Solar Power Generation Forecasting Using a Hybrid LSTM-Linear Model with Multi-Head Attention

Hyeonseok Jin[†] · David J. Richter[††] · MD Ilias Bappi[††] · Kyungbaek Kim[†††]

ABSTRACT

Due to the negative impact on the environment, the demand for solar energy, which can effectively replace fossil fuels, is increasing. In order to operate solar energy efficiently, deep learning has been used recently to predict future power generation, but it is still a challenge to provide accurate predictions because the power generation greatly depends on external factors such as weather and time of day. In this paper, we propose a hybrid model that combines LSTM and Linear models using Multi-Head Attention to provide more accurate predictions. The proposed model can improve prediction accuracy and learn richer time series features because each Linear model can capture trends and irregularities in time series features. We conducted extensive experiments, and the results showed that the proposed model outperformed other prediction models by about 10%, and the ablation study confirmed that combining three models based on Multi-Head Attention is the most effective way to consider trends and irregularities.

Keywords : Solar Power Prediction, Power Generation, Prediction, Time series, Deep Learning

멀티 헤드 어텐션이 적용된 하이브리드 LSTM-Linear 모델을 이용한 태양광 발전량 예측

진 현 석[†] · David J. Richter[††] · MD Ilias Bappi[††] · 김 경 백[†††]

요 약

환경에 미치는 부정적인 영향으로 인해 화석 연료를 효과적으로 대체할 수 있는 태양광 에너지에 대한 수요가 증가하고 있다. 태양광 에너지를 효율적으로 운영하기 위해 최근 딥러닝을 사용하여 미래의 발전량을 예측하고 있으나, 발전량은 날씨, 시간대와 같은 외부 요인에 크게 의존하기 때문에 정확한 예측을 제공하는 것은 여전히 도전적인 과제이다. 본 논문에서는 보다 정확한 발전량 예측을 제공하기 위해 멀티 헤드 어텐션을 사용하여 LSTM과 선형 모델을 결합한 하이브리드 모델을 제안한다. 제안된 모델은 각 선형 모델이 시계열 특징의 추세와 불규칙성을 포착할 수 있기 때문에 예측 정확도를 개선하고 더 풍부한 시계열 특징을 학습할 수 있다. 광범위한 실험을 수행한 결과, 제안된 모델은 다른 예측 모델보다 약 10% 정도 성능이 우수하였으며, 절제 연구를 통해 멀티 헤드 어텐션에 기반하여 세 가지 모델을 결합하는 것이 추세와 불규칙성을 고려하는 데 가장 효과적인 방법임을 확인하였다.

키워드 : 태양광 발전량 예측, 발전량, 예측, 시계열, 딥러닝

1. Introduction

Climate change and global warming are critical global challenges that demand immediate and innovative solutions. Traditional energy sources such as coal and oil significantly contribute to environmental degradation, emphasizing the urgent need for sustainable energy alternatives. Renewable energy sources, including solar, wind, and water, have emerged as essential components

* Corresponding Author : Kyungbaek Kim(kyungbaekkim@jnu.ac.kr)

for a sustainable future. Solar energy, in particular, has gained widespread adoption due to its potential for clean and efficient power generation. However, its reliance on external factors such as solar radiation, time of day, and weather conditions introduces variability and unpredictability in power generation, posing challenges for effective integration into the power grid [1]. The International Energy Agency (IEA) estimates that photovoltaic (PV) systems will achieve a peak power capacity of 1200 GW by the end of 2023 [2]. These systems, commonly installed in open areas, on rooftops, and in solar power plants, offer a promising path to reducing dependence on traditional energy sources. Despite their advantages, the intermittent nature of solar power generation disrupts regular operations, complicating energy distribution and grid stability. For instance, fluctuating solar radiation and environmental conditions can lead to energy deficits or surpluses, necessitating advanced forecasting systems to optimize energy storage and usage [3]. Reliable forecasting of solar power generation is essential for ensuring grid stability and maximizing the benefits of renewable energy. Unlike traditional power plants that offer predictable energy outputs, solar energy systems depend on dynamic environmental conditions, making accurate forecasting crucial for energy management. Effective prediction of energy generation enables better planning for storage during surplus periods and ensures energy availability during low-generation intervals [4].

Recent advancements in deep learning have demonstrated significant potential in addressing the challenges of solar power forecasting. Time series forecasting techniques such as Long Short-Term Memory (LSTM) networks, linear models, and attention mechanisms have proven effective in capturing complex patterns and irregularities in solar power data. These approaches surpass traditional methods like physical and statistical models, which often struggle with non-linearities and variability in data [5]. To address these challenges, this study proposes a hybrid deep learning model combining LSTM and linear models with a multi-head attention mechanism. The hybrid model leverages the strengths of each component: LSTM networks capture short-term dependencies and irregular patterns, linear models identify long-term trends, and the multi-head attention mechanism enhances feature extraction by focusing on critical temporal relationships. By integrating these

techniques, the proposed model achieves superior forecasting performance, offering a reliable solution for managing the complexities of solar power generation [6]. Extensive experiments conducted in this study demonstrate that the hybrid model outperforms existing forecasting approaches by approximately 10% in prediction accuracy. Additionally, ablation studies validate the effectiveness of integrating multi-head attention with LSTM and linear models, highlighting its ability to capture trends and irregularities in time series data. These findings underscore the potential of the proposed hybrid approach to advance solar energy forecasting and enhance its integration into the energy grid [7-8].

The main contributions of this research work are as follows:

- We propose a novel hybrid model that integrates LSTM networks and linear models with a multi-head attention mechanism to enhance solar power generation forecasting.
- The hybrid model effectively captures both short-term irregularities and long-term trends in time series data, addressing the inherent variability of solar power generation.
- By leveraging the multi-head attention mechanism, the model learns richer temporal features and focuses on critical patterns in historical data for improved prediction accuracy.
- Extensive experiments demonstrate that the proposed hybrid model outperforms existing state-of-the-art prediction methods by approximately 10% in forecasting accuracy.
- Ablation studies validate the effectiveness of combining LSTM, linear models, and attention mechanisms, highlighting the contributions of each component to the overall performance.

The rest of this article is organized as follows: Section 2 reviews the related works, followed by a comprehensive description of background knowledge and the proposed methodology in Sections 3 and 4. Section 5 presents the experimental results and analysis, while Section 6 providing key insights and discusses future directions. Finally, the last section, Section 7 concludes the study.

2. Related Works

With the unpredictability of green energy, e.g. solar power, when compared with energy sources such as nuclear, oil, and coal for example, the need for methods that can forecast the power output becomes clear. As such, this problem has received a lot of attention in recent years, some of which will be introduced in this section.

Prema et al. [9] provide a lot of insight into the field of solar power generation prediction in their review paper, with a focus on datasets, models used, as well as common performance metrics. The paper reviews not only solar power, but also wind power applications. The paper identifies irradiance as the feature with the greatest impact on prediction performance. The authors state that deep learning technologies have shown great results recently, but point out that short-term prediction (up to 1 or 2 days) is recommended. Lastly the study found that the Mean Absolute Error (MAE), Mean Absolute Percentage Error (MAPE), and Symmetric Mean Average Percentage Error (SMAPE) are the most commonly used metrics.

Shin et al. [10] performed solar power generation forecasting using deep learning techniques on a dataset that was collected in Mokpo in Korea. The data was collected over 2 years between 2013 and 2015. The dataset included solar irradiance and solar radiation and the models trained on said data were based on ANFIS, DNN, Recurrent Neural Network (RNN) and LSTM. The authors' results showcase that the RNN model managed to perform better than the other models used in this work based on the metrics used (RMSE, MAE, BIAS, and COR).

Park et al. [11] analyze the performance of BiLSTM and SHAP-based XAI techniques in solar power forecasting. This work also uses power data obtained in Korea, from different plants across the country. The integration of XAI allows researchers to better understand the trained models and, therefore, also allows them to update the model based on that, overcoming the common "black-box" problem that Deep Learning (DL) models often pose.

Khan et al. [12] incorporated attention mechanisms into their DL model. The dual stream network utilizes both CNN and GRU methods, allowing the model to capture both spatial and temporal relations. This is then combined with a self attention block to further learn relations present in the data. Three datasets were used during the experiments and MSE, RMSE, and MAE indicate that the proposed model achieves the highest performance under the author's conditions.

Polasek et al. [13] also present a hybrid model to solve the problem of solar power forecasting. In this case a U-Net structure in combination with the ResNet CNN model were used. Pre-processing techniques were applied to solve issues present in uncertain datasets. The dataset, which was published as part of this paper, contains 16 power plants' data over a year's time. After evaluation, the authors present their hybrid model as the best performing model on the dataset.

Shamshirband et al. [14] survey the field in their paper. The paper also reviews wind energy. Based on the data provided in the paper, it seems as if wind power has garnered more research interest. The authors stress the importance of enough and high quality data to train reliable models. Common metrics, methods and workflows are presented in the paper, as are metrics. According to the authors, RMSE is recommended as a metric.

Dairi et al. [15] make use of variational Auto Encoder (VAE) based deep learning models. This model was compared to RNN, LSTM, Bidirectional LSTM, Convolutional LSTM network (ConvLSTM), GRU, stacked autoencoder (SAE), restricted Boltzmann machine (RBM), logistic regression and support vector regression (SVM). The VAE based methods, as proposed by the authors, managed to outperform all other models in their experiments. Also noteworthy is the pre-processing applied to the dataset, as the authors used normalization to range all input values between 0 and 1.

Chang et al. [16] propose TESDL. TESDL stands for Traditional Encoder Single Deep Learning. TESDL was compared to SVM, CNN and LSTM-CNN in the paper and has proven to perform better. This paper too states the difference between long and short-term forecasting, with short-term forecasting being much more feasible and as such much more accurate. TESDL managed to generate good results until about 50 steps ahead, after which it started to fall off.

Zazoum et al. [17] did not use DL methods, but rather studied ML for the field. Different variations of SVM and gaussian process regression (GPR) were included in this work, with Matern 5/2 GPR achieving the best performing model. Cubic SVM, on the other hand, was the worst performing model. The results obtained in this work showcase that ML models, which are model more light-weight than computationally much heavier DL

methods, can also generate well performing predictors.

Alharkan et al. [18] present a dual stream CNN-LSTM model with self-attention. The CNN model excels in learning spatial relations, the LSTM excels in learning temporal relations and the self-attention block is meant to learn relations within the dataset. The DKASC Alice Spring dataset was used in this work to train the model and CNN, LSTM, GRU, CNNLSTM, CNNGRU for comparison. The comparative study showed that the proposed model (DSCLANet) managed the highest performance.

Alkandari et al. [19] built a Ensemble model for solar power generation prediction that uses DL methods in combination with statistical methods to rely on the strength of both approaches. The Machine Learning and Statistical Hybrid Model (MLSHM) uses LSTM, GRU and Auto LSTM as the DL methods of choice. Results show that the ensemble methods managed to outperform the DL methods without ensembleing.

Shouman et al. [20] use and compare neural networks, SVM, Random Forest (RF), and gradient boosting to forecast solar power generation. To pre-process the data, the authors used normalization as well as feature selection methods to identify and use only the most useful features, while discarding features that might negatively impact the models performance. Here MSE and RMSE were used for evaluation. In their use-case comparative study gradient boosting managed to outperform the other models.

Lee et al. [21] used a dataset recorded in Taiwan that they then normalized for pre-processing. The data was collected for about a year and 3 months between late 2018 and early 2020. The dataset was then tested on RNN, LSTM and GRU models, both uni and bi-directional for each of them. The results have shown that in this scenario the Bidirectional-GRU model managed to perform better than any of the other models tested.

Barrera et al. [22] test a number of different DL models to predict future solar power generation. For this dense layers were compared with stratified layers, models with 1 hidden layer were compared to models with 2, and global radiation was compared to decomposed radiation. The model stratified layers, 2 hidden layers and decomposed radiation managed to perform better than the other models tested.

Kim et al. [23] showcase a multi-step approach to power prediction where they first train a model to predict the solar irradiance, to then use said predicted irradiance value to train and predict a solar power generation model

with it. Random Forest Regression (RFR), Support Vector Regression (SVR), and gradient boosting were tested for this task. The authors explain that unexpected weather events can cause high inaccuracy in a model's prediction, even if the models otherwise achieves very high accuracy and even models that are otherwise weather aware. RFR was the best performing model trained in this work.

3. Background

In this section, we introduce the concepts and features of LSTM [24], Temporal Convolutional Network (TCN) [25], and Linear models [26], which are widely used in the field of time series prediction.

3.1 LSTM

RNN have difficulty capturing long-term time series features due to gradient vanishing when the length of input data becomes too long. LSTM were proposed to solve this problem and effectively solve it by utilizing the Cell and Gate structure. LSTM consist of a Cell state that updates and transmits information, an Input gate that stores new information, a Forget gate that removes unimportant information, and an Output gate that determines the information to be output. The Cell state C_{t-1} is updated by passing the information h_{t-1} of the previous time point and the input value x_t of the current time point through the Forget gate and the Input gate, and the updated Cell state C_t and the Hidden state h_t are passed through the Output gate to generate the output value of the current time point.

3.2 TCN

LSTM overcomes the limitations of existing RNN and shows innovative performance in the time series field, but
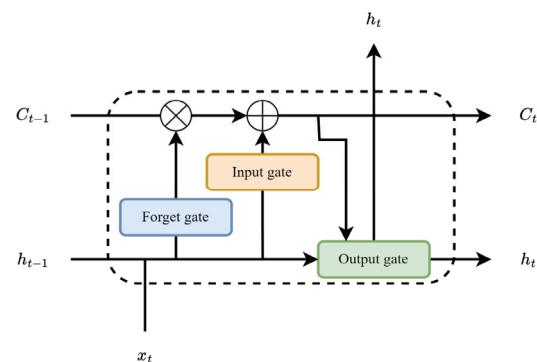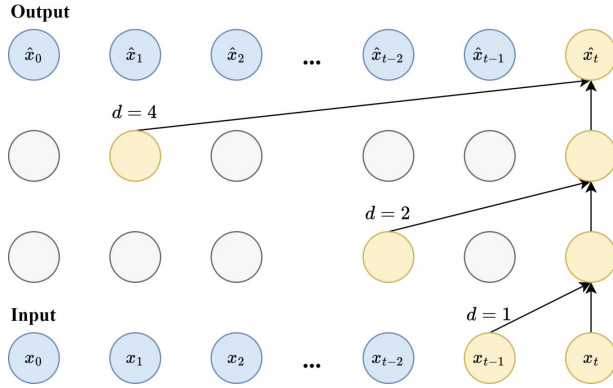

Fig. 1. LSTM Structure
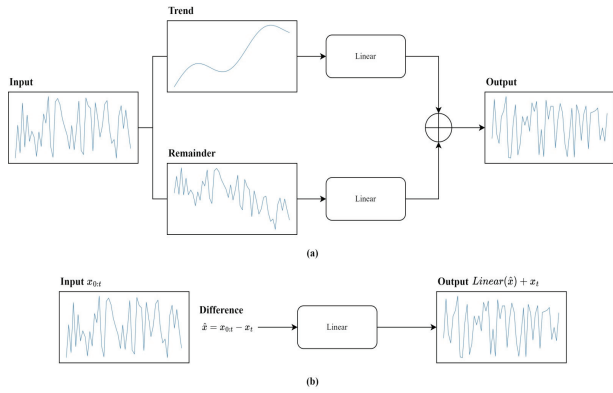
Fig. 2. TCN Structure

Fig. 3. Linear models structure, (a): DLinear, (b): NLinear

Fig. 4. Correlation analysis between power generation and weather data

it has the disadvantage of high computational cost due to its complex structure and taking a long time due to its sequential nature, which requires waiting until the inference of the previous time point is completed to perform the inference of the next time point. TCN performs parallel operations based on Convolutional Neural Networks (CNNs) and overcomes these shortcomings. In addition, it can capture long-term time series features by adopting Causal padding that uses only data up to the current time point and Dilation structure that gradually increases the interval of the filter, and it enables flexible expansion by increasing the filter size or interval. We will compare our proposed model to TCN.

3.3 Linear Models

These models were proposed to question the validity of Transformer [27] based time series prediction models relatively recently, and effectively learn time series features using a simple linear layer consisting of only Fully-Connected Layers (FC). DLinear decomposes the entire input value into a way to return the average value
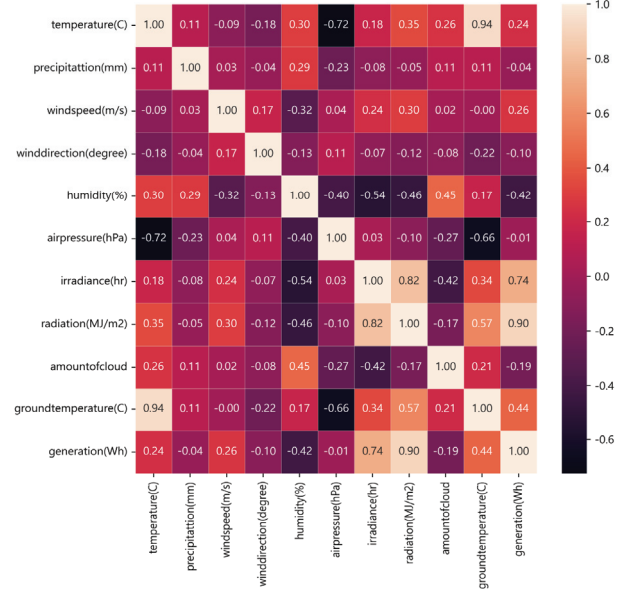
for a specific period and learns each of them to effectively predict data with a trend. NLinear operates by learning the difference by subtracting the last step from the input value and then adding it back to restore it, and effectively predicts irregular data.

4. Methods

In order to make more accurate solar power generation predictions, we need to consider various external variables that are closely related to solar power generation [28], and we need to be able to capture rapid changes in data [9]. In this paper, we perform predictions using solar power generation data and weather data. We also propose a new hybrid model for accurate solar power generation predictions by merging them.

4.1 Data Collection and Preprocessing

In this paper, we used solar power generation data [29] and weather data [30] in the Ulsan area. Solar power generation data is provided by the Korea Public Data Portal, and includes power generation data collected every hour from April 1, 2022 to June 30, 2024. Weather data is provided by the Korea Meteorological Administration, and includes temperature, precipitation, wind speed, wind direction, humidity, airpressure, irradiance, radiation, amount of cloud, and ground temperature during the
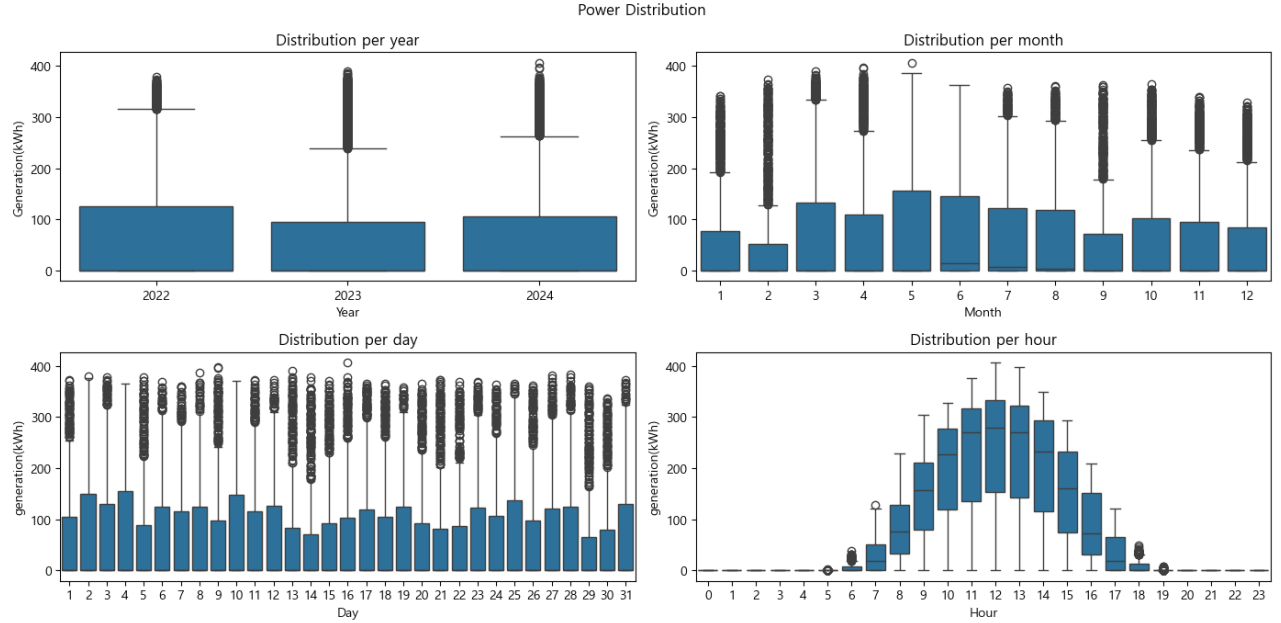
Power Distribution

Fig. 5. Distribution of power generation by time interval

same period.

First, we measured the correlation between power generation and weather variables to remove unnecessary variables in Fig. 4. radiation, irradiance, surface temperature, humidity, etc. showed high correlations, but precipitation and air pressure showed low correlations and were removed.

Next, we performed the task of imputing the missing values in the data. As shown in Fig. 5, the distribution of power generation is greatly affected by the sun, so the distribution varies the most by hour. Therefore, we impute the missing values by the median of each hour to maintain the distribution of the original data.

Finally, we reflect the continuity of wind direction and time data. For wind direction, it has a scalar value from 0 to 360, and if used as is, 0 and 360 mean the same direction, but there is a big difference. Time also has a scalar value from 0 to 24, and yesterday 23:00 and today 1:00 represent adjacent dates and times, but there is a big difference. We use (1), and (2) to express the scalar value as a continuous coordinate vector and reflect the continuity.

$$
\begin{aligned}
Wind_x &= \cos(W_S * (W_D * \pi/180)) \\
Wind_y &= \sin(W_S * (W_D * \pi/180))
\end{aligned}
\tag{1}
$$

$$
\begin{aligned}
Day_x &= \cos(T_S * (2\pi/SoD)) \\
Day_y &= \sin(T_S * (2\pi/SoD))
\end{aligned}
\tag{2}
$$

W_S and W_D represent wind speed and wind direction,

respectively, and T_S is a timestamp representing the time in seconds since 00:00 (UTC) on January 1, 1970. Seconds of the Day (SoD) is a value representing the time elapsed from midnight to the present time, expressing a day in seconds.

4.2 Hybrid Model Structure

The structure of the hybrid model proposed in this paper is as shown in Fig. 6. In order to accurately perform the challenging task of power generation prediction, we learn sequential features as well as extract and reflect trends and irregular features. The input data $x_{0:t} \in \mathbb{R}^{B \times T \times N}$, which consists of N weather data and power generation for T hours with batch size B, is learned and features are extracted using three backbone models: LSTM, DLinear, and NLinear. DLinear learns the trend of the input data and extracts features reflecting it, and NLinear learns differences to learn and extract irregular time series features. LSTM learns sequential features and generates an overall prediction value. Each prediction can be expressed as follows (3):

$$
\begin{aligned}
O_N &= NLinear(x_{0:t}) \\
O_D &= DLinear(x_{0:t}) \\
O_L &= LSTM(x_{0:t})
\end{aligned}
\tag{3}
$$

In addition, to effectively integrate the output of each backbone model, the Multi-Head Attention (MHA)
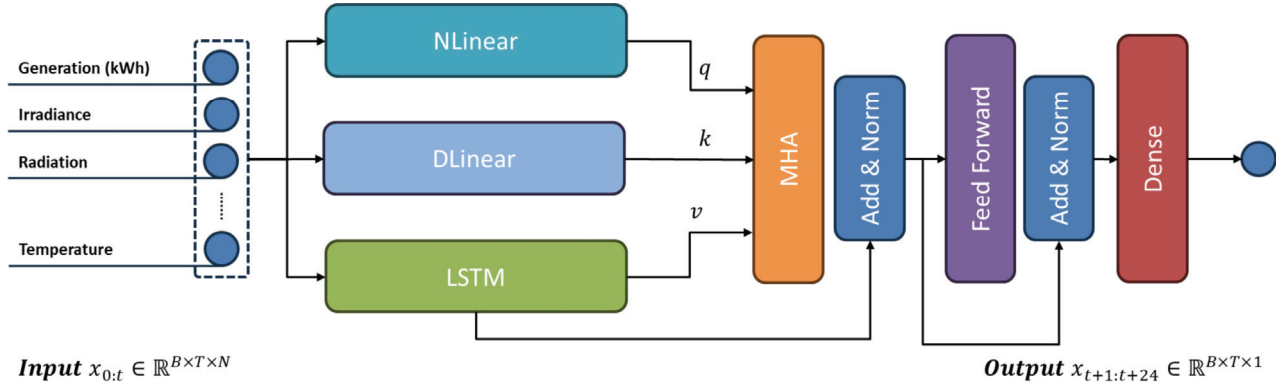
Fig. 6. Hybrid model structure

structure is used to distribute it to k parallel heads, each of which performs self-attention tasks. Each head computes a dimension of size d_k, and MHA can be expressed as follows (4):

$$d_k = d_{model}/k$$

$$MHA(Q,K,V) = softmax(\frac{QK}{\sqrt{d_k}}V^T) \quad (4)$$

The outputs O_N and O_D obtained from NLinear and DLinear generate an attention matrix based on the trend and irregular features of the input data, and are multiplied by the output O_L of LSTM to suppress unnecessary features and allow the model to focus on important features.

After that, the attention-reflected features are refined through the Feed Forward Network (FFN) based on Multi Layer Perceptron (MLP), and the last Dense layer compresses the features with the size of N into a single power feature. The final output of the model $x_{t+1:t+24} \in \mathbb{R}^{B \times T \times 1}$ can be expressed as follows (5):

$$O_{Att} = LayerNorm(O_L + MHA(O_N, O_D, O_L))$$

$$O_{FFN} = LayerNorm(O_{Att} + FFN(O_{Att})) \quad (5)$$

$$x_{t+1:t+24} = Dense(O_{FFT})$$

The proposed hybrid model can reflect the trend and irregular features captured from linear models in the prediction of LSTM, thereby richly reflecting the features of time series data. In addition, it suppresses unnecessary features by using self-attention technique, and allows the model to focus on important features, enabling robust and accurate prediction.

5. Experiments

In order to demonstrate the superiority of the hybrid model proposed in this paper, we conducted comparative experiments with other time series forecasting models. In addition, to ensure the robustness of the experimental results, we conducted a ablation study on the prediction length and the combining method.

5.1 Implementation Details

The experiments were conducted on Python 3.9.13, Tensorflow 2.10.1 and Keras 2.10.0 with Intel(R) Core(TM) i7-10700 CPU @ 2.90GHz, RTX 3070 environment in Windows 10 WSL. The batch size is set to 256, and training is performed with the AdamW optimizer with a learning rate of 0.0005 for 100 iterations and the MSE loss function.

The dataset is configured to predict the future for 24 hours based on 24-hour input data using the sliding window technique [31], and the value range is normalized to 0 to 1 through Min-Max normalization. Of the total 19,680 data, 17,520 data for 2 years are used for training, and the remaining 2,160 data are divided in half and used for validation and testing, respectively.

5.2 Metrics

Evaluation is performed with actual value range by performing denormalization on the predicted values. In addition to Root Mean Squared Error (RMSE) and MAE, mMAPE [32] and R^2 are also used as metrics to provide more intuitive indicators.

RMSE and MAE represent the difference between the predicted value and the actual value, respectively, and along with mMAPE, the lower the value, the better the

prediction quality. mMAPE represents the degree of error in the predicted value as a percentage between 0 and 100. R^2 represents the maximum explanatory power of the model, and the higher the value, the better the explanatory power. Each metric for the actual value y and the predicted value $\hat{y}$ can be expressed as (6):

$$RMSE = \frac{1}{N}\sqrt{\sum_{i=1}^{N}(y_i - \hat{y}_i)^2}$$

$$MAE = \frac{1}{N}\sum_{i=1}^{N}\left|y_i - \hat{y}_i\right|$$

$$mMAPE = \begin{cases} \dfrac{100}{N}\sum_{i=1}^{N}\left|y_i - \hat{y}_i\right|, \text{if } |y|_{max} < 1 \\ \dfrac{100}{N}\sum_{i=1}^{N}\dfrac{\left|y_i - \hat{y}_i\right|}{|y|_{max}}, else \end{cases} \quad (6)$$

$$R^2 = 1 - \frac{\sum_{i=1}^{N}(y_i - \hat{y}_i)^2}{\sum_{i=1}^{N}(y_i - \overline{y})^2}$$

5.3 Experimental Results

The quantitative comparison results with other models measured using test data are shown in Table 1. The comparison results show that the hybrid model proposed in this paper achieved the best results in all four metrics, and showed about 10% improved prediction quality compared to the single-structure model, proving the superiority of the proposed structure.

In addition, as shown in the Fig. 7, the hybrid model's predicted values and actual values were compared for arbitrary test data, and it was confirmed that it could generate predictions similar to the actual values.

Table 1. Quantitative comparison results

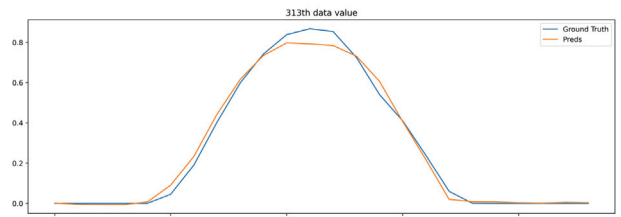| Models	RMSE↓	MAE↓	mMAPE↓	R^2↑
LSTM [24]	53.32	28.64	7.05	0.78
TCN [25]	57.76	33.43	8.22	0.74
DLinear [26]	52.53	27.89	6.86	0.79
NLinear [26]	52.39	28.73	7.07	0.79
Hybrid model (Proposed)	50.68	26.12	6.43	0.80

Fig. 7. Prediction results of the hybrid model

5.4 Ablation Study

Additionally, we conduct ablation studies to test the performance of different concatenation methods for different prediction lengths. Table 2 shows the performance changes by concatenation methods and prediction lengths. The common method of concatenating each output does not significantly improve performance because it does not extract valid features well, and it also increases complexity by doubling the number of dimensions. On the other hand, when concatenating based on self-attention techniques, it shows higher performance than a single model, proving that it is useful to reflect both trends and irregularities using the outputs of three models.

Table 2. Ablation study on predict length and merging methods
(NL: NLinear, DL: DLinear, MHA: Multi-Head Attention)

Methods	Predict Length (hours)											
	1			6			12			24		
	RMSE	MAE	mMAPE	RMSE	MAE	mMAPE	RMSE	MAE	mMAPE	RMSE	MAE	mMAPE
LSTM [24]	30.30	17.80	4.38	43.11	24.11	5.93	49.25	27.66	6.80	53.32	28.64	7.05
NLinear [26]	29.42	17.88	4.40	42.10	23.44	5.76	49.47	27.26	6.70	52.39	28.73	7.07
Concatenate (NL, LSTM)	31.55	17.92	4.41	44.90	24.65	6.06	50.16	26.97	6.63	53.16	28.51	7.01
Concatenate (DL, NL, LSTM)	33.19	17.38	4.22	46.27	24.57	6.04	50.99	27.42	6.74	57.28	31.10	7.65
MHA (NL, LSTM)	26.99	15.64	3.64	39.39	21.24	5.22	47.08	25.46	6.26	50.94	27.52	6.77
MHA (DL, NL, LSTM) (Proposed)	25.68	13.48	3.31	39.26	20.84	5.12	46.13	24.71	6.08	50.68	26.12	6.43

6. Future Work

Future work could be conducted by extending the dataset to also include weather data and power generation data from power plants in other regions, both within Korea and maybe even beyond. The model could be extended to also accept positional information and with that spatially aware DL blocks, so that a single model could be trained to handle all the different locations, therefore training a model that could potently even, to some extend, generalize in a way where it can predict the power for new plants without additional training. Another possible future direction would be the attempt to predict even further into the future, a task that is not trivial by any means. One could, for example, consider utilizing real live weather forecast data to provide the model with more relevant and more reliable weather data for future timesteps, so that the results can be based on those.

7. Conclusion

In this study, we tackled the challenges of forecasting solar power generation, which is critical for the efficient use of renewable energy. We proposed a hybrid model that combines LSTM and linear models with a multi-head attention mechanism. This approach effectively captures both short-term variations and long-term trends in time series data, addressing the inherent unpredictability of renewable energy sources. Our experimental results show that the proposed model outperforms existing prediction models by approximately 10%, achieving an RMSE of 50.68, MAE of 26.12, mMAPE of 6.43, and an R^2 value of 0.80. These metrics confirm the model's ability to deliver more accurate and reliable solar power forecasts compared to other methods. By improving prediction accuracy, this hybrid approach can enhance the integration of renewable energy into power grids, enabling better energy management and storage planning.

References

[1] D. Gielen, F. Boshell, D. Saygin, M. D. Bazilian, N. Wagner, and R. Gorini, "The role of renewable energy in the global energy transformation," *Energy strategy reviews*, Vol.24, pp.38-50, 2019.

[2] I. U. Khalil, A. U. Haq, and N. ul Islam, "A deep learning-based transformer model for photovoltaic fault forecasting and classification," *Electric Power Systems Research*, Vol.228, pp.110063, 2024.

[3] International Energy Agency (IEA), World energy investment 2020 [Internet], https://www.iea.org/reports/world-energy-investment-2020.

[4] M. Talaat, T. Said, M. A. Essa, and A. Hatata, "Integrated MFFNN-MVO approach for PV solar power forecasting considering thermal effects and environmental conditions," *International Journal of Electrical Power & Energy Systems*, Vol.135, pp.107570, 2022.

[5] J. A. Qadourah, "Energy and economic potential for photovoltaic systems installed on the rooftop of apartment buildings in Jordan," *Results in Engineering*, Vol.16, pp.100642, 2022.

[6] I. U. Khalil et al., "Comparative analysis of photovoltaic faults and performance evaluation of its detection techniques," *IEEE Access*, Vol.8, pp.26676-26700, 2020.

[7] V. R. Vemuri and R. D. Rogers, "Artificial neural networks-forecasting time series," in *IEEE Computer Society Press*, 1994.

[8] R. I. Mukhamediev et al., "Review of artificial intelligence and machine learning technologies: classification, restrictions, opportunities and challenges," *Mathematics*, Vol.10, No.15, p.2552, 2022.

[9] V. Prema, M. S. Bhaskar, D. Almakhles, N. Gowtham, and K. U. Rao, "Critical review of data, models and performance metrics for wind and solar power forecast," *IEEE Access*, Vol.10, pp.667-688, 2021.

[10] D. Shin, E. Ha, T. Kim, and C. Kim, "Short-term photovoltaic power generation predicting by input/output structure of weather forecast using deep learning," *Soft Computing*, Vol.25, pp.771-783, 2021.

[11] S. Park, S. Jung, J. Moon, and E. Hwang, "Explainable photovoltaic power forecasting scheme using BiLSTM," *KIPS Transactions on Software and Data Engineering*, Vol.11, No.8, pp.339-346, 2022.

[12] Z. A. Khan, T. Hussain, and S. W. Baik, "Dual stream network with attention mechanism for photovoltaic power forecasting," *Applied Energy*, Vol.338, pp.120916, 2023.

[13] T. Polasek and M. Čadík, "Predicting photovoltaic power production using high-uncertainty weather forecasts," *Applied Energy*, Vol.339, pp.120989, 2023.

[14] S. Shamshirband, T. Rabczuk, and K. W. Chau, "A survey of deep learning techniques: Application in wind and solar energy resources," *IEEE Access*, Vol.7, pp.164650-164666, 2019.

[15] A. Dairi, F. Harrou, Y. Sun, and S. Khadraoui, "Short-term forecasting of photovoltaic solar power production using variational auto-encoder driven deep learning approach," *Applied Sciences*, Vol.10, No.23, p.8400, 2020.

[16] R. Chang, L. Bai, and C. H. Hsu, "Solar power generation prediction based on deep learning," *Sustainable energy technologies and assessments*, Vol.47, p.101354, 2021.

[17] B. Zazoum, "Solar photovoltaic power prediction using different machine learning methods," *Energy Reports*, Vol.8, pp.19-25, 2022.

[18] H. Alharkan, S. Habib, and M. Islam, "Solar power prediction using dual stream CNN-LSTM architecture," *Sensors*, Vol.23, No.2, p.945, 2023.

[19] M. AlKandari and I. Ahmad, "Solar power generation forecasting using ensemble approach based on deep learning and statistical methods," *Applied Computing and Informatics*, Vol.20, No.3/4, pp.231-250, 2024.

[20] E. R. M. Shouman, "Solar power prediction with artificial intelligence," in *Advances in Solar Photovoltaic Energy Systems*, IntechOpen, ch.2, 2024.

[21] C. H. Lee, H. C. Yang, and G. B. Ye, "Predicting the performance of solar power generation using deep learning methods," *Applied Sciences*, Vol.11, No.15, p.6887, 2021.

[22] J. M. Barrera, A. Reina, A. Maté, and J. C. Trujillo, "Solar energy prediction model based on artificial neural networks and open data," *Sustainability*, Vol.12, No.17, p.6915, 2020.

[23] S. G. Kim, J. Y. Jung, and M. K. Sim, "A two-step approach to solar power generation prediction based on weather data using machine learning," *Sustainability*, Vol.11, No.5, p.1501, 2019.

[24] S. Hochreiter, "Long Short-term memory," *Neural Computation MIT-Press*, 1997.

[25] S. Bai, J. Z. Kolter, and V. Koltun, "An empirical evaluation of generic convolutional and recurrent networks for sequence modeling," *arXiv preprint arXiv:1803.01271*, 2018.

[26] A. Zeng, M. Chen, L. Zhang, and Q. Xu, "Are transformers effective for time series forecasting?," in *Proceedings of the AAAI conference on artificial intelligence*, Vol.37, No.9, pp.11121-11128, 2023.

[27] A. Vaswani, "Attention is all you need," *Advances in Neural Information Processing Systems*, 2017.

[28] G. de Freitas Viscondi and S. N. Alves-Souza, "A Systematic Literature Review on big data for solar photovoltaic electricity generation forecasting," *Sustainable Energy Technologies and Assessments*, Vol.31, pp.54-63, 2019.

[29] Public Data Portal, Solar power generation data by day and branch [Internet], https://www.data.go.kr/data/15099650/fileData.do.

[30] Korea Meteorological Administration, Automated Synoptic Observing System (ASOS) [Internet], https://data.kma.go.kr/data/grnd/selectAsosRltmList.do.

[31] C. S. J. Chu, "Time series segmentation: A sliding window approach," *Information Sciences*, Vol.85, No.1-3, pp.147-173, 1995.

[32] K. H. Shin, C. Kim, S. H. Nam, S. J. Park, and S. S. Yoo, "Estimation method of predicted time series data based on absolute maximum value," *Journal of Energy Engineering*, Vol.27, No.4, pp.103-110, 2018.

Hyeonseok Jin
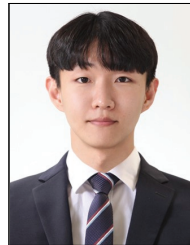
https://orcid.org/0009-0002-8271-6792
e-mail : ggyo003@jnu.ac.kr
He received his B.S. degree in Computer and Information Engineering from Chonnam National University, in 2023. He is currently pursuing the M.S degree in AI Convergence at Chonnam National University. His research interests include AI, deep learning, spatiotemporal learning, time-series analysis.

David J. Richter

https://orcid.org/0000-0001-5413-6710
e-mail : david_richter@jnu.ac.kr
He received the B.S. degree in computer science from Kempten University of Applied Sciences, in 2019, and the M.S. degree in computer information technology from Purdue University Northwest, in 2021. He is currently pursuing the master's degree in AI convergence with Chonnam National University. His research interests include AI, machine and deep learning, reinforcement learning, genetic programming.

MD Ilias Bappi

https://orcid.org/0009-0000-7616-3074

e-mail : i_bappi@jnu.ac.kr

He received his B.S. degree in Computer Science and Engineering from Daffodil International University, in 2017. He worked as a Software Engineer at Samsung Electronics, over five years until 2023. Currently, he is pursuing a master's degree in AI Convergence at Chonnam National University. His research interests including machine learning, deep learning, continual learning, and their applications in agriculture and healthcare.

Kyungbaek Kim

https://orcid.org/0000-0001-9985-3051

e-mail : kyungbaekkim@jnu.ac.kr

He received the B.S., M.S., and Ph.D. degrees in electrical engineering and computer science from Korea Advanced Institute of Science and Technology (KAIST), in 1999, 2001, and 2007, respectively. He is currently a Professor with the Department of AI Convergence, Chonnam National University. Previously, he was a Postdoctoral Researcher with the Department of Computer Sciences, University of California, Irvine. His research interests include intelligent distributed systems, SDN/SDI, big data platform, GRID/cloud systems, social networking systems, AI applied cyber-physical systems, blockchain, and other issues of distributed systems.