Forecasting Air Quality Using Massive-Scale WSN Based on Convolutional LSTM Network

Karisma Trinanda Putra
Department of Electrical Engineering
Universitas Muhammadiyah Yogyakarta
Yogyakarta, Indonesia
karisma@ft.umy.ac.id

Prayitno
Department of Electrical Engineering
Politeknik Negeri Semarang, Indonesia
Semarang, Indonesia
prayitno@polines.ac.id

Eko Fajar Cahyadi
Department of Telecommunication Engineering
Institut Teknologi Telkom Purwokerto
Purwokerto, Indonesia
ekofajarcahyadi@ittelkom-pwt.ac.id

Ardia Suttyawati Mamonto
Department of Electrical Engineering
Universitas Muhammadiyah Yogyakarta
Yogyakarta, Indonesia
ardia.s.ft17@mail.umy.ac.id

Sunneng Sandino Berutu
Department of Informatics
Universitas Kristen Immanuel Yogyakarta
Yogyakarta, Indonesia
sandinoberutu@ukrimuniversity.ac.id

Syahri Muharom
Department of Electrical Engineering
Institut Teknologi Adhi Tama Surabaya
Surabaya, Indonesia
syahrimuharom@itats.ac.id

***Abstract*— PM2.5 is ultra-light micro-grained particles and dangerous air pollution that threatens public health. A real-time wireless sensor network (WSN) that measure air pollution is a solution to increase public awareness about the long-term impact of PM2.5 exposure. However, in a massive-scale WSN-based air pollution monitoring system, there are numerous noisy and low-concentration periods in the raw PM2.5 dataset, which may lead to unreliable causality predictions. This paper addresses the problem of optimizing sensor acquisition of a wireless sensor network to reconstruct and predict spatiotemporal PM concentrations data using ConvLSTM network. This prediction model is built by combining convolution network and long short-term memory network. The dataset is gathered from air quality WSNs that are already installed across Taiwan. Using the last 48-hour records, the next hour PM2.5 concentration is predicted. RMSE is used to evaluate the prediction accuracy. The results reveal that the ConvLSTM network achieves better performance than those using the LSTM network and regression analysis with RMSE of 1.31, 2.59, and 16.34, respectively.**

Keywords*— *PM2.5, ConvLSTM network, convolutional network, long-short term memory network.

I. Introduction

In recent decades, with the rapid growth of industrialization and traffic, air pollution has become a major problem in modern cities, especially in Asia. Air pollution has a major impact on the health of urban people. A report by WHO estimates that microparticles less than 2.5 m in size (*i.e.,* PM2.5) are responsible for 7 million premature deaths per year and over 92% of the world's population is estimated to breathe toxic air quality [1]. Meanwhile, a study that tracked 1.2 million populations from 1982 until 2008 reveals that every 10 micrograms per cubic meter increase in PM2.5 concentrations was related to an increase up to 27% in lung cancer mortality [2]. Therefore, PM2.5 is one of the most dangerous air pollutants that degrades air quality in big cities with long-term effects that threaten public health.

To combat air pollution, several countries have developed air quality monitoring sites to monitor air pollution propagation using wireless sensor network technologies. Today, the WSN technology is growing with thousands of cheap nodes scattered to monitor many atmospheric variables simultaneously [3]. After the invention of the IoT concept [4], a large number of air quality monitoring stations are installed in various locations with the number doubling every year. This increase was due to the implementation of WSN nodes by non-governmental organizations, the private sector, and the general public. Data from this monitoring location can be accessed in general as a reminder to the public about the condition of their living environment. It is recorded that thousands of air quality monitoring sensors called AirBox are installed in major cities in Taiwan. As shown in Fig. 1, these environmental data can be accessed widely by the public through websites or smartphones. As a result, people can be more aware of the surrounding conditions because they can also take part in installing or monitoring environmental pollutions. For example, incense burning activities contribute to Taiwan's air pollution problem, especially during the Lunar New Year period. In addition, two places of worship in major Taiwanese cities, *i.e.,* Longshan Temple and Xingtiangong Temple, have removed the activity and replaced it by installing AirBox monitoring systems [5].

However, in large-scale environmental prediction systems, data sparsity is a major challenge that impacts the performance of the prediction system [6]. In air pollution datasets, the data sparsity is numerous noisy and low-concentration periods in measurements [7]. For big dataset in the spatiotemporal space, the computational complexity of constructing a prediction model is very high. Recently, more and more researchers are using machine learning technology to overcome the challenges of data sparsity [8]. Several advances in machine learning technology (*e.g.,* deep learning [9], [10]) were introduced to train predictive models more accurately. These models extract connections between different data and can be used to process environmental data to make predictions more accurate by considering both local and global temporal patterns. For example, the authors [11] used RNN to forecast weather conditions for the next few days. Meanwhile, the previous research investigated a multilayer LSTM-based approach to predict air quality [12]. However, in the existing studies, prediction models were trained without taking into account both spatial and temporal relationships between WSN's nodes, especially among adjacent nodes.

2021 1st International Conference on Electronic and Electrical Engineering and Intelligent System (ICE3IS) | 978-1-6654-0546-1/21/$31.00 ©2021 IEEE | DOI: 10.1109/ICE3IS54102.2021.9649763

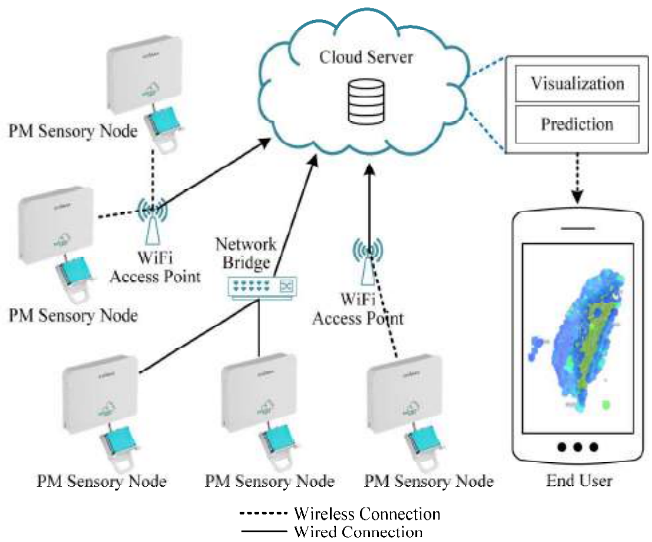

Fig. 1. The Airbox utilizes both wired and wireless connections provides environmental data acquisition based on WSN technology to monitor PM2.5 concentrations in real-time. Thousands of sensory nodes are connected to the server via the Internet and the data can be publicly accessed by end-users.

A training and inference model named convolutional long-short term memory (ConvLSTM) is proposed to predict environmental data that are recorded from a WSN. Different from global temporal prediction models, this study inherits the basic idea of spatiotemporal datasets [13]. But, compared with previous works, the proposed ConvLSTM considers data correlation using spatial sampling so as to reduce the number of model parameters while maintaining the efficient number of active nodes. The proposed scheme overcomes the problem of data sparsity which may lead to unreliable causality predictions. This scheme also considers computational complexity of constructing a spatiotemporal prediction model. This study first presents a feature extraction method that divides the area into a set of regions. Every region represents spatial position of several sensory nodes. Then, Gaussian sampling is used to extract the features inside spatiotemporal dataset. Finally, the ConvLSTM network is used to predict and visualize the dataset into the propagation map.

The remainder of this paper is organized as follows. Section II reviews the previous works related prediction system. Section III reviews the basic approach while Section IV elaborates in detail the proposed approach. Section V examines the results. Furthermore, a discussion in provided in Section VI. Finally, Section VII summarizes the conclusion.

II. Previous Works

Forecasting is the science of predicting future events. This can be done by recording historical data and projecting it into the future. Scientists forecast future events by forming mathematical models, for example predicting air pollution by analyzing the recorded data from the past few days. By forecasting environmental conditions, people can see the real-time condition and predictions so that they are aware of changing their habits that have the potential to increase air pollution. Various studies regarding forecasting technology have been carried out to increase the model accuracy. There are several related studies that have been conducted by researchers including prediction of stock movements [14], [15], prediction of tourist visits [16], and handwritten Chinese character recognition [17].

Research conducted by [14] predicted the direction of the daily stock market index. The main topic of this research is the ability to predict the price direction on the next day from the Japanese stock market index. The predictions were made using an artificial neural network (ANN) model that was optimized using genetic algorithms (GA). Meanwhile, the long-short term memory (LSTM) was used to predict stock movements on the next day [15]. Basic trading stock data was used in making prediction models. In this study, researchers compared the LSTM model with three other models, namely moving average (MA), exponential moving average (EMA), and support vector machine (SVM). The performance of the four models was evaluated using the Root Mean Square Error (RMSE). In addition, the study predicted tourist visits [16] using a neural network model. This study examined the prediction of tourist visits with the Recurrent Neural Network Long Short Term Memory (RNNLSTM) approach using multi-time steps. Predictive model performance was evaluated using RMSE.

Generally, ANN derivative method (*i.e.,* LSTM) is more powerful to use as time-series forecasting systems because they provide better performance. However, this model does not give better performance in multivariate forecasting cases with many input dimensions [18]. LSTM only records temporal patterns in a dataset without extracting spatial patterns that may be formed from spatial relationships between variables. Therefore, this study initiated the convolution layer and long-short term memory layer to generate better forecasting systems.

III. Metodhology

As shown in Fig. 2, the ConvLSTM utilizes 2 kinds of networks, *i.e.,* convolutional (Conv) layer and long-short term memory (LSTM) layer. First a review of convolutional neural network is presented. In the end, the working principle of long-short term memory is described. ConvLSTM combine these networks to enable processing of high-dimensional input data.

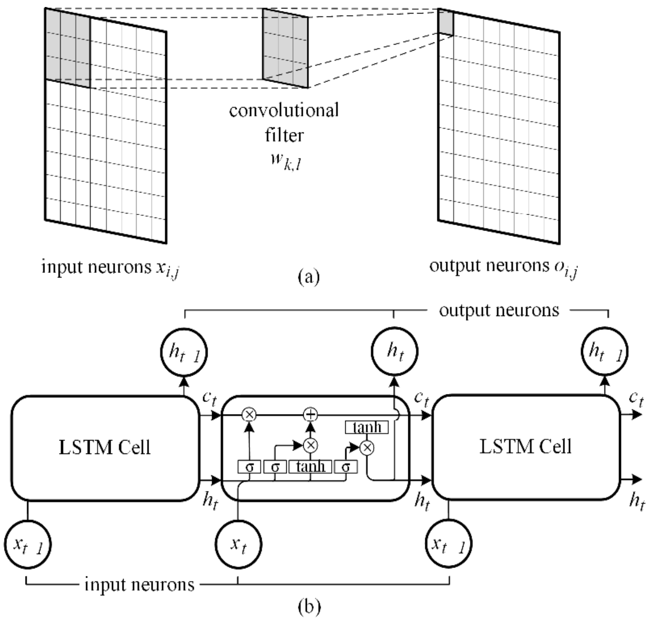

Fig. 2. Components that constructs the convLSTM model *i.e.,* (a) convolutional network and (b) long-short term memory network.

A. Convolutional Neural Network

Convolutional Neural Network (CNN) is one of the layers in Deep Learning that can take input as an image representation, determine the weights and biases of the patterns learned from a dataset. CNN allows for much lower pre-processing compared to other classification algorithms. CNN has the ability to study the spatial characteristics of a 2-dimensional or more input, so that pre-processing is replaced directly by CNN as a feature extraction module. In Fig. 2 (a), it can be seen that the CNN utilizes the convolution process by moving a convolution kernel/filter (w) of a certain size to an input image (x). Then, the computer gets new representative information (o) from the result of multiplying that part of the image with the filter used, shown in Equation (1).

$$o[t] = (x \times w)[t] = \sum_{a=-\infty}^{a=\infty} x[a]\, w[a+t] \qquad (1)$$

B. Long-short Term Memory Network

The LSTM cell at high levels is very similar to that of RNN cell. The LSTM consists of three parts, as shown in the figure below and each part performs an individual function. These three parts are known as the gate. Successively, they are the Forgot gate, the Input gate, and the last one is the Output gate. LSTM has a hidden state where (h_{t-1}) represents the hidden state of the previous timestamp and (h_t) is the hidden state of the current timestamp. In addition, LSTM also has cell states represented by (c_{t-1}) and (c_t) for the previous and current timestamps, respectively.

As presented in Fig. 2 (b), the working principle of the LSTM is, first of all, the information is filtered by Forget gate (f_t) whether to keep the previous timestamp or forget it. Then, the sigmoid function is applied. Meanwhile, the input gate (i_t) is used to measure the importance of the new information carried by the input. Furthermore, the sigmoid function is also applied. Finally, the new information ($\tilde{c}_t$) that needs to be passed to the cell state is a function of the hidden state on the previous timestamp and the current input (x_t). The activation function here is tanh and the biases for the respective gates (x) is b. Equation (2–4) are mathematical equations of the working principle of LSTM components.

$$f_t = \zeta\left(W_f \cdot (h_{t-1}, x_t) + b_f\right) \qquad (2)$$

$$i_t = \sigma\left(W_i \cdot (h_{t-1}, x_t) + b_i\right) \qquad (3)$$

$$\tilde{c}_t = tanh\left(W_c \cdot (h_{t-1}, x_t) + b_c\right) \qquad (4)$$

IV. PREDICTION MODEL

First, a description of how to build a global model that covers a huge number of environmental sensory nodes of WSN is presented. Finally, the working principle of PM2.5 prediction model that utilizes ConvLSTM is described.

A. Global Temporal Model

A global temporal model has the advantage of efficient use of the number of parameters that construct the prediction model. In global model, only the temporal pattern is recorded by the prediction model. Spatial patterns are not extracted from a dataset. A collection of datasets, which were originally spatiotemporal, were arranged according to the sequence of events. Spatial patterns will be lost from compiling this dataset, but the number of parameters required to represent global inputs will be minimized. This model can be implemented using a one-dimensional LSTM network. This LSTM network sequence can be arranged in several layers as needed to produce appropriate convergence in the training process. As shown in Fig. 3, the datasets are collected and arranged according to the time series of events. Therefore, this model describes global pattern inside the dataset without considering the spatial relationship between nodes.

B. The Proposed Spatiotemporal Model

Spatiotemporal data analysis is a research that started because of the development and application of new computational techniques that allow analysis of databases in large numbers and dimensions. The spatiotemporal model appears when the data collected is bound to the space and time domain [7]. An event in the spatiotemporal data describes the spatial and temporal phenomena that exist at time t and location x. An example is the pattern of air pollution concentrations in an area within one year. By analyzing the pattern, there are times when the pattern repeats itself over short periods or years. Meanwhile, the unique geographical position of each region also increases the differences in observational variables, causing spatiotemporal analysis to be more objective and consistent than global temporal analysis.

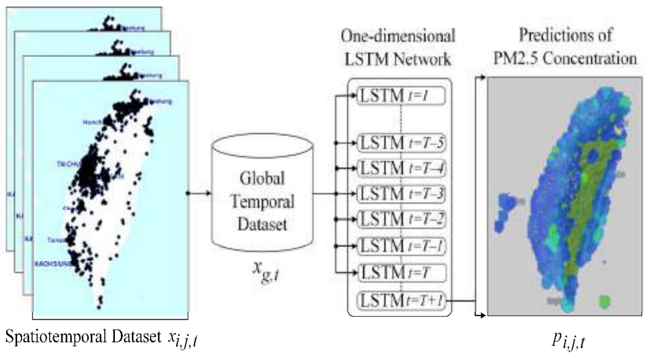

Fig. 3. The global temporal model minimized the number of parameters to build the prediction model at cost of a loss in spatial correlation in the dataset.

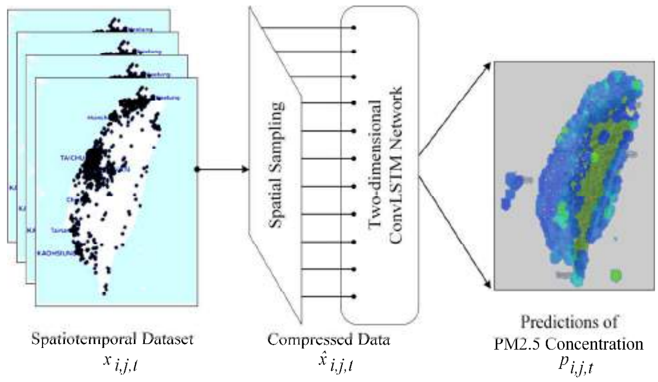

Fig. 4. The proposed scheme supports the training of spatiotemporal dataset using the Gaussian sampling as a filter and ConvLSTM as a feature extraction technique to construct the PM2.5 predictions system.

With the increasing size of the data and dimensions that are processed, the computational process will also increase. The main challenge of the spatiotemporal data is the amount of data that is computationally expensive to the prediction model. Although computing capabilities in recent decades have increased rapidly, the efficiency of data use which leads to the use of computing power also needs to be considered. The proposed scheme provides a workflow guide that can be used to process spatiotemporal big data with compressive sensing techniques and training techniques with neural networks. Fig. 4 shows the stages of analyzing the spatiotemporal dataset obtained from the Airbox sensor network. Data in the space and time domains are filtered to provide more uniform results, eliminate noise caused by data sparse, and shorten training parameters.

Furthermore, the proposed model consists of 2 components, *i.e.,* spatial sampling and ConvLSTM. The ConvLSTM replaces the need of matrix approximation to restore the compressed data. The ConvLSTM can be used simultaneously to reconstruct the data and to predict the patterns inside the spatiotemporal dataset.

1) Spatial Sampling

A matrix approximation is a minimization problem in mathematics, in which the cost function calculates the fit between a given matrix (the data) and an optimized matrix. One of the matrix approximation method is low-rank matrix approximation (LRMA). The LRMA is commonly used in a massive WSN to efficiently record data without wasting resources using a small number of samples. In the LRMA, several nodes are randomly selected to minimize the number of nodes that are involved during global data recording. Therefore, a spatial sampling procedure is used. By using the spatial sampling procedure, the global patterns can be recorded effectively without losing their fidelity. First, at the data acquisition stage, a spatial sampling procedure is performed to select several sensing nodes randomly from period $t = 1$ to T. This study uses Gaussian sampling with $d = 0.3$. It means that only 30% of the total nodes are simultaneously used for data acquisition in one sampling period. To represent uniform spatiotemporal space, the dataset is arranged into 256x256 sections. Each section has several nodes according to its geographical location. Not all sections will be recorded, but the data recording is done randomly with a capacity of 30% of 65,536 sections

According to the LRMA theory [19], an optimized sampling scheme that achieves a good representation of global recording process is performed using uniformly random sampling. Therefore, a Gaussian sampling is used during the data acquisition. Meanwhile, at the data recovering stage, a singular value decomposition (SVD) is used to reconstruct the data from the compressed dataset. However, in the current method (*i.e.,* LRMA via SVD), the procedure does not explore temporal correlation inside the spatiotemporal dataset [20]. The SVD only takes into account the spatial correlation in an incomplete matrix. It is possible that spatiotemporal features can provide deeper knowledges for training the prediction model. This could be the contributing factor to the current setback, such as low data recovery accuracy, especially in the spatiotemporal data processing. Therefore, to improve the accuracy of the current data acquisition procedure, this study utilizes the ConvLSTM network as shown in Fig. 5. This scheme minimizes the communication and calculation process during data generation while simultaneously providing data reconstruction and data prediction at $T+1$. Finally, this paper provides two prediction models which are the conventional scheme (*i.e.,* SVD+LSTM) and the proposed scheme (*i.e.,* ConvLSTM network).

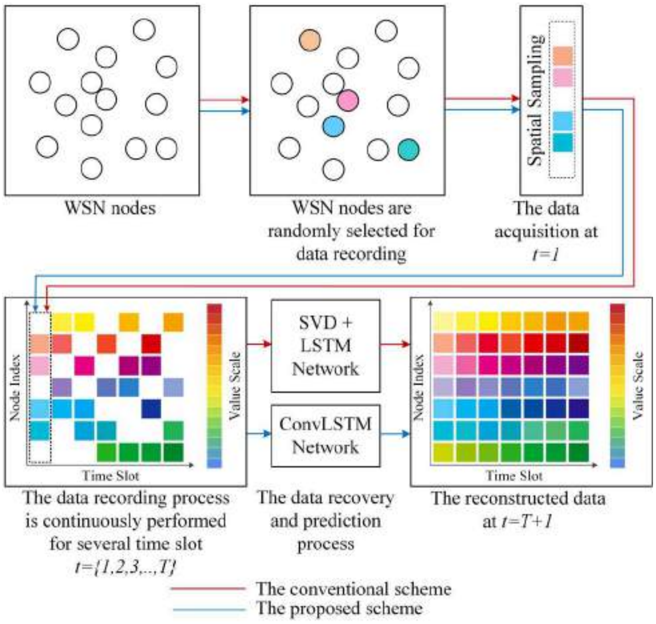

Fig. 5. The experimental models compare the conventional scheme (i.e., SDV+LSTM) with the proposed ConvLSTM model.

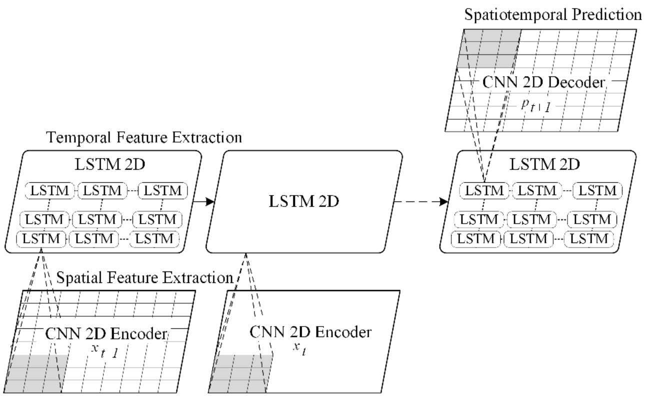

Fig. 6. A series of ConvLSTMs is used to extract spatiotemporal features from 2D input sequentially. The CNN encoder learns spatial features while the LSTM records temporal features. In the end, the CNN decoder represent the spatiotemporal feature into the prediction of PM2.5 concentrations.

2) ConvLSTM Network

A convolutional neural network (CNN) works with two dimensional input data. However, they can be used with multi-dimensional input data. The working principle of CNN is to convolute a number of regions alternately against a filter or kernel. This convolution operation makes this network known as CNN. Convolution is a linear operation that involves multiplying a set of weights (*i.e.,* a 2D filter) by an input. The use of a filter that is smaller than the input is intentional because it allows the same filter (set of weights) to be multiplied by the input array multiple times at different points on the input. In particular, the filters are applied systematically, moving to each overlapping part, alternating from left to right, from top to bottom.

Meanwhile, convolutional long-short term memory imitates the basic principle of CNN where input is convoluted first and then a 2D LSTM network is performed to extract temporal pattern from series of 2D input. In the end, the spatial data collected over successive time periods are represented as a spatiotemporal prediction map using the CNN decoder. In such cases, the prediction model that considers the spatial and temporal correlation produces the best performance. The architecture inside the ConvLSTM is shown in Fig. 6. The main idea of this architecture is from CNN-LSTM Autoencoder. The autoencoder is a neural network model that has the same input and output, except the feature extraction is performed both spatially and temporally so that it can be used to process spatiotemporal datasets, *e.g.,* PM2.5 concentrations over time.

V. RESULT

A dataset downloaded from the Airbox system was used in this study, covering 76.45 MB of data [21]. The data was recorded from 3574 sensory nodes spread across Taiwan within January 2021. The experiment was carried out only by extracting PM2.5 data from the dataset. This PM2.5 dataset is spatiotemporal data with a timestamp index and position coordinates. Then, the data is localized and represented in a sequence of 2-dimensional images in sequence with a size of 256x256. After that, these images are processed using the ConvLSTM model. The system performance assessment is carried out by observing the losses during training phase using mean absolute error (MAE) parameters. After the learning process reaches convergence, then the RMSE is calculated from the output validation generated by the inference system. RMSE calculates the similarity of the predicted signal with the ground truth signal. In the end, a qualitative assessment is shown by visually comparing the shape of the predicted results between the proposed scheme and other conventional methods.

The assessment matrix used to measure the performance of the training process is the MAE, formulated in Equation (5) below.

$$MAE = \frac{1}{T}\sum_{t=1}^{T}\left|o_{target(i)} - o_{predicted(i)}\right| \tag{5}$$

Meanwhile, quantitatively, the prediction results were evaluated using the RMSE, described in Equation (6).

$$RMSE = \sqrt{\sum_{t=1}^{T}\frac{\left(o_{target(i)} - o_{predicted(i)}\right)^2}{T}} \tag{6}$$

where $o_{target(i)}$ is the ground truth data, $o_{predicted(i)}$ is the prediction of PM2.5, and i is index timestamp of the data.

Based on the experiments, the MAE training and validation data were obtained as shown in Fig. 7. It can be seen that the training process can run optimally with a validation value that follows the training MAE value. These values continue to decrease until the epoch (ε) reaches 100 iterations. The MAE at ε = 100 epochs of the SVD+LSTM method is higher than that of using the ConvLSTM. It means that the ConvLSTM provides better training performance with a level of similarity that is closer to the ground truth data. To provide visual evidence, then the model is tested on an inference engine which only calculates the forward propagation process to predict the forecasting results. Visually, the forecasting results of the proposed method can be seen in Fig. 8. The data generated by the proposed method is better and closer to the ground truth data than those using the SVD+LSTM. It can be noted that this method can successfully reconstruct and predict the signal only using 30% concurrent nodes of the total nodes in one period.

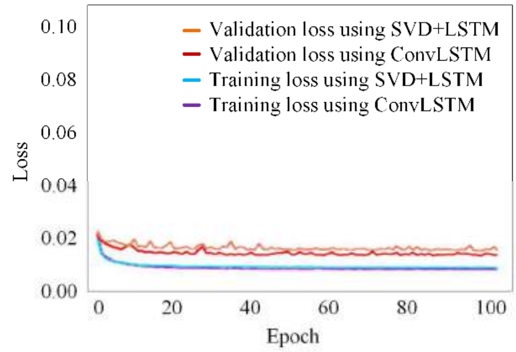

Fig. 7. Diagnostic line plot showing training progress of the proposed model.

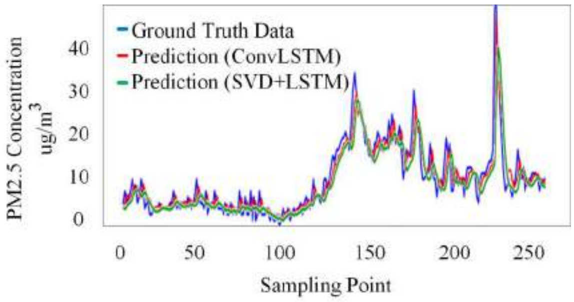

Fig. 8. Comparison of prediction results i.e., achieved using the spatiotemporal ConvLSTM model (red), SVD+LSTM model (green), with the ground truth data (blue).

TABLE I. SPECIFICATION OF THE EXPERIMENTAL MODELS

Prediction Model	Spatial Sampling	Feature Extraction	Number of Parameters	RMSE at ε=100
Regression analysis input: 48	Full	Temporal	10	16.34
Global LSTM input: 48	Full	Temporal using 2 hidden layers (48 LSTM cells each layer)	28,273	2.59
SVD+LSTM input: 48	Random using Gaussian Sampling *d=0.3*	Temporal using 2 hidden layers (48 LSTM cells each layer)	28,273	1.74
The proposed ConvLSTM network input: 256x256	Random using Gaussian Sampling *d=0.3*	Spatial and temporal using 2 hidden layers of ConvLSTM 2D kernel size = 3x3 filter = 32	18,177	1.31

This paper also provides architectural comparison of several prediction models, namely the classical regression model [22], the global LSTM model [14], the SVD+LSTM model [17], and the proposed ConvLSTM model. The test specifications are shown in Table 1. The first three models use temporal feature extraction while the last one uses spatiotemporal feature extraction. The SVD+LSTM model uses an LSTM network with 48 LSTM cells and one output neuron that represents the predicted results at T+1. Meanwhile, the ConvLSTM model uses two layers of 2D CNN networks on the encoder part. This part has a configuration of input, kernel size, and filter as much as 256x256, 3x3, 32, respectively. Furthermore, the 2D LSTM module is used in a row with 488 strands. The last strand that serves as the prediction output is processed using a CNN 2D decoder. This decoder module reverses the pattern representation inside the network so that it is converted to a PM2.5 propagation map that has the same size as the input. From the experiments, the number of parameters generated by the regression model, the global LSTM model, the SVD+LSTM model, and the proposed ConvLSTM model is 10, 28,273 28,273, and 18,177 parameters, respectively. By using a smaller number of parameters compared to the temporal-based prediction system, the proposed model produces the best accuracy with the smallest RMSE of 1.31. The larger the number of parameters, the longer the training time. However, at the implementation stage, only the inference engine is used so that the prediction system can be executed in a much shorter time.

VI. Conclusion

The discovery and prediction of air pollution patterns are very important to reduce the health risks from the long impacts of PM2.5 exposure. In the literature on finding and predicting air quality patterns through neural networks, previous research has only focused on finding temporal patterns without considering spatial patterns that might be used as clues to improve model performance. This paper complements the shortcomings of previous studies, which establish the complete structure of PM2.5 concentration discovery and prediction through the proposed 2-stage architecture, *i.e.,* spatial sampling and ConvLSTM. Through the two-stage architecture, this proposed scheme can analyze and predict PM2.5 concentrations with periodic patterns spatially and temporally. The results reveal that the proposed ConvLSTM is superior in generating a better performance of prediction system than those using global LSTM model and regression analysis with RMSE of 1.31, 2.59, and 16.34, respectively.

References

[1] "WHO releases country estimates on air pollution exposure and health impact." https://www.who.int/news/item/27-09-2016-who-releases-country-estimates-on-air-pollution-exposure-and-health-impact (accessed Jul. 26, 2021).

[2] M. C. Turner, D. Krewski, C. A. Pope, Y. Chen, S. M. Gapstur, and M. J. Thun, "Long-term ambient fine particulate matter air pollution and lung cancer in a large cohort of never-smokers," *Am. J. Respir. Crit. Care Med.*, vol. 184, no. 12, pp. 1374–1381, 2011.

[3] W. Widyasmoro, "Two-Tiered Network Design Architecture for Collision Avoidance Protocol in Wireless Sensor Networks," *J. Electr. Technol. UMY*, vol. 2, no. 4, pp. 153–160, 2018.

[4] D. Susanto, F. Mujaahid, R. Syahputra, and K. T. Putra, "Open Source System for Smart Home Devices Based on Smartphone Virtual Assistant," *J. Electr. Technol. UMY*, vol. 3, no. 1, 2019.

[5] "BBC science show Click presents Taiwan invention the AirBox | Taiwan News | 2019-11-07 17:31:00." https://www.taiwannews.com.tw/en/news/3812388 (accessed Jul. 26, 2021).

[6] K. T. Putra *et al.*, "Federated Compressed Learning Edge Computing Framework with Ensuring Data Privacy for PM2.5 Prediction in Smart City Sensing Applications," *Sensors 2021, Vol. 21, Page 4586*, vol. 21, no. 13, p. 4586, Jul. 2021.

[7] J. Y. Zhu *et al.*, "pg-Causality: Identifying Spatiotemporal Causal Pathways for Air Pollutants with Urban Big Data," *IEEE Trans. Big Data*, vol. 4, no. 4, pp. 571–585, Jul. 2017.

[8] F. D. Syahfitra, R. Syahputra, and K. T. Putra, "Implementation of Backpropagation Artificial Neural Network as a Forecasting System of Power Transformer Peak Load at Bumiayu Substation," *J. Electr. Technol. UMY*, vol. 1, no. 3, pp. 118–125, Dec. 2017.

[9] L. M. Azizah, S. F. Umayah, and F. Fajar, "Deteksi Kecacatan Permukaan Buah Manggis Menggunakan Metode Deep Learning dengan Konvolusi Multilayer," *Semesta Tek.*, vol. 21, no. 2, pp. 230–236, Dec. 2018.

[10] D. O. Oyewola, A. F. Augustine, E. G. Dada, and A. Ibrahim, "Predicting Impact of COVID-19 on Crude Oil Price Image with Directed Acyclic Graph Deep Convolution Neural Network," *J. Robot. Control*, vol. 2, no. 2, pp. 103–109, Mar. 2021.

[11] C. Yu, H. Ahn, and J. Seok, "Coordinate-RNN for error correction on numerical weather prediction," *Int. Conf. Electron. Inf. Commun. ICEIC 2018*, vol. 2018–January, pp. 1–3, Apr. 2018.

[12] T. Li, M. Hua, and X. Wu, "A Hybrid CNN-LSTM Model for Forecasting Particulate Matter (PM2.5)," *IEEE Access*, vol. 8, pp. 26933–26940, 2020.

[13] X. Piao, Y. Hu, Y. Sun, B. Yin, and J. Gao, "Correlated spatio-temporal data collection in wireless sensor networks based on low rank matrix approximation and optimized node sampling," *Sensors (Switzerland)*, vol. 14, no. 12, pp. 23137–23158, 2014.

[14] T. Gao, Y. Chai, and Y. Liu, "Applying long short term momory neural networks for predicting stock closing price," *Proc. IEEE Int. Conf. Softw. Eng. Serv. Sci. ICSESS*, vol. 2017–November, pp. 575–578, Apr. 2018.

[15] Q. Mingyue, L. Cheng, and S. Yu, "Application of the Artifical Neural Network in Predicting the Direction of Stock Market Index," *Proc. - 2016 10th Int. Conf. Complex, Intelligent, Softw. Intensive Syst. CISIS 2016*, pp. 219–223, Dec. 2016.

[16] A. A. Rizal and S. Soraya, "Multi Time Steps Prediction dengan Recurrent Neural Network Long Short Term Memory," *MATRIK J. Manajemen, Tek. Inform. dan Rekayasa Komput.*, vol. 18, no. 1, pp. 115–124, Nov. 2018.

[17] Y. Yang *et al.*, "Accelerating and compressing LSTM based model for online handwritten Chinese character recognition," *Proc. Int. Conf. Front. Handwrit. Recognition, ICFHR*, vol. 2018–August, pp. 110–115, Dec. 2018.

[18] S. Yang, X. Yu, and Y. Zhou, "LSTM and GRU Neural Network Performance Comparison Study: Taking Yelp Review Dataset as an Example," *Proc. - 2020 Int. Work. Electron. Commun. Artif. Intell. IWECAI 2020*, pp. 98–101, Jun. 2020.

[19] E. J. Candès and B. Recht, "Exact matrix completion via convex optimization," *Found. Comput. Math.*, vol. 9, no. 6, pp. 717–772, 2009.

[20] H. Chen *et al.*, "A spatiotemporal estimation method for hourly rainfall based on F-SVD in the recommender system," *Environ. Model. Softw.*, vol. 144, p. 105148, Oct. 2021.

[21] "Civil IoT Taiwan-Data Service Platform - Civil IoT." https://ci.taiwan.gov.tw/dsp/en/environmental_en.aspx (accessed Sep. 01, 2021).

[22] Y. Xu and S. Lan, "Additive calibration model for the monitoring data of PM2.5 and PM10 based on ARIMA and multiple linear regression," *Proc. - 2019 Int. Conf. Econ. Manag. Model Eng. ICEMME 2019*, pp. 302–306, Dec. 2019.