

Research Article

Deep Learning Enhanced Solar Energy Forecasting with AI-Driven IoT

Hangxia Zhou ¹, Qian Liu ¹, Ke Yan ², and Yang Du ³

¹Key Laboratory of Electromagnetic Wave Information Technology and Metrology of Zhejiang Province, College of Information Engineering, China Jiliang University, Hangzhou, China 310018

²National University of Singapore, 4 Architecture Drive, Singapore 117566

³College of Science and Engineering, James Cook University, Cairns, QLD 4870, Australia

Correspondence should be addressed to Qian Liu; liuqian235@cjl.u.edu.cn

Received 14 April 2021; Accepted 7 June 2021; Published 19 June 2021

Academic Editor: Xiao Zhang

Copyright © 2021 Hangxia Zhou et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Short-term photovoltaic (PV) energy generation forecasting models are important, stabilizing the power integration between the PV and the smart grid for artificial intelligence- (AI-) driven internet of things (IoT) modeling of smart cities. With the recent development of AI and IoT technologies, it is possible for deep learning techniques to achieve more accurate energy generation forecasting results for the PV systems. Difficulties exist for the traditional PV energy generation forecasting method considering external feature variables, such as the seasonality. In this study, we propose a hybrid deep learning method that combines the clustering techniques, convolutional neural network (CNN), long short-term memory (LSTM), and attention mechanism with the wireless sensor network to overcome the existing difficulties of the PV energy generation forecasting problem. The overall proposed method is divided into three stages, namely, clustering, training, and forecasting. In the clustering stage, correlation analysis and self-organizing mapping are employed to select the highest relevant factors in historical data. In the training stage, a convolutional neural network, long short-term memory neural network, and attention mechanism are combined to construct a hybrid deep learning model to perform the forecasting task. In the testing stage, the most appropriate training model is selected based on the month of the testing data. The experimental results showed significantly higher prediction accuracy rates for all time intervals compared to existing methods, including traditional artificial neural networks, long short-term memory neural networks, and an algorithm combining long short-term memory neural network and attention mechanism.

1. Introduction

Photovoltaic power generation has the advantages of low carbon consumption, adaptive to various applications, and low installation and maintenance costs, which is known as a sustainable energy source [1]. Because of different weather conditions, PV panels often cannot stably output electrical power from solar energy. While integrating the PV power to the power grid, the grid is seriously influenced. The stability of the entire grid will be greatly reduced. With the wireless sensor network, it is possible for deep learning techniques to forecast solar energy generation and consequently stabilize the smart grid systems [2]. Therefore, artificial intelligence- (AI-) driven internet of things (IoT) technology becomes one of the key technologies of solving this problem [3, 4].

Along with the fast development of IoT technology, the extended deep learning methods nowadays are capable of performing short-term time series data forecasting, such as PV power generation, anomaly detection, and energy consumption forecasting, with considerable high forecasting accuracy [5–7].

In recent years, with the fast development of AI-driven IoT technology, the applications of deep learning technologies have been extended to various fields [8, 9], such as digital twinning [10], computer security [11], cyberphysical systems [12], transportation systems [13], and air quality forecasting [14]. Jiménez-Pérez and Mora-López [15] proposed a forecasting system simulating global solar irradiance forecasts every hour. The system can be separated into two phases. The first phase is clustering. The original dataset was divided

into groups by the k -means clustering algorithm. Each group represents data of different weather types. Machine learning techniques, such as decision trees, artificial neural networks, and support vector machines, are employed to perform forecasting in the second phase. Yang et al. [16] proposed a hybrid deep learning method based on weather types for PV power output forecasting with a timestep at 1 hour. The proposed method has three steps: classification, training, and prediction. The classification step involves a self-organizing map (SOM) [17] and Learning Vector Quantization (LVQ) [18]. In the training and prediction step, the fuzzy training method is used to select the most appropriate deep learning model for prediction. Han et al. [19] proposed an alternative multimodel PV power interval prediction method that considers the seasonal property of the PV power and absolute power deviation in the prediction process. Van-Deventer et al. [20] proposed a support vector machine model based on a genetic algorithm (GASVM). The GASVM model classifies the historical weather data using an SVM classifier initially, and later, it is optimized by the genetic algorithm using an ensemble technique. Malvoni et al. [21] proposed to use wavelet decomposition and principal component analysis to decompose meteorological data and treat it as predictive inputs. A time series forecasting method, named as GLSSVM (Group Least Square Support Vector Machine), which combines the Least Square Support Vector Machines (LS-SVM) and Group Method of Data Handling (GMDH), was applied to the measured weather data.

More recently, deep learning techniques, such as the convolutional neural network (CNN), long short-term memory (LSTM), and deep belief networks (DBNs), are widely applied to the PV system energy generation forecasting problem. Srivastava and Lessmann [22] proposed to adopt the original LSTM neural network for PV energy generation forecasting, and the proposed algorithm is hardly generalizable. Jian et al. [23] proposed an online sequential extreme learning machine with a forgetting mechanism model for PV energy generation prediction. The normalized data is input into a two-layer CNN for feature extraction and then merged into LSTM by fusion layer for PV power prediction. Zhou et al. [24] proposed a LSTM neural network combining the attention mechanism (AM) for PV energy generation forecasting. The AM is used to adaptively perceive feature information from the time series data. Chen et al. [25] proposed a cloud shadow model using computer vision techniques to forecast real cloud coverage nature for PV systems and enhanced the PV energy generation forecasting results consequently. Chang et al. [26] introduced a virtual inertia control based on PV load forecasting results, showing real-world applications of the PV energy generation forecasting techniques.

In this paper, we proposed a novel PV power forecasting framework combining clustering and deep learning technologies. The entire framework can be divided into three stages. In the first stage, according to the impact of the training data in the previous paragraph, we divide the training data into four clusters using the SOM algorithm, mimicking the four seasons. The reason for using the SOM algorithm as the experimental clustering algorithm is described in 2.2.2, and

the reason for simulating the data into four seasons for the experiment is described in 2.2.1. In the second stage, use AM, CNN, and LSTM to build the model. The third stage is the forecasting stage. According to the month of the testing data, the most suitable forecasting model is selected to predict the PV energy generation for the next time stamp. Experimental results show the superior performance of the proposed method over the existing works. Compared to the existing methods tackling on the same problem, the main contributions of the proposed method are summarized in the following points.

1.1. The Raw Data Is Processed by Clustering Techniques for More Accurate Forecasting Performance. Based on the data collected from the eastern region of China, the solar irradiance patterns vary between different seasons. The data is first clustered into four classes and trained using four different LSTM neural networks, respectively, to enhance the forecasting accuracy.

1.2. CNN and LSTM Are Combined for a More Sophisticated Deep Learning Framework. Both CNN and LSTM are popular deep learning techniques for forecasting problems. In this study, the two techniques are combined to produce better forecasting results for the PV energy generation forecasting problem.

1.3. The AM Technique Is Adopted to Further Enhance the Forecasting Performance. The AM is an important extension for the traditional LSTM neural network. It forces the LSTM neural networks to pay more attention to the features that are more relevant to the forecasting outputs. According to the experimental results, the AM technique greatly improves the forecasting results produced by LSTM neural networks.

2. Methodology

The dataset employed in this study was collected by a PV power station located in Shaoxing city in the eastern part of China. This solar energy generation dataset was collected from October 2014 to September 2018, in a time interval of 7.5 min. The data from 2014 to 2016 is taken as the training dataset, and the data from 2017 to 2018 is used for testing. Remote sensors were utilized to record the PV module temperature, the current, voltage, frequency, phases, and PV power every 7.5 minutes. Since the power station uses three-phase inverter equipment, the data contains the PV module temperature, three AC currents (alternating current 1, alternating current 2, and alternating current 3), three AC voltages (AC voltage 1, AC voltage 2, and AC voltage 3), two DC currents (direct current 1 and direct current 2), two DC voltages (DC voltage 1 and DC voltage 1), frequency, phases, and PV power. The frequency and phase do not change with the PV power, so the frequency and phase data are not used in the experiments. It is very difficult to obtain huge amount of historical weather data with a time interval of 7.5 min at the same location from the weather bureau. We do not add weather data to the experimental data and find a data preprocessing method to effectively improve the predictive ability of the model.

We calculated correlation coefficients between other factor data and PV power. The correlation of the multifactor combination on the PV power is shown in Table 1. A combined correlation test method for calculating the modified RV coefficient is introduced here. Modified RV coefficient (RV_{mod}) is a correlation analysis method based on matrix calculation. The equation for calculating the modified RV coefficient is

$$RV_{\text{mod}} = \frac{\text{tr}(JK)}{\sqrt{\text{tr}(J)^2 \text{tr}(K)^2}}, \quad (1)$$

where $J = M \cdot M^T - \text{diag}(M \cdot M^T)$; $K = N \cdot N^T - \text{diag}(N \cdot N^T)$; M represents the matrix of influencing factors, including alternating current 2 (AC2) and alternating current 3 (AC3); N represents PV power output matrix; $\text{diag}(\cdot)$ is a function that takes out the diagonal elements of the matrix; $\text{tr}(\cdot)$ is a function that takes the sum of the elements on the diagonal of the matrix; the range of RV_{mod} is $(-1, 1)$, when the RV_{mod} is closer to -1 or 1 ; and the correlation is higher between the factor and the power output.

For the coefficients AC2 and coefficients AC3 which are very close, we specially carry out experiments and compare the experimental results of the datasets of AC2 and AC3. The experimental results show that, although the coefficient difference is not large, the prediction accuracy of AC2 is significantly higher than AC3. In this study, we extract AC2 data as experiment data.

2.1. Clustering of the Original Data. In the clustering stage, the AC2 data from 2014 to 2016 is the training dataset and used for clustering. Considering the different numbers of days per month and the lack of PV data, we selected 15 days of complete data from each month and made each month AC2 data to a dataset as $[x_1, x_2, \dots, x_{15}]$; the dataset indicates that a month's AC2 data, x , contains 192 AC2 real-time data for one day. The clustering experiment uses two-year AC2 data. We merged the datasets with the same month label. The SOM algorithm is used to perform the clustering task. The raw data was clustered into 4 clusters.

The SOM model structure is shown in Figure 1. The learning steps of the SOM algorithm are as follows.

In the first step, the weight vector corresponding to each neuron in the competition layer was initialized, normalizing the current input mode vector X and the weight vector corresponding to the neuron. Secondly, the neuron node d with the smallest Euclidean distance between X_c (input vector) and w_j (connection weight vector) is selected as the winning neuron node. As a result, the weights were updated accordingly. Following Equation (3), the weight of the winning neuron node d and other nodes in its domain is updated, where $\eta(t)$ is the learning rate of the t step in the range of $(0, 1)$ and $h_{d\bar{d}j}(t)$ is the neighborhood function. $h_{d\bar{d}j}(t)$ generally uses a Gaussian function, as shown in Equation (4). The specific adjustment rules are as shown in Equations (5) and (6).

$$\|X_c - w_j\| = \min \{d_{jc}\}, \quad (2)$$

$$W_{ij}(t+1) = W_{ij}(t) + \eta(t)h_{d_j}(t)(X_i - W_{ij}(t)), \quad (3)$$

$$h_{d_j}(t) = \exp\left(-\frac{d_{d_j}^2}{2r^2(t)}\right), \quad (4)$$

$$r(t+1) = \text{INT}\left(\left(r(t) - 1\right) \times \left(1 - \frac{t}{T}\right)\right) + 1, \quad (5)$$

$$\eta(t+1) = \eta(t) - \frac{\eta(0)}{T}, \quad (6)$$

where the distance between the neuron d and the neuron j is $d_{d_j}^2$, $r(t)$ is the neighborhood radius, INT is the rounding function, and T is the learning frequency.

2.2. Deep Learning Models. In the training stage, the 4 clusters of data are inputted into a model for training, and 4 forecasting models are trained. The training process follows the training method of time series. We display the training data of each category as $[y_1, y_2, y_3, \dots, y_n]$, where y represents AC2 data and n represents the number of AC2 data contained in each category. We set timestep i , $[y_1, y_2, \dots, y_i]$, as a new training set, and y_{i+1} as the target value; then, we put the new training set into the training model, and the model parameters are adjusted during training so that the predicted output is constantly close to y_{i+1} . Next, $[y_2, y_3, \dots, y_{i+1}]$ is inputted into the training model and adjusts the model parameters so that the predicted output is continuously close to y_{i+2} , and so on, until the target value is y_n ; end the model training and get the forecasting model.

The training model structure of this experiment consists of two convolutional kernel layers of CNN, one layer of LSTM with attention mechanism, and a fully connected (FC) layer. The overall flowchart of the proposed framework is depicted in Figure 2.

This proposed deep learning model (CNN-ALSTM) is a hybrid deep learning model to extract features from the raw data and perform forecasting using the LSTM neural network. The CNN layer is employed to extract the useful features from the time series data, representing additional latent information of the data, which has the potential to improve the prediction accuracy. The experimental results show that the CNN layer contains one 16×1 convolution kernel layer and one 32×1 convolution kernel convolution layer that optimizes the prediction performance.

The feature vector obtained from the second layer of CNN was inputted into the LSTM layer for prediction. Each element of the feature vector has correspondence to one of the 32 units in the LSTM layer. The attention mechanism puts higher weight to the feature quantities that are significantly related to the current output. Last, the output vector of the attention mechanism is processed by a FC layer using the unfolding operation. The predicted value of the AC2 at the next moment is output.

LSTM is very suitable for prediction experiments of time series data. Existing works show superior forecasting performance combining CNN and LSTM for various applications

TABLE 1: Correlation coefficient of PV data factors.

PV data factor name	Correlation coefficient
DC voltage 1 (V)	0.19
DC voltage 2 (V)	0.20
Direct current 1 (A)	0.13
Direct current 2 (A)	0.13
AC voltage 1 (V)	0.10
AC voltage 2 (V)	0.11
AC voltage 3 (V)	0.14
Alternating current 1 (A)	0.28
Alternating current 2 (A)	0.992
Alternating current 3 (A)	0.998
Alternating current 2 (A) and alternating current 3 (A) combination	0.73
PV module temperature	0.16

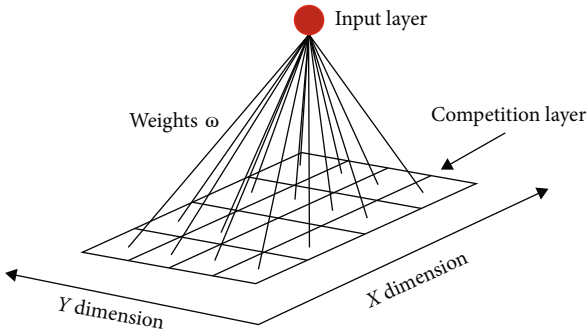


FIGURE 1: Structure of the SOM model.

[23, 27–29]. CNN helps LSTM to better extract the characteristics of experimental data. The attention mechanism is a process of assigning weights. Using the attention mechanism, more accurate weightage values are assigned to the LSTM output vector to improve the prediction capability of the model.

The predictive power is obtained by the inverse normalization according to Equation (7).

$$\text{Pr}_p = \text{Pr}_{Iac2} * (P_{\max} - P_{\min}) + P_{\min}, \quad (7)$$

where Pr_p is the predicted value of power and Pr_{Iac2} is the predicted value of AC2.

2.2.1. Long Short-Term Memory Neural Network. The LSTM model structure is shown in Figure 3.

The appearance of LSTM cell structure effectively resolves the gradient explosion/vanishing problems. There are four important elements in the flowchart of the LSTM model: cell status, input gate, forget gate, and output gate. The input, forget, and output gates are used to control the update, maintenance, and deletion of information contained

in cell status. The forward computation process can be denoted as

$$\begin{aligned} f_t &= \sigma(W_f \cdot [h_{t-1}, x_t] + b_f), \\ i_t &= \sigma(W_i \cdot [h_{t-1}, x_t] + b_i), \\ O_t &= \sigma(W_o \cdot [h_{t-1}, x_t] + b_o), \\ \tilde{C}_t &= \tanh(W_C \cdot [h_{t-1}, x_t] + b_C), \\ C_t &= f_t \cdot C_{t-1} + i_t \cdot \tilde{C}_t, \\ h_t &= O_t \cdot \tanh(C_t), \end{aligned} \quad (8)$$

where W_f , W_i , and W_o are the weight matrix of the forgetting gate, input gate, and output gate, respectively; b_f , b_i , and b_o are the offset items of the forgetting gate, the input gate, and the output gate, respectively; σ is the sigmoid activation function; \tanh is the hyperbolic tangent activation function.

2.2.2. The Attention Mechanism. The attention mechanism is a brain signal processing mechanism peculiar to human vision. Human vision quickly scans the global image to obtain the target area that needs attention and ignores other areas of useless information. The attention mechanism algorithm has been successfully implemented and applied to model training [26] and other related fields.

The model proposed in this paper uses the LSTM hidden layer output vector $H = \{h_1, h_2, \dots, h_t\}$ as the input of the attention mechanism, and the attention mechanism will find the attention weight α_i of h_i , which can be calculated as shown in

$$\begin{aligned} e_i &= \tanh(W_h h_i + b_h), \\ \alpha_i &= \frac{\exp(e_i)}{\sum_{i=1}^t \exp(e_i)}, \end{aligned} \quad (9)$$

where W_h is the weight matrix of h_i and b_h is the bias.

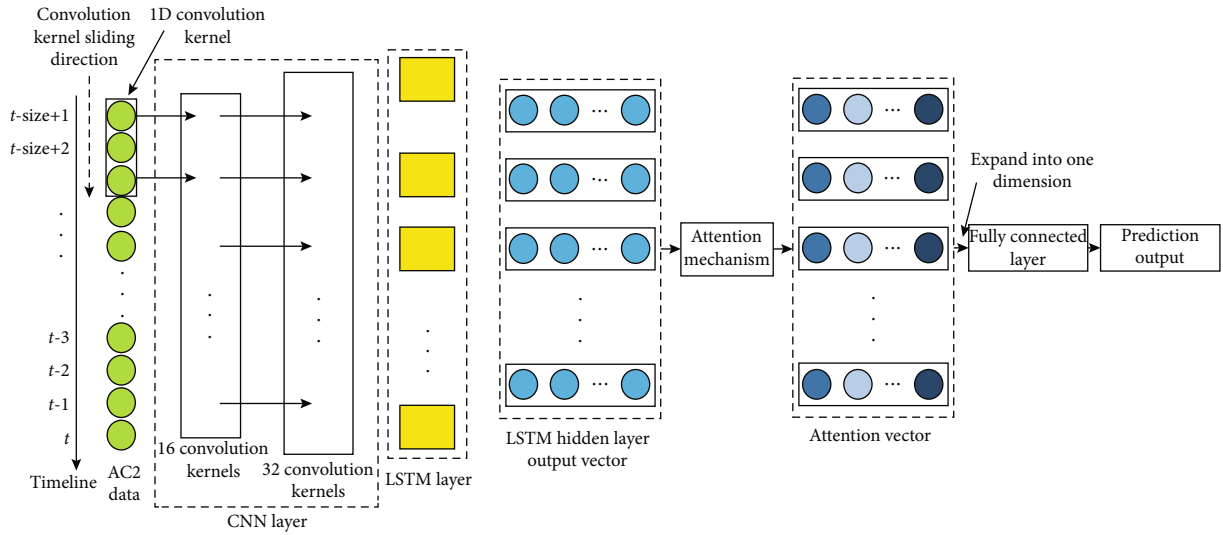


FIGURE 2: The overall flowchart of the proposed algorithm.

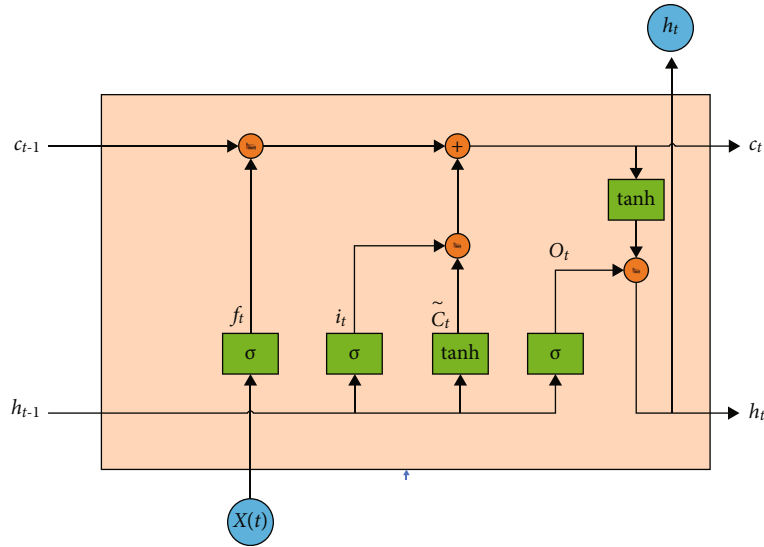


FIGURE 3: Structure of LSTM.

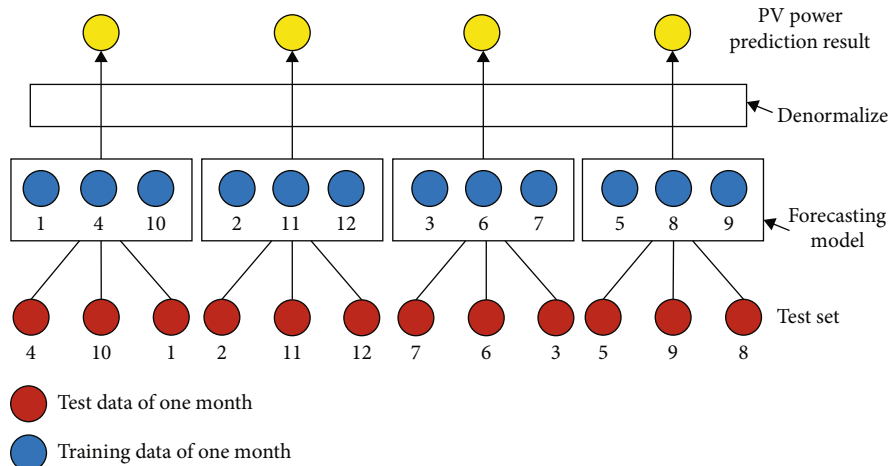


FIGURE 4: Structure of prediction stage. The numbers under test data and training data represent months.

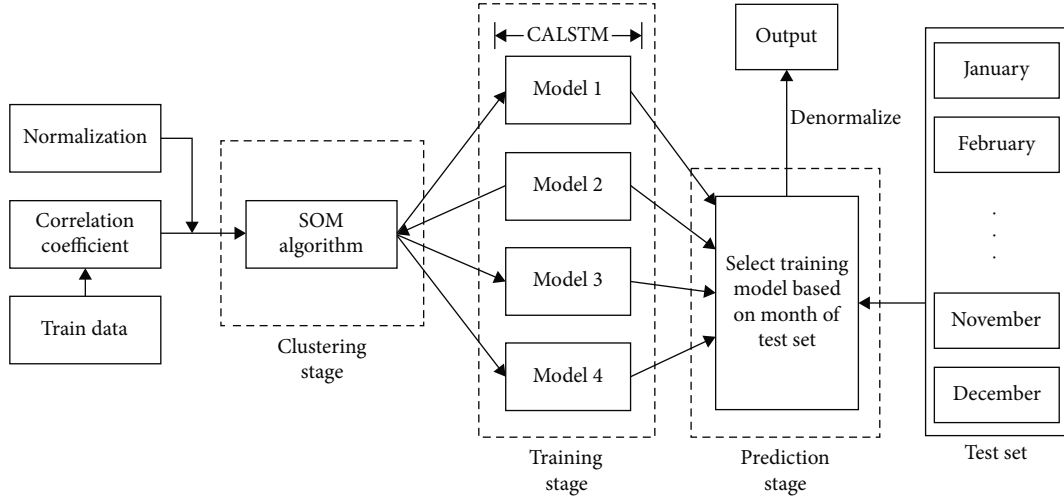


FIGURE 5: Structure of the proposed hybrid model.

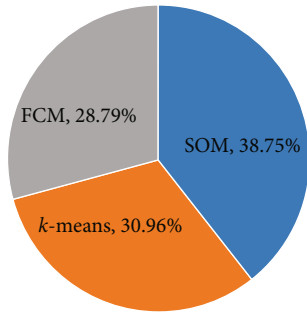


FIGURE 6: Accuracy of different clustering algorithms in the training phase.

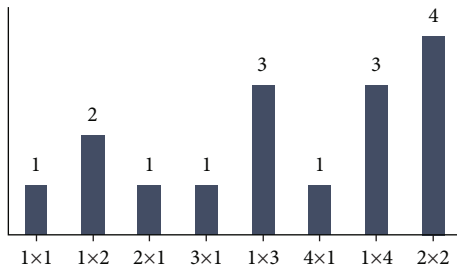


FIGURE 7: Horizontally, the X-axis represents the size of the competition layer, and vertically, the Y-axis represents the number of output neurons.

The values of W_h and b_h vary during the ALSTM training process. The range of e_i is $(-1,1)$. The attention vector $H' = \{h_1', h_2', \dots, h_t'\}$ can be obtained by multiplying attention weight α_i and h_i :

$$h_i' = \alpha_i \cdot h_i. \quad (10)$$

The attention mechanism is implemented as a custom layer where the parameters are optimized using RMSProp backpropagation [33, 34].

2.3. The Overall Forecasting Framework. There are four deep learning models trained in the proposed framework (Figure 4). In the testing stage, an appropriate forecasting model is selected based on the grouping result of the testing data sample. The proposed hybrid deep learning model structure is shown in Figure 5. As a necessary preprocessing step, we used correlation coefficient to determine inputs for prediction. The weather data was employed by the SOM algorithm to cluster inputs into four categories. Each category is trained by the proposed hybrid deep learning framework. In particular, the CNN is used to extract the time series features of AC2 data; the LSTM neural network further extracts high latitude features in the data. The attention mechanism is used to assign different attention weights to the output elements of the LSTM hidden layer. In the testing phase, according to the month of the test set, we chose the appropriate model to predict the testing result (Figures 4 and 5).

The process of data classification method determination and hyperparameter determination is also described, including the reason for simulating data in four seasons, the selection of clustering algorithm, the settings of SOM competition layer, and the selection of SOM classified result.

2.3.1. The Selection of Clustering Algorithm. The original idea of the clustering is to map the raw data into four different groups, representing the four seasons. These four groups of data are trained and tested using the proposed CNN-ALSTM structure separately. Different clustering algorithms are tested in the section. In the training phase, we analyze clustering algorithms, including k -means, FCM, and SOM (Figure 6), for clustering the training dataset into four clusters. Based on the experimental results on the training data, SOM is more accurate than most of the above-mentioned traditional clustering algorithms (Figure 6) according to the original season labels of the raw data.

The size of the SOM competition layer has been evaluated through a series of experiments as shown in Figure 7. The X-axis represents the size of the competition layer, and

TABLE 2: Predicting MAPE results for one year of data.

		2017.10	2017.11	2017.12	2018.1	2018.2	2018.3	2018.4	2018.5	2018.6	2018.7	2018.8	2018.9	Overall
7.5 min	MLP	22.15	20.07	23.05	25.62	18.10	24.97	26.00	23.19	28.39	25.08	22.18	20.03	23.11
	LSTM	23.87	21.42	20.96	28.76	16.63	28.96	20.99	31.19	40.02	20.21	22.18	21.73	24.61
	ALSTM	21.64	22.44	21.44	28.16	23.52	24.55	18.43	17.82	32.79	25.46	29.05	19.78	23.32
	Prop.	19.95	16.76	20.79	25.01	16.90	19.29	15.84	16.83	17.21	22.20	19.88	18.69	18.82
15 min	MLP	30.19	24.41	26.42	28.58	20.63	37.38	47.40	40.80	53.70	38.39	31.30	27.10	34.17
	LSTM	23.42	21.47	23.74	27.51	17.74	23.62	22.63	21.21	28.95	28.72	19.00	18.64	22.75
	ALSTM	20.30	23.08	22.40	28.38	20.40	29.96	29.69	20.53	68.64	47.60	24.06	22.49	28.82
	Prop.	22.69	20.32	26.51	28.04	21.055	28.17	28.61	25.12	18.23	25.44	23.87	20.89	23.79
30 min	MLP	33.09	35.15	36.98	37.68	28.49	33.44	28.85	28.92	30.17	35.92	32.78	30.65	32.57
	LSTM	37.75	41.62	38.86	38.42	41.45	36.07	36.34	30.06	37.90	39.01	32.59	40.96	37.39
	ALSTM	25.67	29.86	37.76	42.77	33.89	30.41	57.01	28.64	30.09	30.71	37.50	35.79	34.18
	Prop.	32.62	29.59	29.32	32.88	23.50	33.38	34.87	31.34	45.62	30.19	31.84	32.62	32.39

TABLE 3: Predicting RMSE results for one year of data.

		2017.10	2017.11	2017.12	2018.1	2018.2	2018.3	2018.4	2018.5	2018.6	2018.7	2018.8	2018.9	Overall
7.5 min	MLP	1.39	1.58	2.09	2.16	1.23	1.07	0.63	0.72	0.69	1.25	1.10	1.45	1.32
	LSTM	1.45	1.63	2.16	2.20	1.29	1.10	0.63	0.81	0.78	1.30	1.15	1.50	1.37
	ALSTM	1.60	1.75	2.16	2.36	1.33	1.10	0.69	0.81	0.71	1.34	1.17	1.63	1.42
	Prop.	1.44	1.52	2.08	2.32	1.23	1.10	0.63	0.65	0.70	1.36	1.24	1.40	1.30
15 min	MLP	1.42	1.60	2.12	2.19	1.22	1.12	0.77	0.82	0.82	1.29	1.16	1.47	1.37
	LSTM	1.63	1.94	2.51	2.44	1.45	1.16	0.69	0.83	0.73	1.56	1.17	1.61	1.53
	ALSTM	1.54	1.90	2.38	2.49	1.74	1.19	0.71	0.77	0.96	1.61	1.25	1.69	1.55
	Prop.	1.57	1.78	1.47	1.86	1.47	1.21	0.75	0.90	0.87	1.45	1.29	1.62	1.40
30 min	MLP	2.23	2.74	2.91	3.29	1.84	1.56	0.94	1.22	1.10	1.96	1.83	2.03	2.02
	LSTM	3.10	3.95	4.13	3.72	3.52	2.01	1.46	1.64	1.19	2.32	2.52	3.51	2.89
	ALSTM	1.80	2.95	3.78	3.91	2.66	1.19	0.94	0.99	0.76	2.06	1.46	3.18	2.22
	Prop.	2.26	2.29	2.95	3.12	1.71	1.92	1.00	0.94	0.96	1.71	1.69	2.71	2.04

TABLE 4: Predicting MAE results for one year of data.

		2017.10	2017.11	2017.12	2018.1	2018.2	2018.3	2018.4	2018.5	2018.6	2018.7	2018.8	2018.9	Overall
7.5 min	MLP	0.829	0.88	1.15	1.36	0.70	0.61	0.39	0.42	0.39	0.72	0.67	0.83	0.72
	LSTM	0.89	0.95	1.21	1.47	0.74	0.67	0.35	0.52	0.53	0.74	0.69	0.91	0.78
	ALSTM	1.049	1.12	1.22	1.56	0.84	0.62	0.39	0.45	0.44	0.83	0.77	1.05	0.83
	Prop.	0.85	0.84	1.25	1.55	0.78	0.57	0.32	0.40	0.33	0.81	0.80	0.88	0.70
15 min	MLP	0.91	0.92	1.17	1.38	0.69	0.73	0.58	0.58	0.61	0.84	0.75	0.91	0.82
	LSTM	1.05	1.25	1.61	1.61	0.97	0.67	0.41	0.49	0.42	0.97	0.70	1.02	0.90
	ALSTM	0.95	1.24	1.42	1.65	1.25	0.76	0.47	0.44	0.78	1.16	0.82	1.12	0.95
	Prop.	1.022	1.20	1.53	1.55	1.16	0.74	0.49	0.54	0.42	0.81	0.83	0.91	0.85
30 min	MLP	1.58	2.02	2.09	2.29	1.35	1.03	0.55	0.73	0.57	1.26	1.30	1.47	1.32
	LSTM	2.18	2.98	3.00	2.73	2.78	1.27	0.86	0.97	0.67	1.51	1.73	2.60	1.90
	ALSTM	1.19	2.24	2.99	3.16	2.23	0.76	0.77	0.66	0.46	1.39	1.12	2.43	1.47
	Prop.	1.64	1.66	2.07	2.29	1.30	1.39	0.67	0.64	0.68	1.10	1.23	2.13	1.38

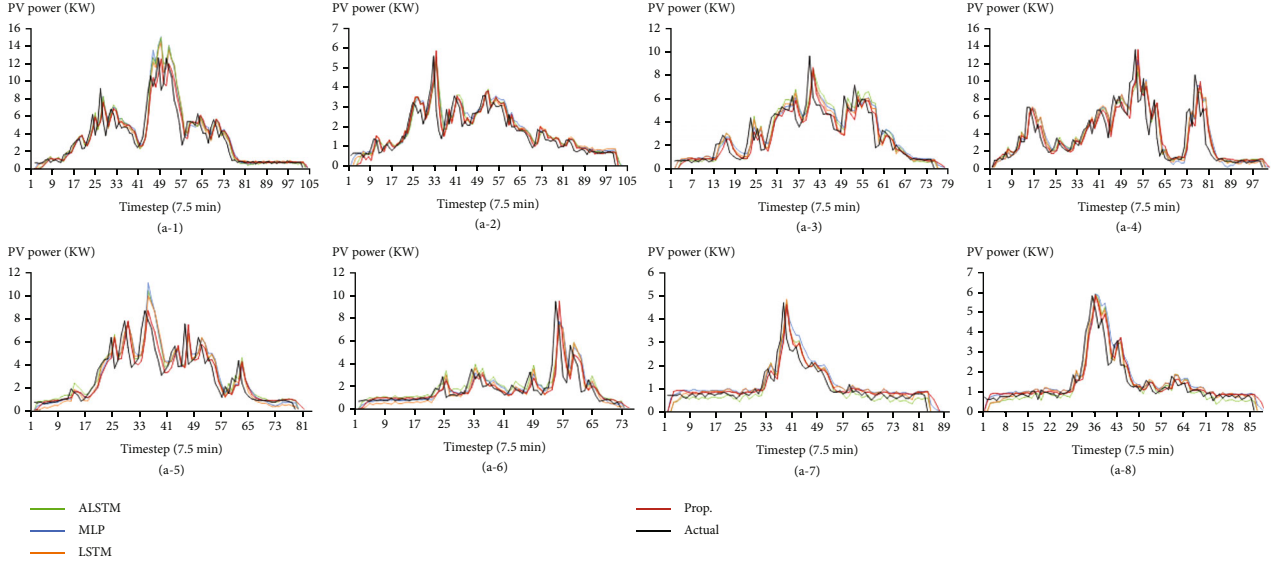


FIGURE 8: PV power forecast result in 7.5 min ahead. (a-1) and (a-2) are the January 5th and January 6th, (a-3) and (a-4) are the March 1st and March 3rd, (a-5) and (a-6) are the June 6th and June 8th, and (a-7) and (a-8) are the October 3rd and October 4th.

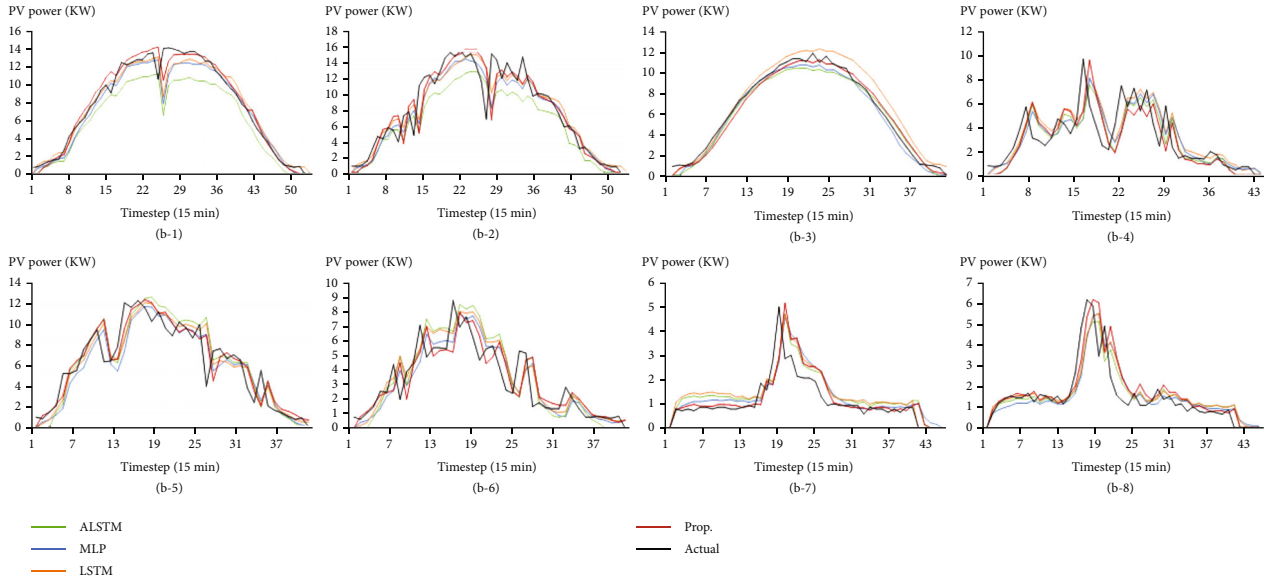


FIGURE 9: PV power forecast result in 15 min ahead. (b-1) and (b-2) are the January 14th and January 15th, (b-3) and (b-4) are the May 1st and May 2nd, (b-5) and (b-6) are the July 1st and July 2nd, and (b-7) and (b-8) are the October 3rd and October 4th.

the Y-axis represents the number of output neurons. The result shows that the most suitable competition layer for SOM is 2×2 .

The clustering results of the SOM algorithm have a potential relationship with the prediction effect of the forecasting model. The number of iterations of the SOM algorithm is set between 500 and 1000, and the number of times is debugged in hundred. The batch size is set to 15 or 30.

3. Experimental Process and Results

In this study, the average absolute percentage error MAPE, the mean absolute error MAE, and the root mean square

error RMSE are used to evaluate the prediction ability of the model. The detailed equations of three error metrics are formulated in

$$\begin{aligned} \text{MAPE} &= \frac{1}{n} \sum_{i=1}^n \frac{|x_{\text{model},i} - x_{\text{actual},i}|}{x_{\text{actual},i}}, \\ \text{RMSE} &= \sqrt{\frac{1}{n} \sum_{i=1}^n (x_{\text{model},i} - x_{\text{actual},i})^2}, \\ \text{MAE} &= \frac{1}{n} \sum_{i=1}^n |x_{\text{model},i} - x_{\text{actual},i}|. \end{aligned} \quad (11)$$

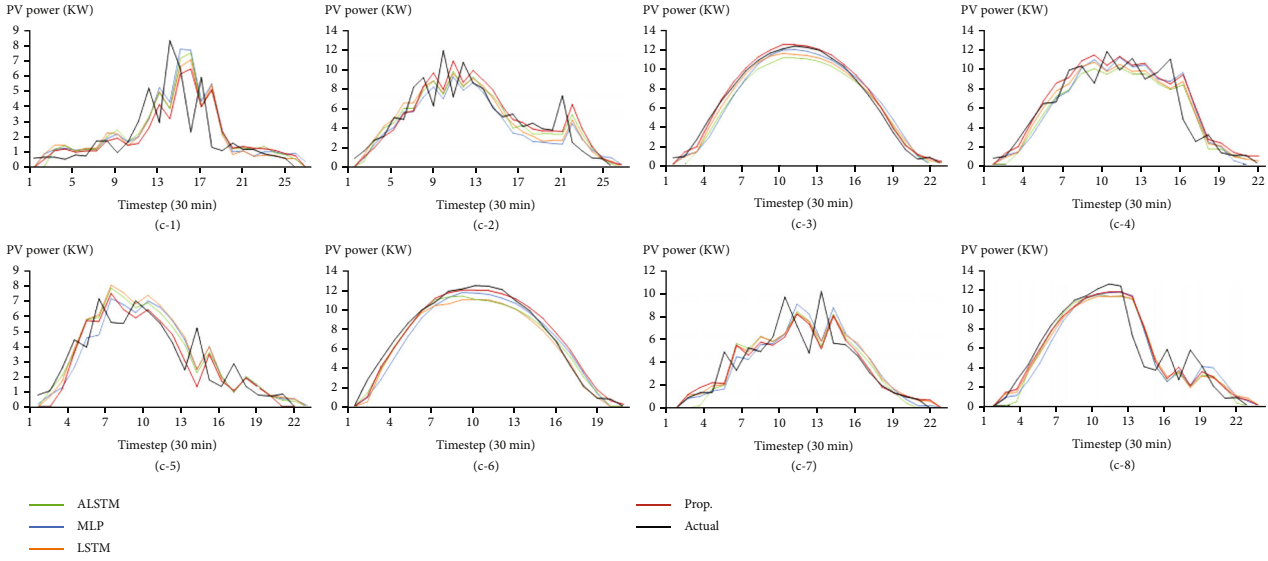


FIGURE 10: PV power forecast result in 30 min ahead. (c-1) and (c-2) are the January 9th and January 11th, (c-3) and (c-4) are the May 30th and May 31th, (c-5) and (c-6) are the August 2nd and July 9th, and (c-7) and (c-8) are the October 6th and October 12th.

The experiment compares the proposed model, our previous work, that combines the LSTM and attention, together with the MLP and LSTM model alone. The most relevant factor obtained in this paper is AC2 as input to the training model. The resulting MAPE results are shown in Table 2. The resulting RMSE results are shown in Table 3. The resulting MAE results are shown in Table 4.

In the case of less training data, using the clustering phase and prediction phase mentioned in this paper will reduce the prediction accuracy. When the time interval is 30 min, the amount of data is reduced to a quarter of the data volume with a time interval of 7.5 min. From the comparison of different time intervals, when the interval is 30 min, the prediction accuracy drops. On the contrary, in the case of more training data, the prediction accuracy can be effectively improved.

The experiment randomly selects two days of the month as a display of the predicted results. As shown in Figures 8–10, according to the law of PV power generation, the power generation in the early evening to the early morning of the next day is 0, which is not shown in Figures 8–10.

In Figures 8–10, (a-1) and (a-2), (b-1) and (b-2), and (c-1) and (c-2) are forecasting results of two days in winter; (a-3) and (a-4), (b-3) and (b-4), and (c-3) and (c-4) are forecasting results of two days in spring; (a-5) and (a-6), (b-5) and (b-6), and (c-5) and (c-6) are forecasting results of two days in summer; (a-7) and (a-8), (b-7) and (b-8), and (c-7) and (c-8) are forecasting results for two days in autumn. For winter, in Figure 10, deviations of the predicted power curve are found for 30 min ahead predictions. Satisfactory prediction performances in the winter season for 7.5 min ahead and 15 min ahead forecasts are achieved according to Figures 8 and 9.

Through the experimental clustering, training, and prediction stage, the accuracy of the forecasting model has a significant outperformance over the compared models when the amount of training data is sufficient. The clustering stage

divides the original PV data into four different clusters. And each clustered data is trained by a hybrid deep learning framework combining CNN and LSTM. The sophisticated deep learning framework obviously enhances the forecasting performance over singular methods, such as MLP and LSTM. The AM technique already shows a great influence of the forecasting results in the experiments with ALSTM. The performance improvement is further enlarged for the proposed methods. According to Tables 2–4, the proposed model has a good performance on the RMSE indicator. Compared with the RMSE indicator, the MAE indicator reflects the actual situation of the prediction value error. The proposed model has the better performance than other models in the MAE indicator for intervals within 7.5 min. It is noted that the proposed model has significantly higher accuracy in the prediction of intervals within 7.5 min and has no obviously high accuracy compared with the other individual model in the prediction of intervals within 15 min and 30 min.

4. Conclusion and Discussion

Photovoltaic power generation prediction is of great significance for maintaining grid security and coordinating resource utilization. In the era of big data, it is possible for AI-driven IoT technology to perform accurate solar energy generation forecasting based on historical solar energy data [24, 30–32]. This paper proposes a hybrid deep learning method based on weather categories. Unlike traditional models, this experiment includes correlation analysis and clustering calculation of data, which effectively improves the generalization ability of the model and improves the prediction accuracy. The training algorithm employs the CNN algorithm, which can extract the data features more effectively and get more potential information in the data. Secondly, the attention mechanism is applied to the LSTM model, focusing on the extracted important features. In the prediction stage, the month of test set is used to determine

the forecasting model, which improves the accuracy of the prediction. This forecasting model is superior to the traditional algorithm model in predicting performance, and it also shows outstanding results in predicting other data types.

The experimental results shown in Section 3 show that the proposed method outperforms the traditional time series data prediction methods, such as MLP, LSTM, and ALSTM for PV system energy generation forecasting. Three evaluation metrics are used to show the superior performance, including RMSE, MAPE, and MAE. According to the results collected in Tables 2–4, the proposed method achieves higher forecasting accuracy for short-term and very short-term forecasting, e.g., the 7.5 min advanced forecasting. For longer-term forecasting, such as time intervals of 15 and 30 min, the forecasting performance advantage decreases.

One of the future works of this study is to extend the existing work for more generalized datasets, i.e., achieving acceptable forecasting results for longer-term forecasting of the PV system energy generation. Another future work direction is to apply the proposed framework towards a broader range of time series data applications in other fields, such as air quality forecasting [14, 35] and energy consumption forecasting [36].

Data Availability

Access to data is restricted. The solar irradiance data used in this study is collected by a local company in Shaoxing, China. The data is confidential and is only available for the company's internal usage.

Conflicts of Interest

The authors declare no conflict of interest.

Authors' Contributions

Hangxia Zhou and Yang Du performed the conceptualization and validation; Hangxia Zhou contributed to the methodology and investigation; Ke Yan helped in the formal analysis and project administration; Yang Du participated in the data curation and wrote and prepared the original draft; Qian Liu wrote, reviewed, and edited the manuscript and helped in the supervision. All authors have read and agreed to the published version of the manuscript.

Acknowledgments

This research was funded by the Public Welfare Research Project of Zhejiang Province, China, grant number LGF18F020017, and in part by the National Natural Science Foundation of China under grant number 61850410531.

References

- [1] P. Nema, R. K. Nema, and S. Rangnekar, "A current and future state of art development of hybrid energy system using wind and PV-solar: a review," *Renewable and Sustainable Energy Reviews*, vol. 13, no. 8, pp. 2096–2103, 2009.
- [2] E. Lorenz, T. Scheidsteiger, J. Hurka, D. Heinemann, and C. Kurz, "Regional PV power prediction for improved grid integration," *Progress in Photovoltaics: Research and Applications*, vol. 19, no. 7, pp. 757–771, 2011.
- [3] X. Zheng and Z. Cai, "Privacy-preserved data sharing towards multiple parties in industrial IoTs," *IEEE Journal on Selected Areas in Communications*, vol. 38, no. 5, pp. 968–979, 2020.
- [4] Z. Cai and Z. He, "Trading private range counting over big IoT data," in *In 2019 IEEE 39th international conference on distributed computing systems (ICDCS)*, pp. 144–153, Dallas, TX, USA, 2019.
- [5] X. Zhou, W. Liang, S. Shimizu, J. Ma, and Q. Jin, "Siamese neural network based few-shot learning for anomaly detection in industrial cyber-physical systems," *IEEE Transactions on Industrial Informatics*, no. 8, pp. 5790–5798, 2021.
- [6] W. Kong, Z. Y. Dong, Y. Jia, D. J. Hill, Y. Xu, and Y. Zhang, "Short-term residential load forecasting based on LSTM recurrent neural network," *IEEE Transactions on Smart Grid*, vol. 10, no. 1, pp. 841–851, 2019.
- [7] X. Zhou, Y. Hu, W. Liang, J. Ma, and Q. Jin, "Variational LSTM enhanced anomaly detection for industrial big data," *IEEE Transactions on Industrial Informatics*, vol. 17, no. 5, pp. 3469–3477, 2020.
- [8] M. Qiu, S. Y. Kung, and Q. Yang, "Editorial: IEEE transactions on sustainable computing, special issue on smart data and deep learning in sustainable computing," *IEEE Transactions on Sustainable Computing*, vol. 4, no. 1, pp. 1–3, 2019.
- [9] K. Yan, L. Liu, Y. Xiang, and Q. Jin, "Guest editorial: AI and machine learning solution cyber intelligence technologies: new methodologies and applications," *IEEE Transactions on Industrial Informatics*, vol. 16, no. 10, pp. 6626–6631, 2020.
- [10] X. Zhou, X. Xu, W. Liang et al., "Intelligent small object detection based on digital twinning for smart manufacturing in industrial CPS," *IEEE Transactions on Industrial Informatics*, pp. 1–1, 2021.
- [11] Z. Cai, Z. He, X. Guan, and Y. Li, "Collective data-sanitization for preventing sensitive information inference attacks in social networks," *IEEE Transactions on Dependable and Secure Computing*, vol. 15, no. 4, pp. 577–590, 2016.
- [12] Z. Cai and X. Zheng, "A private and efficient mechanism for data uploading in smart cyber-physical systems," *IEEE Transactions on Network Science and Engineering*, vol. 7, no. 2, pp. 766–775, 2018.
- [13] Z. Cai, X. Zheng, and J. Yu, "A differential-private framework for urban traffic flows estimation via taxi companies," *IEEE Transactions on Industrial Informatics*, vol. 15, no. 12, pp. 6492–6499, 2019.
- [14] N. Jin, Y. Zeng, K. Yan, and Z. Ji, "Multivariate air quality forecasting with nested LSTM neural network," *IEEE Transactions on Industrial Informatics*, pp. 1–1, 2021.
- [15] P. F. Jiménez-Pérez and L. Mora-López, "Modeling and forecasting hourly global solar radiation using clustering and classification techniques," *Solar Energy*, vol. 135, pp. 682–691, 2016.
- [16] H. T. Yang, C. M. Huang, Y. C. Huang, and Y. S. Pai, "A weather-based hybrid method for 1-day ahead hourly forecasting of PV power output," *IEEE Transactions on Sustainable Energy*, vol. 5, no. 3, pp. 917–926, 2014.
- [17] J. Vesanto and E. Alhoniemi, "Clustering of the self-organizing map," *IEEE Transactions on Neural Networks*, vol. 11, no. 3, pp. 586–600, 2000.

- [18] S. Bashyal and G. K. Venayagamoorthy, "Recognition of facial expressions using Gabor wavelets and learning vector quantization," *Engineering Applications of Artificial Intelligence*, vol. 21, no. 7, pp. 1056–1064, 2008.
- [19] Y. Han, N. Wang, M. Ma, H. Zhou, S. Dai, and H. Zhu, "A PV power interval forecasting based on seasonal model and non-parametric estimation algorithm," *Solar Energy*, vol. 184, pp. 515–526, 2019.
- [20] W. VanDeventer, E. Jamei, G. S. Thirunavukkarasu et al., "Short-term PV power forecasting using hybrid GASVM technique," *Renewable Energy*, vol. 140, pp. 367–379, 2019.
- [21] M. Malvoni, M. G. De Giorgi, and P. M. Congedo, "Forecasting of PV power generation using weather input data-preprocessing techniques," *Energy Procedia*, vol. 126, pp. 651–658, 2017.
- [22] S. Srivastava and S. Lessmann, "A comparative study of LSTM neural networks in forecasting day-ahead global horizontal irradiance with satellite data," *Solar Energy*, vol. 162, pp. 232–247, 2018.
- [23] X. Jian, H. Gu, and R. Wang, "A short-term photovoltaic power prediction model based on dual-channel CNN and LSTM," *Electric Power Science and Engineering*, vol. 35, no. 5, pp. 7–11, 2019.
- [24] H. Zhou, Y. Zhang, L. Yang, Q. Liu, K. Yan, and Y. Du, "Short-term photovoltaic power forecasting based on long short term memory neural network and attention mechanism," *IEEE Access*, vol. 7, pp. 78063–78074, 2019.
- [25] X. Chen, Y. Du, E. Lim, H. Wen, K. Yan, and J. Kirtley, "Power ramp-rates of utility-scale PV systems under passing clouds: module-level emulation with cloud shadow modeling," *Applied Energy*, vol. 268, p. 114980, 2020.
- [26] J. Chang, Y. Du, X. Chen, E. G. Lim, and K. Yan, "Forecasting based virtual inertia control of PV systems for islanded microgrid," in *In 2019 29th Australasian universities power engineering conference (AUPEC)*, pp. 1–6, Nadi, Fiji, 2019.
- [27] X. Zhou, Y. Li, and W. Liang, "CNN-RNN based intelligent recommendation for online medical pre-diagnosis support," *IEEE/ACM Transactions on Computational Biology and Bioinformatics*, vol. 18, no. 3, pp. 912–921, 2021.
- [28] N. Jin, J. Wu, X. Ma, K. Yan, and Y. Mo, "Multi-task learning model based on multi-scale CNN and LSTM for sentiment classification," *IEEE Access*, vol. 8, pp. 77060–77072, 2020.
- [29] K. Yan, X. Wang, Y. Du, N. Jin, H. Huang, and H. Zhou, "Multi-step short-term power consumption forecasting with a hybrid deep learning strategy," *Energies*, vol. 11, no. 11, p. 3089, 2018.
- [30] D. L. Fernandes, R. Siqueira-Batista, A. P. Gomes et al., "Investigation of the visual attention role in clinical bioethics decision-making using machine learning algorithms," *Procedia Computer Science*, vol. 108, pp. 1165–1174, 2017.
- [31] J. Campbell, H. B. Amor, H. M. Ang, and G. Fainekos, "Traffic light status detection using movement patterns of vehicles," in *2016 IEEE 19th International Conference on Intelligent Transportation Systems (ITSC)*, pp. 283–288, Rio de Janeiro, Brazil, 2016.
- [32] X. Zhou, W. Liang, I. Kevin, K. Wang, R. Huang, and Q. Jin, "Academic influence aware and multidimensional network analysis for research collaboration navigation based on scholarly big data," *IEEE Transactions on Emerging Topics in Computing*, vol. 9, no. 1, pp. 246–257, 2018.
- [33] Z. Cai, Z. Xiong, H. Xu, P. Wang, W. Li, and Y. Pan, "Generative adversarial networks: a survey towards private and secure applications," *ACM Computing Surveys*, vol. 54, no. 2, pp. 1–38, 2021.
- [34] K. Yan, A. Chong, and Y. Mo, "Generative adversarial network for fault detection diagnosis of chillers," *Building and Environment*, vol. 172, p. 106698, 2020.
- [35] Z. Zhang, Y. Zeng, and K. Yan, "A hybrid deep learning technology for PM 2.5 air quality forecasting," *Environmental Science and Pollution Research*, pp. 1–14, 2021.
- [36] K. Yan, W. Li, Z. Ji, M. Qi, and Y. Du, "A hybrid lstm neural network for energy consumption forecasting of individual households," *IEEE Access*, vol. 7, pp. 157633–157642, 2019.