electronics

MDPI

Article

A Multi-Step-Ahead Photovoltaic Power Forecasting Approach Using One-Dimensional Convolutional Neural Networks and Transformer

Jihoon Moon

Department of AI and Big Data, Soonchunhyang University, Asan 31538, Republic of Korea; jmoon22@sch.ac.kr

Abstract: Due to environmental concerns about the use of fossil fuels, renewable energy, especially solar energy, is increasingly sought after for its ease of installation, cost-effectiveness, and versatile capacity. However, the variability in environmental factors poses a significant challenge to photovoltaic (PV) power generation forecasting, which is crucial for maintaining power system stability and economic efficiency. In this paper, a novel muti-step-ahead PV power generation forecasting model by integrating single-step and multi-step forecasts from various time resolutions was developed. One-dimensional convolutional neural network (CNN) layers were used for single-step forecasting to capture specific temporal patterns, with the transformer model improving multi-step forecasting by leveraging the combined outputs of the CNN. This combination can provide accurate and immediate forecasts as well as the ability to identify longer-term generation trends. Using the DKASC-ASA-1A and 1B datasets for empirical validation, several preprocessing methods were applied and a series of experiments were conducted to compare the performance of the model with other widely used deep learning models. The framework proved to be capable of accurately predicting multi-step-ahead PV power generation at multiple time resolutions.

Keywords: renewable energy; photovoltaic power generation; multi-time resolution forecasting; time-series forecasting; transformer model; one-dimensional convolutional neural network

check for updates

Citation: Moon, J. A Multi-Step-Ahead Photovoltaic Power Forecasting Approach Using One-Dimensional Convolutional Neural Networks and Transformer. *Electronics* **2024**, *13*, 2007. https://doi.org/10.3390/electronics13112007

Academic Editor: Francisco Javier Ruiz-Rodríguez

Received: 1 April 2024
Revised: 9 May 2024
Accepted: 20 May 2024
Published: 21 May 2024

Copyright: © 2024 by the author. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (https://creativecommons.org/licenses/by/4.0/).

1. Introduction

Interest in renewable energy generation technologies is growing as concerns about climate change and fossil fuel depletion increase [1]. Among these methods, photovoltaic (PV) energy is widely recognized as clean, abundant, and economically beneficial [2]. Because of these advantages, PV power generation systems are becoming increasingly popular among renewable energy sources [3]. However, the main drawback of PV power generation systems is their unpredictable and intermittent nature [4]. Since PV power is highly dependent on solar irradiance, variations in the availability of solar energy have the potential to seriously affect the ability of the grid to operate consistently and economically. By using accurate PV power generation forecasting technology that can accurately capture changes in PV generation, these problems can be mitigated by reducing the uncertainty in dynamic power generation [5].

The two primary approaches used in the past to develop prediction models for PV power generation have been physical and statistical methods [6,7]. The physical approaches utilize mathematical calculations to define the physical state of dynamic behavior and meteorological conditions. When the weather is stable, these approaches work well, but in dynamic weather conditions, the prediction accuracy can be compromised. The relationship between meteorological factors and PV power generation data is analyzed by statistical approaches, which are developed through the analysis of historical parameters and prediction models [8]. The predictive accuracy of these approaches depends on the forecast horizon and input data. Regression, autoregressive, and other techniques are widely used [9,10]. Because the input parameters are continuously improved and past

data are considered during training, these approaches can provide reliable performance. Their predictive accuracy is greater for linear data, but they face challenges with nonlinear complex data. Other limitations include the need for sufficient explanatory parameters for the prediction model and the requirement that the autoregressive integrated moving average (ARIMA) model remains stationary.

Recently, prediction solutions for PV power generation based on traditional machine learning, which considers external factors, have been developed [11–13]. The challenge of performance degradation caused by nonlinearity between input and output, a limitation seen in statistical and physical approaches, is being addressed by machine learning-based prediction models. Machine learning techniques, such as support vector regression (SVR), decision tree (DT), and extreme learning machine (ELM), are commonly used to build prediction models [14–16]. However, unbalanced PV power generation patterns under varying weather conditions can negatively impact the effectiveness of machine learning-based models. These methods also have certain drawbacks, such as poor performance with increasing numbers of data and prediction steps, complexity due to numerous features, and a heavy reliance on feature engineering, which increases the size and complexity of the model. It has been suggested that the instability of these models may render them less suitable for decision-making, challenging their suitability for PV power generation forecasting [17].

An improved statistical approach based on neural networks (NNs) was proposed by Chen et al. [18] for 24 h solar power prediction. Almonacid et al. [19] proposed an artificial NN (ANN)-based method to predict PV power output one hour in advance. They employed an ANN to predict air temperature and solar irradiance, which was then fed into another ANN to predict PV power. To predict PV power, Vaz et al. [20] used ANNs, taking measurements from nearby PV systems as inputs along with meteorological data. They noted that model performance decreased, and the root mean square error (RMSE) increased to 24% when the forecasting time horizon was extended. For a day-ahead forecast of solar PV power, Yang et al. [21] integrated NNs with a pattern sequence identification approach. This method relies on developing a separate prediction model for each type of pattern sequence. However, due to its dependence on the extracted attributes, the model is optimized for day-ahead prediction. To predict PV power generation, Wang et al. [22] evaluated three deep learning (DL) techniques and provided insights for selecting the most effective network for real-world applications. They concluded that the prediction of PV power generation benefits significantly from DL techniques.

For nonlinear data, DL-based models often perform well, but the volume and complexity of the input data significantly influence prediction performance. When categorized by prediction horizon, PV power generation predictions can be divided into one-step and multi-step predictions. A multi-step prediction is a prediction for several points in the future or a long sequence, while a one-step prediction is for only one point in the future. Contrary to previous statements, PV power generation patterns may not always be predicted more accurately using multi-step prediction than single-step prediction. This is because, regardless of the technique, predictions tend to face the problem of increasing errors as the number of prediction steps increases [23]. Weather forecasting is a critical component for prediction because various weather elements affect PV power output. However, relying solely on DL techniques such as convolutional neural networks (CNNs), long short-term memory (LSTM), and gated recurrent units (GRUs) for predictions is insufficient due to the dynamic nature of PV time series data.

To address this, in this research, a multi-time-forecast horizon PV power generation forecasting approach is presented. This approach improves PV power generation estimates by merging forecasts from multiple forecast horizons. To enhance the effectiveness of multi-step PV power generation forecasting, this multi-time resolution forecasting model unites single-step and multi-step forecasting. Multi-step forecasting is adept at identifying trends in long sequences, while single-step forecasting precisely predicts the amount of electricity generated at a specific time. By integrating single-step forecasting over the same physical

time horizon with low temporal resolution with multi-step forecasting, which forecasts many points with high temporal precision, the aim is to enhance the performance of multi-step forecasting. This approach aims to provide accurate estimates at each step in addition to revealing extended sequences of PV power generation patterns. To validate the approach, historical weather, solar radiation, and PV power generation data are collected and the models are tested using datasets relevant to the objectives of each model. Investigations include data exploration, comparisons with existing models, and further research.

The rest of this paper is organized as follows: In Section 2, reviews each model in the proposed framework. In Section 3, reports the results of the experiment, and in Section 4, this study is concluded.

2. Materials and Methods

This section presents a multi-time resolution PV power generation forecasting model based on a one-dimensional CNN (1D-CNN) and transformer. Using two different time resolutions over the same horizon, single-step and multi-step forecasts are generated and combined. The process starts with an overview of data acquisition and preprocessing. Following this, the single-step and multi-step forecasting models are explored and how the two forecasts are combined is detailed.

2.1. Data Acquisition and Preprocessing

First, the DKASC-ASA-1A and 1B datasets were collected; they are publicly available datasets and comprise detailed records from two PV power plants initiated in 2009 [24]. Dataset 1A relates to a plant with a 10.5 kW capacity, while dataset 1B relates to a plant with a 23.4 kW capacity. Both employ mono-Si technology with dual-axis tracking. The data for these datasets are meticulously logged at 5 min intervals, capturing a wide range of measurements. These measurements include environmental timestamps and various weather conditions, as well as numerous internal operational factors. Data from 1 January 2016 to 31 December 2023 was examined. Data from the last two years were intentionally excluded from the analysis to ensure that the proposed model would accurately predict recent trends and situations. The strategic use of the most recent information for validation emphasizes the researcher's efforts to create models that are relevant and timely for current and future applications by presenting models that are grounded in historical data but are also capable of predicting future conditions.

From 1 January 2016 to 31 December 2023, the DKASC-ASA-1A and 1B datasets carefully analyzed, and it was found that the missing data were distributed across a number of variables, each representing a different percentage of the total data points. For Active Energy Delivered Received (AEDR), Current Phase Average (CPA), and Active Power (AP), both datasets were missing approximately 1.15%, corresponding to 9669 and 9665 data points, respectively. Wind Speed (WS) had the largest data deficit, with an astounding 77.4% of the dataset missing, or 651,586 points. The Weather Temperature Celsius (WTS) and Weather Relative Humidity (WRH) datasets were missing about 1.15% or 9675 data points each, a figure that also applies to Global Horizontal Radiation (GHR) and Diffuse Horizontal Radiation (DHR) with 9676 missing entries each. Wind Direction (WD) and Weather Daily Rainfall (WDR) each had about 1.15% of the data missing, corresponding to 9673 and 9670 data points, respectively. Notably, Radiation Global Tilted (RGT) was missing 11.24% or 94,603 data points, while Radiation Diffuse Tilted (RDT) had a deficit of about 9.27%, corresponding to 78,003 data points.

The timestamp information was converted from date and time to sine and cosine values, as shown in Algorithm 1, to preserve the cyclical nature of time [25]. This coding approach allows for the model to effectively capture the periodicity inherent in days and minutes, which is critical for recognizing patterns and making accurate predictions over time.

Algorithm 1: Cyclical Feature Encoding of Timestamp Data

Input: Dataset with 'timestamp' column.
Output: Dataset with added 'Date_Sine,' 'Date_Cosine,' 'Time_Sine,' and 'Time_Cosine' features.
Import necessary libraries.
Load dataset into a DataFrame.
Convert 'timestamp' column to datetime objects.
For each 'timestamp' in the DataFrame do
  Calculate day of the year and minute of day.
  If year is a leap year then
    Set total days in year to 366.
  Else
    Set total days in year to 365.
  EndIf
  Calculate 'Date_Sine' and 'Date_Cosine' using day of year.
  Calculate 'Time_Sine' and 'Time_Cosine' using minute of day.
  Append new features to DataFrame.
EndFor
Return updated DataFrame.

Any erroneous records with negative values for CPA and AP were corrected during the data cleansing process by setting them to zero. This was a critical step in maintaining the integrity of the dataset, particularly for these attributes where negative values are not possible. It was recognized that generation capacity and metrics such as weather conditions, RGT, and RDT inherently have positive qualities and therefore cannot be negative. The outliers were identified using the process outlined in Algorithm 2 and were initially left blank to accurately reflect their status as anomalies pending further treatment. In this instance, the WDR attribute was not subjected to the outlier identification process outlined in Algorithm 2, as its values primarily reflect zero rainfall.

Algorithm 2: Replacement of Outliers with NAs

Input: CSV file containing the dataset.
Output: Dataset with outliers in specific variables replaced with NAs.
Load the CSV file into a data frame '*total_data*'.
Define the target variables for outlier replacement in '*interpolate_vars*'.
For each variable in '*interpolate_vars*' do
  Calculate the first and third quartiles and the interquartile range (IQR).
  Define cutoffs for outliers as 1.5 × IQR below the first quartile and above the third quartile.
  For each data point in the variable do
    If the data point is **below the lower cutoff** or **above the upper cutoff**, then
      Replace the data point with NA.
    If the data point is **below the 10th percentile**, then
      Replace the data point with NA.
    EndFor
  EndFor
  Return the modified data frame '*total_data*'.

In order to effectively manage data quality, it was determined that the variable WS, which contained a significant number of missing entries, should be removed. For all other variables, except RGT and RDT, linear interpolation was used to impute missing values. For the variables RGT and RDT, a missing value imputation model was constructed to accurately impute their missing values [26]. In addition, the variable AEDR was removed from the dataset to ensure clarity and focus in the analysis. Furthermore, to prevent data leakage, the values for CPA, RGT, and RDT were replaced with data from one hour earlier. This approach ensures the integrity of the predictive modeling process. Table 1 presents the statistical analysis of each attribute from the DKASC-ASA-1A and 1B datasets after preprocessing has been completed.

Table 1. Statistics for each attribute of the post-processed DKASC-ASA-1A and 1B datasets.

Attribute	Mean	Min.	25%	50%	75%	Max.	Unit
AP (1A)	2.01	0	0	0	3.97	9.91	kW
AP (1B)	3.78	0	0	0	7.20	23.42	kW
WTC	21.97	7.67	14.91	22.05	28.68	46.31	°C
WRH	36.74	11.6	18.39	30.34	50.39	97.59	%
GHR	270.75	1.44	2.31	11.39	541.59	1341.97	W/m^2
DHR	43.77	0.78	1.35	8.34	74.79	181.05	W/m^2
WD	66.32	12.01	28.17	39.16	94.42	204.32	°
WDR	0.34	0	0	0	0	66.00	mm
$Date_{Sine}$	0	−1	−0.71	0	0.71	1	–
$Date_{Cosine}$	0	−1	−0.71	0	0.71	1	–
$Time_{Sine}$	0	−1	−0.71	0	0.71	1	–
$Time_{Cosine}$	0	−1	−0.71	0	0.71	1	–
CPA_{H-1}	4.65	0	0.30	0.68	9.44	21.57	A
RGT_{H-1}	290.64	0	2.72	12.18	586.33	1515.75	W/m^2
RDT_{H-1}	58.75	0	1.64	10.01	83.14	766.95	W/m^2

In this study, the dataset was prepared by dividing the available data into a training set of six years and a test set of the last two years, representing a 75:25 split. After imputing missing values, data normalization techniques were applied. Specifically, each variable (except the dependent variable AP) in the training set was normalized to a range between 0 and 1 using the min–max normalization technique, as described in Equation (1):

$$X_{norm.} = (X - X_{min.})/(X_{max.} - X_{min.}) \quad (1)$$

The normalization parameters determined from the training set were then applied to the test set to ensure consistency in data processing between the two sets. This approach allows for a valid comparison between the proposed model and alternative models, thereby improving the accuracy of predictions made one hour in advance.

2.2. PV Power Prediction Model

2.2.1. One-Dimensional Convolution

One method for learning local trend features is 1D convolution. To extract features from local input patches, 1D convolution uses a convolution technique that promotes data efficiency and representation modularity [27]. However, it is important to clarify that while 1D convolution is well suited for sequence processing, its application in computer vision typically involves 2D or 3D convolution due to the spatial nature of image data [28]. By reducing multivariate time series processing parameters, the local awareness and weight sharing capabilities of 1D convolution can increase learning efficiency. This is because each subsequence undergoes the same input modification. Temporal translation invariance [29] allows for a pattern learned at one point in a sequence to be recognized at a different location later.

Sequence fragments within the window size can be learned by processing the input data using the convolution window of the convolution layer. Figure 1 illustrates the process of 1D convolution. The 1D input is processed through the 1D convolution layer to produce the 1D output. This approach is effective because it allows for the detection of these subsequences anywhere in the entire time series, capturing the local trend change aspects of the multivariate time series over time. The one-dimensional input time series is effectively analyzed in this way.

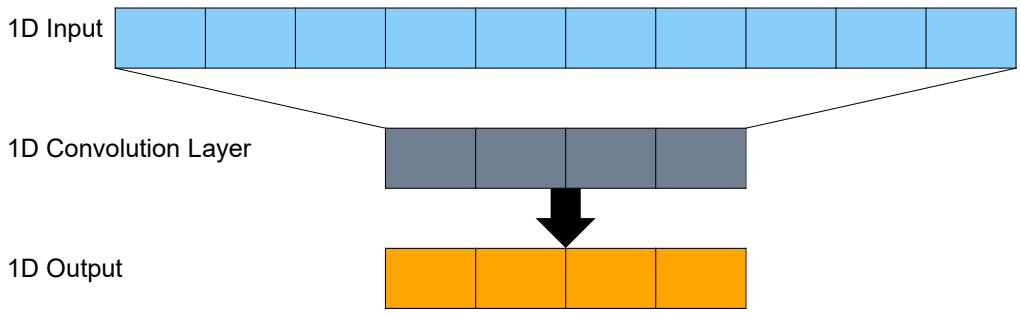

Figure 1. Schematic of the one-dimensional (1D) convolution process from input to output layers.

2.2.2. Transformer Architecture

The LSTM model's gradient vanishing and gradient explosion issues, as well as long-distance dependencies, may all be successfully mitigated by adding an attention mechanism. However, it does not fully address the aforementioned problems. In addition, problems such as low efficiency and excessive computational complexity may persist. To more effectively address these challenges, Vaswani et al. [30] presented the transformer model in 2017 that eschews the LSTM in favor of an all-encompassing attention mechanism. Numerous studies have demonstrated that when it comes to capturing long-term dependencies, transformer models generally outperform LSTM models [31]. Additionally, they have been shown to provide better prediction accuracy in fields such as speech recognition and machine translation when compared to LSTM [32].

The transformer model is critical for understanding relationships within sequences of data, significantly improving the ability to process data over recurrent models. The transformer model is composed of six layers of encoders and decoders, as depicted in Figure 2. The encoder layers work to process input data through self-attention mechanisms that allow for the model to weigh the importance of different parts of the input data. Each encoder layer comprises a feedforward layer and a multi-head attention layer. The decoder includes three sub-layers: the feedforward layer, the multi-head attention layer, and the masked multi-head attention layer. These layers in the decoder work to generate predictions by focusing on relevant parts of the encoded input as well as previously generated outputs. After every sublayer, a residual connection and layer normalization are applied. This helps the model maintain stability and speeds up training by avoiding the vanishing gradient problem. The multi-head attention method splits the three parameters (query, key, and value) into multiple parts, yet the total parameter count remains constant. This division into multiple "heads" allows for the model to simultaneously attend to information from different representational subspaces at different locations. The attention weight is then computed by mapping each set of partitioned parameters to a distinct subspace of the high-dimensional space, enabling focused analysis on different input components. This parallel processing is key to the efficiency of the model and its ability to handle complex patterns in the data. Finally, the attention information from all subspaces is aggregated after several parallel computations.

Multi-head attention is designed to identify connections between different types of input data, encoding multiple relationships and nuances as attention is allocated differently across different subspaces. This aspect endows the transformer model with significant functionality. Multiple independent heads focus on different aspects of the data (such as global and local contexts) to extract richer and more nuanced features.

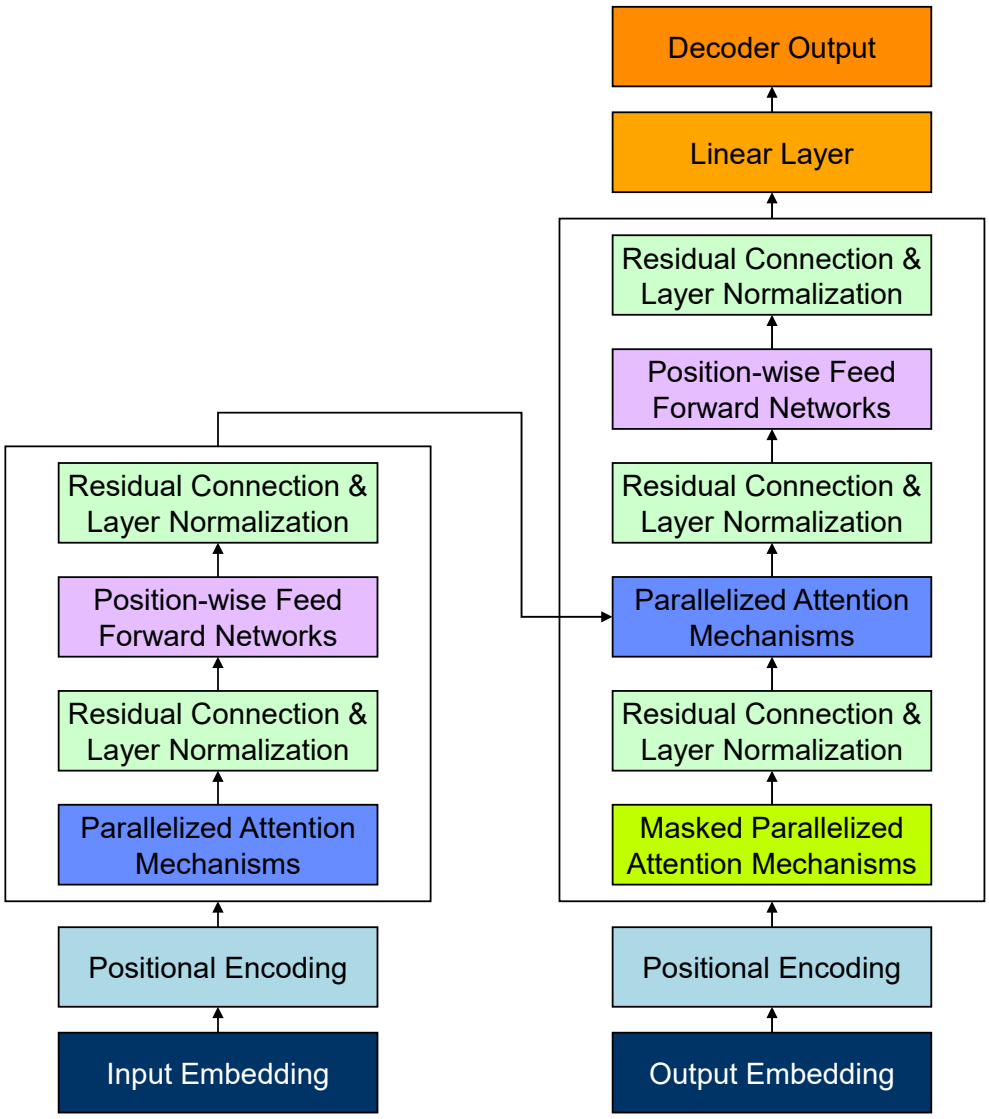

Figure 2. Transformer architecture flowchart depicting encoder and decoder processing layers.

2.2.3. Multi-Resolution PV Prediction

A model with a 1 h temporal resolution for one-step PV power generation prediction was built, ensuring that only nondependent variables were used to prevent data leakage. This approach minimized the risk of incorporating future information that could invalidate the model's predictions. This model was based on a 1D-CNN and incorporated timestamps (i.e., $\text{Date}_{\text{Sine}}$, $\text{Date}_{\text{Cosine}}$, $\text{Time}_{\text{Sine}}$, and $\text{Time}_{\text{Cosine}}$), weather conditions (i.e., WTC, WRH, GHR, DHR, WD, and WDR), and historical data (i.e., $\text{CPA}_{\text{H}-1}$, $\text{RGT}_{\text{H}-1}$, and $\text{RDT}_{\text{H}-1}$) from both the DKASC-ASA-1A and 1B datasets. These inputs were carefully selected based on their relevance and independence from future data. To adapt the proposed model to practical time frames, the 5 min interval data were aggregated into hourly values. Specifically, cumulative values were used for AP and $\text{CPA}_{\text{H}-1}$, summing 12 data points (e.g., from 10:05 a.m. to 11:00 a.m.). In contrast, the remaining variables—timestamp, weather information, $\text{RGT}_{\text{H}-1}$, and $\text{RDT}_{\text{H}-1}$—were adjusted to use the corresponding values at the hour mark (e.g., 11:00 a.m.). This approach ensured that the forecast was based solely on the information available at the time of the forecast, thereby maintaining the integrity of the forecast. The 1D-CNN model processed this information to produce a prediction output, which was referred to as PV_{h+1}, ensuring the reproducibility and

reliability of the prediction process. The 1D-CNN acted as a feature extractor, transforming the input data into a form suitable for further processing by the transform encoder.

The integration of single-step and multi-step forecasting was a novel approach that improved the accuracy of the model by combining short-term precision with long-term contextual understanding. Single-step and multi-step PV power generation forecasting were integrated to build a multi-time resolution forecasting model. This multi-time resolution approach provided a detailed forecast at different future intervals, providing a more granular understanding of expected power generation. The one-step forecast used a one-hour forecast horizon, while the multi-step forecast used a 5 min time resolution over a 1 h time horizon. The 5 min resolution was particularly important for applications where short-term variability was critical. The last two hours of data were sent to the multi-step forecaster. A 1D convolutional layer was used to extract features from the data and reduce dimensionality, as the input sequence length of the encoder could be extensive. Dimensionality reduction was an essential step in managing the complexity and computational load of the model. It was then subjected to positional encoding to represent temporal features. Position encoding was essential for the model to capture sequence order, a critical aspect for time series prediction. Figure 3 shows the general layout of the proposed model.

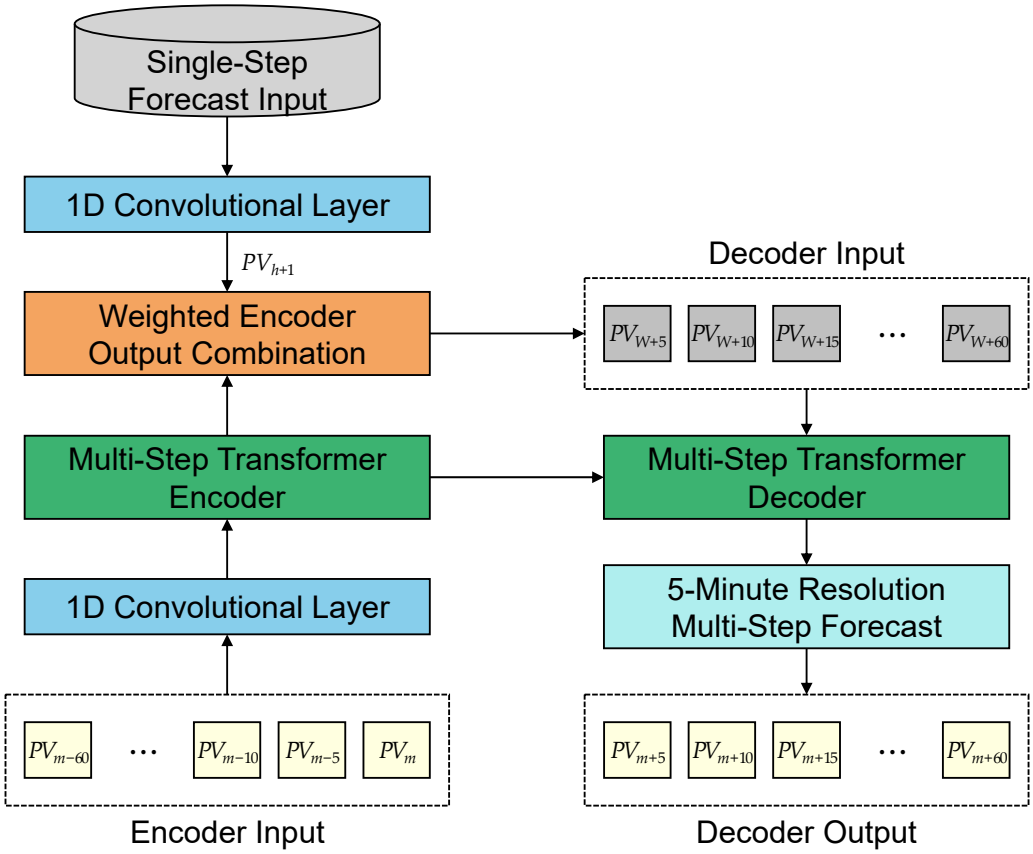

Figure 3. Integrated forecasting architecture combining single-step and multi-step predictions with 1D convolutional and transformer models.

The first multi-step prediction was then performed using a transformer encoder. The transformer encoder was able to capture complex dependencies across different time steps. To merge the two predictions, a combination of single-step prediction and a weighted partitioning on the encoder outputs was used. The weighting process was designed to balance the contribution of each time step to the final prediction. The decoder received the weighted partition values as input, and the two predictions were integrated to enhance the multi-step prediction. Integration was a form of ensemble learning that improved the robustness and accuracy of the prediction. The weighted division value was calculated

using Equation (2), where EO represents the one-step prediction value at the ith point, which is used to compute the contribution of each encoder output to the final prediction:

$$\text{Weighted Division at } i = EO_i/(\Sigma \text{ from } j = 1 \text{ to } n \text{ of } EO_j) \tag{2}$$

The weighted division normalizes the influence of each encoder output, ensuring that no single step has a disproportionate influence on the multi-step prediction. By concatenating the weighted division values and the weather data and then feeding them into the decoder, a more accurate multi-step prediction was achieved. This process effectively combined the extracted features with external data to refine the final output.

3. Experimental Results

This section outlines the assessment measures and a comparative study of several models. The technical resources utilized in this research included a GeForce RTX 2080 Super GPU (NVIDIA, Santa Clara, CA, USA), a Core i7 processor (Intel, Santa Clara, CA, USA), Windows 10 (Microsoft, Redmond, WA, USA), Python 3.9 (Python Software Foundation, Wilmington, DE, USA), and the PyTorch DL framework (Facebook, Menlo Park, CA, USA). To begin, the outcomes of the proposed model are contrasted with the actual solar PV power generation data from the DKASC database. This comparison is conducted with a focus on the forecasted power versus the actual power for the specified time interval. Secondly, the proposed model is contrasted with the most recent forecasting models and the findings of an ablation study are presented.

The mean absolute error (MAE) and RMSE measures are used to evaluate each model's performance. These are the general metrics used in the forecasting model evaluation procedures used by the baseline. The MAE (Equation (3)) is the average absolute difference between predicted and actual values. The RMSE (Equation (4)) calculates the average square root difference between actual and predicted model values.

$$MAE = (1/n) \times \text{sum}(|y_i - yhat_i|) \tag{3}$$

$$RMSE = sqrt((1/n) \times \text{sum}((y_i - yhat_i)^2)) \tag{4}$$

Here, y_i refers to the actual values observed; $yhat_i$ signifies the predicted values by the model; n is the total number of observations used in the evaluation.

Upon evaluation, the MAE and RMSE metrics were calculated using only non-zero actual values of solar PV power generation to assess the performance of the proposed model in comparison to other forecasting models. This method ensures that the performance metrics reflect the accuracy of predictions where actual generation occurred, providing a focused analysis of model effectiveness in practical scenarios.

3.1. Comparative Analysis of Hourly PV Power Generation Forecasting

To ensure that the proposed model can accurately predict recent trends and situations, the data were divided into two sets: the training set, which included data from 2016 to 2021, and the test set, which included data from 2022 to 2023. Some preliminary experiments were then conducted to determine whether the 1D-CNN model was the most effective for forecasting hour-ahead hourly PV power generation.

Figure 4 presents a comparative analysis of the actual hourly solar PV power generation against the 1D-CNN model's predictions for the DKASC-ASA-1A and 1B datasets. The dataset encompasses the pivotal summertime period in the Southern Hemisphere, specifically from 25 December 2022 to 7 January 2023. This period falls within the testing phase and coincides with the austral summer in Australia, resulting in elevated solar intensity and power generation.

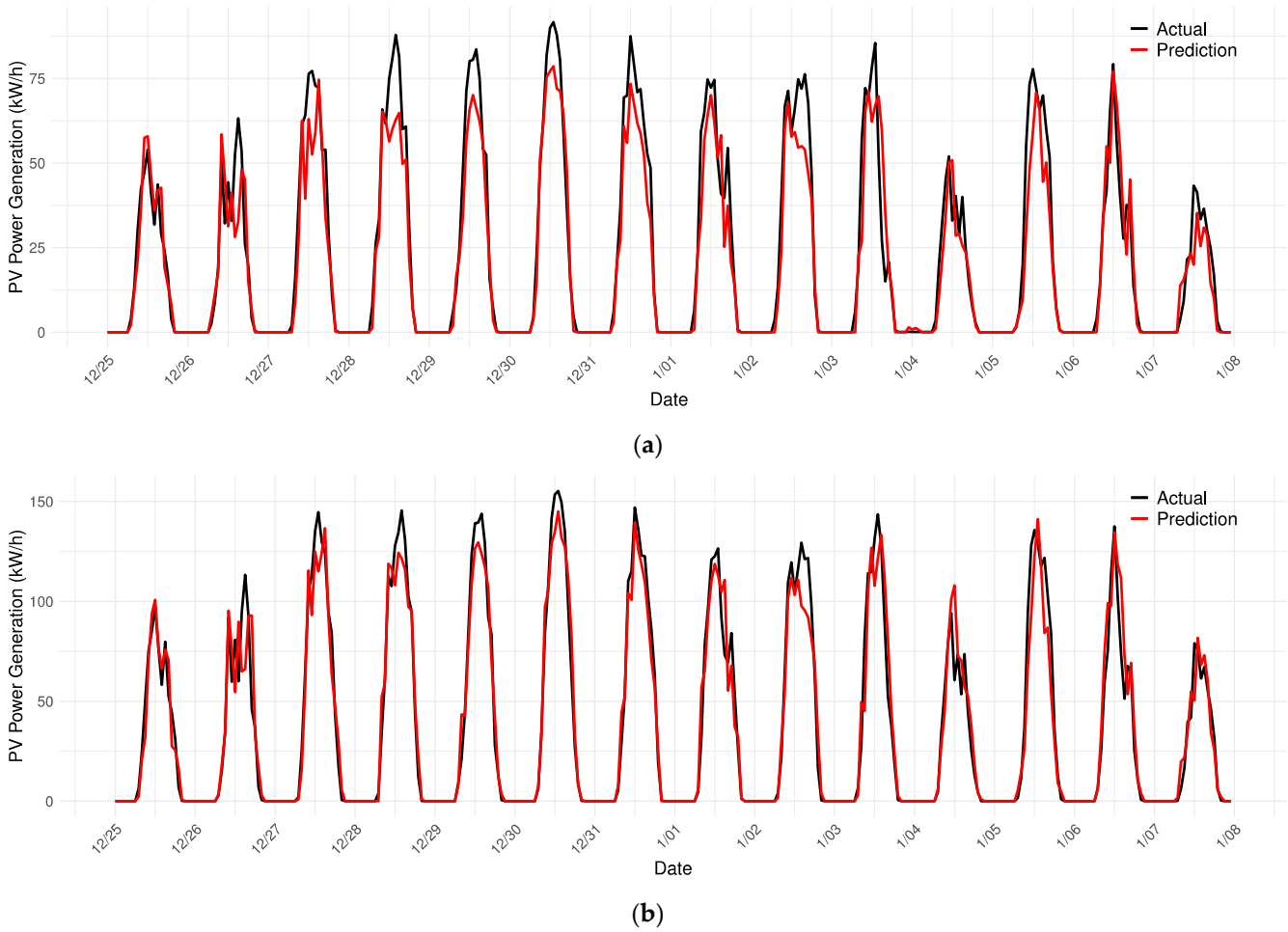

Figure 4. Performance analysis of the 1D convolutional neural network model: 25 December 2022 to 7 January 2023. (**a**) DKASC-ASA-1A; (**b**) DKASC-ASA-1B.

The figure displays the prediction performance of the 1D-CNN model over multiple days, demonstrating the potential for reliable hour-ahead hourly solar PV power generation forecasting. Despite the inherent challenges posed by the irregular weather patterns characteristic of the summer months, this comparison demonstrates the 1D-CNN model's ability to adapt to the high energy output typically observed in Australia's hot summers.

In order to facilitate a performance comparison between the DL and 1D-CNN models, a number of different NN architectures were considered, including deep neural network (DNN), LSTM, bidirectional LSTM (Bi-LSTM), GRU, and bidirectional GRU (Bi-GRU). The hyperparameter settings of the DL models are as follows: the activation function was set to Leaky rectified linear unit (LeakyReLU), the number of feedforward neurons was set to 128, and two hidden layers were selected. Adam was used as the optimization algorithm. The learning rate was set to 0.0001, and the batch size was set to 24.

Table 2 compares the prediction performance of DL models for hour-ahead hourly PV power generation forecasting. The results indicate that the 1D-CNN model exhibits the lowest values for both MAE and RMSE in both datasets. Therefore, it can be concluded that the 1D-CNN model provides more accurate predictions than the other forecasting models. This evidence supports the claim that the 1D-CNN architecture is effective for accurate hourly solar PV power generation forecasting.

Table 2. Mean absolute error (MAE) and root mean square error (RMSE) comparisons for hourly solar photovoltaic (PV) power generation forecasting models (unit: kWh).

Model	DKASC-ASA-1A		DKASC-ASA-1B	
	MAE	RMSE	MAE	RMSE
DNN	14.88	20.39	11.83	15.88
LSTM	14.90	20.81	12.28	17.49
Bi-LSTM	15.69	21.68	12.85	17.89
GRU	14.68	20.21	11.75	16.61
Bi-GRU	14.87	21.10	11.78	16.18
1D-CNN	14.17	18.51	10.58	14.41

3.2. Comprehensive Ablation Study and Performance Analysis

Figure 5 presents the actual 5min PV power generation in comparison to the proposed PV power generation forecasting model predictions for the DKASC-ASA-1A and 1B datasets. In contrast to the hourly forecasts in Figure 4, Figure 5 shows the 5 min interval forecasts from 30 December 2022 to 2 January 2023. The proposed model includes forecasts from one step ahead (5 min) up to twelve steps ahead (one hour). This figure illustrates the prediction performance of different forecast intervals on actual PV power generation during the pivotal summer period in the Southern Hemisphere.

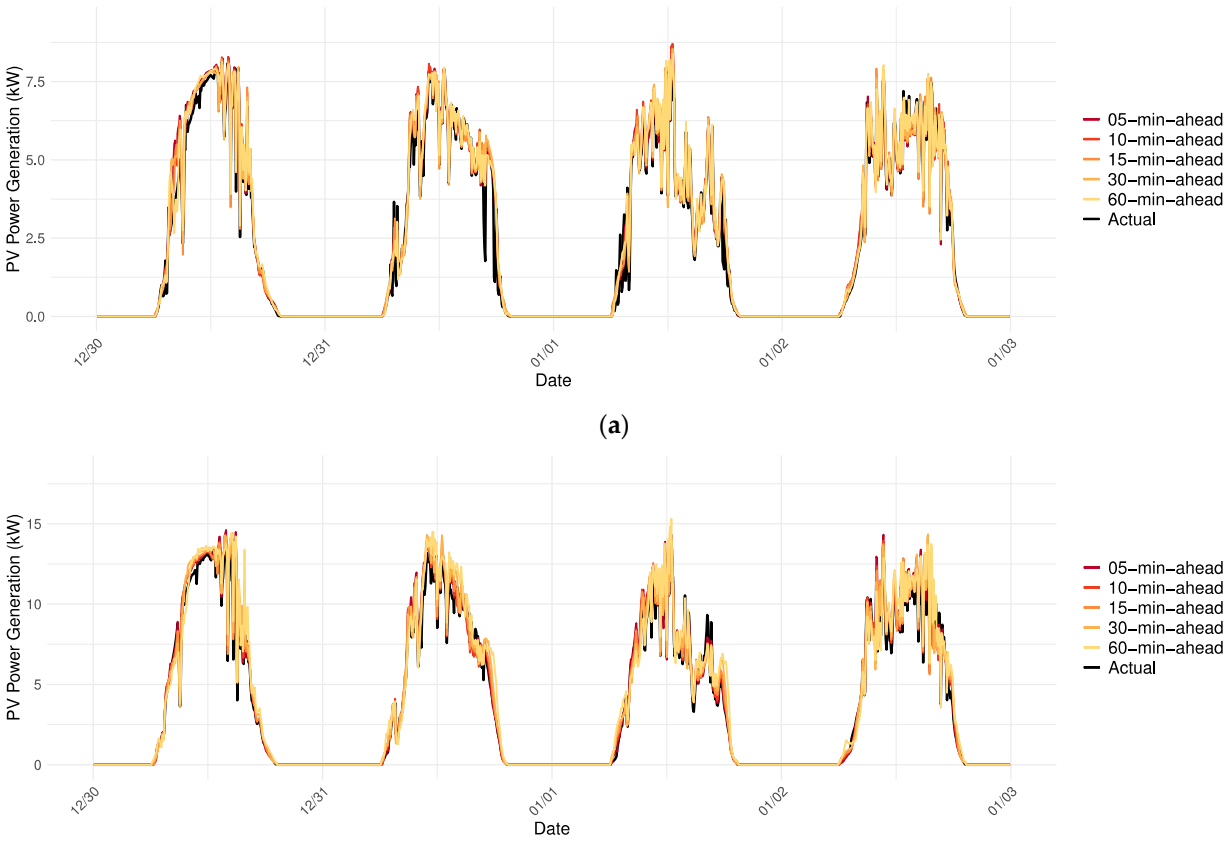

Figure 5. Multi-step photovoltaic (PV) power generation forecasting results using the proposed model: 30 December 2022 to 2 January 2023. (**a**) DKASC-ASA-1A; (**b**) DKASC-ASA-1B.

The proposed model captured the distinct patterns of solar PV power generation throughout this period, maintaining a high degree of accuracy despite the challenging fluctuations due to varying solar intensities. The predictions align closely with the actual data, demonstrating the model's ability to reliably forecast both the sharp rises and declines

characteristic of the summer season. By capturing the day-to-day peaks and troughs as well as the fine fluctuations at 5 min intervals, the model offers a comprehensive insight into short-term solar PV power generation.

To demonstrate the effectiveness of integrating both transformer and 1D-CNN models with single-step forecasting, a comprehensive ablation study was conducted, detailed in Table 3. This table presents three distinct model configurations, which elucidate the rationale behind the chosen approach. Case 1 is a 5 min interval PV power generation forecasting model based on only 1D-CNN. This case can serve as a baseline to ascertain the predictive capabilities of a purely convolutional approach. Case 2 is a model combining transformer and 1D-CNN architectures without single-step forecasting. This case allows for the examination of the impact of not including that crucial feature. Case 3 is the proposed model, which integrates both transformer and 1D-CNN architectures while enhancing prediction performance by including single-step forecasting. This ablation study was essential for the identification of the most effective components and configurations for accurate PV power generation forecasting.

Table 3. Ablation study configurations.

Case	Configuration
Case 1	1D-CNN-only forecasting model
Case 2	Transformer and 1D-CNN combined model without single-step forecasting
Case 3 (ours)	Transformer and 1D-CNN combined model enhanced with single-step forecasting

To ascertain the potential advantages of the proposed model, a comprehensive evaluation was conducted, as detailed in Tables 4 and 5. The results from Table 4 (DKASC-ASA-1A dataset) and Table 5 (DKASC-ASA-1B dataset) show that the proposed model (Case 3) consistently outperformed the transformer-only and 1D-CNN-only models in terms of both MAE and RMSE across a multi-step-ahead PV power generation forecasting.

In the DKASC-ASA-1A dataset (Table 4), the proposed model achieved an average MAE of 1.243 and RMSE of 1.813, significantly outperforming Case 1 and Case 2. This trend was consistent across all forecast steps, demonstrating the proposed model's robustness in delivering accurate predictions even with varying forecast horizons.

In the DKASC-ASA-1B dataset (Table 5), the proposed model also demonstrated superior performance, with an average MAE of 0.770 and RMSE of 1.125. The performance gap widened notably at higher forecasting steps, with the proposed model providing a substantial improvement over Case 1 and Case 2.

Table 4. Ablation study results on the DKASC-ASA-1A dataset (unit: kW).

Step	MAE			RMSE		
	Case 1	Case 2	Case 3	Case 1	Case 2	Case 3
1	1.279	1.272	1.227	1.856	1.838	1.789
2	1.272	1.273	1.236	1.841	1.837	1.803
3	1.277	1.269	1.232	1.846	1.833	1.800
4	1.284	1.267	1.235	1.856	1.831	1.802
5	1.275	1.269	1.242	1.848	1.836	1.813
6	1.277	1.276	1.247	1.853	1.846	1.819
7	1.290	1.282	1.249	1.866	1.852	1.823
8	1.295	1.273	1.243	1.871	1.840	1.814
9	1.292	1.274	1.241	1.869	1.840	1.812
10	1.284	1.274	1.243	1.862	1.841	1.814
11	1.276	1.270	1.256	1.854	1.833	1.829
12	1.278	1.268	1.264	1.853	1.836	1.834
Avg.	1.282	1.272	1.243	1.856	1.839	1.813

Table 5. Ablation study results on the DKASC-ASA-1B dataset (unit: kW).

Step	MAE			RMSE		
	Case 1	Case 2	Case 3	Case 1	Case 2	Case 3
1	0.535	0.494	0.416	0.808	0.760	0.675
2	0.607	0.577	0.491	0.910	0.880	0.777
3	0.679	0.652	0.546	1.007	0.976	0.842
4	0.749	0.719	0.603	1.103	1.060	0.909
5	0.814	0.786	0.674	1.190	1.141	0.996
6	0.869	0.846	0.750	1.258	1.211	1.094
7	0.922	0.907	0.823	1.325	1.287	1.188
8	0.974	0.973	0.881	1.385	1.370	1.267
9	1.021	1.037	0.942	1.442	1.456	1.347
10	1.058	1.095	0.996	1.487	1.527	1.417
11	1.106	1.151	1.037	1.542	1.595	1.467
12	1.149	1.193	1.081	1.590	1.639	1.517
Avg.	0.874	0.869	0.770	1.254	1.242	1.125

The statistical significance of the proposed model's superiority was validated through the Friedman test, which was performed separately for each dataset and evaluation metric. The Friedman test is a nonparametric statistical test employed to identify discrepancies in performance across a range of models or conditions [33]. In the DKASC-ASA-1A dataset, the p-value for MAE was 1.54×10^{-5}, and the RMSE p-value was 6.14×10^{-6}. In the DKASC-ASA-1B dataset, the p-value for MAE was 8.84×10^{-5}, and for RMSE, it was also 8.84×10^{-5}. These results demonstrate the robustness of the proposed model in terms of predictive accuracy, particularly with the combined transformer and 1D-CNN architecture when integrating single-step prediction.

3.3. Model Performance: Comparison with the State of the Art

A comparison was conducted between the prediction performance of the proposed model and that of recent state-of-the-art DL models, including transformer [30], CNN-LSTM [22], LSTM-CNN [34], GRU-CNN [35], and echo state network (ESN)-CNN [36]. To ensure the reliability of the results, a comprehensive review of the original papers was conducted to replicate the DL architectures in their original form and adapt them to the experimental environment. The hybrid DL approaches—CNN-LSTM, LSTM-CNN, GRU-CNN, and ESN-CNN—employ feature extraction at the front end and prediction at the back end. Previous studies have demonstrated that these models outperform single DL models in terms of prediction performance when used with similar datasets. Consequently, these models were selected as benchmark models to validate the effectiveness of the proposed model, given their proven capability in providing accurate predictions.

Furthermore, post-processed datasets and identical hyperparameters were employed in all experiments to guarantee a fair comparison. To ensure consistency, the training set (2016–2021) and the test set (2022–2023) were aligned. Predictions for PV power generation were made at 5 min intervals, ranging from 5 min to one hour ahead (12-step prediction). This approach ensured that each model was implemented consistently and evaluated under the same conditions. Consequently, accurate and reliable comparative results can be provided. The experimental results are presented in Tables 6 and 7.

In Table 6, the proposed model achieved an average MAE of 1.243 and RMSE of 1.789, demonstrating an advanced predictive capability in terms of RMSE, where it consistently outperformed all other forecasting models, including transformer, CNN-LSTM [22], GRU-CNN [35], LSTM-CNN [34], and ESN-CNN [36]. However, in terms of MAE, the proposed model was slightly less accurate than ESN-CNN (MAE 1.241), but still showed an improvement over the remaining models. These results highlight the robustness of this approach, especially in minimizing large errors captured by the RMSE.

In Table 7, the proposed model achieved the lowest MAE (0.770) and RMSE (1.125), significantly outperforming the transformer and other benchmark models in both metrics. The MAE was significantly lower than that of ESN-CNN (0.790), the second best performing model, highlighting the predictive accuracy of the proposed model. Similarly, the RMSE showed a significant improvement and was consistently lower than all other forecasting models.

Table 6. MAE and RMSE comparisons for predicting hour-ahead PV power generation using 5 min interval multi-step forecasts (steps 1 to 12) on the DKASC-ASA-1A dataset (unit: kW).

Step	MAE						RMSE					
	Transformer	CNN-LSTM	LSTM-CNN	GRU-CNN	ESN-CNN	Ours	Transformer	CNN-LSTM	LSTM-CNN	GRU-CNN	ESN-CNN	Ours
1	1.246	1.242	1.242	1.246	1.231	1.227	1.805	1.820	1.818	1.835	1.804	1.789
2	1.256	1.236	1.239	1.242	1.225	1.236	1.809	1.815	1.811	1.829	1.793	1.803
3	1.257	1.239	1.232	1.243	1.228	1.232	1.813	1.817	1.798	1.830	1.796	1.800
4	1.255	1.243	1.242	1.241	1.231	1.235	1.810	1.823	1.808	1.826	1.803	1.802
5	1.258	1.244	1.247	1.246	1.233	1.242	1.813	1.821	1.815	1.832	1.805	1.813
6	1.259	1.250	1.251	1.257	1.244	1.247	1.813	1.828	1.819	1.840	1.822	1.819
7	1.261	1.260	1.253	1.269	1.251	1.249	1.814	1.839	1.827	1.849	1.831	1.823
8	1.269	1.258	1.252	1.263	1.248	1.243	1.823	1.838	1.822	1.844	1.829	1.814
9	1.263	1.263	1.256	1.266	1.249	1.241	1.816	1.841	1.834	1.849	1.829	1.812
10	1.256	1.258	1.256	1.256	1.252	1.243	1.810	1.836	1.831	1.839	1.834	1.814
11	1.259	1.262	1.257	1.251	1.246	1.256	1.816	1.839	1.832	1.832	1.821	1.829
12	1.256	1.266	1.249	1.253	1.251	1.264	1.812	1.840	1.821	1.836	1.821	1.834
Avg.	1.258	1.252	1.248	1.253	1.241	1.243	1.813	1.830	1.820	1.837	1.816	1.789

Table 7. MAE and RMSE comparisons for predicting hour-ahead PV power generation using 5 min interval multi-step forecasts (steps 1 to 12) on the DKASC-ASA-1B dataset (unit: kW).

Step	MAE						RMSE					
	Transformer	CNN-LSTM	LSTM-CNN	GRU-CNN	ESN-CNN	Ours	Transformer	CNN-LSTM	LSTM-CNN	GRU-CNN	ESN-CNN	Ours
1	0.492	0.458	0.467	0.487	0.447	0.416	0.737	0.691	0.747	0.765	0.718	0.675
2	0.574	0.553	0.545	0.561	0.506	0.491	0.851	0.841	0.860	0.861	0.792	0.777
3	0.659	0.635	0.612	0.637	0.572	0.546	0.969	0.959	0.951	0.951	0.892	0.842
4	0.734	0.717	0.671	0.699	0.642	0.603	1.069	1.070	1.024	1.036	0.985	0.909
5	0.806	0.791	0.738	0.761	0.702	0.674	1.158	1.162	1.108	1.114	1.063	0.996
6	0.863	0.853	0.807	0.818	0.767	0.750	1.228	1.238	1.189	1.187	1.139	1.094
7	0.928	0.907	0.872	0.877	0.833	0.823	1.314	1.303	1.274	1.260	1.218	1.188
8	0.981	0.962	0.924	0.930	0.890	0.881	1.378	1.367	1.341	1.328	1.288	1.267
9	1.033	1.021	0.982	0.982	0.948	0.942	1.441	1.442	1.426	1.393	1.369	1.347
10	1.079	1.073	1.031	1.034	1.001	0.996	1.496	1.501	1.487	1.462	1.431	1.417
11	1.125	1.121	1.086	1.080	1.054	1.037	1.553	1.560	1.547	1.518	1.496	1.467
12	1.157	1.160	1.121	1.127	1.116	1.081	1.588	1.603	1.585	1.574	1.573	1.517
Avg.	0.869	0.854	0.821	0.833	0.790	0.770	1.232	1.228	1.212	1.204	1.164	1.125

In order to ascertain the statistical significance of the proposed model's performance, the Wilcoxon signed-rank test, a nonparametric statistical method, was employed [37]. This test compares two related groups in order to determine whether one group exhibits systematically higher values than the other. The Wilcoxon signed-rank test offers a robust means of evaluating prediction performance by analyzing the rankings of individual scores rather than relying on assumptions about data distribution. The procedure entails ranking the differences between pairs of related observations in ascending order. After assigning ranks, the sum of ranks for positive and negative differences is calculated, which is denoted as $R+$ and $R-$. Then, the test statistic W is computed as the smaller of these two sums, expressed as Equation (5).

$$W = \min(R+, R-). \tag{5}$$

The p-value can then be obtained by comparing W to its distribution under the null hypothesis, which assumes no significant difference between the two groups.

Table 8 presents the results of the Wilcoxon signed-rank test, which uses 12-step MAE and RMSE values to assess the performance of various models across the DKASC-ASA-1A and DKASC-ASA-1B datasets. This table compares the proposed model with recent state-of-the-art DL models.

Table 8. Wilcoxon signed-rank test results for different PV power generation forecasting models.

Model	DKASC-ASA-1A		DKASC-ASA-1B	
	MAE	RMSE	MAE	RMSE
Transformer [30]	0.00146	0.85010	0.00049	0.00049
CNN-LSTM [22]	0.00049	0.00049	0.00049	0.00049
LSTM-CNN [34]	0.04255	0.03418	0.00049	0.00049
GRU-CNN [35]	0.01221	0.00049	0.00049	0.00049
ESN-CNN [36]	0.33936	0.38037	0.00049	0.00049

The results indicate that the proposed model exhibits statistically significant improvements over several benchmark models, including CNN-LSTM [22], LSTM-CNN [34], and GRU-CNN [35], with p-values below 0.05. However, the differences between the proposed model and the ESN-CNN [36] or transformer [30] models are not statistically significant, as their p-values are above 0.05. The transformer model demonstrates inconsistent results across two datasets. In the DKASC-ASA-1A dataset, the proposed model exhibited a statistically significant difference for MAE but not for RMSE. Conversely, in the DKASC-ASA-1B dataset, the proposed model demonstrated a significant difference for both metrics. In contrast, the ESN-CNN model exhibited no statistically significant difference in the DKASC-ASA-1A dataset but did demonstrate statistical significance in the DKASC-ASA-1B dataset.

In conclusion, the proposed model consistently outperforms many advanced deep learning models in predictive accuracy. Although it sometimes has a slightly higher MAE than the ESN-CNN model, it still provides reliable multi-step forecasts for PV power generation, which helps stabilize the grid and improve energy management. However, since 1D convolution struggles to capture spatial patterns and transformer models are complex to compute, future research could adopt the parallel pooling approach [38]. This would facilitate a more comprehensive understanding of complex weather patterns and enhance prediction accuracy across diverse forecast horizons. In future research, the researcher intends to investigate the potential for refining the parallel pooling approach to enhance the model's capacity to capture complex weather patterns and improve predictive accuracy.

4. Conclusions

This study presented a multi-time resolution solar PV power generation forecasting approach that employed both one-step and multi-step forecasting methods integrated into a final model that served as a comprehensive solution for maintaining a solar PV power system planning and maintenance schedule. The proposed model was distinguished by its exceptional predictive accuracy and advanced pattern recognition capabilities. Experiments conducted to evaluate the performance of the framework demonstrated that the proposed model significantly improved predictive accuracy, achieved lower MAE and RMSE compared to existing DL-based forecasting models, and more accurately identified solar PV power generation patterns.

Notwithstanding the favorable outcomes, the proposed approach is not without inherent limitations. A principal limitation lies in the model's reliance on the quality and completeness of the input data, which may influence performance under varying environmental conditions. Furthermore, the computational complexity involved in integrating multiple forecasting methods presents a substantial challenge for real-time application scenarios.

Future efforts will concentrate on optimizing the computational efficiency of the forecasting model to enhance its applicability in real-time forecasting scenarios. Furthermore, the researcher intends to incorporate a wider array of environmental factors to enhance the accuracy of the predictions. Additionally, focus will be placed on improving the quality of the data through the use of advanced preprocessing techniques and on investigating the model's adaptability to other forms of renewable energy sources beyond solar PV.

Funding: This study was supported by BK21 FOUR (Fostering Outstanding Universities for Research) (No. 5199990514663) and the Soonchunhyang University Research Fund.

Data Availability Statement: The dataset supporting this study's findings is available from the DKA Solar Centre and can be accessed online at https://dkasolarcentre.com.au/download?location=alice-springs (accessed on 1 April 2024).

Acknowledgments: The author wishes to thank Hwimyeong Ha of LG Energy Solution, Ltd. for his expert insights. His contributions were invaluable to the development and potential application of the technology discussed in this paper. The author would also like to express sincere gratitude to the reviewers and editors for their insightful and valuable feedback, which helped in improving this work.

Conflicts of Interest: The author declares no conflicts of interest.

References

1. Ameur, A.; Berrada, A.; Loudiyi, K.; Aggour, M. Forecast modeling and performance assessment of solar PV systems. *J. Clean. Prod.* **2020**, *267*, 122167. [CrossRef]
2. Mitrašinović, A.M. Photovoltaics advancements for transition from renewable to clean energy. *Energy* **2021**, *237*, 121510. [CrossRef]
3. Gandhi, O.; Kumar, D.S.; Rodríguez-Gallegos, C.D.; Srinivasan, D. Review of power system impacts at high PV penetration Part I: Factors limiting PV penetration. *Sol. Energy* **2020**, *210*, 181–201. [CrossRef]
4. Liu, L.; Zhao, Y.; Chang, D.; Xie, J.; Ma, Z.; Sun, Q.; Yin, H.; Wennersten, R. Prediction of short-term PV power output and uncertainty analysis. *Appl. Energy* **2018**, *228*, 700–711. [CrossRef]
5. Gu, B.; Shen, H.; Lei, X.; Hu, H.; Liu, X. Forecasting and uncertainty analysis of day-ahead photovoltaic power using a novel forecasting method. *Appl. Energy* **2021**, *299*, 117291. [CrossRef]
6. Wu, Y.-K.; Huang, C.-L.; Phan, Q.-T.; Li, Y.-Y. Completed review of various solar power forecasting techniques considering different viewpoints. *Energies* **2022**, *15*, 3320. [CrossRef]
7. Fara, L.; Diaconu, A.; Craciunescu, D.; Fara, S. Forecasting of energy production for photovoltaic systems based on ARIMA and ANN advanced models. *Int. J. Photoenergy* **2021**, *2021*, 6777488. [CrossRef]
8. Antonanzas, J.; Osorio, N.; Escobar, R.; Urraca, R.; Martinez-de-Pison, F.J.; Antonanzas-Torres, F. Review of photovoltaic power forecasting. *Sol. Energy* **2016**, *136*, 78–111. [CrossRef]
9. Louzazni, M.; Mosalam, H.; Khouya, A.; Amechnoue, K. A non-linear autoregressive exogenous method to forecast the photovoltaic power output. *Sustain. Energy Technol. Assess.* **2020**, *38*, 100670. [CrossRef]
10. AlShafeey, M.; Csáki, C. Evaluating neural network and linear regression photovoltaic power forecasting models based on different input methods. *Energy Rep.* **2021**, *7*, 7601–7614. [CrossRef]
11. Konstantinou, M.; Peratikou, S.; Charalambides, A.G. Solar photovoltaic forecasting of power output using LSTM networks. *Atmosphere* **2021**, *12*, 124. [CrossRef]
12. Ziane, A.; Necaibia, A.; Sahouane, N.; Dabou, R.; Mostefaoui, M.; Bouraiou, A.; Khelifi, S.; Rouabhia, A.; Blal, M. Photovoltaic output power performance assessment and forecasting: Impact of meteorological variables. *Sol. Energy* **2021**, *220*, 745–757. [CrossRef]
13. Lateko, A.A.; Yang, H.-T.; Huang, C.-M. Short-term PV power forecasting using a regression-based ensemble method. *Energies* **2022**, *15*, 4171. [CrossRef]
14. Das, U.K.; Tey, K.S.; Seyedmahmoudian, M.; Idna Idris, M.Y.; Mekhilef, S.; Horan, B.; Stojcevski, A. SVR-based model to forecast PV power generation under different weather conditions. *Energies* **2017**, *10*, 876. [CrossRef]
15. Wang, J.; Li, P.; Ran, R.; Che, Y.; Zhou, Y. A short-term photovoltaic power prediction model based on the gradient boost decision tree. *Appl. Sci.* **2018**, *8*, 689. [CrossRef]
16. Ramkumar, G.; Sahoo, S.; Amirthalakshmi, T.; Ramesh, S.; Prabu, R.T.; Kasirajan, K.; Samrot, A.V.; Ranjith, A. A short-term solar photovoltaic power optimized prediction interval model based on FOS-ELM algorithm. *Int. J. Photoenergy* **2021**, *2021*, 3981456. [CrossRef]
17. Alcañiz, A.; Grzebyk, D.; Ziar, H.; Isabella, O. Trends and gaps in photovoltaic power forecasting with machine learning. *Energy Rep.* **2023**, *9*, 447–471. [CrossRef]
18. Chen, C.; Duan, S.; Cai, T.; Liu, B. Online 24-h solar power forecasting based on weather type classification using artificial neural network. *Sol. Energy* **2011**, *85*, 2856–2870. [CrossRef]

19. Almonacid, F.; Pérez-Higueras, P.; Fernández, E.F.; Hontoria, L. A methodology based on dynamic artificial neural network for short-term forecasting of the power output of a PV generator. *Energy Convers. Manag.* **2014**, *85*, 389–398. [CrossRef]
20. Vaz, A.; Elsinga, B.; Van Sark, W.; Brito, M. An artificial neural network to assess the impact of neighbouring photovoltaic systems in power forecasting in Utrecht, the Netherlands. *Renew. Energy* **2016**, *85*, 631–641. [CrossRef]
21. Yang, H.-T.; Huang, C.-M.; Huang, Y.-C.; Pai, Y.-S. A weather-based hybrid method for 1-day ahead hourly forecasting of PV power output. *IEEE Trans. Sustain. Energy* **2014**, *5*, 917–926. [CrossRef]
22. Wang, K.; Qi, X.; Liu, H. A comparison of day-ahead photovoltaic power forecasting models based on deep learning neural network. *Appl. Energy* **2019**, *251*, 113315. [CrossRef]
23. Cui, Z.; Zhou, Y.; Guo, S.; Wang, J.; Xu, C.-Y. Effective improvement of multistep-ahead flood forecasting accuracy through encoder-decoder with an exogenous input structure. *J. Hydrol.* **2022**, *609*, 127764. [CrossRef]
24. DKA Solar Centre. Data Download; DKA Solar Centre: Alice Springs. Available online: https://dkasolarcentre.com.au/download?location=alice-springs (accessed on 1 April 2024).
25. Oh, J.; So, D.; Jo, J.; Kang, N.; Hwang, E.; Moon, J. Two-Stage Neural Network Optimization for Robust Solar Photovoltaic Forecasting. *Electronics* **2024**, *13*, 1659. [CrossRef]
26. Jung, S.; Moon, J.; Park, S.; Rho, S.; Baik, S.W.; Hwang, E. Bagging Ensemble of Multilayer Perceptrons for Missing Electricity Consumption Data Imputation. *Sensors* **2020**, *20*, 1772. [CrossRef]
27. Kim, J.; Moon, J.; Hwang, E.; Kang, P. Recurrent inception convolution neural network for multi short-term load forecasting. *Energy Build.* **2019**, *194*, 328–341. [CrossRef]
28. Xiao, Z.; Xu, X.; Xing, H.; Luo, S.; Dai, P.; Zhan, D. RTFN: A robust temporal feature network for time series classification. *Inform. Sci.* **2021**, *571*, 65–86. [CrossRef]
29. Kim, J.; Kang, H.; Kang, P. Time-series anomaly detection with stacked Transformer representations and 1D convolutional network. *Eng. Appl. Artif. Intell.* **2023**, *120*, 105964. [CrossRef]
30. Vaswani, A.; Shazeer, N.; Parmar, N.; Uszkoreit, J.; Jones, L.; Gomez, A.N.; Kaiser, Ł.; Polosukhin, I. Attention is all you need. In Proceedings of the Advances in Neural Information Processing Systems 30 (NIPS 2017), Long Beach, CA, USA, 4–9 December 2017; p. 30.
31. Jin, Y.; Hou, L.; Chen, Y. A time series transformer based method for the rotating machinery fault diagnosis. *Neurocomputing* **2022**, *494*, 379–395. [CrossRef]
32. Di Gangi, M.A.; Negri, M.; Turchi, M. Adapting transformer to end-to-end spoken language translation. In Proceedings of the INTERSPEECH, Graz, Austria, 15–19 September 2019; pp. 1133–1137.
33. So, D.; Oh, J.; Jeon, I.; Moon, J.; Lee, M.; Rho, S. BiGTA-Net: A Hybrid Deep Learning-Based Electrical Energy Forecasting Model for Building Energy Management Systems. *Systems* **2023**, *11*, 456. [CrossRef]
34. Wang, K.; Qi, X.; Liu, H. Photovoltaic power forecasting based LSTM-Convolutional Network. *Energy* **2019**, *189*, 116225. [CrossRef]
35. Hussain, A.; Khan, Z.A.; Hussain, T.; Ullah, F.U.M.; Rho, S.; Baik, S.W. A Hybrid Deep Learning-Based Network for Photovoltaic Power Forecasting. *Complexity* **2022**, *2022*, 7040601. [CrossRef]
36. Khan, Z.A.; Hussain, T.; Haq, I.U.; Ullah, F.U.M.; Baik, S.W. Towards efficient and effective renewable energy prediction via deep learning. *Energy Rep.* **2022**, *8*, 10230–10243. [CrossRef]
37. Moon, J.; Kim, J.; Kang, P.; Hwang, E. Solving the Cold-Start Problem in Short-Term Load Forecasting Using Tree-Based Methods. *Energies* **2020**, *13*, 886. [CrossRef]
38. Korkmaz, D. SolarNet: A hybrid reliable model based on convolutional neural network and variational mode decomposition for hourly photovoltaic power forecasting. *Appl. Energy* **2021**, *300*, 117410. [CrossRef]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.