep learning applications.

Jitae Shin received his B.S. degree from Seoul National University in 1986 and his M.S. degree from the Korea Advanced Institute of Science and Technology (KAIST) in 1988. After working at Korea Electric Power Corp. (KEPCO) and the Korea Atomic Energy Research Institute (KAERI), he returned to study and received M.S. and Ph.D. degrees in electrical engineering from the University of Southern California, Los Angeles, in 1998 and 2001, respectively. Currently, he is a professor in the School of Electronic and Electrical Engineering of Sungkyunkwan University, Suwon, Republic of Korea. His current research interests include image/video signal processing, video transmission over wireless/mobile communication systems, and deep learning applications.