Solar PV Generation Prediction Based on Multisource Data Using ROI and Surrounding Area

Bowoo Kim, *Student Member, IEEE*, and Dongjun Suh, *Member, IEEE*

***Abstract*— This study introduces a lightweight hybrid solar photovoltaic (PV) generation prediction model operating on 1-h intervals, utilizing remote sensing data to enhance power grid management. Multisource remote sensing data, including spatial features from infrared satellite images and temporal data from various hourly recorded datasets, capture spatiotemporal characteristics. The model defines and synthesizes regions of interest (ROI) and surrounding areas of ROI (ROI_{surr}) within satellite images to reduce computational load. Integration of image and numerical weather prediction (NWP) process modules ensures accurate prediction. Comparative analysis against five machine learning algorithms shows significant improvements, with up to a 33.7% decrease in mean absolute error (MAE) and a 19.51% decrease in root mean square error (RMSE). Additionally, the model consistently meets ASHRAE Guideline 14 standards and outperforms single-source data models. Experimentation highlights the effectiveness of smaller ROIs in enhancing predictive accuracy, demonstrating adaptability to climate variations. This lightweight multisource remote sensing-based hybrid model promises to guide smart grid operations and sustainable power grid systems, advancing remote sensing applications in renewable energy management.**

***Index Terms*— Convolutional neural network (CNN)-long short-term memory (LSTM), deep learning, lightweight, multisource data, region of interest (ROI), satellite image, solar photovoltaic (PV) generation forecasting, spatiotemporal.**

I. Introduction

THE proportion of renewable energy integrated into primary power grids has demonstrated a progressive uptrend, with projections indicating accelerated growth by 2030 [1]. Furthermore, renewable energy generation, which includes solar and wind power, is expected to increase to up to 85% of the global power composition by 2050. Consequently, congruent with these forecasts, the effective operation and management of renewable energy-dependent grid systems are imperative for the displacement of fossil fuels [2]. A fundamental obstacle lies in discerning the optimal equilibrium between power provision and demand leveraging renewable sources.

Manuscript received 21 March 2024; accepted 18 April 2024. Date of publication 22 April 2024; date of current version 23 May 2024. This work was supported by the National Research Foundation of Korea (NRF) grant funded by the Korea government (No. NRF-2021R1A5A8033165) and the Korea Institute of Energy Technology Evaluation and Planning (KETEP) and the Ministry of Trade, Industry & Energy (MOTIE) of the Republic of Korea (No. 20224000000150). *(Corresponding author: Dongjun Suh.)*

The authors are with the Department of Convergence and Fusion System Engineering, Kyungpook National University, Sangju 37224, South Korea (e-mail: kbw5913@knu.ac.kr; dongjunsuh@knu.ac.kr).

Digital Object Identifier 10.1109/TGRS.2024.3392337

From an environmental and economic perspective, solar photovoltaic (PV) generation is a superior renewable energy source. Therefore, large-scale solar PV power plant farms are being constructed globally [3], [4], with a projected increase of approximately 10% in 2021 compared to that in 2020 and further growth expected in the future [5]. Compared to other power generation facilities, solar PV generation can better preserve the surrounding environment and incurs lower installation and operation costs. Moreover, this energy resource is inexhaustible, as it harnesses direct current electricity through the photoelectric effect triggered by the absorption of sunlight in solar panels. Nonetheless, the foremost influencer of solar PV generation, irradiance, remains susceptible to uncontrollable external meteorological variables, such as cloud cover [6], [7] and particulate matter (PM) [8], [9], thereby presenting significant challenges to ensuring a consistent power supply. Thus, it is essential to develop a precise solar PV generation prediction model to maximize the economic effect of solar PV generation and establish an efficient and stable power supply plan simultaneously.

There are various methods for predicting solar PV generation based on the prediction time and other techniques. These methods can be categorized into three types. The first type is a method based on the predictive technique. Direct prediction predicts solar PV generation using external independent parameters, and indirect prediction first predicts the irradiance through external independent parameters and then predicts the solar PV generation through the predicted irradiance [10]. The second type is based on prediction time. These methods can be further divided into ultrashort-term prediction from seconds to minutes [11], [12], short-term prediction of up to 48–72 h [13], [14], medium-term prediction from days to weeks [15], [16], and long-term prediction from months to years [17]. In employing these methods, selecting an appropriate prediction period is crucial to achieving optimal performance. For instance, ultrashort-term prediction is suitable for applications demanding precision, such as microgrid power system control, while long-term prediction is more apt for large-scale control tasks such as power supply planning. The third type is based on predictive parameters. The most prevalent approach for predicting solar PV generation is numerical weather prediction (NWP) [18]. It relies on continuously measured data of external meteorological factors obtained either directly from meteorological administrations or from areas surrounding solar PV power plants [19]. Furthermore, certain prediction methods utilize conventional statistical techniques such as ARIMA and

SARIMA [20], [21], [22]. With the recent advancements in computational capabilities, particularly with GPUs, a plethora of machine learning algorithms such as deep neural networks and long short-term memory (LSTM) models are increasingly being employed [23], [24], [25], [26], [27]. Since NWP relies on temporally measured data, it effectively captures the temporal dynamics of solar PV generation. However, the movement of clouds driven by wind can significantly impact the amount of irradiance, resulting in pronounced variability in generated solar PV power. Accordingly, a range of remote sensing methodologies has been employed in forecasting solar PV generation, with numerous studies focusing on techniques, such as optical flow or convolutional neural network (CNN)-based algorithms for this purpose. These approaches primarily rely on remote sensing data and utilize indices or motion vectors extracted from sky or satellite images to incorporate the spatial characteristics of clouds into the prediction process [28], [29], [30].

In numerous previous studies employing predictive parameter methods, researchers have predominantly relied on either NWP or satellite data as single-source inputs. While NWP data can accurately forecast solar PV generation by utilizing climate information surrounding the solar PV power plant over time, the prediction range is limited. As previously mentioned, relying solely on numerically recorded wind direction and speed data has its limitations in predicting the impact of cloud movement driven by the wind. In contrast, satellite data can provide a broader range of predictions by incorporating numerous surface characteristics, which can address the challenge of collecting precise data related to the surrounding area due to a shortage of weather observation stations. However, it is often infeasible to utilize accurate and detailed meteorological data regarding temperature and humidity conditions in the vicinity of a solar PV power plant. To overcome this limitation, researchers have sought to enhance prediction accuracy by integrating remote sensing techniques with NWP, thus considering the spatiotemporal characteristics of external parameters relevant to solar PV generation [31], [32], [33], [34], [35]. By combining multiple sources of data instead of solely relying on NWP or remote sensing data individually, researchers can simultaneously integrate spatiotemporal characteristics at various scales.

Therefore, this study integrates remote sensing techniques with NWP data, external meteorological parameters observed at the solar PV power plant location, to utilize multisource data including the spatiotemporal characteristics of PM moving by clouds and wind over time. Further, a region of interest (ROI) and the surrounding area of the region of interest (ROI_{surr}) are set in the satellite image to process the vast quantity of multisource data effectively. Finally, we propose a deep learning-based hybrid spatiotemporal solar PV generation prediction model that reflects the spatiotemporal characteristics through the synthesis of the set ROI and ROI_{surr}.

Fig. 1 illustrates the overall framework of this study. The spatiotemporal hybrid solar PV generation prediction model designs an image process module based on an algorithm that can predict cloud movement [36] and PM movement according to the wind and incorporates the NWP process module that processes the numerical data. The prediction accuracy is increased by intensely reflecting the spatial characteristics of the external weather parameters in the region where the solar PV power plant is located, and the model computational complexity is decreased by setting and synthesizing the ROI and ROI_{surr} of the satellite image to learn the solar PV generation prediction model. The contributions of this study are given as follows.

1) Multisource data, combining NWP data and satellite images, are employed to integrate remote sensing techniques with NWP data, thereby facilitating the incorporation of spatiotemporal characteristics for precise prediction of solar PV generation. The ROI and ROI_{surr} are set based on the area of the solar PV power plant to emphasize the spatial information in the area. The data collection and preprocessing are described in Section III.
2) The design of the image process module, which combines an encoder–decoder structure with a CNN-LSTM-based ensemble algorithm, effectively captures the movement of clouds and PM induced by wind. Two identical image process modules are trained separately on clouds and PM. The ROI and ROI_{surr} effectively capture spatial characteristics in satellite images using the proposed image synthesis layer (ISL) within the image processing network. The ISL enables the network to adapt to different satellite image resolutions and sensor characteristics, making it more robust and versatile in handling diverse satellite imagery. This flexibility allows the network to effectively capture the spatial characteristics of the ROI and its surrounding area, enhancing its ability to accurately delineate and analyze relevant regions in satellite images. The NWP process module learns the numerical data from the area around the solar power plant and reflects the time-series characteristics of the data that are observed at hourly intervals. Finally, the output values from both the image and NWP process modules are merged and utilized as multisource data input for the proposed model. As a result, forecasts for solar PV generation that consider their spatiotemporal peculiarities are achievable. A detailed description of the proposed model is provided in Section IV.
3) Setting the ROI and ROI_{surr} within satellite images aims to enhance prediction accuracy and reduce model complexity. Experiments comparing the model using the original image with the model using the proposed ROI image were performed with various machine learning algorithms to analyze the influences of ROI and ROI_{surr}. An additional experiment was performed to select the optimal ROI size. Finally, the performance of the prediction model was verified through k-fold cross-validation for quantitative analysis. The performance evaluation and comparative experiments are described in Section V.

II. Problem Formulation

The proposed prediction model uses various input factors to predict solar PV generation. The entire timeline is defined as the union of the historical time set t_h and future time set t_f with a time step of 1 h: $t_h \cup t_f = \{t_1, t_2, \ldots, t_h\} \cup$

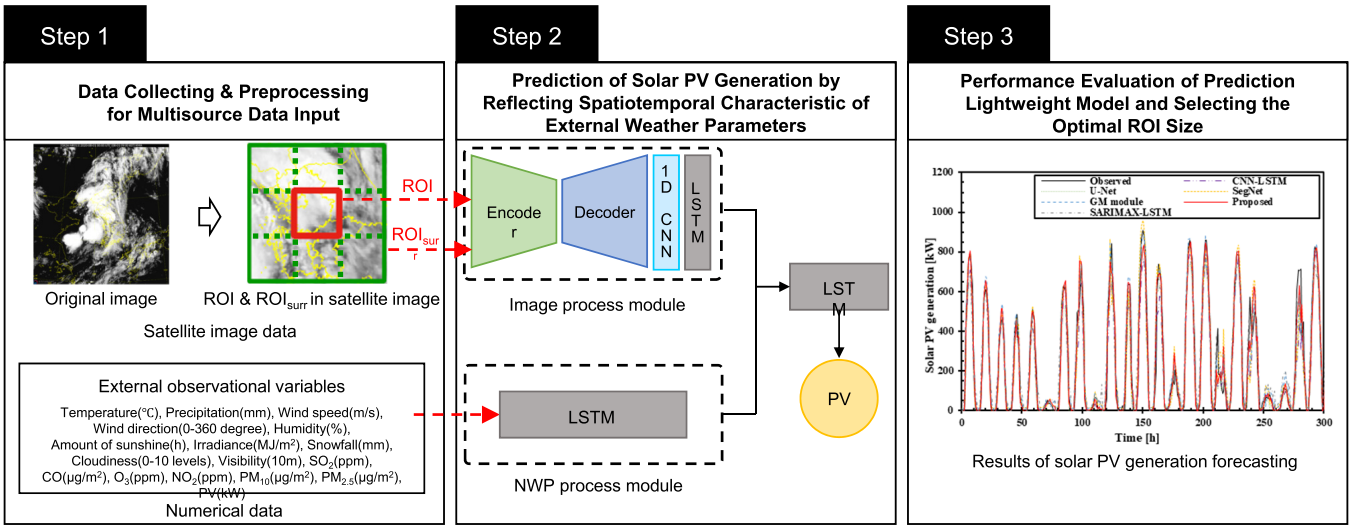

Fig. 1. Research framework.

$\{t_{h+1}, t_{h+2}, \ldots, t_{h+f}\}$. The equation for forecasting 1 h ahead solar PV generation $Y'^{\mathrm{PV}}_{t_f}$ is as follows:

$$Y'^{\mathrm{PV}}_{t_f} = f\left(X^{\mathrm{Image}}_{t_h}, X^{\mathrm{NWP}}_{t_h}\right) \tag{1}$$

$$\begin{cases} X^{\mathrm{Image}}_{t_h} = \left[X^{\mathrm{Cloud}}_{t_h}, X^{\mathrm{PM}}_{t_h}\right] \\ X^{\mathrm{NWP}}_{t_h} = \left[X^{\mathrm{MF}}_{t_h}, X^{\mathrm{PF}}_{t_h}, Y^{\mathrm{PV}}_{t_h}\right] \end{cases} \tag{2}$$

where $f(\cdot)$ is the artificial neural network architecture that performs nonlinear calculations for predicting solar PV generation in the proposed model. $X^{\mathrm{Image}}_{t_h}$ and $X^{\mathrm{NWP}}_{t_h}$ denote the prediction output of the image and the NWP process modules that use multisource data as the input. The image process module processes data with satellite image data consisting of cloud images $X^{\mathrm{Cloud}}_{t_h}$ and PM images $X^{\mathrm{PM}}_{t_h}$ as inputs. At this time, $X^{\mathrm{Cloud}}_{t_h}$ and $X^{\mathrm{PM}}_{t_h}$ can be expressed as the output of the ROI and $\mathrm{ROI_{surr}}$ images in the satellite image, which are combined through the ISL. The NWP process module processes data with the external meteorological factors $X^{\mathrm{MF}}_{t_h}$, air pollutant factors $X^{\mathrm{PF}}_{t_h}$, and historical solar PV generation $Y^{\mathrm{PV}}_{t_h}$ as inputs.

III. Satellite and Numerical Data

Consecutive satellite images and numerical data measured at 1 h intervals were used to train the solar PV generation prediction model to reflect the spatiotemporal information of each climate factor. The experimental period was set from August 1, 2020, to July 31, 2021, and a solar PV power plant located in Incheon, South Korea, was selected as the experimental subject. Given that solar PV generation is inactive at night, the experimental timeframe was limited to 12 h (7:00 A.M. to 7:00 P.M.).

A. Satellite Images

The experimental satellite images to reflect the spatial characteristics of clouds and PM were captured by GK2A and provided by the National Meteorological Satellite Center (NMSC) in Korea [37]. GK2A operates as a geostationary-orbit meteorological satellite positioned approximately 36 000 km above the equator. It can produce a total of 52 meteorological products through 16 channels, and it enables faster monitoring and preparation for meteorological disasters by capturing images at up to 2-min intervals. GK2A captures images of the Korean Peninsula with a spatial resolution of 2 km for infrared channels and 0.5 km for visible channels. For this experiment, two infrared satellite images featuring clouds and PM were utilized.

Cloud movement influenced by wind significantly impacts solar PV generation, as shadow occlusion results in lower temperatures compared to the Earth's surface. These effects can be discerned through infrared satellite imagery. For that reason, this study utilized near-infrared (1.38 μm) cirrus images of the Korean Peninsula, measuring 900 $\times$ 900 pixels, provided by the NMSC [38]. Similarly, the spatial characteristics of PM suspended in the atmosphere were reflected through infrared (12.3 μm) satellite images [39]. Both types of images were captured at 1 h intervals. For the interpolation of missing data, we improved the reliability and accuracy of the interpolation process by incorporating data occurring 30 min after the missing data point, in addition to the standard linear interpolation method, instead of relying solely on linear interpolation using data before and after the missing data point.

Previous studies [31], [32] that have utilized existing satellite data have made use of secondary outputs self-generated from satellite units, such as atmospheric motion vector data, cloud optical thickness data, aerosol optical depth data, and insolation data. To mitigate the time-consuming process of generating secondary outputs from satellite data, this study opted to utilize infrared satellite images as a more efficient alternative. By leveraging infrared satellite images, the data calculation time was reduced, enabling faster processing and forecasting analysis.

B. Numerical Data

The data on external meteorological factors and air pollutants that affect solar PV generation prediction were provided by the KMA [40] and Air Korea [41]. A total of 16 types of external observational data composed of numerical values (10 types of meteorological data including temperature and

TABLE I
EXTERNAL OBSERVATIONAL VARIABLES OF THIS RESEARCH

Data	Variables
Input	Temperature(℃), Precipitation(mm), Wind speed(m/s), Wind direction(0-360 degree), Humidity(%), Amount of sunshine(h), Irradiance(MJ/m^2), Snowfall(mm), Cloudiness(0-10 levels), Visibility(10m), SO_2(ppm), CO(μg/m^2), O_3(ppm), NO_2(ppm), PM_{10}(μg/m^2), $PM_{2.5}$(μg/m^2), PV_1h_previous(kW)
Output	PV_1h_ahead(kW)

humidity, and 6 types of air pollutants including PM_{10} and $PM_{2.5}$) were used. Furthermore, the solar PV generation data of a solar PV power plant located in Incheon, which was the experimental subject, were provided through the Open Data Portal [42]. The external observational variables utilized in the experiment are listed in Table I.

The weather and air pollutant observatory stations are located 10 and 3 km away from the solar PV power plant, respectively. The missing time periods in the measured numerical data were interpolated using the k-NN algorithm with data obtained 36 h before and after the missing time point. The k-NN algorithm is renowned for its simplicity and its inherent quality of refraining from imposing rigid assumptions upon the data. Moreover, it exhibits a notable capability to adeptly discern and encapsulate local patterns and oscillations in the dataset by exclusively considering the nearest neighbor data points when formulating predictions. In weather data analysis, where significant daily variances are infrequent, a temporal interpolation approach was adopted. This approach incorporates a comprehensive window of surrounding data, spanning three consecutive days (equivalent to 36 h) both before and after the missing data point. Additionally, three temporal parameters representing the month, day, and time, along with a parameter representing the power generated 1 h in the past, were included to enhance the prediction performance of solar PV generation through NWP. This augmentation strengthens the time-series characteristics of the external observational data.

IV. PROPOSED SOLAR PV GENERATION FORECASTING MODEL

In previous studies [31], [32] utilizing multisource data, numerical data such as wind direction and speed were extracted from satellite images and used as input for the model. However, extracting 1-D data from 2-D satellite images fails to fully exploit the inherent characteristics of the images, representing a limitation of previous studies utilizing multisource data. Therefore, to fully utilize the 2-D information contained in satellite images, this study applied the CNN-LSTM algorithm.

A. CNN-LSTM

CNN is used in various applications, such as natural image classification, image recognition, and motion prediction through computer vision [43], [44]. CNN-based networks excel at extracting raw data features from original images used for learning. Compared to the existing multilayer perceptron, the CNN maintains the shape of the input and output data of each layer and performs convolution through an image patch instead of a single pixel. Therefore, CNN can reduce noise and enhance efficiency using pooling layers, which aggregate and amplify the features extracted from the image [45].

The recurrent neural network (RNN) algorithm compensates for the inability of the artificial neural network to capture the time relationships between the data [46]. The RNN uses the analysis results of the previous layer as an input value for the current layer by sharing weights between layers, making it suitable for processing sequential data, such as time-series data. However, the RNN forgets the most deleted input information because its sensitivity to information in the initial step decreases as learning progresses. Moreover, issues such as gradient vanishing and gradient explosion arise as the time step between data points increases, leading to challenges in reflecting long-term dependencies. LSTM, which is a type of RNN, was proposed to solve the long-term dependency problem of RNNs [47]. Unlike the single hidden layer of the RNN, LSTM comprises a cell state that stores and passes key input information and three gates that can add and remove various information from the data; this allows LSTM to identify long-term relationships in the data accurately by passing and learning only important information [48].

Although a CNN achieves excellent performance when spatial characteristics are reflected, it has difficulty identifying related information in consecutively captured images because it uses only one frame at a time as input. The LSTM can properly identify the features of consecutive temporal data, but it is disadvantageous for processing input data such as consecutively captured image frames. Thus, this study applied a CNN-LSTM-based ensemble algorithm to develop a hybrid solar PV generation prediction model that reflects the spatiotemporal characteristics of external meteorological variables. The CNN-LSTM algorithm, a modification of the regions with CNN features (R-CNN) [49], combines the advantages of both CNNs and LSTMs, making it well-suited for optimizing spatial features and time-series characteristics present in consecutively captured images [50], [51]. In this approach, spatial patterns are identified and extracted through the CNN, and these extracted patterns' variability over time is subsequently reflected through the LSTM to predict solar PV generation. However, standard LSTMs do not inherently incorporate spatial information, as they employ full connections for input-to-state and state-to-state transitions when processing spatiotemporal data. To circumvent this limitation, all inputs were combined with the CNN and used as a 3-D tensor with two dimensions of columns and rows.

B. Semantic Segmentation

Semantic segmentation or dense prediction predicts the labels of all pixels in an image. As one of the most important areas of computer vision, semantic segmentation can fully understand photos rather than simply classify them, leading to its use in diverse applications, such as autonomous driving and inferring the relationships between objects. SegNet was proposed to classify structures relating to autonomous

driving such as roads, buildings, cars, and pedestrians in pixel units [52].

SegNet largely consists of encoders and decoders. The encoder network adopts a similar architecture to VGG16, a top performer in the 2014 ImageNet competition [53]. However, it is lighter than other algorithms because it omits the fully connected layer of VGG16 and uses only 13 convolution layers. Furthermore, it performs max pooling during encoding and stores max pooling indices that represent the location information. The decoder network of SegNet is constructed by replacing the max pooling layers in the encoder with upsampling layers. These upsampling layers receive the stored max pooling indices from the encoder to perform upsampling. This approach enhances boundary delineation and enables end-to-end training; thereby, reducing the number of algorithm parameters. The encoder and decoder of SegNet are similar to the autoencoder equation and are expressed by (3). $f_\theta(\alpha)$ denotes an encoder block, and $g_{\theta'}(\beta)$ denotes a decoder block

$$\begin{aligned}\beta &= f_\theta(\alpha) = s(W\alpha + b)\\ \gamma &= g_{\theta'}(\beta) = s\left(W\beta + b'\right).\end{aligned} \tag{3}$$

In semantic segmentation tasks, U-Net is often used along with SegNet. A study investigated predicting solar PV generation using satellite images [54]. However, U-Net performs convolution after transferring the entire feature map from the encoder to the decoder instead of pooling the indices, which increases the model size and memory usage. SegNet does not have the bridge network of U-Net and uses 10 times smaller training parameters [55]. Thus, in this study, SegNet, which has excellent computational efficiency, was set as a baseline model to use four model inputs through two satellite image types, improved, and applied to the solar PV generation prediction model.

C. ROI and ROI_{surr} Synthesis Process Methods

R-CNN-based such as Fast R-CNN [56], Faster R-CNN [57], and Mask R-CNN [58] are predominantly employed for segmentation tasks. These models operate by establishing multiple ROI to detect objects within an image. Subsequently, these ROIs are projected according to the size of the feature map and processed through ROI pooling. The feature vectors that have undergone ROI pooling are aggregated into a fully connected layer. While this approach demonstrates impressive performance in object detection and ROI classification within an image, it encounters challenges in learning the spatial association and connectivity between individual objects or ROI.

In various fields, numerous studies focusing on synthesis methods using multisource data as model inputs have typically employed basic operations such as addition, concatenation, and fully connected layers for each data type. However, these methods often exhibit limitations, as they can be relatively coarse and may overlook specific features due to their simplistic approach of merely adding or connecting output values from each layer. To address these limitations, MGM-GAN proposed a gate mergence (GM) module to synthesize a multimodal MRI effectively [59]. As different modalities

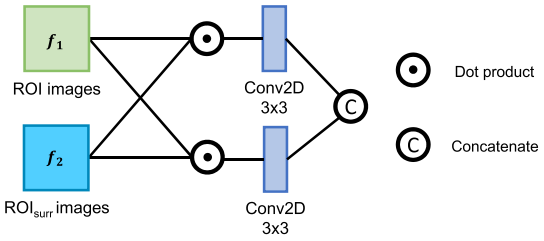

Fig. 2. Structure of ISL.

encompass distinct feature information across different locations, the GM module is employed for the automatic learning of weights associated with various modalities at multiple locations. This is achieved by training a combined weight matrix through a convolution block, with each channel corresponding to a single modality. Consequently, the GM module enables adaptive enhancement of valuable information such as edges or textures, while effectively suppressing unnecessary information such as noise.

This study aimed to learn the spatial characteristics through the efficient synthesis of the ROI and ROI_{surr} in satellite images. To achieve this, an ISL was developed, drawing inspiration from the projection-based method [60], [61] and the GM module. The structure of the ISL is depicted in Fig. 2. In the ISL, the ROI and ROI_{surr} are synthesized through the dot product. It is noteworthy that the result value of the dot product operation varies depending on the order of operations. Therefore, to enhance the quality of results, convolution and concatenate layers are applied after each of the two operation sequences. The ISL can be expressed as follows:

$$A * B = f_{\text{Conv}}(A \cdot B) \oplus f_{\text{Conv}}(B \cdot A) \tag{4}$$

where $f_{\text{Conv}}(\cdot)$, $\cdot$, and $\oplus$ represent the convolutional layer, dot product, and concatenate operation, respectively.

D. Detailed Information and Training Process of Proposed Lightweight Model

A total of five input channels, including spatiotemporal characteristics, were used through multisource data to enhance the prediction performance. In training, regarding the spatial characteristics of the satellite image, if the original satellite image is used unchanged, the range of the image includes wider information than the point where the solar PV power plant is located. Therefore, the ROI was set based on the solar PV power plant, and the ROI_{surr} was set to the same size along the ROI and then synthesized [31], [32]. This approach facilitates effective training on the movement effects of clouds and PM while reducing computational complexity. Fig. 3 depicts the ROI and ROI_{surr} in the satellite image in red and green, respectively.

Fig. 4 presents the detailed structure of the image process module in the proposed model. The image process module proposed in this study incorporates an ISL based on an autoencoder architecture, such as the SegNet-based algorithm [36], to distinguish between cloud and ground-based images. This enables the proposed model to effectively learn spatiotemporal characteristics from two satellite images capturing clouds and PM captured by infrared rays. Thus, the synthesis of ROI and

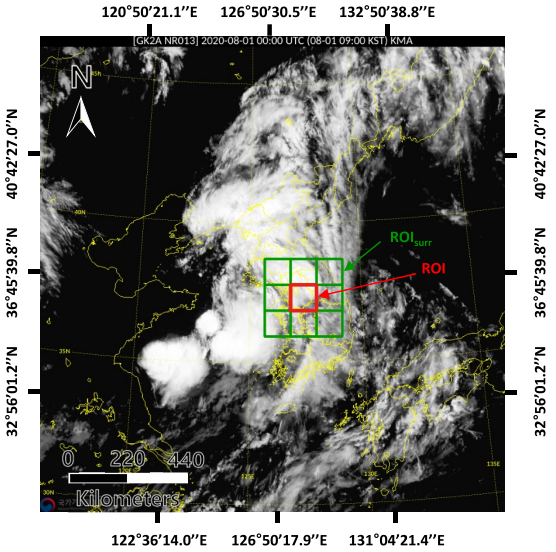

Fig. 3. ROI (red) and ROI_{surr} (green) set in infrared cloud satellite image.

ROI_{surr} through the image process module enables the efficient training of the two images. Furthermore, the CNN-LSTM-based ensemble algorithm including ISL and autoencoder was applied to strengthen the spatiotemporal characteristics of continuously captured satellite images. The encoders and decoders of VGG16 and SegNet have the same input and output sizes. However, weights of the same size were extracted through a max pooling layer in the encoder part because the ROI and ROI_{surr} in the satellite image had different sizes. Four convolutional and three max pooling layers were applied to the encoder block, and three convolutional and three upsampling layers were applied to the decoder block. All convolutional layers used ReLU as the activation function. The synthesis of ROI and ROI_{surr} occurred within the ISL in the encoder part and subsequently passed through the decoder part, expressed as follows:

$$X_t^{\text{Cloud}} = g_{\theta'}\left(f_\theta\left(X_t^{\text{Cloud}_{\text{ROI}}} * X_t^{\text{Cloud}_{\text{ROI}_{\text{surr}}}}\right)\right)$$
$$X_t^{\text{PM}} = g_{\theta'}\left(f_\theta\left(X_t^{\text{PM}_{\text{ROI}}} * X_t^{\text{PM}_{\text{ROI}_{\text{surr}}}}\right)\right). \quad (5)$$

Thereafter, the output values were passed through the CNN-LSTM, and training was performed in each cloud and PM image process module.

If the ROI and ROI_{surr} of each image of clouds and PM reflected the spatial information of the surrounding context through the output value through the image process module, numerical data were used to reflect the information of the external meteorological factors, including time-series characteristics affecting solar PV generation. This new input was utilized as multisource data within the NWP process module. The NWP process module employed LSTM to emphasize the time-series characteristics of the numerical data measured at 1-h intervals. Finally, after combining the results of the image and NWP process modules through the element-wise multiplication layer, learning was conducted using a new LSTM to predict solar PV generation 1 h into the future. To prevent overfitting issues as the model layers deepened, a dropout layer was applied.

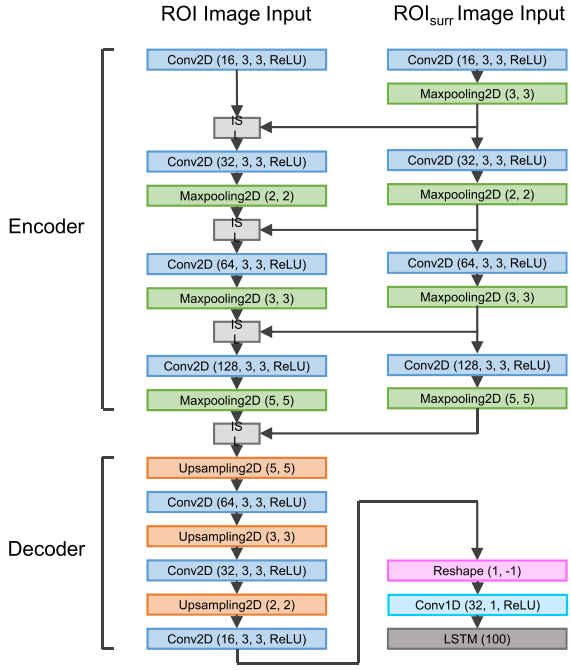

Fig. 4. Structure of image process module in the proposed model.

The platform for all experiments was a high-performance server built on Ubuntu 18.04 with Intel(R) i9-10900X CPU@3.70 GHz, 256 GB RAM, NVIDIA TESLA V100 2-way GPU. Python 3.6.13, Tensorflow 1.15.0, and Keras 2.1.0 were configured for the experimental environment.

V. Experiments Results and Discussions

A. Experimental Setup

The entire experimental period was divided into three groups to train the spatiotemporal hybrid solar PV generation prediction model: training, validation, and evaluation. To account for the significant influence of seasonal variations, particularly in South Korea, data from the 1st to the 24th of each month were allocated to the training set, while data from the 25th to the 30th were assigned to the test set. Additionally, data of equivalent size to the test set were randomly extracted from the training set to form the validation set.

The ROI and ROI_{surr} were delineated within the satellite image to examine the surrounding influences on the solar PV power plant's location. Two comparison experiments were performed to identify the influences of the ROI and ROI_{surr}: the original satellite image was directly learned (original image), and the ROI and surrounding region were used together (ROI image). For the original image experiment, the 900 × 900 pixel image was modified to 450 × 450 pixels to reduce the computational complexity and memory usage. Various machine learning algorithms were employed as baseline models for the optimal solar PV generation analysis prediction. In addition to the CNN-LSTM algorithm, which is the most basic machine learning technique using images, this study also compared SegNet and U-Net, which are the most commonly used semantic segmentation approaches, with the proposed model. The sizes of the ROI and ROI_{surr} were set to 30 × 30 and 90 × 90, respectively. Furthermore, two additional

experiments with varied sizes of ROI and ROI_{surr} were performed to analyze the prediction performance based on the sizes. The sizes of the ROI were set to 5×5 and 50×50, and those of the ROI_{surr} were set to 15×15 and 150×150 for the comparative analysis.

B. Performance Evaluation Indicators

This study considered five different performance evaluation indicators to verify the performance of the proposed spatiotemporal hybrid solar PV generation prediction model. These metrics including the mean absolute error (MAE), root mean square error (RMSE), and SMAPE were determined. Although the mean absolute percentage error (MAPE) is frequently used to assess the prediction performance because it enables absolute evaluation, it cannot be used if the actual observed value is zero because it returns an infinite value. Given that solar PV generation measurements near sunrise and sunset may occasionally register as zero, the MAPE's applicability in this context is limited. To address this limitation, the study adopted SMAPE for evaluating the absolute prediction performance of the proposed model. Unlike MAPE, SMAPE calculates the error relative to the average of the actual and predicted values, thereby avoiding issues with zero values. Moreover, SMAPE is expressed in terms of ratios rather than absolute numbers, facilitating the comparison of model efficacy across various datasets and models. The expressions of MAE, RMSE, and SMAPE are shown as follows:

$$\mathrm{MAE} = \frac{1}{n}\sum_{i=1}^{n}\left|y'_i - y_i\right| \tag{6}$$

$$\mathrm{RMSE} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}\left(y'_i - y_i\right)^2} \tag{7}$$

$$\mathrm{MAPE}(\%) = \frac{1}{n}\sum_{i=1}^{n}\left|\frac{y_i - y'_i}{y_i}\right| \tag{8}$$

$$\mathrm{SMAPE}(\%) = \frac{1}{n}\sum_{i=1}^{n}\frac{\left|y_i - y'_i\right|}{|y_i| + \left|y'_i\right|}. \tag{9}$$

In addition to the aforementioned evaluation metrics, further validation of the prediction model was conducted based on two standards outlined in the ASHRAE Guideline 14 [62], which establishes minimum acceptable performance levels and assists energy managers in determining standardized energy and demand savings. These standards include the mean bias error (MBE) and coefficient of variation of the RMSE [CV (RMSE)]. The expressions for the MBE and CV (RMSE) are as follows:

$$\mathrm{MBE}(\%) = \frac{1}{n}\sum_{i=1}^{n}\left(y'_i - y_i\right) \tag{10}$$

$$\mathrm{CV(RMSE)}(\%) = \frac{1}{\bar{y}}\sqrt{\frac{1}{n}\sum_{i=1}^{n}\left(y'_i - y_i\right)^2}. \tag{11}$$

The ASHRAE Guideline 14 specifies two standards for the MBE and CV (RMSE) in terms of monthly and hourly predictions. For hourly prediction, the MBE should be within $\pm 10\%$, and the CV (RMSE) should be within 30%. The average value was calculated by applying the absolute value to the resultant value because the MBE may produce negative values.

Deep learning models typically demand substantial computational resources for both training and inference tasks; as the model size increases, the amount of computation increases proportionally, and accordingly, the required power consumption and training time increase. Hence, the development of lightweight technologies, which require minimal storage and computational resources while maintaining high inference speeds, is essential. Performance comparisons of model weight reduction are primarily based on the number of parameters used and the amount of computation. In this study, to compare the lightweight performance of the proposed model, a number of parameters and floating point operations (FLOPs) were used for the analysis.

C. Experimental Results of Proposed Lightweight Model

Performance comparisons with existing algorithms were conducted to evaluate the influence of the ROI and ROI_{surr} in satellite images on the proposed lightweight model. A total of 12 experiments were conducted for 6 models. Table II summarizes the experimental results. While no significant difference in performance was evident between the original model group and the ROI and ROI_{surr} model group based on SMAPE, an improvement in prediction performance was observed across all models for the ROI and ROI_{surr} based on MAE and RMSE metrics. This result demonstrates that using satellite images to predict solar PV generation instead of the unmodified original image is more effective for applying an ROI based on the solar PV power plant and intensively learning the relevant section. The prediction performance likely decreased when using the unmodified original image in the prediction model for learning due to the excessive inclusion of unnecessary spatial characteristics of the location region of the solar PV power plant. The overall prediction performance of the models increased in the order of SARIMAX-LSTM, U-Net, CNN-LSTM, SegNet, the GM module, and the proposed lightweight model. The SARIMAX-LSTM model suffers from performance degradation due to the extraction of 2-D satellite image data into 1-D data, unlike other models that utilize the full 2-D image data. Notably, the proposed lightweight model achieved the best performance in all performance evaluation indicators except for SMAPE. While the GM module outperformed slightly in terms of SMAPE, the difference was only 0.017%. According to the ASHRAE Guideline 14, the model utilizing the original image of SegNet and ROI and ROI_{surr} of SARIMAX-LSTM, SegNet, GM module, and the proposed lightweight model satisfied the MBE criteria. However, only the GM module and proposed lightweight model using the ROI and ROI_{surr} satisfied the CV (RMSE) criteria. Remarkably, the proposed lightweight model using the ROI and ROI_{surr} achieved an overwhelmingly high performance of 0.027% for the MBE.

The results of model weight reduction indicate that all models, except for SARIMAX-LSTM, achieved significant weight reduction when utilizing the ROI and ROI_{surr} compared

TABLE II
COMPARISON OF EXPERIMENTAL RESULTS USING ORIGINAL IMAGE AND ROI AND ROI_{surr} IMAGE

	SARIMAX-LSTM [32]		CNN-LSTM-based [50]		SegNet-based [52]		U-Net-based [54]		GM module-based [59]		Proposed	
	Original	ROI & ROI_{surr}	Original	ROI & ROI_{surr}	Original	ROI & ROI_{surr}	Original	ROI & ROI_{surr}	Original	ROI & ROI_{surr}	Original	ROI & ROI_{surr}
MAE [kW]	76.858	68.831	61.597	59.027	67.290	57.083	66.250	61.542	56.574	55.514	59.731	**50.957**
RMSE [kW]	101.179	94.770	89.418	89.118	94.876	88.601	96.404	90.516	84.810	83.395	89.528	**81.434**
SMAPE [%]	33.922	32.354	32.643	31.633	30.629	29.122	29.370	36.708	28.884	**28.880**	28.996	28.897
MBE [%]	14.036	5.119	10.840	11.635	9.504	1.971	13.874	11.701	6.496	6.050	13.713	**0.027**
CV (RMSE) [%]	36.089	33.803	31.894	31.787	33.841	31.603	34.386	32.286	30.251	29.746	31.934	**29.046**
Params [M]	1.209	1.213	207.628	2.245	26.142	0.547	13.163	**0.386**	212.775	10.645	207.989	3.314
FLOPs [M]	90.871	17.174	414.866	3.871	52.043	0.774	26.082	**0.448**	425.221	41.906	415.654	6.299

TABLE III
COMPARISON OF EXPERIMENTAL RESULTS OF MODEL USING SINGLE-SOURCE DATA AND PROPOSED MODEL USING MULTISOURCE DATA

	NWP	Satellite	Proposed
MAE	65.490	100.185	**50.957**
RMSE	94.570	127.772	**81.434**
SMAPE	31.005	37.844	**27.897**
MBE	10.014	1.804	**0.027**
CV (RMSE)	33.732	45.575	**29.046**

to using the original image. The model with the greatest weight reduction was CNN-LSTM, the parameters were 98.919%, and the FLOPs were 99.067% lightweight. The proposed lightweight model also achieved substantial weight reduction, with parameters reduced by 98.407% and FLOPs reduced by 98.485%. Thus, in all experimental cases, when the ROI and ROI_{surr} were used, a weight reduction of at least 90.145% was achieved, which is higher than that achieved by the models using the original image.

A comparative experiment was conducted with a model using single-source data to determine the importance of using multisource data in the solar power generation prediction model. The structure of the single-source data model mirrored that of the proposed model, incorporating NWP and image process modules. Table III presents the comparison results between the proposed model using multisource data and the model using single-source data, revealing a discernible disparity in performance.

Fig. 5 shows the results that were extracted from January 30th and February 25th to 27th among the results from November 30th, 2020, to March 30th for the ROI and ROI_{surr} model groups. Although all models demonstrated predictions closely aligned with the actual observed solar PV generation data, the proposed model exhibited the most accurate results. All models reliably predicted solar PV generation when the solar power was uniformly generated in sunny weather. However, inaccuracies were evident on days characterized by significant fluctuations in external meteorological factors, as observed during the second peak on February 25th. Accordingly, an additional experiment was performed to predict solar PV generation more accurately.

D. Experimental Results According to ROI Size

An experiment to set the optimal ROI size was conducted to increase the proposed model prediction accuracy. The ROI

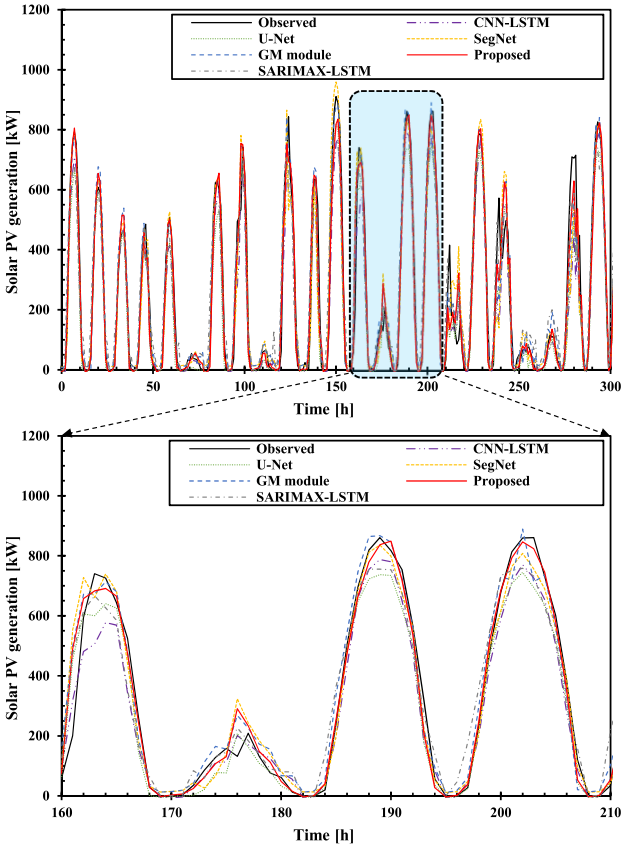

Fig. 5. Results of solar PV generation prediction model using ROI image and ROI_{surr} image.

in the satellite image reflects spatial information on external meteorological factors in the region of the solar PV power plant. If the ROI is too diminutive, it might inadequately capture spatial characteristics. Conversely, an excessively large ROI could lead to superfluous learning of irrelevant information akin to utilizing the original image, potentially compromising performance. In the proposed lightweight model, the ROI and ROI_{surr} were initially set at 30 × 30 and 90 × 90, respectively, encompassing an area spanning 60 × 60 km and 180 × 180 km around the solar PV power plant.

Five new sets of ROI sizes were established based on the previous ROI dimensions, encompassing both larger and smaller configurations (ROI: 5 × 5, 10 × 10, 20 × 20, 40 × 40, and 50 × 50; ROI_{surr}: 15 × 15, 30 × 30, 60 × 60, 120 × 120, and 150 × 150). To determine the optimal ROI size, prediction models were assessed for each ROI dimension using

TABLE IV
RESULTS OF SIXFOLD CROSS-VALIDATION BY ROI SIZE

Models	MAE	RMSE	SMAPE (%)	MBE (%)	CV (RMSE) (%)
ROI 5 × 5	49.383	75.734	28.754	5.800	27.939
ROI 10 × 10	**47.531**	**73.416**	**28.066**	**2.382**	**27.186**
ROI 20 × 20	47.895	75.219	28.619	2.549	27.872
ROI 30 × 30	49.235	75.208	28.765	5.547	27.581
ROI 40 × 40	50.331	78.691	29.503	2.929	29.111
ROI 50 × 50	49.333	75.150	29.396	3.762	27.632

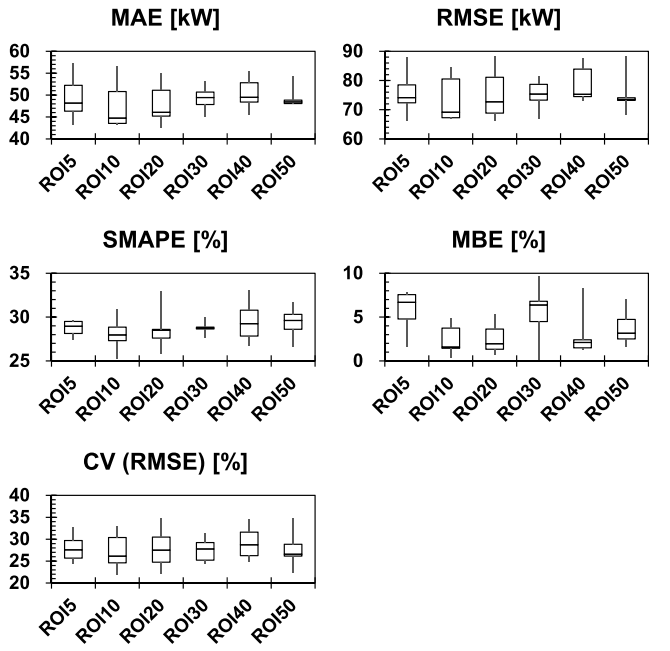

Fig. 6. Box and whisker plot of sixfold cross-validation by ROI size.

k-fold cross-validation. In the previous experiment, the test set period was designated as five days from the 25th to the 30th of each month within the entire experimental period. Therefore, each month was divided into six equal parts of five days each; the k-value in the k-fold cross-validation was set to 6 and the experiment was performed identically using the proposed lightweight model. Table IV displays the results of the sixfold cross-validation tests, while Fig. 6 depicts a box and whisker plot illustrating the model performance across different ROI sizes.

According to the analysis results, all six models satisfied the ASHRAE Guideline 14 standard. Excluding ROI 5 × 5 and ROI 40 × 40 from the set of performance evaluation metrics, an increase in prediction accuracy is observed as the size of the ROI diminishes. Particularly for ROI 5 × 5, limitations in feature extraction capacity are evident due to the small input image size within the encoder segment of the proposed model. Additionally, the loss of information that occurs as the model progresses to deeper layers is believed to have a notable impact. Therefore, underscores the importance of precisely defining the ROI size, centered on the region of the solar PV power plant in the satellite image, for accurate solar PV generation prediction.

E. Universally Applicable Input Data for the Proposed Model

The proposed model integrates remote sensing data, satellite images, and numerically recorded meteorological data as multisource inputs. Satellite images are provided in various formats depending on the source organization and the satellite used. They are typically available in image formats such as PNG or data array formats such as NetCDF, BUFR, etc. When satellite images are provided in NetCDF format, conversion to image format can be achieved using dedicated for data visualization. In cases where dedicated software is unavailable, conversion can still be carried out using various libraries or tools. For images provided directly, they can be used as input for the proposed model without further processing. However, if the data are in array format, conversion to image format is necessary before applying it to the model. Typically, pixel values in image data range from 0 to 255, applicable to both grayscale and color images. To convert data arrays into image data, each value in the array can be divided by 255, normalizing the data to a range between 0 and 1. Following this method, the data can be effectively applied to the proposed model, regardless of the type of satellite data provided.

The most accurate and reliable numerically recorded meteorological data for solar PV forecasting are obtained directly from measurements taken at solar PV power plants. However, acquiring such data for model development purposes can pose challenges due to various factors, including security concerns and data storage cycles. As an alternative, weather observation data provided by organizations such as the Meteorological Administration or other relevant authorities can be utilized. These observations typically cover the areas surrounding solar PV power plants and offer valuable insights into meteorological conditions that influence solar PV generation. While not as precise as direct measurements from solar PV plants, leveraging this meteorological data remains a practical and effective approach for solar PV generation forecasting.

Most artificial satellites typically cover highly large areas at once during observation, rather than focusing solely on one region or country. Geostationary satellites such as GK2A, employed in experiments, not only prioritize capturing South Korea but also include neighboring countries, enabling the use of the proposed model. Furthermore, it is anticipated that analysis will be feasible in new countries as well, based on the results of this study, by leveraging satellite and numerical data from those respective countries.

VI. CONCLUSION

The proposed lightweight spatiotemporal hybrid solar PV generation prediction model operates on 1-h intervals, facilitating stable and efficient management and power supply planning for power grid systems. To compensate for the shortcomings of existing single-source data forecasting, multisource data were used to reflect the spatiotemporal characteristics. Spatial features were derived from two infrared satellite images, while temporal characteristics were captured through 16 types of numerical data recorded hourly. Weight reduction in the prediction model was achieved by setting the ROI and ROI_{surr} in the satellite image. Furthermore, the image and

NWP process modules were used to process the multisource data. The image process module reflects spatial characteristics by synthesizing the ROI and ROI_{surr} through the ISL and strengthens the time-series characteristics of continuously captured satellite images with the CNN-LSTM-based ensemble algorithm. The output value of the image process module is combined with that of the NWP process module, and the time-series characteristics of the observed weather variables are integrated to finally predict the amount of solar PV generation.

A comparative analysis was performed with the baseline model of five machine learning algorithms to evaluate the prediction performance of the proposed lightweight model. In a total of 12 cases, the prediction performance was improved when the ROI image was used compared to when the original image was used; furthermore, the performance improved by up to 33.7% and 19.51% based on MAE and RMSE, respectively. Moreover, the proposed lightweight model consistently satisfied all standards outlined in the ASHRAE Guideline 14. The performance of the proposed model was also the best compared to the model utilizing single-source data. To further comprehend the impact of the ROI size in the satellite image, an experiment aimed at determining the optimal ROI size for the proposed lightweight model was conducted. According to the verification results obtained through sixfold cross-validation for a total of five ROIs, a smaller ROI size yielded better performance. To further comprehend the impact of the ROI size in the satellite image, an experiment aimed at determining the optimal ROI size for the proposed lightweight model was conducted.

The proposed lightweight multisource data-based hybrid spatiotemporal solar PV generation prediction model offers a versatile solution for adapting to climate variations by providing accurate and precise forecasts of solar PV generation. In the future, this model holds promise in guiding the operation of smart grids, particularly when integrated with energy storage systems. Furthermore, it could play a pivotal role in designing future power grid systems that prioritize power prosumers and emphasize sustainability and efficiency. Through its ability to harness multisource data and incorporate spatiotemporal characteristics, the model is poised to contribute significantly to the advancement of renewable energy utilization and grid management strategies.

REFERENCES

[1] Korea Ministry of Trade, Industry and Energy. (2014). *Renewable Energy Statistics 2013*. [Online]. Available: http://www.motie.go.kr

[2] (2021). *New Energy Outlook 2021*. [Online]. Available: https://about.bnef.com/new-energy-outlook/

[3] Q.-T. Phan, Y.-K. Wu, Q.-D. Phan, and H.-Y. Lo, "A novel forecasting model for solar power generation by a deep learning framework with data preprocessing and postprocessing," *IEEE Trans. Ind. Appl.*, vol. 59, no. 1, pp. 220–231, Jan. 2023, doi: 10.1109/TIA.2022.3212999.

[4] H. S. Jang, K. Y. Bae, H.-S. Park, and D. K. Sung, "Solar power prediction based on satellite images and support vector machine," *IEEE Trans. Sustain. Energy*, vol. 7, no. 3, pp. 1255–1263, Jul. 2016.

[5] IEA. (2020). *Renewables 2020*. [Online]. Available: https://www.iea.org/reports/renewables-2020

[6] X. Zhang et al., "Solar irradiance prediction interval estimation and deterministic forecasting model using ground-based sky image," in *Proc. IEEE/IAS 58th Ind. Commercial Power Syst. Tech. Conf. (I&CPS)*, May 2022, pp. 1–8, doi: 10.1109/ICPS54075.2022.9773822.

[7] K. Hu, S. Cao, L. Wang, W. Li, and M. Lv, "A new ultra-short-term photovoltaic power prediction model based on ground-based cloud images," *J. Cleaner Prod.*, vol. 200, pp. 731–745, Nov. 2018, doi: 10.1016/j.jclepro.2018.07.311.

[8] Y. Yin, Z. Li, X. Li, X. Wu, and Y. Yang, "Ultra-short term photovoltaic power prediction," in *Proc. IEEE 2nd Int. Conf. Electron. Technol., Commun. Inf. (ICETCI)*, May 2022, pp. 645–649, doi: 10.1109/ICETCI55101.2022.9832053.

[9] J. Kong, S. Oh, J. Jung, and I. Lee, "Improving the accuracy of photovoltaic generation forecasting by considering particulate matter variables," in *Proc. IEEE Power Energy Soc. Gen. Meeting (PESGM)*, Aug. 2020, pp. 1–5, doi: 10.1109/PESGM41954.2020.9282054.

[10] V. H. Wentz, J. N. Maciel, J. J. Gimenez Ledesma, and O. H. Ando Junior, "Solar irradiance forecasting to short-term PV power: Accuracy comparison of ANN and LSTM models," *Energies*, vol. 15, no. 7, p. 2457, Mar. 2022.

[11] A. N.-L. Huynh, R. C. Deo, D.-A. An-Vo, M. Ali, N. Raj, and S. Abdulla, "Near real-time global solar radiation forecasting at multiple time-step horizons using the long short-term memory network," *Energies*, vol. 13, no. 14, p. 3517, Jul. 2020, doi: 10.3390/en13143517.

[12] E. Collino and D. Ronzio, "Exploitation of a new short-term multimodel photovoltaic power forecasting method in the very short-term horizon to derive a multi-time scale forecasting system," *Energies*, vol. 14, no. 3, p. 789, Feb. 2021, doi: 10.3390/en14030789.

[13] X. G. Agoua, R. Girard, and G. Kariniotakis, "Short-term spatio-temporal forecasting of photovoltaic power production," *IEEE Trans. Sustain. Energy*, vol. 9, no. 2, pp. 538–546, Apr. 2018, doi: 10.1109/TSTE.2017.2747765.

[14] P. M. Kumar, R. Saravanakumar, A. Karthick, and V. Mohanavel, "Artificial neural network-based output power prediction of grid-connected semitransparent photovoltaic system," *Environ. Sci. Pollut. Res.*, vol. 29, no. 7, pp. 10173–10182, Feb. 2022, doi: 10.1007/s11356-021-16398-6.

[15] H. Li, Z. Ren, Y. Xu, W. Li, and B. Hu, "A multi-data driven hybrid learning method for weekly photovoltaic power scenario forecast," *IEEE Trans. Sustain. Energy*, vol. 13, no. 1, pp. 91–100, Jan. 2022, doi: 10.1109/TSTE.2021.3104656.

[16] T. Hove, "A method for predicting long-term average performance of photovoltaic systems," *Renew. Energy*, vol. 21, no. 2, pp. 207–229, Oct. 2000, doi: 10.1016/s0960-1481(99)00131-7.

[17] D. P. Larson, L. Nonnenmacher, and C. F. M. Coimbra, "Day-ahead forecasting of solar power output from photovoltaic plants in the American Southwest," *Renew. Energy*, vol. 91, pp. 11–20, Jun. 2016, doi: 10.1016/j.renene.2016.01.039.

[18] K. Mahmud, S. Azam, A. Karim, S. Zobaed, B. Shanmugam, and D. Mathur, "Machine learning based PV power generation forecasting in Alice springs," *IEEE Access*, vol. 9, pp. 46117–46128, 2021, doi: 10.1109/ACCESS.2021.3066494.

[19] G. G. Kim et al., "Prediction model for PV performance with correlation analysis of environmental variables," *IEEE J. Photovolt.*, vol. 9, no. 3, pp. 832–841, May 2019, doi: 10.1109/JPHOTOV.2019.2898521.

[20] Y. Li, Y. Su, and L. Shu, "An ARMAX model for forecasting the power output of a grid connected photovoltaic system," *Renew. Energy*, vol. 66, pp. 78–89, Jun. 2014, doi: 10.1016/j.renene.2013.11.067.

[21] E. G. Kardakos, M. C. Alexiadis, S. I. Vagropoulos, C. K. Simoglou, P. N. Biskas, and A. G. Bakirtzis, "Application of time series and artificial neural network models in short-term forecasting of PV power generation," in *Proc. 48th Int. Universities' Power Eng. Conf. (UPEC)*, Sep. 2013, pp. 1–6, doi: 10.1109/UPEC.2013.6714975.

[22] D. Yang and Z. Dong, "Operational photovoltaics power forecasting using seasonal time series ensemble," *Sol. Energy*, vol. 166, pp. 529–541, May 2018, doi: 10.1016/j.solener.2018.02.011.

[23] D. Yu, W. Choi, M. Kim, and L. Liu, "Forecasting day-ahead hourly photovoltaic power generation using convolutional self-attention based long short-term memory," *Energies*, vol. 13, no. 15, p. 4017, Aug. 2020, doi: 10.3390/en13154017.

[24] M. Abdel-Nasser and K. Mahmoud, "Accurate photovoltaic power forecasting models using deep LSTM-RNN," *Neural Comput. Appl.*, vol. 31, no. 7, pp. 2727–2740, Oct. 2019, doi: 10.1007/s00521-017-3225-z.

[25] M. Massaoudi et al., "An effective hybrid NARX-LSTM model for point and interval PV power forecasting," *IEEE Access*, vol. 9, pp. 36571–36588, 2021, doi: 10.1109/ACCESS.2021.3062776.

[26] F. Mei, Q. Wu, T. Shi, J. Lu, Y. Pan, and J. Zheng, "An ultrashort-term net load forecasting model based on phase space reconstruction and deep neural network," *Appl. Sci.*, vol. 9, no. 7, p. 1487, Apr. 2019, doi: 10.3390/app9071487.

[27] A. Mellit, A. M. Pavan, E. Ogliari, S. Leva, and V. Lughi, "Advanced methods for photovoltaic output power forecasting: A review," *Appl. Sci.*, vol. 10, no. 2, p. 487, Jan. 2020, doi: 10.3390/app10020487.

[28] H. Yang, L. Wang, C. Huang, and X. Luo, "3D-CNN-based sky image feature extraction for short-term global horizontal irradiance forecasting," *Water*, vol. 13, no. 13, p. 1773, Jun. 2021, doi: 10.3390/w13131773.

[29] P. Wood-Bradley, J. Zapata, and J. Pye, "Cloud tracking with optical flow for short-term solar forecasting," in *Proc. 50th Conf. Austral. Sol. Energy Soc.*, 2012, pp. 2–7. [Online]. Available: http://stg.anu.edu.au/publications/assets/inproc/woodbradley-auses-2012.pdf

[30] Q. Paletta and J. Lasenby, "Convolutional neural networks applied to sky images for short-term solar irradiance forecasting," 2020, *arXiv:2005.11246*.

[31] B. Kim and D. Suh, "A hybrid spatio-temporal prediction model for solar photovoltaic generation using numerical weather data and satellite images," *Remote Sens.*, vol. 12, no. 22, p. 3706, Nov. 2020, doi: 10.3390/rs12223706.

[32] B. Kim, D. Suh, M.-O. Otto, and J.-S. Huh, "A novel hybrid spatio-temporal forecasting of multisite solar photovoltaic generation," *Remote Sens.*, vol. 13, no. 13, p. 2605, Jul. 2021, doi: 10.3390/rs13132605.

[33] T. Yao et al., "Intra-hour photovoltaic generation forecasting based on multi-source data and deep learning methods," *IEEE Trans. Sustain. Energy*, vol. 13, no. 1, pp. 607–618, Jan. 2022, doi: 10.1109/TSTE.2021.3123337.

[34] H. Jiang et al., "Geospatial assessment of rooftop solar photovoltaic potential using multi-source remote sensing data," *Energy AI*, vol. 10, Nov. 2022, Art. no. 100185, doi: 10.1016/j.egyai.2022.100185.

[35] M. Trigo-González et al., "Photovoltaic power electricity generation nowcasting combining sky camera images and learning supervised algorithms in the Southern Spain," *Renew. Energy*, vol. 206, pp. 251–262, Apr. 2023, doi: 10.1016/j.renene.2023.01.111.

[36] S. Dev, A. Nautiyal, Y. H. Lee, and S. Winkler, "CloudSegNet: A deep network for nychthemeron cloud image segmentation," *IEEE Geosci. Remote Sens. Lett.*, vol. 16, no. 12, pp. 1814–1818, Dec. 2019, doi: 10.1109/LGRS.2019.2912140.

[37] *National Meteorogical Satellite Center*. [Online]. Available: https://nmsc.kma.go.kr/

[38] B. Lee, "GEO-KOMPSAT-2A AMI algorithm theoretical basis document cloud detection," Tech. Rep., 2018.

[39] K.-H. Lee, J.-B. Jee, and K.-T. Lee, "GK-2A AMI algorithm theoretical basis document aerosol detection product (ADP)," Tech. Rep., 2018.

[40] *Korea Meteorolgical Administration*. [Online]. Available: https://data.kma.go.kr/

[41] *Air Korea*. [Online]. Available: https://www.airkorea.or.kr/

[42] *Open Data Portal*. [Online]. Available: https://www.data.go.kr/

[43] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "ImageNet classification with deep convolutional neural networks," *Commun. ACM*, vol. 60, no. 2, pp. 84–90, Jun. 2012.

[44] J. Long, E. Shelhamer, and T. Darrell, "Fully convolutional networks for semantic segmentation," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2015, pp. 3431–3440, doi: 10.1109/CVPR.2015.7298965.

[45] Y.-L. Boureau, J. Ponce, and Y. LeCun, "A theoretical analysis of feature pooling in visual recognition," in *Proc. 27th Int. Conf. Mach. Learn. (ICML)*, 2010, pp. 111–118. [Online]. Available: https://www.di.ens.fr/sierra/pdfs/icml2010b.pdf

[46] M. Biehl, "Supervised sequence labelling with recurrent neural neural networks," *Neural Networks*, vol. 1999, no. 3, p. 160, 2005. [Online]. Available: http://www.amazon.com/Supervised-Labelling-Recurrent-Computational-Intelligence/dp/3642247962

[47] K. Greff, R. K. Srivastava, J. Koutník, B. R. Steunebrink, and J. Schmidhuber, "LSTM: A search space Odyssey," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 28, no. 10, pp. 2222–2232, Oct. 2017, doi: 10.1109/TNNLS.2016.2582924.

[48] Y. Yu, J. Cao, and J. Zhu, "An LSTM short-term solar irradiance forecasting under complicated weather conditions," *IEEE Access*, vol. 7, pp. 145651–145666, 2019, doi: 10.1109/ACCESS.2019.2946057.

[49] R. Girshick, J. Donahue, T. Darrell, and J. Malik, "Rich feature hierarchies for accurate object detection and semantic segmentation," 2013, *arXiv:1311.2524*.

[50] 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. Neural Inf. Process. Syst.*, 2015, doi: 10.1155/2018/6184713.

[51] C. Ding, G. Wang, X. Zhang, Q. Liu, and X. Liu, "A hybrid CNN-LSTM model for predicting $PM_{2.5}$ in Beijing based on spatiotemporal correlation," *Environ. Ecological Statist.*, pp. 503–522, Apr. 2021, doi: 10.1007/s10651-021-00501-8.

[52] V. Badrinarayanan, A. Kendall, and R. Cipolla, "SegNet: A deep convolutional encoder-decoder architecture for image segmentation," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 39, no. 12, pp. 2481–2495, Dec. 2017.

[53] K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," in *Proc. 3rd Intl. Conf. Learning Represent. (ICLR)*, 2015, pp. 1–14.

[54] O. Ronneberger, P. Fischer, and T. Brox, "U-Net: Convolutional networks for biomedical image segmentation," in *Medical Image Computing and Computer-Assisted Intervention—MICCAI*, 2015, doi: 10.1007/978-3-319-24574-4.

[55] P. K. Jain, N. Sharma, A. A. Giannopoulos, L. Saba, A. Nicolaides, and J. S. Suri, "Hybrid deep learning segmentation models for atherosclerotic plaque in internal carotid artery B-mode ultrasound," *Comput. Biol. Med.*, vol. 136, Sep. 2021, Art. no. 104721, doi: 10.1016/j.compbiomed.2021.104721.

[56] R. Girshick, "Fast R-CNN," in *Proc. IEEE Int. Conf. Comput. Vis. (ICCV)*, Dec. 2015, pp. 1440–1448.

[57] S. Ren, K. He, R. Girshick, and J. Sun, "Faster R-CNN: Towards real-time object detection with region proposal networks," in *Proc. Adv. Neural Inf. Process. Syst.*, 2015. [Online]. Available: https://github.com/

[58] K. He, G. Gkioxari, P. Dollár, and R. Girshick, "Mask R-CNN," 2017, *arXiv:1703.06870*.

[59] B. Zhan, D. Li, X. Wu, J. Zhou, and Y. Wang, "Multi-modal MRI image synthesis via GAN with multi-scale gate mergence," *IEEE J. Biomed. Health Informat.*, vol. 26, no. 1, pp. 17–26, Jan. 2022, doi: 10.1109/JBHI.2021.3088866.

[60] T. Miyato and M. Koyama, "CGANs with projection discriminator," 2018, *arXiv:1802.05637*.

[61] W. He et al., "InSituNet: Deep image synthesis for parameter space exploration of ensemble simulations," *IEEE Trans. Vis. Comput. Graphics*, vol. 26, no. 1, pp. 23–33, Jan. 2020, doi: 10.1109/TVCG.2019.2934312.

[62] *ASHRAE Guideline 14-2002 Measurement of Energy and Demand Savings*, ANSI/ASHRAE Standard 8400, 2002, p. 170.

Bowoo Kim (Student Member, IEEE) received the M.S. degree from the Department of Convergence and Fusion System Engineering, Kyungpook National University (KNU), Sangju, South Korea, where he is currently pursuing the Ph.D. degree with the Department of Convergence and Fusion System Engineering.

His research interests include machine learning and deep learning, photovoltaic generation prediction, and energy storage systems.

Dongjun Suh (Member, IEEE) received the Ph.D. degree from the Department of Civil and Environmental Engineering, Korea Advanced Institute of Science and Technology (KAIST), Daejeon, South Korea, in 2014.

From 2015 to 2018, he was a Senior Researcher at the Korea Institute of Science and Technology Information (KISTI). From 2014 to 2015, he was a Research Assistant Professor at the KAIST Institute for Information Technology Convergence. He is currently an Associate Professor with the Department of Convergence and Fusion System Engineering, Kyungpook National University (KNU), South Korea. His research interests include big data and smart control systems, employing predictive analytics through machine learning and deep learning techniques.