Review

A Review on Neural Network Based Models for Short Term Solar Irradiance Forecasting

Abbas Mohammed Assaf [1,*], Habibollah Haron [1,*], Haza Nuzly Abdull Hamed [1], Fuad A. Ghaleb [1], Sultan Noman Qasem [2] and Abdullah M. Albarrak [2]

1 Faculty of Computing, Universiti Teknologi Malaysia, Johor Bahru 81310, Malaysia; haza@utm.my (H.N.A.H.); abdulgaleel@utm.my (F.A.G.)

2 Department of Computer Science, College of Computer and Information Sciences, Imam Mohammad Ibn Saud Islamic University (IMSIU), Riyadh 11432, Saudi Arabia; snmohammed@imamu.edu.sa (S.N.Q.); amsbarrak@imamu.edu.sa (A.M.A.)

* Correspondence: assaf.abbas@graduate.utm.my (A.M.A.); habib@utm.my (H.H.)

Abstract: The accuracy of solar energy forecasting is critical for power system planning, management, and operation in the global electric energy grid. Therefore, it is crucial to ensure a constant and sustainable power supply to consumers. However, existing statistical and machine learning algorithms are not reliable for forecasting due to the sporadic nature of solar energy data. Several factors influence the performance of solar irradiance, such as forecasting horizon, weather classification, and performance evaluation metrics. Therefore, we provide a review paper on deep learning-based solar irradiance forecasting models. These models include Long Short-Term Memory (LTSM), Gated Recurrent Unit (GRU), Recurrent Neural Network (RNN), Convolutional Neural Network (CNN), Generative Adversarial Networks (GAN), Attention Mechanism (AM), and other existing hybrid models. Based on our analysis, deep learning models perform better than conventional models in solar forecasting applications, especially in combination with some techniques that enhance the extraction of features. Furthermore, the use of data augmentation techniques to improve deep learning performance is useful, especially for deep networks. Thus, this paper is expected to provide a baseline analysis for future researchers to select the most appropriate approaches for photovoltaic power forecasting, wind power forecasting, and electricity consumption forecasting in the medium term and long term.

Keywords: Attention Mechanism; Convolutional Neural Network; deep learning; Generative Adversarial Network; hybrid model; solar irradiance forecasting; Long Short-Term Memory

1. Introduction

Recently, research on alternative energy resources has grown in popularity because traditional energy sources are environmentally unfriendly and non-sustainable. Moreover, there is an ever-growing global energy crisis, since economic and technological advancements are highly dependent on the availability of energy that is essential to industrialization and urbanization worldwide. Conversely, the progressive growth of global inhabitancy has exacerbated the severity of the global energy shortages. According to [1], electricity demand is anticipated to grow by up to 70%. In the 20th century, fossil fuels were acknowledged as the main sources of electrical energy generation, and they continue to play an important role today. Despite this fact, the prolonged consumption of fossil fuel supplies, which are already in short supply, adversely affects global climate change (greenhouse effect or global warming) [2,3] and is hazardous to global health [4].

More recently, solar energy has been widely utilized in photovoltaic (PV) power generation. The global PV power capacity may exceed 1700 GW by 2030 [5]. Thus, the generation of PV power is known as another promising renewable energy source because it offers more benefits, including reducing the use of fossil fuels and helping power system

operators to achieve their peak load demand [6]. However, the forecastability of solar PV output is heavily influenced by weather conditions, such as rainy, cloudy, and sunny days; sudden weather changes; snowy days; and other weather types, which present challenges for system administrators. Henceforth, PV power generation must be reliable in terms of accuracy for effective grid function [7,8]. Thus, commercial electric power companies are having difficulties in delivering reliable and safe electricity to their clients because electric power systems require accurate forecasting models to plan their operations. The demand patterns for electricity also include time, the economy, and social and environmental factors [9,10]. Solar irradiance forecasting assists in the incorporation of solar PV plants into the electricity grid, the scheduling of energy storage systems, and energy transmission optimization, which reduces energy loss [11]. It also reduces reserve capacity and cost generation, thus preventing electrical energy system disruption [12], making it easier to predict PV power generation. Over the years, solar irradiance forecasting has gained a lot of interest from both academics and businesses due to the abundant and limitless nature of solar energy that allows it to provide sustainable power globally. On average, the Earth receives $1367 \, W/m^2$ of solar radiation per day, which can yield $1.74 \times 10^{17} \, W$ in a year. Figure 1 illustrates the potential distribution of solar energy globally, inferencing the limitless solar energy supply on Earth. Thus, solar energy is shown to be the best alternative option to ensure the availability of energy sources for industrial, commercial, and residential use [13].
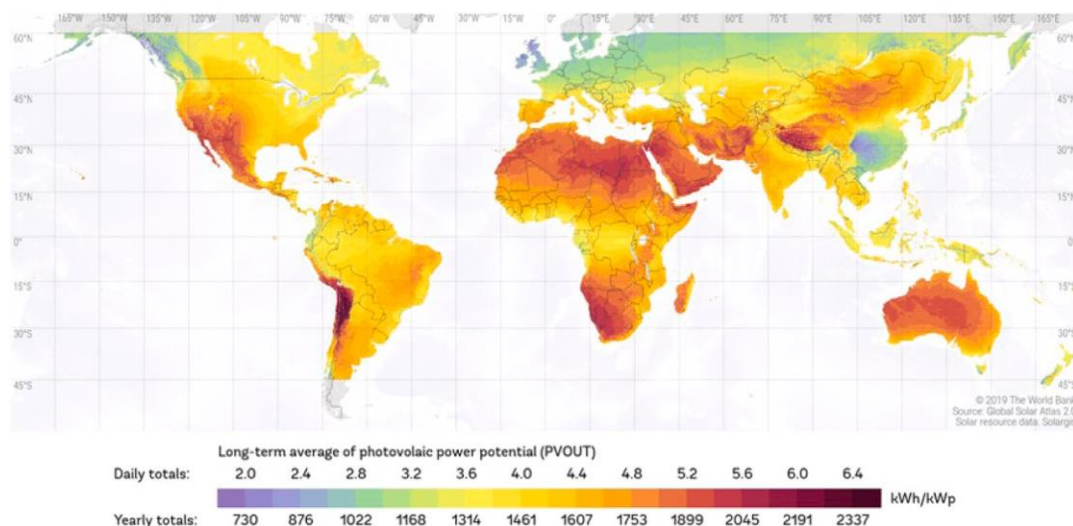

Figure 1. Global distribution of solar resources across the map [14].

The main types of power forecasting models are (1) short-term models, which forecast up to 1 day/week in advance; (2) medium-term models, which forecast up to 1 year in advance; and (3) long-term models, which forecast more than 1 year in advance and are mainly utilized to develop electricity power distribution and power supply systems [15–17]. Sudden changes in solar irradiance, known as ramp events, are relevant for extremely short-term and very short-term forecast horizons. Furthermore, if changes in solar irradiance happen too quickly and strongly, PV power could become less reliable and reduce in quality. Therefore, the results of short-term forecasting may be utilized to determine the most effective PV power ramp rates [18]. Moreover, forecasting for the long term and medium term enables operational enhancement and market engagement [19]. Day-ahead solar irradiance estimates have been shown to enhance yearly energy consumption for commercial operations in building micro grids. Therefore, solar energy forecasts need to be tailored based on specific applications and adjusted using a suitable forecasting approach.

The available literature discusses the solar irradiance forecasting of photovoltaic systems (PV), which uses image-based models for GHI forecasting. According to [20],

this method performs admirably when it is used for solar irradiance forecasting in a targeted region. The high temporal and geographical resolutions of sky images enable the capture of cloud motion. In contrast to empirical methods, image-based models extract cloud information from the sky image dataset to ensure the accuracy of GHI forecasting. However, the drawbacks of image-based models include limited access to image datasets, expensive image capture equipment, and expensive image processing [21]. There are four ways to conduct PV power and solar irradiance forecasting: physical approaches, statistical models, artificial intelligence-based techniques, and hybrid approaches. The physical approaches use satellite images and the Numerical Weather prediction (NWP) model to forecast solar irradiance [22]. The NWP model determines the physical state and solar irradiance using a dynamic atmospheric model. Moreover, the NWP model incorporates both geospatial information and meteorological data [23]. Even though physical models are commonly utilized to forecast atmospheric dynamics, their complexity is a major setback. This problem occurs due to the fact that physical models require higher computational resources and more time to handle massive datasets. Consequently, the use of physical models for estimating short-term solar irradiation is considered ineffective [24]. According to the literature, several other statistical models have been utilized, including autoregressive moving averages [25] and dynamic systems with coupled autoregressive features [26]. Moreover, statistical models are effective when using continuous time-series data, even though these data are not constant because of the effects of season and weather changes. Although the use of statistical models is an excellent option for forecasting solar irradiation over time series, they are unable to incorporate non-linearity into the data, hence the minimal accuracy of forecasts [27]. Moreover, it is challenging to forecast solar energy using physical and statistical methods due to the massive amount of weather data involved.

There has been a significant increase in interest in solar radiation forecasting research. By examining the data presented in Figure 2 from the Web of Knowledge website (WoS), it is evident that the number of published papers in this field has experienced seven-fold growth over the past decade (2010–2023) compared to the previous decade (2000–2010). Notably, within the last three years alone (2020–2023), there have been 1300 papers published, compared to the 2159 papers published between 2010 and 2019.
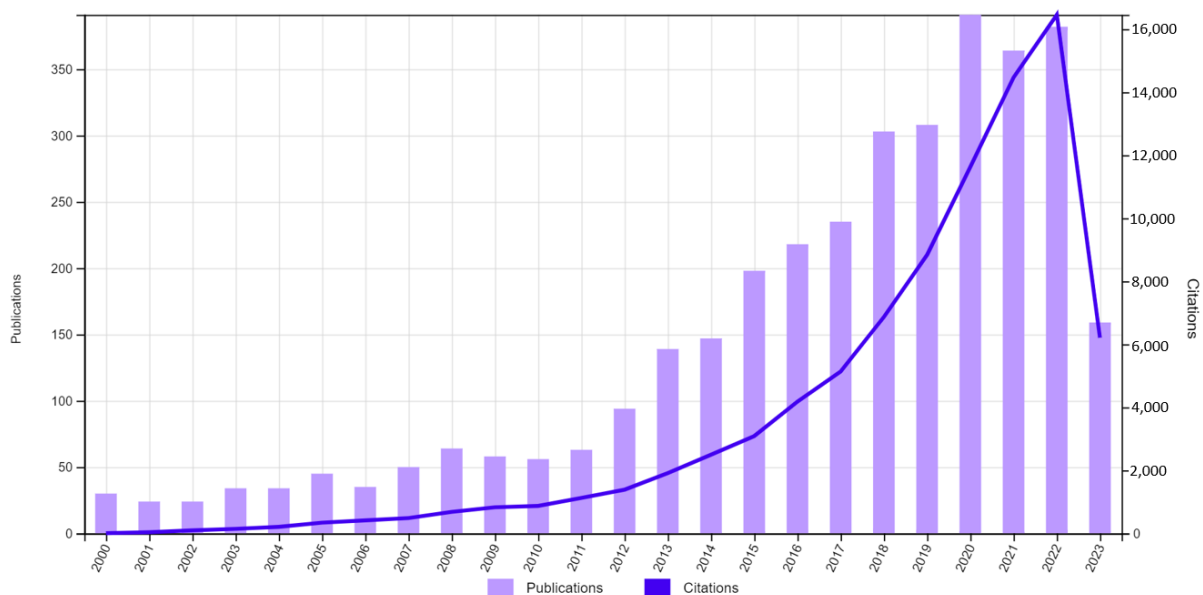

Figure 2. Publication frequency and citations report according to WoS website.

Recently, the problem of forecasting solar irradiance was addressed using machine learning algorithms, which have shown superior outcomes compared to the statistical and physical approaches related to quality and dependability. One of the superior aspects

of the machine learning method is the fact that it can extract non-complex features from large amounts of high-dimensional data [28], as well as be utilized to extract non-complex features. In [29], the time-dependent characteristics of solar irradiance are identified, and an ensemble-learning-based solar radiation forecasting model is provided. Moreover, types of artificial intelligence techniques used include the extreme learning machine (ELM) [30], the support vector machine (SVM) [31], the wavelet transform [32], deep learning [33], the artificial neural network (ANN) [34], and ensemble learning [35]. Ref. [36] utilized ANN and SVM to forecast solar radiation at two different locations in Saudi Arabia on a tilted surface.

The advancements in data acquisition and storage technologies have made it possible for meteorological stations and photovoltaic power plants to collect massive amounts of data samples. Nonetheless, establishing a research question is still the best way to use these samples. While some machine learning models are unable to handle high-dimensional inputs or large datasets effectively, others even deal with them at all. However, the theory of deep learning is rapidly developing, especially in unsupervised feature extraction [37]. Moreover, deep learning models represent the most suitable approach for solar irradiance forecasting, especially when it involves a complex and substantial amount of data. In contrast, machine learning models' efficiency decreases as the quantity of input data increases. Ref. [38] investigated the impact of altering the input size on deep learning and ML models. The findings demonstrated that the performance of traditional machine learning models was stagnant after a specified quantity of data, whereas deep learning models continued to be enhanced when they received more training data. Deep learning models thrive in a variety of areas, including image processing, pattern extraction, classification, and forecasting, since they are efficient in learning complex patterns data without human expertise. For example, building thermal load forecasting using deep learning models was utilized in [39]. Before that study, the deep learning model was introduced by [40] for short-term wave energy forecasting. Likewise, deep learning models have been extensively used for various forecasting applications, such as wind speed [41], PV power [42], solar irradiance [43], and others. The known problems associated with Conventional Neural Networks (CNN), such as gradient vanishing and training complexity, can be simply resolved using deep learning networks. To forecast time series, [44] developed a sophisticated neural network, while [45] utilized deep learning to forecast solar radiation at 30 locations in Turkey. Thus, deep learning models are better than machine learning and empirical models in terms of accuracy. This fact is due the structure of the hidden layers; however, training takes more time to complete because there are more parameters to adjust as the architecture becomes more complex. Although deep learning may yield positive outcomes, the training process is typically challenging; therefore, researchers often merge deep learning with feature extraction methods, like Attention Mechanism, to enhance the efficiency of feature extraction. Subsequently, the neural network that was paired with Attention Mechanism was able to perform most effectively, improving neural network interpretability by eliminating irrelevant data and only selecting more important information, as well as differently weighting each feature of the meteorological data.

Ref. [33] proposed Convolutional Neural Networks and Generative Adversarial Networks that categorize the weather into 10 categories. Enervative adversarial networks increase the training dataset for every weather category. The simulation outcomes demonstrate that Generative Adversarial Networks have the potential to build high fineness quality specimens that capture the underlying properties of the original data. A data augment based on Generative Adversarial Networks optimizes all categorization models. Thus, GAN is ideal for data creation to augment and rebalance the training dataset. On the other hand, researchers have rarely used GAN to create time-series data (i.e., hourly GHI and meteorological data). Indeed, there is a significant need for sufficient training data for PV power models, which might come from either an actual dataset or augmentations.

Despite solar energy being the most prominent renewable energy resource, few review papers have been published on this topic. Thus, ref. [46] presented a review analysis of an

ANN-based model used for solar radiation forecasting. Similarly, ref. [47] analyzed a total of 87 journals that implement ANN for solar power forecasting. Before that study, a review analysis of machine learning-based solar radiation forecasting models was conducted [48]. Recent work by [49] investigated solar irradiance forecasting using the SVM application. In 2019, ref. [50] examined the application of 70 sky solar irradiance models. Following that study, ref. [49] provided a review paper on physical models used to predict solar irradiation. Next, ref. [13] presented deep learning-based solar radiation forecasting models.

According to statistics, deep learning-based models has been extensively explored for forecasting model, but there is still no available publication that concentrated on investigating the generation of meteorological data and its effect on the forecasting accuracy of PV power. Most previous studies focused on developing forecasting models, which suffer from inaccuracy, particularly for extreme weather types, for which models use small datasets that do not give enough representative training specimens for Deep Neural Networks. As this review study focuses on generating and balancing meteorological data to improve forecasting performance, it also paves the way for future research in a variety of areas, including photovoltaic power forecasting, wind power forecasting, electricity consumption forecasting in the medium term and long term, etc., by expanding and embracing the modern technique suggested in this work through data generation. In contrast, most studies of photovoltaic power, wind power, and energy consumption concentrate on the short-term forecast, disregarding medium- and long-term forecasts because of a lack of training data for models.

In sum, neural networks and deep learning-based models have been extensively explored for solar irradiance forecasting. Although there have been many review papers published recently on solar irradiance forecasting, such as [1–6], there is a lack of a comprehensive review paper that studies the significance of deep learning models for solar irradiance forecasting. A deep analysis of the effects of deep feature extraction on forecasting performance has been conducted, as many past studies on solar irradiance forecasting tended to overlook how to successfully extract features from meteorological data. This review paper focuses on reviewing recent and effective deep learning models that have achieved notable performance. Additionally, this study has covered forecasting impacts that are often neglected by existing reviews. It provides a comprehensive review that covers various aspects of solar energy forecasting related to deep learning models. It can also serve as a valuable reference point for researchers developing deep-learning solar energy solutions.

The main aim of this paper is to address the limitations of existing reviews in terms of covering aspects such as the main variables influencing solar irradiance forecast, including the forecasting horizon, input parameter optimization, weather classification, and others. Furthermore, this paper outlines the application of CNN, RNN, GRU, AM, LSTM, DBN, Attention Mechanism, and GAN for solar irradiance forecasting. It presents factors that influence the accuracy of the solar irradiance forecasting models, in addition to providing a background introduction and discussing prominent deep learning techniques. The review emphasizes the limitations of existing solutions and provides recommendations for future work.

2. The Impact Factor of Solar Power Forecasting

Several factors can influence the effectiveness of solar irradiance forecasting, including forecasting horizon, weather classification, performance evaluation metrics, and input feature optimization. The related factors are further elaborated in the following subsections.

2.1. Forecasting Horizon

Forecasting models for solar energy can be classified based on their forecasting horizon. A power generation operator must be aware of future power generation and consumption demands. The use of solar irradiance forecasting relies heavily on the types of forecasting horizon to ensure its efficient installation of various applications, including the execution

of PV power plants. Specifically, the operation of the grid requires varied time horizons to achieve grid stability, spinning reserve scheduling, and unit commitment [51]. There are three main types of forecasting horizons: (1) short-term, (2) medium-term, and (3) long-term horizons. In addition, limited studies included a fourth type, which is referred to as "extremely short-term forecasting". Currently, there is no general taxonomy available for forecasting horizon; therefore, the list is presented as follows.

- Intra-hour forecasting, also known as nowcasting, is a very short-term forecasting horizon that ranges from 1 min to many minutes in advance and is often used in electricity pricing, bidding, real-time power system dispatch monitoring, and peak load matching [52].
- Short-term forecasting horizons range from 1 or several hours to 1 day or week in advance [2] and are critical for optimal unit commitment, rotating reserve control, and analyzing sales/purchase contracts between multiple enterprises. Therefore, they facilitate the development of a PV-integrated energy management system and increase grid security.
- Medium-term forecasting horizons range from 1 month to 1 year in advance [53] and are efficient for modeling the maintenance schedule of solar power plants developed using transformers and other facilities that incur the least amount of loss [54].
- Long-term forecasting horizons range from 1 year to 10 years in advance. This distance is ideal for designing long-term plans that allow the creation of effective solar power facilities and global management, such as site selection, required to develop a PV power plant and ensure transmission, operation, and distribution of solar energy [55]. However, it is not perfect for long-term forecasting due to its inability to predict long-term weather changes. However, it is still deemed to be the most suitable model for designing schedule plans, setting prices, and selecting site strategies [56,57].

The forecasting horizon is one of the most influential factors for forecasting models in terms of accuracy, even when its input features remain constant [58]. In such a case, ref. [59] utilized SVM to generate a very short-term forecasting model. Their findings included a decrease in forecasting accuracy for a forecasting horizon ranging from 5 to 30 min from 96 to 64.6%, respectively, for similar datasets. The growth of statistical data at shorter time scales complicates the training process rather than improving accuracy. Several interrelated variables affect the accuracy of solar irradiance forecasting, such as very short time scales (including one sec or one min period), cloud conditions, and intensity of solar radiation towards the Earth. It is critical to focus on types of cloud classification and forecasts to achieve effective solar irradiation forecasting. However, existing researchers often do not consider cloud conditions when developing their models, which eventually leads to poor forecast accuracy [60–62].

2.2. Weather Classification

Solar radiation is known as the key feature used in the determination of PV power potential. Various meteorological input features influence the availability of solar irradiance, such as cloud types, temperature, relative humidity, wind speed, pressure, and aerosol index. Based on this fact, weather changes can be considered as the main factor that influences solar irradiance forecasting models in terms of accuracy. This result demonstrates that weather classification is critical for forecasting algorithm performance and durability. According to multiple studies, weather classification is the main factor involved at the short-term solar irradiance forecasting pre-processing stage [63–65]. The lack of data available for model training is considered to be a major roadblock in weather classification. A study reported using 33 weather types that were classified into 10 weather classifications by combining several types of weather into a single type [66]. This limitation can be addressed by classifying weather into three or four types [67,68]. Accordingly, ref. [34] fed surface weather and irradiance data into K-means clustering to analyze variations in cloud conditions. In addition, regime-dependent ANN models were utilized for solar irradiance forecasting and demonstrated better accuracy. Following that study, gew researchers

have concentrated on using Self-Organizing Maps (SOM) to classify meteorological input features into different categories for forecasting models [61,69]. However, before that study, ref. [66] presented four major weather categories based on solar irradiation features for each weather type [66]. In addition, ref. [70] predicted solar irradiance in northeastern Brazil based on an Artificial Neural Network (ANN). The data were divided into two climates—dry and rainy—and the clustering method was used to identify the homogenous climatic zones present.

The comparisons between the GAN-based approach and the other five established forecasting models were presented in [33]. The study adopted 10 weather classes as the input data, and a confusion matrix was utilized for objectivity [71]. The literature makes several comparisons between weather categorization and photovoltaic power forecasting (PVPF) models, making it difficult for PVPF researchers to select the best option. This statistical tool analyses the effectiveness of classifier algorithms using a table layout to determine how frequently a classifier becomes confused and mislabels two weather classes that are close together. These classifiers include CNN with a 1-D convolution layer (CNN1D), CNN with a 2-D convolution layer (CNN2D), multilayer perceptron (MLP), and SVM with a k-nearest neighbor (KNN). The findings showed better accuracy in weather classification and data augmentation, as well as improved data imbalance weather categories using GAN, especially for short-term sample size. The authors suggested GAN-CNN2D for PVPF modeling, although CNN2D achieved the highest scores due to its ability to eliminate non-linear input–output correlations. Correspondingly, the previous literature emphasized the importance of using weather categorization to ensure the effectiveness and accuracy of solar irradiance forecasting models, as weather conditions should always be prioritized for forecasting [52,72,73].

2.3. Model Performance Metrics

The performance evaluation metrics are useful throughout the model development process because it is impossible to analyze how effectively deep learning models would perform without the use of metrics for comparison [6]. However, performance can be affected by many factors, such as the forecasting horizon, model parameters, and site-specific meteorological circumstances. Specifically, the overall performance is assessed based on a comparison between actual and forecasted solar irradiance. The metrics used will provide feedback in terms of forecasting accuracy, allowing models to be fine-tuned to achieve a target degree of precision; a lower score indicates more precise forecasting. Commonly used statistical evolution metrics for evaluation purposes include the following:

- Mean Absolute Error (MAE): this metric demonstrates the mean of the absolute errors among the actual GHI values and the anticipated values, and it provides equal weight distribution to all data inconsistencies, which can be shown in Equation (1) [74]:

$$\text{MAE} = \frac{1}{N} \sum_{i=1}^{n} |Ga_i - Gp_i| \tag{1}$$

where Ga_i and Gp_i denote the actual and predicted GHI values, respectively, while the total number of data points is represented by n.

- Mean bias error (MBE): This metric evaluates the established mean bias of the forecasting model. The MBE adjusts the deviation by subtracting the positive from the negative, as shown in Equation (2). It is considered an unreliable method for performance evaluation; however, it does provide an adequate indication of whether a model is over- or under-estimated [74].

$$\text{MBE} = \frac{1}{N} \sum_{i=1}^{N} (Gp_i - Ga_i) \tag{2}$$

- Mean square error (MSE): This metric is the mean calculation of square differences between the actual and forecasted solar irradiance levels, as shown in Equation (3). Therefore, larger disparities occur [75].

$$\text{MSE} = \frac{1}{N}\sum_{i}^{n}(Gp_i - Ga_i)^2 \qquad (3)$$

- Root mean square error (RMSE): This metric is the calculation of the square root of the average squared variances between actual and forecasted solar irradiance values. RMSE is recognized as being one of the most trusted ways to measure performance because it helps to identify and eliminate data outliers. The metric is expressed via Equation (4) [76]:

$$\text{R} = \sqrt{\frac{1}{N}\sum_{i}^{n}(Gp_i - Ga_i)^2} \qquad (4)$$

- Mean absolute percentage error (MAPE): this metric is the relative average value between MAPE and MAE, which is obtained by dividing the differences between each actual and forecasted observed value by the actual observed value, as shown in Equation (5) [77].

$$\text{MAPE} = \frac{1}{N}\sum_{i=1}^{N}/\frac{Gp_i - Ga_i}{Ga_i}/ \qquad (5)$$

- Normalized RMSE (nRMSE): in general, nRMSE is used to identify overall deviance in larger datasets, as shown in Equation (6) [77].

$$\text{nRMSE} = \frac{\sqrt{\frac{1}{N}\sum_{i}^{n}(Gp_i - Ga_i)^2}}{\overline{Ga_i}} \qquad (6)$$

- Correlation coefficient (R): this metric measures the linear correlation between actual and forecasted solar irradiance [78], and it is calculated as shown in Equation:

$$\text{R} = \frac{\sum_{i}^{n}(Ga_i - \overline{Ga_i})(Gp_i - \overline{Gp_i})}{\sqrt{\sum_{i}^{n}(Ga_i - \overline{Ga_i})\sum_{i}^{n}(Gp_i - \overline{Gp_i})^2}} \qquad (7)$$

$\overline{Gp_i}$ and $\overline{Ga_i}$ represent the average of forecasted and actual GHI, respectively.

- Forecast skill score (FS): the FS score compares the performance of the forecasting model and the benchmark or persistence model, as shown in Equation (8):

$$\text{FS} = 1 - \frac{\text{RMSE}_{\text{model}}}{\text{RMSE}_{\text{simple model}}} \qquad (8)$$

In terms of accuracy, an optimal forecast model will achieve an FS score of 1, whereas a model that exhibits the same forecast error as the benchmark model has a value of 0. The FS score will produce negative results with a larger forecast error [79].

The forecast skill score is represented by FS, while $\text{RMSE}_{\text{model}}$ and $\text{RMSE}_{\text{simple model}}$ represent the RMSE values of the recommended forecasting model and benchmark or persistence model, respectively.

2.4. Optimization of Input Feattures

The forecasting can be improved using the appropriate input selection, both in terms of quantity and type. The presence of large parameters can significantly imbalance the forecasting, whereas redundant or weakly correlated inputs will further complicate the

computation. Hence, optimization techniques are critical to the selection of suitable input features.

Many input optimization strategies for PV output models are described in the literature, including PSO [80], grid-search [81], fruit fly optimization algorithm (FOA) [82], firefly (FF) [8], ant colony optimization (ACO) [83], chaotic ant swarm optimization (CAS) [84], chaotic artificial bee colony algorithm [85], and immune algorithm (IA) [86]. While each algorithm has benefits and drawbacks, the most prominent and efficient technique for improving weights and inputs during forecasting is genetic algorithm (GA)-based optimization. Moreover, it is practical for pairing with ANN. Tao and Chen [87] improved the forecast accuracy of the back propagation neural network (BPNN) model by altering the input weights. Similarly, ref. [88] claimed that GA optimization improved their ANN for the PV power forecasting model. According to [89], PSO is a method of iterative optimization that is comparable to GA. The PSO method starts with a collection of random solutions and discovers the best values iteratively. Compared to GA, it has more convergence and does not require the modification of as many features [90,91]. PSO is used as an alternative for GA in applications such as function optimization, feature selection and neural network training, fuzzy system control, and others; however, it tends to be dismissed in local extrema.

3. The Architecture of Deep Learning Algorithms

We discuss the architecture of prominent deep learning models for solar irradiance forecasting, which include Artificial Neural Networks (ANN), Gated Recurrent Units (GRU), Recurrent Neural Networks (RNN), Deep Belief Networks (DBN), Convolutional Neural Networks (CNN), and Long Short-Term Memory (LSTM), as well as the hybrid models that were discussed in the literature. In this section, detailed information on the architecture and training techniques of related models is provided.

3.1. Artificial Neural Network (ANN)

The structure of the neural network is modeled on the human brain, and it has made significant contributions to the evolution of machine learning technologies. It is considered a simple mathematical model that can be used to address a variety of non-linear issues. Previous researchers reviewed ANN models in forecasting solar energy, such as Ref. [46] for solar radiation forecasting, Ref. [92] for PV applications, and Ref. [93] for solar energy forecasting. They concluded that the ANN model can offer higher accuracy for solar irradiation forecasting than other conventional models. Figure 3 shows the neural network architecture that consists of a multi-layer perceptron [94]. This multi-layer perceptron has aspects known as the input, hidden, and output layers. There are also auxiliary components that are labeled as neurons, weight, bias, and activation functions. These layers communicate with each other such that the input layer receives input values, which are then analyzed via the hidden layer. Next, the output layer gathers data from the hidden layer and decides how to use it. The training process of the neural network is iterative, and it modifies the structure during the learning process to achieve the exact reference or set point as the supervisor [95]. The network's multiple hidden layers allow it to forecast exact feature extraction and the non-linear structure of a model [96]. The findings of the ANN model are compared to those of the ARMA model (Auto Regressive Moving Average) and demonstrate how an ANN model can estimate global solar irradiation.

Equation (9) is a basic mathematical function for ANN:

$$An = \sum_{i=1}^{n} (w_i * I_i) + b \qquad (9)$$

where w_i represents the weight, I_i represents the input, A_n represents the result, b represents the bias, and n represents the number of inputs.
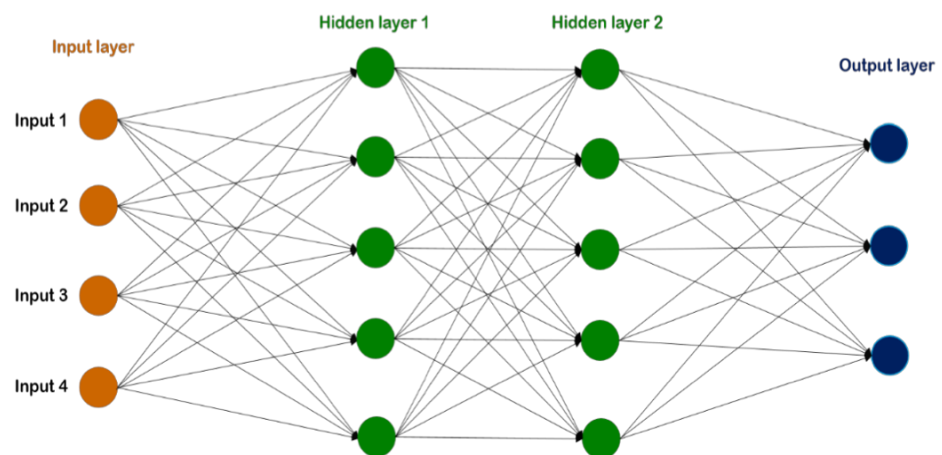
Figure 3. The architecture of a neural network [94].

The activation function is a mathematical equation that governs how neurons produce information, and it is interchangeably known as the transfer function. Hypothetically, it can influence the output value. The activation function types are linear and non-linear. For both input and output layers, a linear activation function produces the exact linear results. However, such a linear connection is insufficient for actual implementations, since problems involving complex information and several elements, which include image, video text, and sound, require more processing. The limits of the linear activation functions can be circumvented using a neural network with a non-linear activation function. Table 1 lists the most used activation functions of neural networks. The leaky ReLU and Rectified Linear Unit (ReLU) are non-linear activation functions, since their slope is not constant for all values. Positive values have a slope of 1, whereas negative values have a slope of 0 [94].

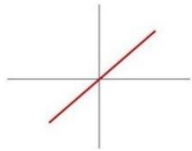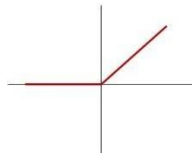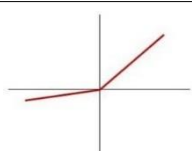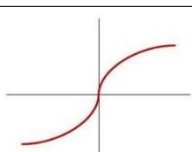**Table 1.** The activation function of the neural network [94].
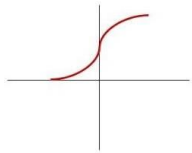
Activation Function	Equation	Plot
Linear	$f(x) = x$	
Relu	$f(x) = \max(0, x)$	
Leaky Relu	$f(x) = \max(0.1 \cdots x, x)$	
Tanh	$f(x) = \tanh(x)$	

Table 1. *Cont.*

Activation Function	Equation	Plot
Sigmoid	$f(x) = \frac{1}{1+e^{-x}}$	

3.2. Convolutional Neural Network (CNN)

CNN is considered to be robust for various applications, since it can learn high-level features from sequences without any human intervention [97]. The CNN algorithm is widely utilized in deep learning. The most notable advantage of CNN is its remarkable ability to capture non-linear information. CNN comprises three layers: the convolutional, pooling, and fully connected layers [98,99]. The most significant component of CNN is the convolution layer, which consists of many convolution kernels used to construct new feature maps. Figure 4 depicts the architecture of CNN, which consists of convolutional, pooling, and fully linked layers [49]. The following tasks are assigned to each layer: The convolution layer is involved in local feature extraction, which is effective when the kernel weights are equally distributed across all input maps. The convolutional layers are utilized to extract spatial patterns from the underlying training sample and its related input variables. The mathematical formula for the convolutional layer is expressed via Equation (10).

$$y^k{}_{i_k} = f((w^k * h)i_j + b_k) \tag{10}$$

where f represents an activation function, * represents a convolutional process operator, and w^k represents the weight of the kernel.
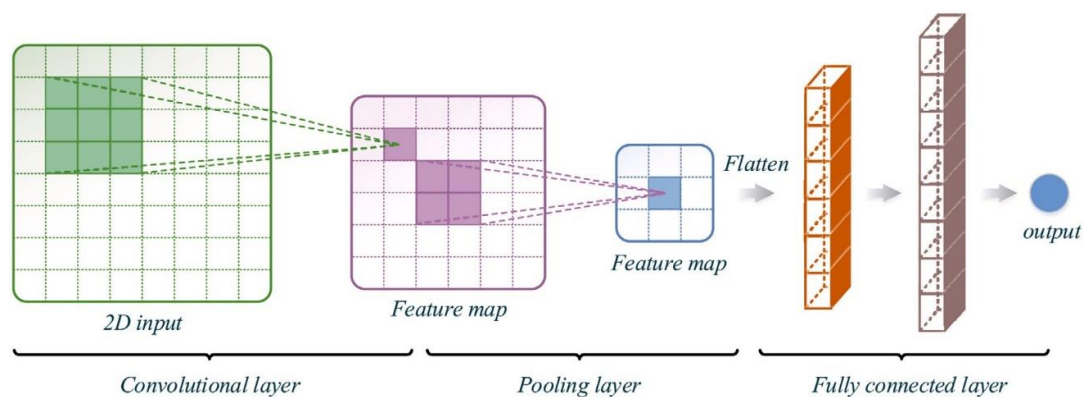
Figure 4. The architecture of CNN [49].

A pooling layer is commonly utilized to minimize the in-plane dimensionality of input maps, hence minimizing the number of learnable parameters and preventing overfitting. Typically, pooling layers can be divided into two types: maximal pooling and average pooling. Both types are used in the procedures outlined, as shown in Equation (11):

$$f(x) = \max(0, x) \tag{11}$$

The fully connected layer is utilized extensively for high-level inference, since it is responsible for transferring the features filtered by the convolutional and pooling layers to the output layer. Next, linear activation and non-linear activation functions are employed for both layers. Finally, a full connection layer is added to the CNN structure to forecast the output using the extracted features [100]. CNN is commonly used in the development of data pattern-based filters and the extraction of hidden features. More importantly, the key

characteristics of Convolutional Neural Networks include local connections and weight sharing [101,102].

3.3. Recurrent Neural Network (RNN)

RNN lies under the classification of Artificial Neural Network and functions substantially effectively for training sequential or time-series data. Time-series data consist of intrinsic temporal information that a Simple Neural Network cannot capture [103]. RNN decomposes sequence data into components and maintains a state to represent the data at different time intervals [104]. Figure 5 highlights the structure of RNN, which is composed of inputs, hidden neurons, and an activation function.
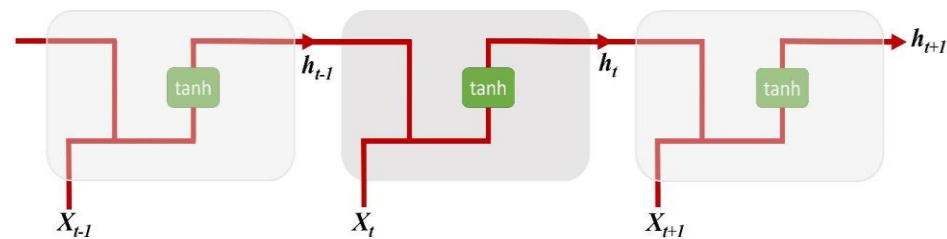

Figure 5. The architecture of RNN [94].

The previously hidden layer (h_t) is presented as shown in Equation (12):

$$h_t = \tan h \ (U \cdot x_t + W \cdot h_t - 1) \tag{12}$$

where x_t denotes the input at time t, h_t denotes the hidden neuron at time t, U denotes the weight of the hidden layer at time t, and W denotes the transition weights of the hidden layer. The input and prior hidden states are merged, as the current and past inputs flow uses the tan$h$ function to provide information. The output is a new hidden state that functions as a Neural Network Memory by storing knowledge from past efforts.

During the training phase, RNN is frequently troubled with gradient vanishing and gradient explosion difficulties. Once backpropagation has been closed at a specific point, the gradient explosion problem can be solved. However, the outcome is suboptimal, since not all the weights are updated. The initialization of weights can help to modify the vanishing gradient to decrease its chances of affecting the model. Moreover, there is another solution to the problem, which uses the LTSM and will be explained in the next section.

3.4. Long Short-Term Memory (LSTM)

The LSTM network is considered to be an extended and improved structure of RNN that can effectively solve its time-series forecasting issues and inability to deal with long-term data dependencies because of vanishing gradient and gradient explosion issues [105]. Therefore, LSTMs are frequently utilized in power systems for time-series forecasting, including load forecasting, demand response, and renewable power generation forecasting [106]. The LTSM network concept was introduced by [43] to resolve the limitation of a vanishing gradient. This approach can be seen in Figure 6, where the LTSM network consists of inputs, outputs, memory cells, and forget gates that work together to regulate information flow. The forget gate separates information into two categories: data that should be deleted and data that should be retained. The input gate is responsible for updating the cells, whereas the output gate determines their next concealed state. The gates use the sigmoid function as their activation function, and they will generate a value between 0 and 1 to selectively enable information to pass through them. The gates can be used to indicate how much data from the current input is allowed to pass through the structure. As previously stated, the whole structure serves as a gate; when it is open, all information is permitted to pass through it (sigmoid output is one). On the other hand, when the

gate is closed (having a sigmoid output of zero), no information can pass through it. The mathematical description and information distribution are described in Equation (13):

- The forget gate, f_t, helps the LSTM to determine which information should be eliminated from the cell state by applying a sigmoid function to the output of the prior h_{t-1} and input data x_t:

$$f_t = (w_f.h_{t-1}, w_f.x_t + b_f) \tag{13}$$

where (.) signifies the sigmoid activation function, w_t signifies the weight matrix, h_{t-1} signifies the prior state, x_t signifies the memory cell of the input vector at time t, and b_f signifies the bias vector.
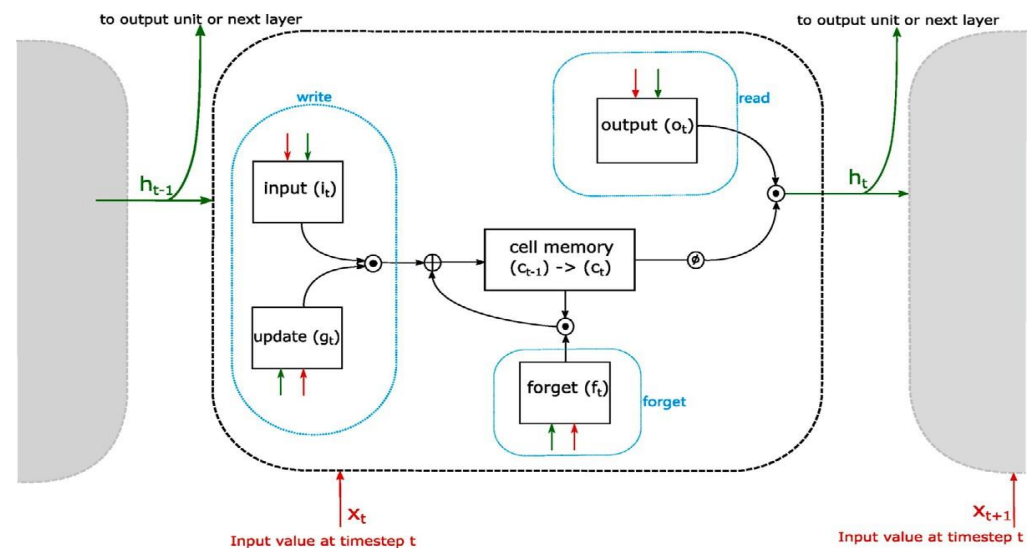
Figure 6. The architecture of the LSTM network [107].

The input gate, i_t, determines which data are stored in the new candidate cell state $\widetilde{C}_t$. The hyperbolic tangent function is symbolized as tan*h* () in Equation (14).

$$\widetilde{C}_t = \tan h\ (w_c.h_{t-1}, w_c.x_t + b_c) \tag{14}$$

$$I_t = (w_i.h_{t-1}, w_i.x_t + b_i) \tag{15}$$

The input gate applies the sigmoid function, (σ), to determine which values to write, and it updates the gate using the tan*h* activation function to generate new cell values.

The new candidate's cell state $\widetilde{C}_t$ and the prior cell state C_{t-1} are incorporated to update the latest cell state, C_t, which is expressed as shown in Equation (16):

$$C_t = I_t * \widetilde{C}_t + f_t * C_{t-1} \tag{16}$$

Finally, the output gate regulates the output of a cell and merges it with a cell state that is activated via the tan*h* function to determine the final output, h_t, which is expressed in the Equations (17) and (18):

$$O_t = \sigma\ (w_o.h_{t-1}, w_o.x_t + b_o) \tag{17}$$

$$h_t = \tan h\ (C_t).\ O_t \tag{18}$$

3.5. Gated Recurrent Unit (GRU)

The GRU is considered to be a less sophisticated RNN [108] than LSTM, which can be more simply applied and computed. GRU and LSTM are equivalent in terms of their capacity to recall and capture long-term dependencies. The computational time is faster and the complexity is lower in GRU compared to LSTM, since there are fewer features [67]. Moreover, GRU can effectively learn long-term dependencies data, which resolves the issue of vanishing gradient that arises when using simple RNN. Furthermore, GRU is deemed as a variant of LSTM due to the similarities between its function mechanism and design. This observation can be seen when both GRU and LSTM use the gate mechanism to control the flow of information by combining the input and forget gates into a single update gate. However, unlike LSTM, GRU has only two gates, which are the update gate (z_t) and the reset gate (r_t). These two gates determine what information from the past should be preserved for the future and what past information should be deleted, respectively. Furthermore, GRU models suffer from delayed convergence and low learning efficiency, resulting in extensive training and even under-fitting. Figure 7 depicts the structure of GRU, and the relationship between its inpute and outpute is presented in Equations (19)–(21):

$$R_t = \sigma\ (U_r \cdot h_{t-1} + b_r + W_r \cdot X_t) \tag{19}$$

$$z_t = \sigma\ (b_z + W_z \cdot X_t + U_z \cdot h_{t-1}) \tag{20}$$

$$A_t = \tan h\ (W_h \cdot X_t + U_h \cdot b_h + (r_t * h_{t-1})) \tag{21}$$

$$h_t = (1 - Z_t)\ *\ h_{t-1} + Z_t * A_t \tag{22}$$
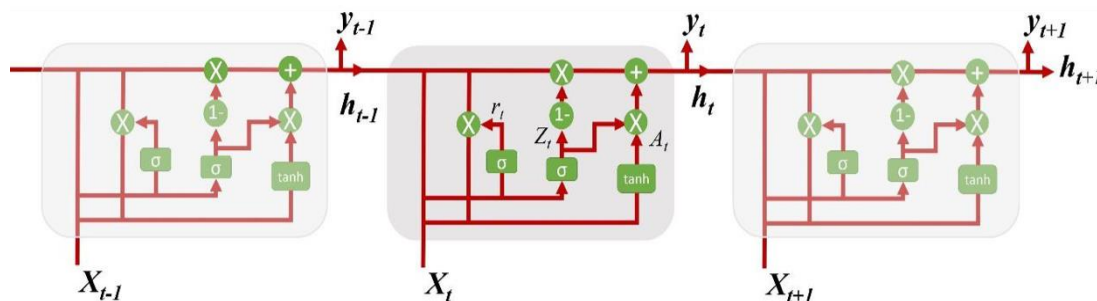
Figure 7. The architecture of the GRU network [94].

At the current time step (*t*), the reset gate is denoted as R_t, the update gate is denoted as z_t, the memory component is denoted as A_t, the activation function is denoted as $\tan h$, and the final memory is denoted as h_t.

3.6. Bidirectional RNN (BRNN)

In general, Normal Recurrent Neural Networks (RNN) are limited in terms of their input data flexibility, since future input data are omitted from the current state. Ref. [109] developed the Bidirectional Recurrent Neural Network (BRNN) to address the limitations of RNNs. BRNNs can be trained using data from both the past and the future at any given time step, since they can receive input from both directions. Figure 8 illustrates the basic structure of the BRNN. The mathematical expression is presented in Equations (23)–(25) [110]:

$$\overrightarrow{h}_t = f\left(W_{x,\overrightarrow{h}}X_t + W_{\overrightarrow{h},\overrightarrow{h}}\overrightarrow{h}_{t-1} + b_{\overrightarrow{h}}\right) \tag{23}$$

$$\overleftarrow{h}_t = f\left(W_{x,\overleftarrow{h}}X_t + W_{\overleftarrow{h},\overleftarrow{h}}\overleftarrow{h}_{t-1} + b_{\overleftarrow{h}}\right) \tag{24}$$

$$y_t = g\left(w_{\overrightarrow{h},y}\overrightarrow{h}_t + W_{\overleftarrow{h},y}\overleftarrow{h}_t + b_y\right) \tag{25}$$
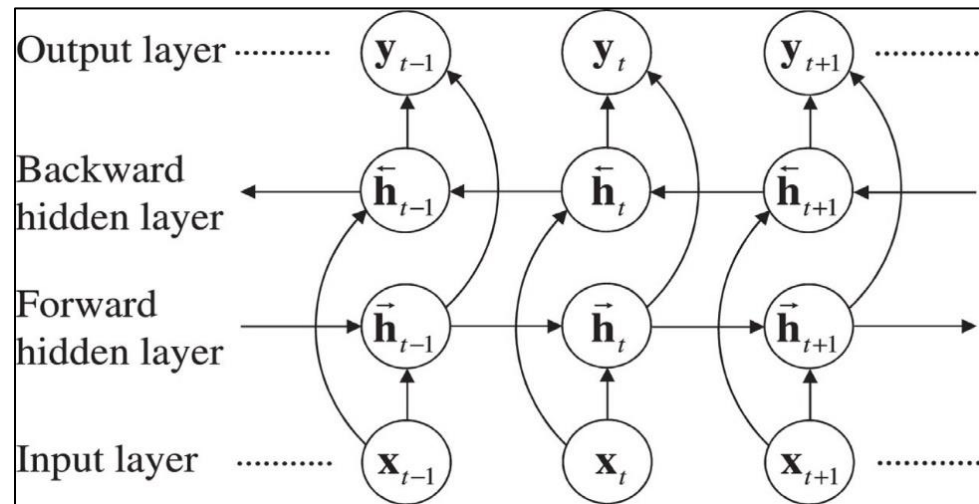

Figure 8. The architecture of BRNN [110].

The input vector at time step t is signified by x_t, the forward hidden layer activation vector at time t is signified by $\overrightarrow{h}$, the backward hidden layer activation vector at time t is signified by $\overleftarrow{h}$, the weight matrix is signified by w, the bias is signified by b, and the output probability is signified by y_t. The probability of the output at time t is signified by y_t. The activation function of each node in the hidden layer is signified by f, while the SoftMax function is signified by g.

3.7. Bi-LSTM Network

The LSTM was modified into the Bi-LSTM, which was presented by [109], to enhance forecasting accuracy by combining forward and backward information from the input sequence. The forward hidden sequence was calculated first, and the reverse hidden sequence was then merged to find the result. The architecture of Bi-LSTM was enhanced by integrating the best features of BRNN and LSTM to improve the accuracy of the state. Furthermore, double processing was beneficial for learning complex temporal correlations.

3.8. Deep Belief Network (DBN)

Geoffrey Hinton invented a Deep Belief Network (DBN) in 2006, and it has since been employed in a variety of applications [111]. DBN represents probabilistic generative graphical models that are developed using probability and unsupervised learning. They solve various typical neural network problems, such as slow convergence and learning rate for local minima due to insufficient input selection. It is also a deep learning network that is built based on the stacking mechanisms of multiple Restricted Boltzmann Machines (RBM). Figure 9 illustrates that each of the stacked RBMs are made up of two layers: (1) the visible layer (v) and (2) the hidden layer (h_i). The hidden layer of each RBM acts as the visible layer for the next RBM in the DBN network. The structural diagram of DBN consists of bidirectional and symmetrical layer connections. A greedy learning algorithm is applied to train the DBN model; this technique involves repetitive training of one layer at a time until reaching a global optimum, which then serves as an input for the RBM. The application of the greedy algorithm is advantageous to the DBN network because it assists in weight distribution at each layer and supports proper network initialization, hence eliminating

the issue of local minima t vanishing, as well as speeding up and optimizing the training process [112]. However, DBN has an expensive computational rate, since it requires the training of multiple RBMs [113].
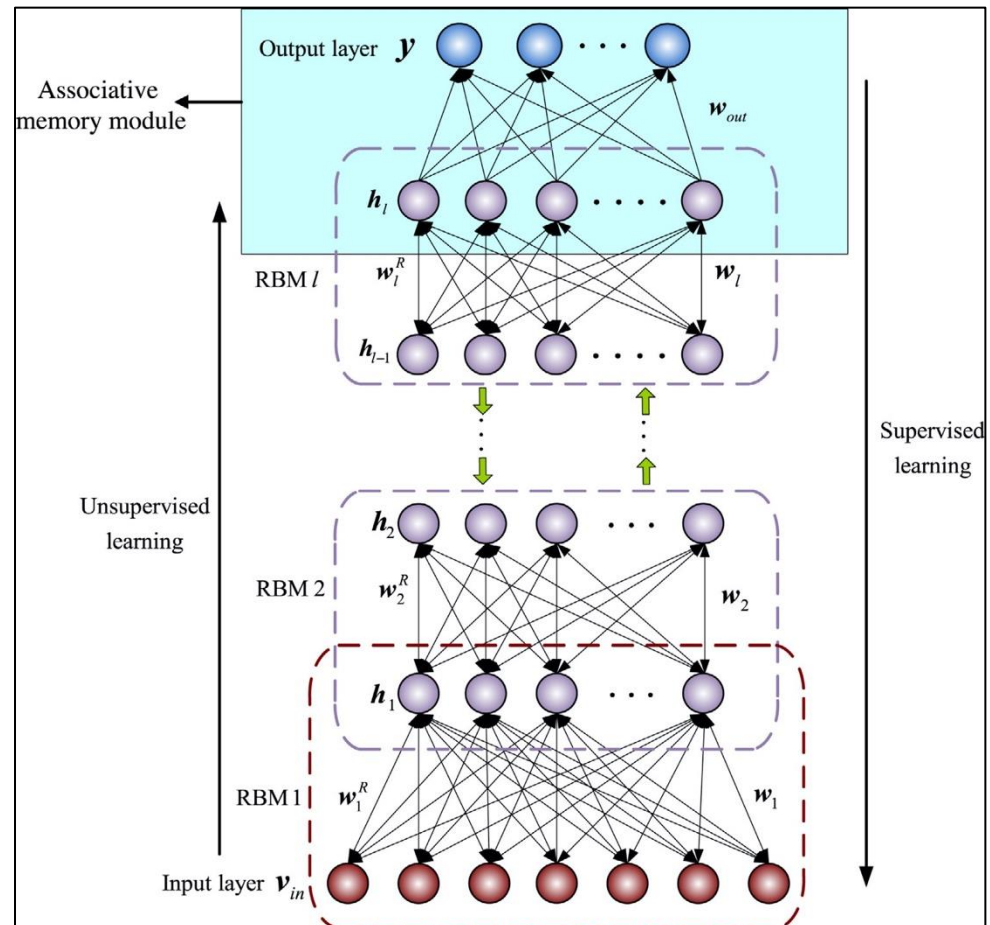

Figure 9. The architecture of DBN [114].

3.9. Attention Mechanism

The Attention Mechanism is an encoder–decoder technique designed specifically for neural machine translation. This technology was developed to facilitate the translation of neural information into machine information using an encoder and a decoder. The input is transformed through the encoder into a fixed-length vector and then translated via the decoder [115]. In 2014, ref. [116] developed the first Attention Mechanism, which was modeled on the human visual Attention Mechanism. The Attention Mechanism acts similarly to the brain in terms of prioritizing specific elements and is beneficial for many domains, including image analysis [117], video analysis [118], machine translation [119], and others. Moreover, it is effective for time-series forecasting, and its diversified range of applications has been acknowledged by academics.

The Attention Mechanism might prioritize more significant input features while overlooking other features. This approach influences the way in which we pay attention to information from the outside world in a wise and sensible way, ignoring irrelevant information and emphasizing relevant information. Hence, the application of the Attention Mechanism allows us to process information more rapidly and accurately. The mathematical description and information distribution are expressed as shown in Equations (26) and (27):

$$e_i = \tan h \ (w_h \ h_i + b_n h_n), \ e_i \in [-1, \ 1] \tag{26}$$

$$a_i = \frac{exp(e_i)}{\sum_{i=1}^{t} exp(e_i)}, \quad \sum_{i=1}^{t} a_i = 1 \tag{27}$$

Self-attention picks up both long- and short-term dependencies and focuses on various aspects of temporal patterns. It is an excellent candidate for time-series forecasting because of these benefits. It is utilized to personalize an attention layer, the parameters of which are specified via an optimization method [119].

The Self-Attention Mechanism comprises a key matrix (K), a value matrix (V), and a query matrix (Q), as shown in Equation (28):

$$T = K = V = Q \tag{28}$$

The basic principle in the self-attention process is the scaled dot-product of attention (SDA). Two stages are involved in the process of determining similarity, which are (1) finding the dot product between Q and matching matrix K and (2) dividing it by the matrix K dimension. To acquire the attention expression, we utilized the SoftMax function to normalize the output before multiplying it by the matrix V. Equation (29) shows how SDA operates.

$$SDA(K, V, Q) = SoftMax(\frac{QK^T}{\sqrt{d_k}}) \, v \tag{29}$$

3.10. Generative Adversarial Networks (GAN)

The GAN, which was invented by the authors of [120], and its versions, including convolutional GAN (Conv-GAN), which was developed by the authors of [121], demonstrate significant potential in producing pictures that are incredibly similar to a given collection of training images. GANs are a robust category of generative models that perform their tasks by implicitly modeling high-dimensional data distributions [122]. In image processing, GAN has shown superiority relative to other generating approaches in terms of its capacity to create realistic synthetic pictures [121,123,124]. In terms of applicability to our study, GAN can be utilized to understand the distribution of meteorological data and augment it. This work has the possibility of supplementing training specimens with a wide range of samples to increase the generalization of deep learning algorithms. The training technique in GAN involves building two antagonistic networks that compete with each other, including the two main neural networks, which are typically the generator (G) and the discriminator (D), that may be trained using a traditional backpropagation approach.

The two networks collaborate to enhance each other, albeit in an aggressive manner. During the training phase, G attempts to learn and make "fake" specimens of input noise Z to trick D, while D endeavors to identify "fake" or "genuine" inputs accurately. The model converges until the discriminator no longer distinguishes between the specimens. It is worth noting that the input noise Z typically keeps track of a Gaussian distribution across G to provide the sample I = G (z), where Z = N (μ, σ^2), and D is a fundamental neural network filter for bilateral categorization. This process provides an excellent opportunity to supplement training specimens with a broad range of data to increase the generalization of deep learning algorithms.

4. Solar Irradiance Forecasting Based on Deep Learning Techniques

In general, this section elaborates the prerequisite for the effective implementation of solar energy systems, which is forecasting horizons. Therefore, it is important to focus on this fact to further achieve accurate solar irradiance forecasting.

4.1. Solar Irradiance Forecasting Model Based on the LSTM Algorithm

The application of the LSTM algorithm in recent research into the development of solar irradiance forecasting models is justified according to its ability to address time series-based forecasting problems. Subsequently, the LSTM model was utilized by [65] to forecast solar

irradiation 1 h in advance at three location sites in the USA, which were located in Hawaii, Atlanta, and New York. The input features that were used to generate the model were temperature, relative humidity, sun zenith angle, cloud type, perceptible water, dew point, wind direction, and clear sky index. According to their analysis, the suggested LSTM model exhibits an RMSE for the sites under consideration in the range of 45.84 and 41.37 W/m^2.

In Seoul, South Korea, ref. [125] utilized LSTM to estimate hourly, daily, and annual solar irradiation. They used the Korea Meteorological Administration (KMA) database to obtain solar data from 2001 to 2017. The model data training was performed using the previous solar irradiance dataset. Following that approach, ref. [43] utilized an LSTM model for hourly solar irradiation forecasting of the following day. Input features considered included the month, hour, and days of the month; wind speed; relative humidity; visibility; dew point; weather conditions; and temperature. To validate the model, the dataset from the Measurement and Instrumentation Data Center (MIDC) was used. Table 2 shows that the proposed model performs better in terms of accuracy than other established models with an RMSE value of 76.245 W/m^2. Figure 10 emphasizes the comparison between real and LSTM forecasted solar irradiance.

Table 2. Comparison results for the proposed models using the MIDC dataset [43].

Algorithm	Testing RMSE
Persistence	209.2509
LR	230.9867
BPNN	133.5313
LSTM	76.245
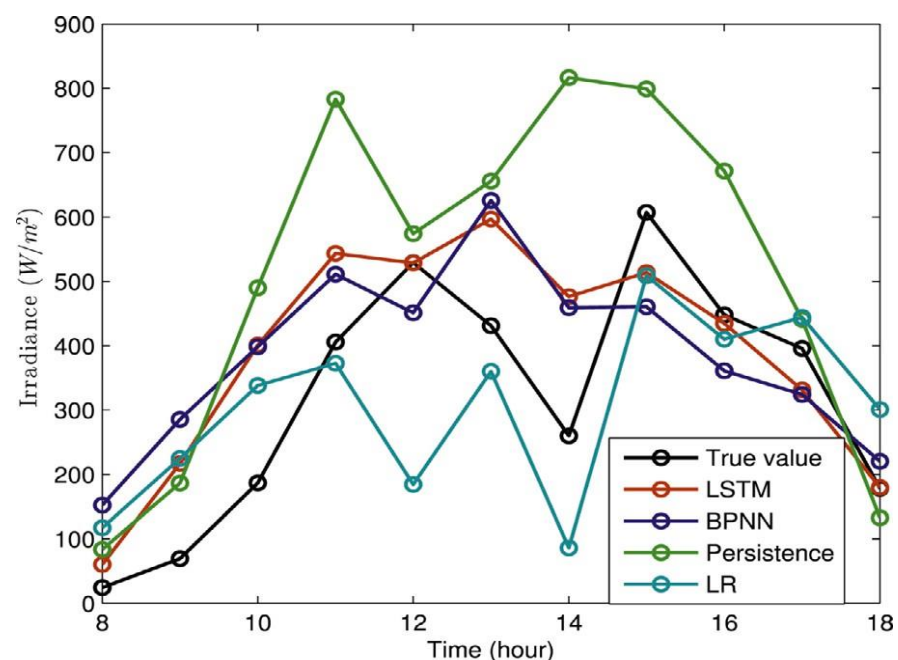
Figure 10. The comparison between actual irradiance and forecasted irradiance [43].

Recently, ref. [126] built a deep LSTM network for solar irradiance forecasting based on various horizons (3/6/24 h ahead) in the dry regions of India. Among the feature inputs considered were dew point, diffuse horizontal irradiance (DHI), global horizontal irradiance (GHI), direct normal irradiance (DNI), wind speed, pressure, temperature, wind direction, and relative humidity. A five-year dataset (2010–2014) from the NSRB in the Thar Desert was used to validate the model. The generated model had MAPE values ranging from 6.79 to 10.47%, which indicate excellent performance. Ref. [107] utilized satellite data to estimate solar irradiance 1 day in advance using an LSTM model. The model was tested

using data from 21 remote sensing sites, 16 of which were in mainland Europe, while 5 were in the United States. The input features considered were maximum temperature, minimum temperature, air pressure, cloud cover, specific humidity, and other meteorological data. According to the empirical results, the generated LSTM outperformed other persistent models tested with a forecast skill score of 52.2%.

Recent work by [127] used LSTM for hourly solar irradiance forecasting for 1 day in advance. The input features considered were temperature, wind speed, sky cover, humidity, and precipitation. The dataset for training was retrieved from the Korea Meteorological Administration, which consists of weather variations recorded at different targeted site locations. It was shown that the model used can generate substantial forecasting with an RMSE of 30 W/m^2. Ref. [128] forecasted solar irradiance 1 h in advance using GRU and LSTM. They used a mono-variate technique, which forecasts solar irradiance using historical time series. Their proposed GRU and LSTM models were superior to other conventional machine learning models for solar irradiation forecasting. Another study by [129] forecasted hourly solar radiation utilizing a deep LSTM network. They trained and tested the proposed model using data from a Canadian solar farm. Figure 11 depicts the forecast results for a variety of weather conditions, including cloudy sky, scattered clouds, clear sky, and few clouds, using the suggested model. In addition, it outperformed the SVR and FFNN.
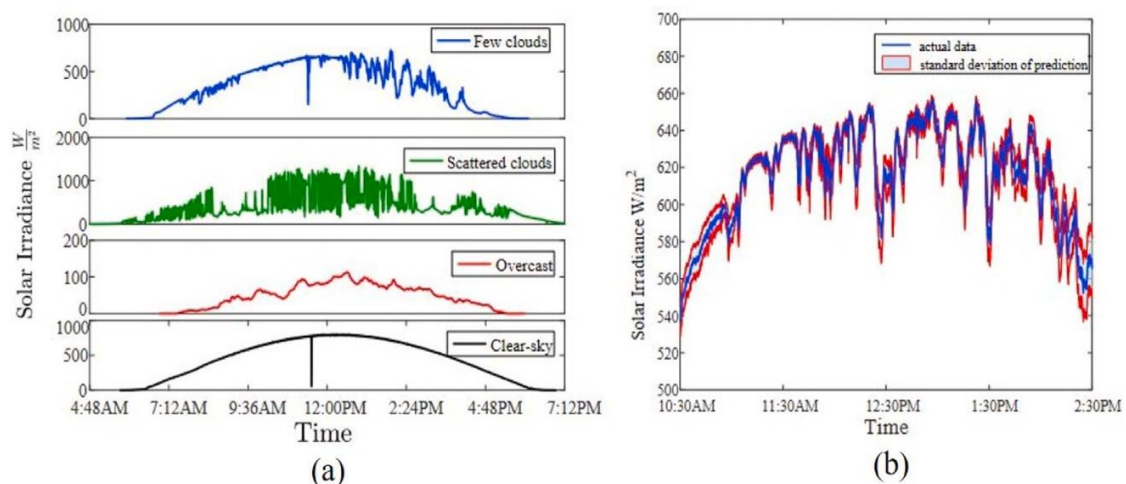

Figure 11. (**a**) Solar irradiation forecasting in various weather situations. (**b**) Daily solar irradiation forecasting [130].

The LTSM model was employed by [131] for hourly solar irradiance forecasting in Johannesburg, South Africa. The LSTM network was trained based on solar radiation, sunlight duration, relative humidity, and temperature. The dataset available from the National Oceanic and Atmospheric Administration (NOAA) consists of a ten-year duration of weather data from 2009–2019. Simulation findings demonstrated that the proposed LSTM network performed satisfactorily compared to the SVR model with an nRMSE of 3.2%. Subsequently, ref. [132] suggested the implementation of two methods based on LSTM for novel solar irradiance forecasting model on an image-based dataset. In this study, different input features were assigned to achieve variations of 5 to 60 min in advance of solar irradiance forecasting. The first method focused on input features, such as solar irradiance 5 min before, current solar irradiance, and center value. In contrast, the second method focused on solar irradiance 5 min earlier, most recent solar irradiance, center value, variance value, the red–blue comparison method, and the three-step search method. Findings indicated that the LSTM model trained using the second method showed superior forecasting.

Before that study, ref. [133] employed LTSM for hourly solar irradiation forecasting using historical and meteorological data from Kharagpur, India. Further, they intended to use

these data for solar PV power output forecasting. For validation purposes, the forecasting models used 15-year data available from 2000–2014 provided by NSRDB. The suggested model demonstrated superior performance compared to the ANN model, achieving an RMSE value of 57.249 W/m^2. Subsequently, ref. [134] utilized the LSTM network to forecast hourly GHI in Islamabad, Pakistan. The data were provided by the meteorological station at NUST Islamabad over 55 months from 2015–2020. The model training was performed on historical GHI data while considering all meteorological feature inputs, such as relative humidity, wind speed and direction, ambient temperature, DNI, and DHI.

A recent study by [135] utilized a new variant of LSTM known as stacked LTSM integrated with principal component analysis (PCA) for solar irradiance forecasting on a 6-month dataset retrieved from a weather station at Morong, the Philippines. Among the input features considered were humidity, station height, wind speed, station temperature, absolute pressure, illuminance, sea level pressure, and ambient temperature. The proposed stacked LSTM model was compared to CNN, Bidirectional-LSTM, and other deep learning models and managed to achieve a performance of R^2 value 0.953 and MAE value 41.738 W/m^2. Previous work by [136] used a deep learning-based univariate LSTM model to conduct short-term solar irradiance forecasting. The model was trained using data on solar radiation collected over two years, starting on 1 January 2014, and originating from a solar monitoring station based at the University of Jaffna's Faculty of Engineering. The outcomes showed that the suggested model outperformed the ARIMA model.

Ref. [137] utilized an LSTM for solar irradiance forecasting based on the range of forecasting horizons (intra-hour and intra-day). The dataset used to validate the model was retrieved from SURFRAD, which consists of data from seven stations located across the United States. The suggested approach outperformed typical machine learning models by 71.5%. For short-term forecasting, ref. [138] utilized the LSTM Neural Network model for solar irradiation forecasting. The experiments were conducted at half-hourly and hourly data intervals, as well as during all four seasons. The model accuracy can be seen to vary from season to season in Table 3. The suggested model demonstrates that the lowest inaccuracy happened in the winter for the half-hourly forecast horizon, since Florida has substantially more rain from May to October, resulting in less predictable PV generation. Regardless, the model frequently showed a gradual increase in inaccuracy for hourly solar irradiance forecasting.

Table 3. Performance of Single Step Ahead of PV Power Forecasting [138].

Season	Hourly Data			Half Hourly Data		
	MAE	RMSE	MBE	MAE	RMSE	MBE
Spring	0.75	1.51	0.02	0.69	1.43	−0.02
Summer	0.98	2.04	0.31	0.87	1.84	−0.05
Autumn	0.70	1.48	0.39	0.55	1.27	0.21
Winter	0.61	1.39	−0.24	0.50	1.23	−0.07

4.2. Solar Irradiance Forecasting Model Based on CNN

Ref. [139] used a 3D-CNN model to forecast direct normal irradiance 10 min ahead of time. The proposed approach derived spatial and temporal features from cloud characteristics by integrating multiple consecutive ground-based cloud pictures. They developed the forecasting algorithm using GBC pictures and DNI data from a two-year period (2013 to 2014) obtained from NREL, and the model was able to achieve a forecasting accuracy of 17.06%. Ref. [33] presented a CNN-based forecasting method based on weather classification, as well as Generative Adversarial Networks (GAN). Initially, 33 different types of weather were integrated and classified into 10 new weather categories for data training. A data augmentation model using GAN was then applied to improve each weather category. Subsequently, the extended dataset that consisted of both original and generated solar irradiance data were used for CNN model training. The authors also used a weather

classifier to evaluate solar brightness and investigated how GAN can increase forecast accuracy. Based on their observation, GAN could generate high-quality samples that precisely matched the unique features of the original data, as opposed to memorization of the training data. Next, ref. [140] developed a Solar Net using 20 layers of deep CNN to calculate the intra-hour global horizontal irradiance (GHI) (i.e., 10-min to one hour ahead with a 10-min pause). Solar Net used a single total sky image (TSI) with no feature engineering or numerical measurements as its input. According to numerical studies based on six years of public data, Solar Net produced multi-step forecasts, resulting in an nRMSE of 8.85% and an accuracy forecasting score of 25.14%, which were greater than those of other compared models.

4.3. Solar Irradiance Forecasting Model Based on GRU

Ref. [141] employed LSTM and GRU up to an hour ahead of solar irradiance forecasting. Both univariate and multivariate forecasting models were developed to utilize exogenous meteorological factors and historical solar irradiance data. These solar data were based on an international airport in Phoenix, Arizona, USA (1 January 2004–31 December 2014). The findings showed that multivariate LSTM and GRU are more accurate than their univariate equivalents. Further, ref. [142] suggested a short-term forecasting horizons solar irradiance forecasting model integrated with GRU and Attention Mechanism, such as 5/10/20/30 min, for use in four climates in the Nevada desert, USA. The dataset used was obtained by the University of Nevada, Las Vegas, USA, from the year 2014 onwards, and it consists of peak wind speed and averaged historical solar irradiation data. A hybrid deep learning model presented by [143] was based on encoder–decoder networks of LSTM, Bidirectional LSTM, RNN, and GRU models for short-term solar irradiance forecasting. The model training was performed on NREL's fifteen-year dataset for Kharagpur, India, which spanned the years 2000 to 2014. A 10-kilo meter grid of meteorological and solar irradiance data were gathered. In addition, the suggested model was evaluated against other previously used machine learning models, such as Feed-Forward Neural Networks (FFNN) and Gradient-Boosted Regression Trees (GBRT), and showed superior performance.

4.4. Solar Irradiance Forecasting Model Based on Other Deep Learning

Ref. [144] introduced a Multi-Reservoir Echo State Network (MR-ESN) and deep Echo State Network architecture-based solar irradiance forecasting model (ESN). The model was designed to forecast a wide range of periods (1 and multiple hours ahead). The California Irrigation Management Information System collects solar irradiance data from Seeley Owens Lake South, Davis, Salinas North, and Blythe NE Markleeville. The suggested MR-ESN model achieved a lower forecasting error than the benchmark models evaluated, such as Elman Neural Networks, BP, and ESN, according to quantitative simulation results (ENN). Furthermore, as illustrated in Figure 12, MR-ESN has a far greater RMSE forecasting accuracy than ESN. Table 4 further illustrates that as the forecasting horizon lengthens, model forecasting accuracy declines.

Table 4. Ratio comparison of ANN and RNN models using a 10-minute data sampling frequency in terms of forecasting accuracy [144].

Forecasting Step	Evaluation Period	RMSE (ESN)	RMSE (MR-ESN)	Reduction (%)
1-step	1	147.258	121.671	17.38
	2	132.787	63.014	52.55
	3	196.057	168.158	14.23
	4	119.763	98.987	17.35
2-step	1	222.684	169.543	23.86
	2	171.536	70.068	59.15
	3	273.373	224.234	17.98
	4	154.126	113.687	26.24

Table 4. *Cont.*

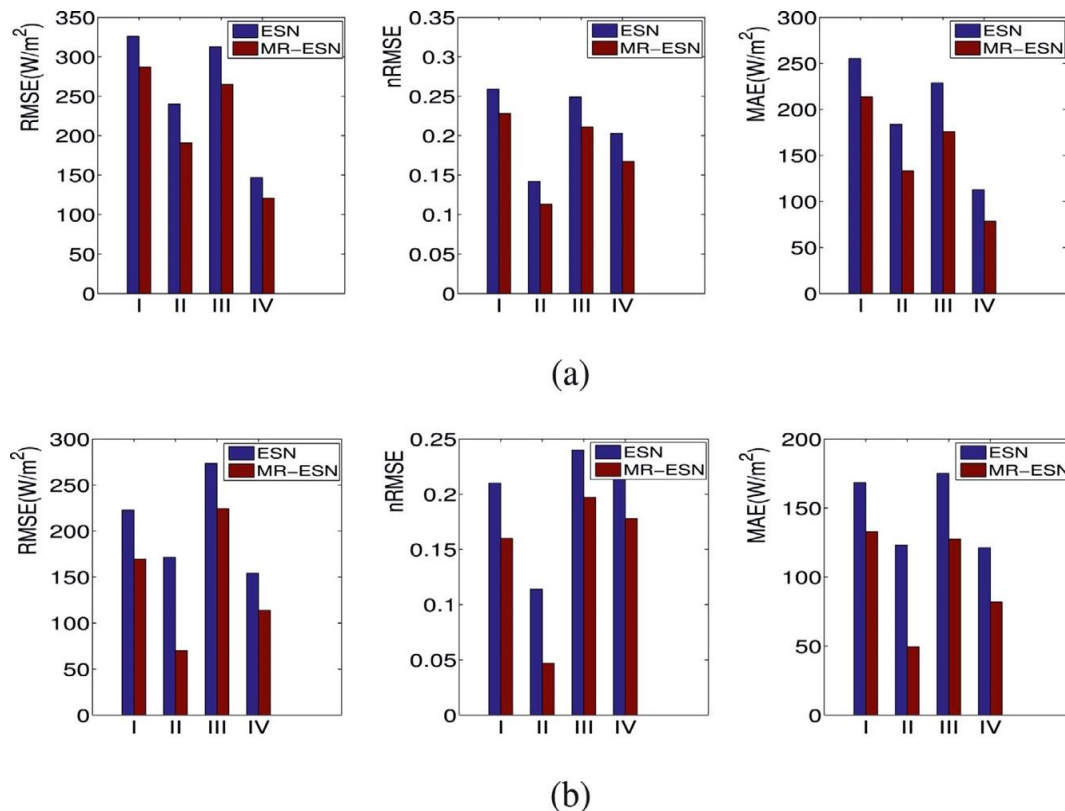| Forecasting Step	Evaluation Period	RMSE (ESN)	RMSE (MR-ESN)	Reduction (%)
3-step	1	235.349	189.584	19.45
	2	156.275	72.064	53.89
	3	286.015	246.643	13.77
	4	158.033	126.020	20.26

Figure 12. Two-Hour-Ahead MR-ESN and ESN estimates for (**a**) Owens Lake South station and (**b**) Blythe NE station [144].

Furthermore, ref. [145] developed RNN using the sliding window technique for short-term solar irradiation forecasting of meteorological data from an Alabama weather station. The forecasting was generated for variations in sampling intervals, such as 1 h, 30 min, and 10 min, based on two factors: outside dry-bulb temperature and time. The model was trained using data from 7 days (22–28 May 2016) and validated on a single day (29 May 2016). According to the results in Table 5, both ANN and RNN have respectable accuracy levels, with RNN significantly exceeding ANN. Further, Figure 13 illustrates the forecasting accuracy of both ANN and RNN models with and without the sliding window techniques.

Table 5. Forecasting accuracy of ANN and models on a 10-minute data sampling frequency [145].

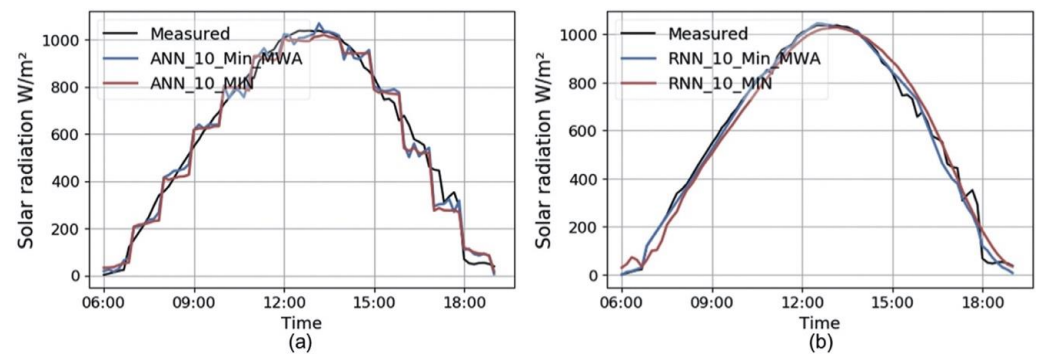| Models	ANN	RNN	Improvement (%)
R^2	0.974	0.983	1%
RMSE	55.7	41.2	26%
CV (RMSE) (%)	9.41	7.64	19%
NMBE (%)	1.73	0.92	47%

Figure 13. Comparison between findings of (**a**) ANN and (**b**) RNN forecasting [145].

Previous work by [45] used a Deep Neural Network (DNN) for solar irradiance forecasting in Turkey. Meteorological input data used included maximum temperature, sunshine duration, cloud type, and lowest temperature, as well as an astronomical feature known as extra-terrestrial rotation. The model training and testing were performed on a 7-year dataset gathered from 34 stations between 2001 and 2007. The results showed highly accurate forecasting results of 0.98 for the coefficient of determination. Further, ref. [146] suggested a new deep learning model integrated with Embedding Clustering (EC) and Deep Belief Networks (DBN) for daily solar irradiance forecasting. The dataset was gathered from 30 meteorological stations located across China and validated based on daily input features (for mean) for wind speed, maximum and minimum dry-bulb temperature, relative humidity, sunshine duration, and global solar radiation. A dataset from the preceding 22 years was used to train the model (1994 to 2015). The suggested hybrid model generated an excellent forecasting score, with an RMSE of 0.282 W/m^2 and an MAE of 0.137 W/m^2. Figure 14 shows the comparative probabilistic density error curves for the models. The improved performance of the developed model emphasizes its global acceptability for solar irradiance forecasting.
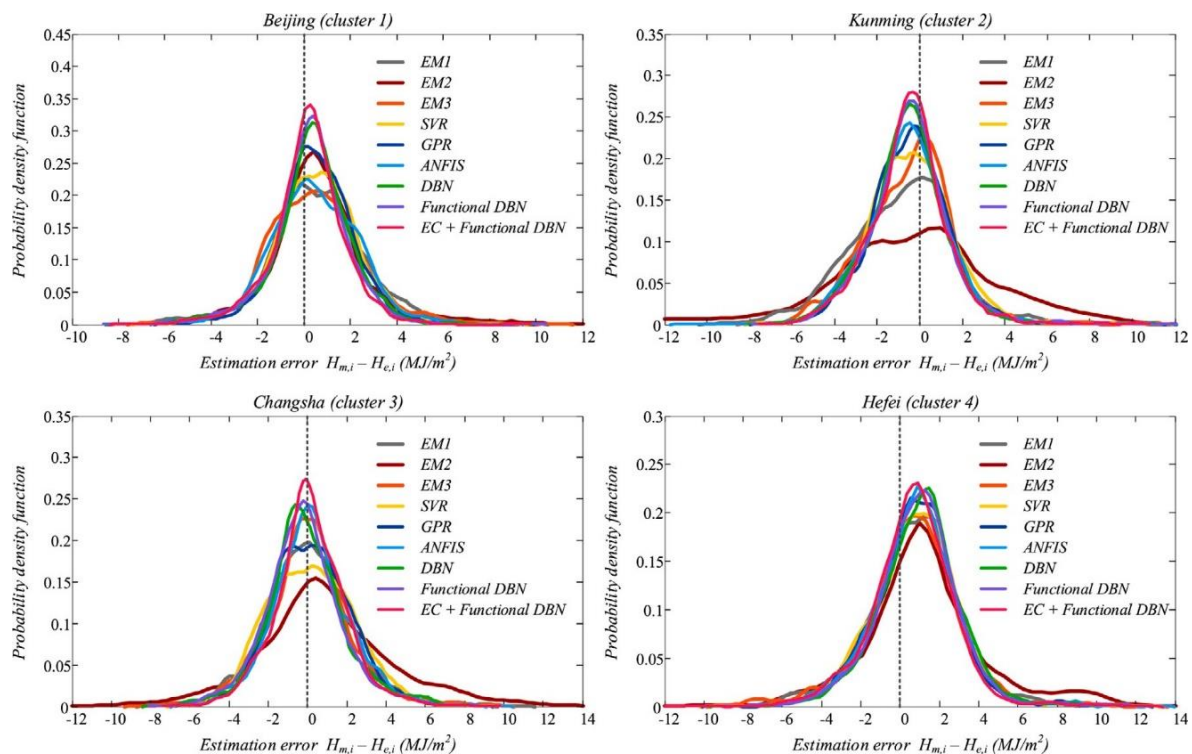

Figure 14. The probability density curves of hourly errors in four Chinese stations [146].

The Hierarchical Multi-Modal (H-MDL) and the Wavelet Decomposition Multi-Modal were introduced as two deep learning-based multi-modal solar irradiance forecasting models [147]. The proposed models were trained on a 12-year dataset gathered via NREL. The dataset included hourly images from all-sky web cameras, as well as data on temperature, turbidity, precipitation, wind speed, zenith angle, DHI, GHI, and DNI. Based on the findings, the novel model outperformed well-known forecasting models, such as ANN and ARIMA. Prior research conducted by [148] forecasted short-term solar irradiance in the Netherlands based on a DNN model. Satellite photographs of previous irradiance values, historical and prospective temperatures, NWP projections, clear-sky irradiance, and relative humidity data for the next 6 h, as well as historical and prospective temperature, NWP projections, clear-sky irradiance, and relative humidity data for the next 6 h, were used to train the model. In the Netherlands, 30 locations were chosen, with 5 serving as model training sites and the remaining 25 serving as model evaluation sites. The results showed superior performance compared to other techniques.

Subsequent work in 2021 by [149] introduced a Comprehensive Ensemble Empirical Mode Decomposition with Adaptive Noise (CEEMDAN), Bi-LSTM, and the Sine Cosine Algorithm (SCA) for solar irradiance forecasting. The time-series data were split into a few periodic intrinsic mode functions (IMFs) using CEEMDAN. Next, the trends of solar irradiation were identified using an auto-correlation function (ACF) and a partial auto-correlation function (PACF). The SCA algorithm was then used to optimize the forecasting of Bi-LSTM. The model was trained using solar radiation data from Dauphin Island, Alabama, USA. The suggested CEN-SCA-Bi-LSTM model outperformed other considered models. As shown in Figure 15, the comparison analyses of all considered models were based on the season ranking, i.e., autumn, winter, summer, or spring.
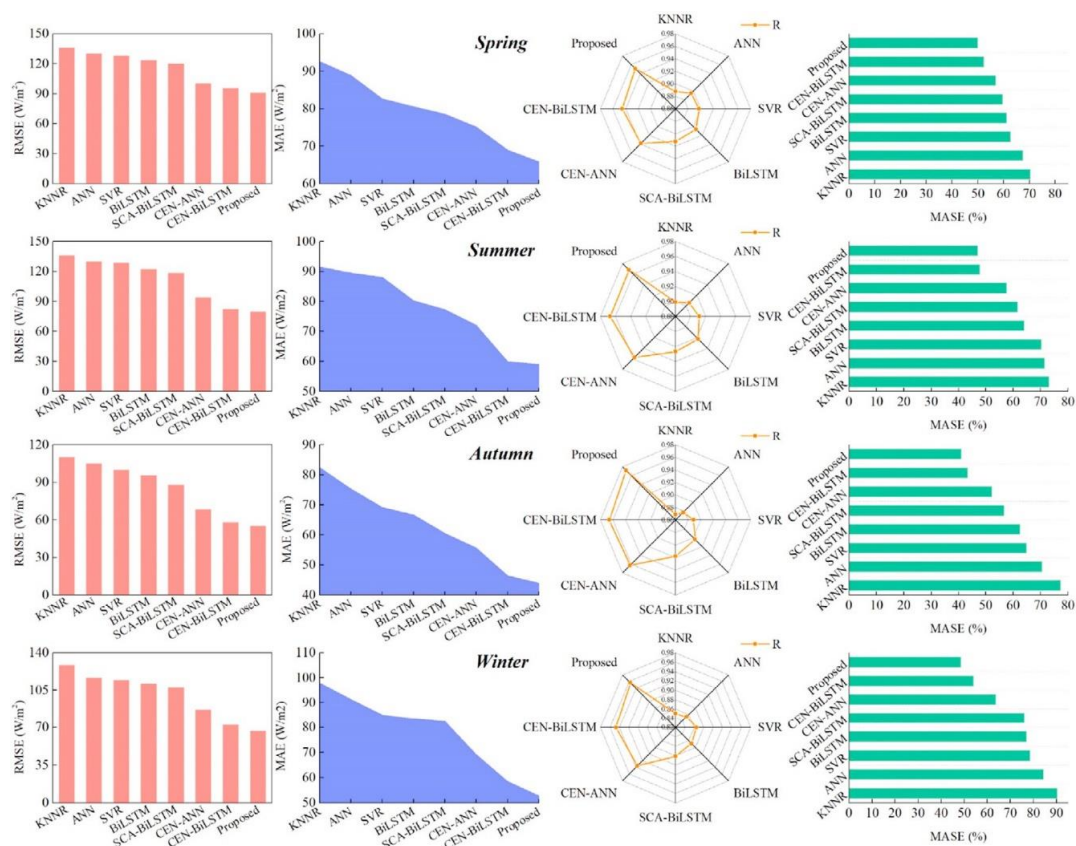

Figure 15. Comparison analysis of all considered models based on four seasons [149].

Prior work conducted in 2020 by [150] utilized two unique variants of LSTM, which were attention LSTM and GRU, for daily solar irradiance forecasting in India. The sug-

gested model was introduced as a Bidirectional LSTM (Bi-LSTM) and was trained using a 36-year dataset (1983–2019) available from the National Aeronautics and Space Administration (NASA). The observed findings indicated superior accuracy performance for Bi-LSTM than other deep learning models. Ref. [151] introduced a sophisticated RNN model integrated with a Chain-Structure Echo State Network (CESN) for hour-ahead solar radiation forecasting. The model was trained with spatial–temporal behavior data and demonstrated rapid learning, as well as a cheaper computational cost than RNN. Based on the findings, the suggested CESN model outperforms backpropagation (BP), ENN, and classical ESN. Further, ref. [152] adopted GAN integrated with CNN for model training to produce realistic future cloud images that were retrieved from the University of Patras, Greece, and consisted of 1.5 million images taken between August 2014 and April 2017. Based on observation, the images produced by the model without the adversarial loss are rather blurry.

4.5. Solar Irradiance Forecasting Model Based on Deep Hybrid Model

Recently, Ref. [153] suggested an integration of CNN and LSTM to establish a framework for forecasting solar irradiance based on Clear-Sky Index that was 1-h ahead. The suggested model was trained on historical solar radiation data and does not consider any outside influences. The model was trained on the NREL solar dataset from Oahu, which consisted of 20 months of data (March 2010 to October 2011). The developed model outperformed the considered baseline models, as evidenced by variations in forecast skill scores of 7.4 to 41%. Before that study, ref. [154] also integrated the CNN and LSTM network (ConvLSTM) for hourly solar irradiation forecasting in the Korean Peninsula. The dataset consisted of 1100 photos taken in a row from 1 April 2011 to 31 December 2015. The simulations observed that the ConvLSTM model generated a higher forecasting accuracy than ANN and RF, with an R^2 value of 0.895 and an RMSE value of 71.334 W/m^2.

An innovative hybrid deep architecture based on (ResNet) and LSTM was developed by [155] for short-time solar irradiance forecasting. The model was trained on an eight-year dataset (from 2000 to 2017) available from the NSRDB. The ResNet-LSTM hybrid model lowered forecasting errors by 52.44 and 17.07%, respectively, more than standalone ResNet and LSTM. Furthermore, [102] integrated the use of LSTM and CNN algorithms to extract spatial and temporal information in sequence. The meteorological data from 23 different sites in California were used for model training, such as DHI, GHI, temperature, relative humidity, cloud cover, and precipitation. The model was tested over a whole year, through all four seasons, and under three various types of meteorological conditions. The forecasts made using the developed LSTM-CNN model are more accurate than individual models, like SP, CNN, LSTM, ANN, SVM, and others. In Alice Springs, Australia, [156] proposed an integration of CNN and LSTM (CLSTM) for both short- and long-term forecasting horizons of the half-hourly solar irradiation forecasting model. The model was evaluated for daily, weekly, and n-month (n = 1 to 8) solar irradiance forecasting horizons using data from a 12-year and and eight-month period (from 1 January 2006 to 31 August 2018). The recommended model obtained a higher percentage, which denoted satisfactory performance. As seen in Figure 16, the CLSTM hybrid model in 1-day forecasting for RMSE, MAE, MAPE, and RMSE obtained a higher percentage, which denoted satisfactory performance.

Ref. [67] used historical time series solar irradiance data to create a new hybrid deep learning model of LSTM, CNN, and wavelet decomposition for day-ahead forecasting of solar irradiance in North Carolina, USA. The model was validated using the partitioned data (into four different weather categories) obtained from NREL and the NOAA Earth System Research Laboratory at two different locations: Desert Rock Station and Elizabeth City State University. Recent work by [157] suggested a new hybrid model using both MLP networks and LSTM that uses two-branch inputs for hour-ahead solar irradiance forecasting. The dataset used is a time series of historical solar irradiation and auxiliary inputs. Findings showed superior performance compared to other machine learning

models, with performances of SVM, BPNN, Random Forest, RNN, and LSTM improved by 19.31, 19.19, 11.68, 20.15, and 13.48%, respectively.
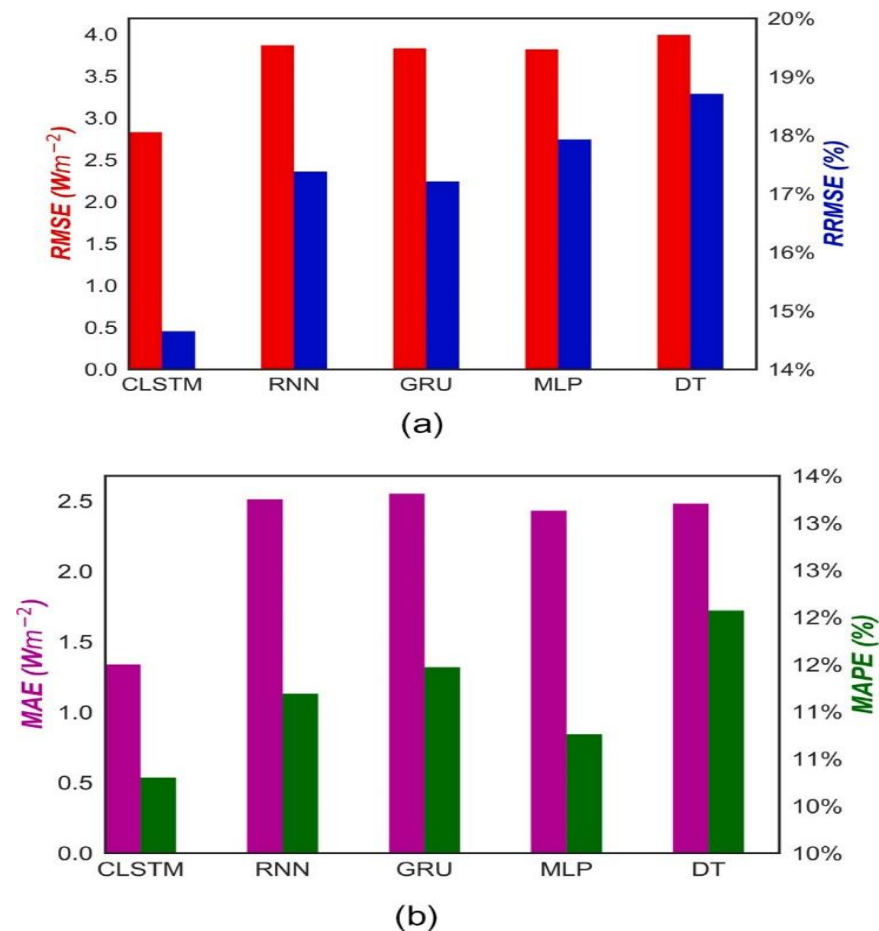

Figure 16. The performance of the CLSTM model for 1-day GSR forecasting (**a**) RMSE (%) and (**b**) MAPE (%) [156].

For hourly GHI forecasting, Kumari [102] developed a hybrid deep learning model that made use of both the LSTM algorithm and the CNN algorithm to extract spatial and temporal information in sequence. The model was trained using the GHI data, in addition to meteorological data from 23 different sites across the state of California, including GHI, temperature, relative humidity, cloud cover, and precipitation data. The model was tested over a whole year, through all four seasons, and under the three types of meteorological conditions. The forecasts made using the developed LSTM-CNN model are more accurate, having a range of about 37–45%, than standalone models, like SP, CNN, LSTM, ANN, SVM, and others.

Ref. [158] suggested an evolutionary attention-based LSTM-based time-series forecasting model to address the problem of attention dispersion that plagues the typical LSTM algorithm. Varying in terms of time steps, the Attention Mechanism can give characteristics to time-series weights. The LSTM-based approach of forecasting has difficulty interpreting complex relationships between features that only use weights. A Self-Attention Mechanism is an Attention Mechanism that extracts deeper information from a row of data. The suggested method aimed to maximize the benefits of the features of the extended input sequences. The Attention Mechanism performed better than the standard LSTM.

Following that approach, ref. [159] proposed a probabilistic forecasting model that integrated residual modeling and an RNN. An LSTM-based point forecasting method was utilized for deterministic forecasting and calculating the residual distributions. The input features considered include temperature, cloud cover, dew point, east sea-level pressure,

relative humidity, wind speed, hour of the day, and month of the year. The model was based on data from the MIDC official website dating back ten years. The suggested model outperformed ELM, RF, SVR, and LSTM in terms of accuracy. Subsequently, ref. [160] utilized six variants of RNN techniques, which are RNN, GRU, Content-Based Attention, Luong Attention, Self-Attention-Based RNN, and LSTM for solar radiation forecasting. A comparative analysis was conducted, using a 37-year dataset, between the two independent sites. The models were trained on historical time-series data of solar radiation with no exogenous characteristics, and forecasts were produced on several forecasting horizons. Based on the findings, it was revealed that the Attention Mechanism significantly improved memory-based RNN.

Another work by [161] suggested a two-stage deep solar radiation forecasting model based on the CNN model to encode a frame from a sky-video to restore a full-sky image, and a two-tier LSTM model was used for monitoring. The model training was performed on a free dataset from two sites in Tucson, Arizona, and Golden, Colorado. Results showed that the suggested model can forecast a range of time ranges of 1–4 h in advance. An ensemble model of (XGBF-DNN) [162] was employed for hourly solar radiation forecasting by integrating Deep Neural Networks and Extreme Gradient Boosting Forest. The suggested hybrid model was evaluated at three sites in India, which were Gangtok, Jaipur, and Delhi, in terms of temperature, pressure, relative humidity, wind speed, clear sky index, wind direction, time of day, and month number. Figure 17 depicts the comparison analysis of the proposed model against benchmarks, such as smart persistence, SVR, RF, DNN, and XGBoost which resulted in superior performance at each site location during the winter, summer, monsoon, and autumn seasons.
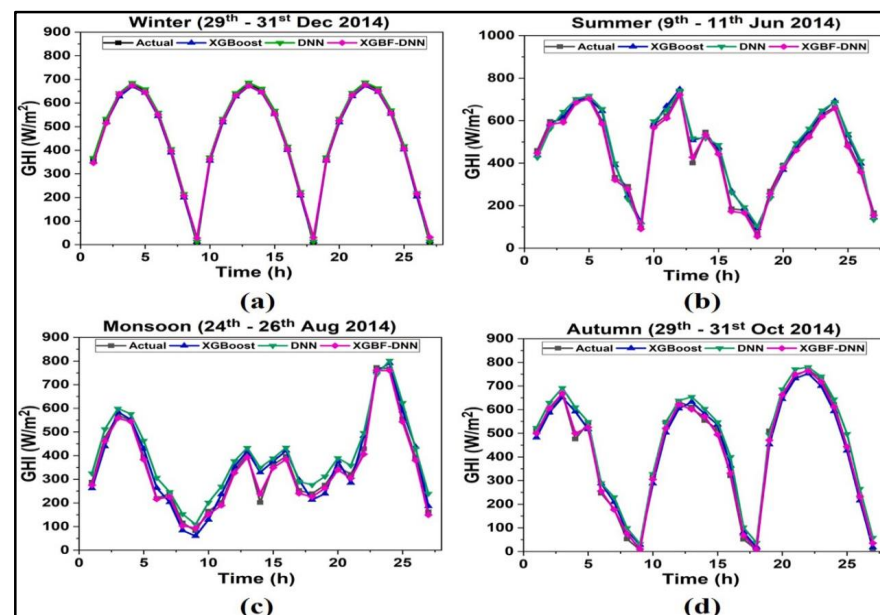

Figure 17. The comparison between performances of actual and predicted GHI in Gangtok for different seasons (**a**) Winter, (**b**) Summer, (**c**) Monsoon and (**d**) Autumn [162].

Recent work by [163] suggested a hybrid deep learning model that utilized both CNN and Bi-LSTM for excellent mid-term solar radiation forecasting. The CNN architecture captured available solar irradiance features, and Bi-LSTM manipulated the time-series data dependencies. The evaluation and comparison of the suggested model were performed against other established DL models at three different geographical areas within similar latitudes that received similar amounts of solar radiation. It was demonstrated that the hybrid DL model was durable and outperformed existing DL models for forecasting mid-term solar radiation in terms of accuracy. A hybrid model of four distinguished deep learning-based techniques was designed by [164] using LTSM and CNN based on Bayesian

Optimization (BO) for short-term PV power generation forecasting. The study analyzed feature selection for both benchmark models and selected five predictive features to be used to directly forecast results (1 to 24 h ahead). The data originated from a PV facility in Shandong, China. While Bi-LSTM and CNN-Bi-LSTM models were more practical for 1-hour forecasting, while LSTM-CNN and CNN-Bi-LSTM models were more suitable for 24-hour forecasting. This work demonstrated that the integration of Bayesian optimized optimal weights offered a satisfactory forecasting performance for commercial PV plants and could minimize error rates by up to 32.80% more than the benchmark model.

Recently, ref. [165] described a multivariate technique for forecasting solar power generation of PV systems in the very short term using the LSTM algorithm. The model is validated using the standard parameter and the Mean Absolute Error (MAE), which was 0.0565. The implementation of LSTM for multivariate extremely short-term solar PV power forecasting was supported by potential results, which outperformed the univariate technique. LTSM and fully connected (FC) layers were introduced into an LSTM-FC deep learning system to optimize forecasting accuracy [166]. Due to the dual-branch input, the model considered was not limited to analysis of how meteorological data affects power generation, but also considered time continuity and periodic dependencies in terms of accuracy performance. The studies used meteorological data, as well as historical continuous and periodic data, which were then combined into different input forms for comparative evaluation of LSTM-FC relative to other benchmark models, like SVM, GBDT, GRNN, FFNN, and LSTM. Meanwhile, the LSTM-FC model outperformed the others in terms of forecasting accuracy, with a RMSE that was 11.79% greater than that of SVM.

Accordingly, ref. [167] investigated the implementation of a hybrid Deep Learning approach by applying RNN, LTSM, and GRU for solar energy forecasting. Real meteorological data (year 2016–2018) from the Errachidia region were used to validate the model. Six metrics, including MAE, MSE, RMSE, ME, R2, and NRMSE, were used in the analysis of the results to measure forecast precision and error margin for real-time photovoltaic forecasts to improve grid management and safety. In terms of cost efficiency, it was found that RNN and LSTM performed much better than GRU because they could preserve long-term dependencies in time-series data. Subsequent work by [168] concentrated on a new deep learning-based multi-step ahead approach using LTSM for feature extraction to enhance the accuracy forecasting of the GHI. The sine–cosine algorithm based on the swarm evolutionary optimization approach was employed to automate the neural network architecture. A three-phase modification model was developed using three datasets gathered from three solar stations in the Eastern USA. This approach was expected to encourage a diversified and reduce premature convergence in the optimization method. The experimental findings demonstrated excellent performance using the MAE, RMSE, and Pearson metrics in comparison to alternative forecasting models.

Ref. [67] used LSTM and GRU for forecasting techniques to estimate the efficiency of learning data in the best, worst, and average circumstances. Table 6 compares the training time of LSTM and GRU in the three scenarios. This table yields the superiority of GRU compared to LSTM because the longest training time for GRU was shorter than the best training time for LTSM.

Table 6. The training time of LSTM vs. GRU [94].

Model	The Best Case (s)	The Worst Case (s)	The Average Case (s)
LSTM	393.01	400.57	396.27
GRU	354.92	379.57	365.40

Further, Table 7 compares the training times of both LSTM and CNN–LSTM. The hybrid model must extract both temporal and spatial aspects of the data [169], which results in longer training times than the LSTM, which takes 983.71 s.

Table 7. Performance of LSTM vs. CNN–LSTM [94].

Model	LSTM (s)	CNN–LSTM (s)
Training time	70.490	983.701

The authors of [170] examined several Deep Neural Network (DNN) models for the one-day forecast of Global Horizontal Irradiance (GHI). The considered DNN models include Gated Recurrent Unit (GRU), Long Short-Term Memory (LSTM), Bidirectional-GRU (Bi-GRU), Bidirectional-LSTM (BiLSTM), one-dimensional Convolutional Neural Network (CNN1D), and various hybrid configurations, like CNN-BiLSTM and CNN-LSTM. For the development and evaluation of the DNN-based models, the authors utilized a dataset provided by NASA, which consisted of daily GHI recordings from 1 January 2000 to 30 June 2020 at a dry site (Hail, Saudi Arabia). The analysis also focused on identifying the features that influenced the accuracy of the models. However, this work did not take into account other weather features, such as air temperature, wind direction, wind speed, pressure, and relative humidity. Only historical values of daily GHI were used to develop the DNN-based models. The outcomes of the study indicated that the DNN models exhibited generally good performance, with a maximum correlation coefficient of 96% achieved for the daily GHI forecasts.

In the Saudi Arabian province, the authors of [171] investigated several models for forecasting solar energy yield. The models considered in the study were Neural Networks (NN), Non-Linear Autoregressive (NAR), Gaussian Process Regression (GPR), Support Vector Machine (SVM), Simple Moving Average (SMA), Simple Average (SA), and Naive (N). The findings demonstrated that all of the models generated accurate estimates for the three years from 2019 to 2021, with the Naive and Simple Moving Average models slightly outperforming the other models.

Ghimire, et al. [172] developed a hybrid forecasting model that combined a Slime-Mould algorithm-based feature selection technique, a Convolutional Neural Network (CNN), a Long-Short-Term Memory Neural Network (LSTM), and a final CNN with a Multilayer Perceptron output (referred to as the SCLC algorithm later). The proposed approach was applied to six solar farms in Queensland, Australia, considering daily temporal horizons in six different time increments. The forecasting model proposed in the study outperformed two Deep Learning models (CNN-LSTM and Deep Neural Network) and three Machine Learning models (Artificial Neural Network, Random Forest, and Self-Adaptive Differential Evolutionary Extreme Learning Machines) across all six selected solar farms. The results demonstrated strong performance metrics (r > 0.9 in all cases) and improvements of more than 10% for daily, monthly, and seasonal periods.

Peng et al. [173] present a PV power forecasting approach for long- and short-term memory networks based on Pearson feature selection. The method analyzed the influence of input features on the variance in PV power using Pearson coefficients, and the model was confirmed using case tests. The results revealed that the intensity of insolation radiation, temperature, and humidity were important effect features in the change in PV power. When LSTM was compared to the BP, RB, and TS algorithms, it was discovered that the proposed approach reduced the forecasting error rate by 1.48, 11.4, and 6.45%, respectively, compared to the LSTM, CNN, and BP.

Ghimire, Deo, Casillas-Pérez, Salcedo-Sanz, Sharma, and Ali [172] developed a hybrid forecasting model that combined a Slime-Mould algorithm-based feature selection technique, a Convolutional Neural Network (CNN), a Long Short-Term Memory Neural Network (LSTM), and a final CNN with a Multilayer Perceptron output (SCLC algorithm later). The suggested approach was applied to six solar farms in Queensland, Australia, at daily temporal horizons in six different time increments. The proposed forecasting model outperformed two Deep Learning (Deep Neural Network, CNN-LSTM) and three Machine Learning (Artificial Neural Network, Random Forest, Self-Adaptive Differential-Evolutionary Extreme Learning Machines) models in all six selected solar farms. The results

showed good performance metrics (r > 0.9 in all cases) and improvements of more than 10% over daily, monthly, and seasonal periods.

Ghimire, et al. [174] suggested a hybrid approach that combines a Convolutional Neural Network (CNN), Long Short-Term Memory (LSTM), and Multi-Layer Perceptron (MLP) to predict Global Sun Radiation (GSR). The proposed method involved utilizing a CNN-LSTM model to extract optimal topological and structural characteristics from the predictive features, followed by an MLP-based predictive model for generating GSR forecasts. Meteorological features were used to predict GSR in six solar farms located in Queensland, Australia. To enhance the efficiency of the proposed CMLP model, a Hybrid–Wrapper feature selection technique based on a Random Forest-Recursive Feature Elimination (RF-RFE) scheme was employed to eliminate redundant predictor features. The performance of the CMLP model was compared to those of seven temperature-based deterministic models based on artificial intelligence. The evaluation demonstrated the excellent performance of the CMLP model across all tested solar energy study sites, spanning daily, monthly, and seasonal scales.

5. Discussion and Analysis of Performance

The performances of existing deep learning-based solar irradiance forecasting models are analyzed and discussed in this section. Currently, the LTSM is deemed to be the most used model for solar irradiance forecasting due to it being one type of Recurrent Neural Network that preserves data for future use, as well as providing extra storage capacity to the network. It is extensively utilized to forecast solar irradiance because it can speed up the convergence of non-linear forecasts and reveal the long-term relationships between time series. In one of the studies, ref. [107] proposed an LSTM model for 1-day ahead solar irradiance forecasting at 21 locations across the USA and Europe based on remote-sensing data. The suggested model demonstrated over 52.2% more accuracy than the smart persistence model. Following that approach, ref. [65] used LTSM for 1-hour ahead solar irradiance forecasting at three locations in the USA. The findings presented the lowest RMSE forecasting of 41.37 W/m^2. Subsequently, different mechanisms and new variants were introduced to enhance the model's performance. Therefore, ref. [150] presented two different variants of LSTM, known as Bi-LSTM and attention-based LSTM, for daily solar irradiance forecasting at two locations in India. The LTSM architecture was composed of a gating mechanism and memory cells, which assisted in learning long-term data dependencies. Thus, it could be asserted that LTSM and its variants performed adequately for solar irradiance time-series data and forecasting accuracy. Before that study, GRU was also extensively applied in the literature [141]. The benefits of GRU is that it is computationally cheaper because it requires less parameters training and memory; hence, faster execution. On the contrary, LTSM is computationally expensive, as it requires multiple training stages, though it generates effective accuracy forecasting. In addition, ref. [49] presented a CNN model for 10-minute ahead direct normal irradiance forecasting. In this study, ground-based cloud images were utilized to extract temporal and spatial features for model training and demonstrated an increase of 17.06% in accuracy forecasting over a smart persistent model. This result occurred due to the structure of CNN, which was made up of convolutional and pooling layers that were particularly efficient in features extraction. Therefore, it worked extremely effectively for image feature extraction (including converting it to two-dimensional data) for solar irradiance data. The connections of CNN and feature-sharing properties minimize the model parameter training and time, and it has a strong ability to extract spatial correlations from weather data, such as cloud cover, and other factors.

Furthermore, the DBN was also widely used for solar irradiance forecasting [146]. The structure of DBN was inherited from the restricted Boltzmann machine, in which a layer-by-layer unsupervised training method was employed. Therefore, it was suitable for extracting input features that were explicitly detectable in renewable energy forecasts. Moreover, another used deep learning model for solar irradiance forecasting was RNN. For example, ref. [175]

introduced a multi-horizon GHI forecasting model using RNN that resulted in an average RMSE of 18.57 W/m^2 for variations in forecasting horizons. The main structure of RNN were the internal feedback and feedforward connections between the neurons. These connections acted as memory functions for the RNN, allowing efficient processing of time-series solar data.

The performance of standalone models is limited in cases where a hybrid model focuses on combining different methods while exploiting the benefits of individual forecasting models to enhance the accuracy of deep learning networks. This observation is true because a single model sometimes may not be able to fully extract necessary features and satisfactorily perform all tasks. Furthermore, a hybrid model is expected to provide better accuracy than single deep learning models. However, it takes longer to train than other models because it needs to extract both spatial and temporal features from the data; therefore, the use of hybrid models was introduced. Several researchers have presented their work, including [49], whose author used a hybrid of model CNN and LSTM. Additionally, they integrated the use of LSTM and CNN algorithms to extract spatial and temporal information in sequences. The model training utilized meteorological data from 23 different sites in California, including DHI, GHI, temperature, relative humidity, cloud cover, and precipitation. The developed LSTM-CNN model was tested throughout a whole year, covering all four seasons and under three types of meteorological conditions. Similarly, [19] presented the use of both LSTM and RNN, [176] developed an integrated GA with the DNN for model optimization (GRU, LSTM, and RNN) of solar irradiance forecasting, and [102] developed an LSTM-CNN model to extract spatial and temporal features. The forecasts produced via the LSTM-CNN model demonstrated higher accuracy than individual models, such as SP, CNN, LSTM, ANN, SVM, and others. In another study, Ghimire, Deo, Casillas-Pérez, Salcedo-Sanz, Sharma, and Ali [172] developed a hybrid forecasting model that combined a Slime-Mould algorithm-based feature selection technique, a Convolutional Neural Network (CNN), a Long Short-Term Memory Neural Network (LSTM), and a final CNN with a Multilayer Perceptron output (referred to as the SCLC algorithm later).

The Attention Mechanism allows deep learning models to engage in enhanced feature extraction for a sequence of inputs. Although deep learning may yield positive outcomes, the training process is typically challenging; therefore, researchers often combine an Attention Mechanism to enhance the efficiency of feature extraction. Such an example can be seen in [177], whose authors used the Attention Mechanism with two LSTM for short-term PV power forecasting. Firstly, the output of both networks was integrated with the Attention Mechanism. Secondly, the integrated output was merged into a fully connected layer for forecasting output. The suggested model was then validated using the dataset of four seasons with variations in forecasting horizons against four other considered models. Based on our observations, the proposed model outperformed other conventional LSTMs, as well as all considered models for more than 15-min solar irradiance forecasting; the neural network that was paired with the Attention Mechanism performed most effectively, improving neural network interpretability by eliminating irrelevant data and only selecting more important information, as well as differently weighting each feature of the meteorological data. In addition, the correlation between features and irradiation in the feature set was analyzed to reduce model complexity and improve forecasting. Ghimire, Nguyen-Huy, Prasad, Deo, Casillas-Perez, Salcedo-Sanz, and Bhandari [174] suggested a hybrid approach that combined a Convolutional Neural Network (CNN), Long Short-Term Memory (LSTM), and Multi-Layer Perceptron (MLP) to predict global sun radiation (GSR). Firstly, a CNN-LSTM extracted optimal topological and structural characteristics inherent in predictive features; an MLP-based predictive model then generated GSR forecasts. Meteorological features were used to predict GSR in six solar farms in Queensland, Australia. To improve the efficiency of the proposed CMLP model, a Hybrid–Wrapper feature selection technique based on a Random Forest–Recursive Feature Elimination (RF-RFE) scheme was employed to eliminate redundant predictor features. Table 8 presents a summary of the related studies.

Table 8. Summary of the main studies related to solar irradiance forecasting using deep learning techniques, including and advantages and disadvantages.

Author	Year	Model	Result	Advantage	Limitation
Mishra and Palanisamy [175]	2018	RNN	Introduced a multi-horizon GHI forecasting model using RNN that results in an average RMSE of 18.57 W/m^2 for variations in forecasting horizons.	The main structure of RNN is the internal feedback and feedforward connections between the neurons. These connections act as memory functions for the RNN, allowing the efficient processing of time-series solar data.	Inability to deal with long-term data dependencies because of vanishing gradient and gradient explosion issues.
Jeon and Kim [127]	2020	GRU	The model generated substantial forecasting with an RMSE of 30 W/m^2.	It is computationally cheaper because it requires lesser parameter training and memory, hence faster execution.	It is a struggle to capture long-term dependencies in sequences.
Srivastava and Lessmann [107]	2018	LSTM	The LSTM model demonstrated 52.2% greater accuracy than the smart persistence model.	It provides extra storage capacity to the network and handles temporal dependencies in data.	It has the capacity to handle temporal dependencies in data, but it may face challenges in capturing and understanding intricate relationships or patterns that span extended periods.
Brahma and Wadhvani [150]	2020	Bi-LSTM, Attention Mechanism.	Bi-LSTM outperforms other deep learning models.	It uses double processing, which is beneficial for learning complex temporal correlations.	Its computational time is more than that of LSTM because the forward hidden sequence is calculated first, and the reverse hidden sequence is then merged to find the result.
Zang, Liu, Sun, Cheng, Wei and Sun [49]	2020	CNN	CNN demonstrated an increase of 17.06% in accuracy forecasting compared to the smart persistent model.	It is superior at extracting spatial correlations from weather data, such as cloud cover and other features; CNN has a strong ability to extract non-linear features.	Though it effectively captures and understands intricate relationships or patterns that span over extended periods, it may face challenges in handling temporal dependencies in data.
Zang, Cheng, Ding, Cheung, Wang, Wei and Sun [146]	2020	DBN	DBN illustrated an excellent forecasting score with an RMSE of 0.282 W/m^2 and an MAE of 0.137 W/m^2.	It is suitable for extracting input features for renewable energy forecasts because it employs layer-by-layer unsupervised training.	DBN is primarily designed for modeling static data distributions and does not naturally handle sequential or temporal data.

Table 8. *Cont.*

Author	Year	Model	Result	Advantage	Limitation
Kumari and Toshniwal [102]	2021	CNN and LSTM	LSTM-CNN model is more accurate in a range of about 37–45% compared to the benchmark models.	It handles temporal and spatial dependencies in time-series data and extracts non-linear features in data.	It takes longer to train than a single model because it needs to extract both spatial and temporal features from the data.
Brahma and Wadhvani [150]	2020	Attention Mechanism	The neural network that was paired with the Attention Mechanism was able to perform most effectively.	The Attention Mechanism supports deep learning models to engage in enhanced feature extraction for a sequence of inputs.	The pairwise calculations required by Attention Mechanisms can result in increased computational complexity, leading to longer training and inference times.
Ghimire, Deo, Casillas-Pérez, Salcedo-Sanz, Sharma and Ali [172]	2022	Hybrid forecasting model (CNN, LSTM, and Multilayer-Perceptron).	The results demonstrated strong performance metrics (r > 0.9 in all cases) and improvements of more than 10% for daily, monthly, and seasonal periods.	The hybrid model focuses on combining different methods while exploiting the benefits of individual forecasting models to enhance the accuracy of deep learning networks.	It takes longer to train than other models because it needs to extract both spatial and temporal features from the data.
Wang, Zhang, Liu, Yu, Pang, Duić, Shafie-Khah and Catalão [33]	2019	Generative algorithm (GAN)	The simulation outcomes demonstrated that Generative Adversarial Networks have the potential to build high fineness quality specimens that catch the underlying properties of the original data.	It makes more training data available by adding new data devised from training data.	Training GAN can be computationally expensive and resource-intensive, particularly for larger and more complex models.

Deep learning techniques have robust performance, especially when dealing with spatial and temporal features. However, small datasets may not offer appropriate or adequate representative training data, which may result in model overfitting. They are not sufficient when dealing with extreme weather types because they do not have sufficient data for training. Data augmentation is a way to make deep learning, and especially deep networks, work better. This process makes more training data available by adding new data made of either old data or new copies of old data that have been changed. Such an example can be seen in [33], whose authors used Convolutional Neural Networks and Generative Adversarial Networks to categorize the weather into ten categories. Based on observation, the Generative Adversarial Networks have the potential to build high fineness-quality samples that catch the underlying properties of the original data and better classification performance than benchmark models. Thus, GAN is ideal for data creation to augment and rebalance the training dataset.

Table 8 presents a summary of the main related studies, including their advantages and disadvantages. This study elaborated several popular deep learning models, such as LSTM, CNN, GAN, and an Attention Mechanism, focusing on their application in solar irradiance forecasting. The study also highlighted improvements in model training by increasing the training data and addressing the issue of overfitting. Additionally, the table includes information about the architecture, strengths, and limitations associated with these models.

6. The Challenge and Future Direction of Solar Energy Forecasting Using Deep Learning

Deep learning has already been used to make solar energy projections, and there has been a lot of research into this topic. Deep learning-based forecasting models, on the other hand, confront the two below real-world issues. Solving these issues will further increase the accuracy of the deep learning forecasting model [178].

Theoretical issues: Theoretical concerns related to deep learning are generally exhibited in two areas: statistics and calculating abilities. Any non-linear function can be represented by a shallow or deep network. The deep learning model outperforms the shallow model in non-linear representation. When it comes to solar and renewable energy forecasts, it is necessary to understand how difficult predicting samples can be, as well as how many training samples and how much processing power are required to build the deep learning network required to train these forecasting samples.

Modeling problems: These problems are related to dealing with large volumes of data. Accordingly, the extracted feature data from large-scale forecasting samples will become more relevant in parallel to the evolution of deep learning. The main objective of deep learning is to improve forecasting accuracy by learning more relevant information directly and spontaneously. The mechanism of deep learning models consists of hidden neurons with up to 6, 7, or even 10 layers, which allow more efficient feature learning than other shallow learning models. Moreover, deep learning makes it easier to learn data from solar energy time series through layer-by-layer feature learning. Despite this fact, developing an efficient hierarchical model with strong feature learning is challenging. Thus, the key future directions for deep learning-based solar energy forecasting models are highlighted as follows:

1. Data augmentation: The use of data augmentation techniques make deep learning, and especially deep networks, work more effectively; therefore, it is always recommended to develop novel methods for generating data that will better serve this purpose. This work also paves the way for future research in a variety of areas, including photovoltaic power forecasting, wind power forecasting, and electricity consumption forecasting in the medium- and long-term, etc., by expanding and embracing the modern technique suggested in this work through data generation; in contrast, most studies of photovoltaic power, wind power, and energy consumption concentrate on

the short-term forecast, disregarding medium- and long-term forecasts due to a lack of training data for models.

2. Probability forecasting: Many studies on deterministic renewable energy forecasting have been published. However, there is still limited work performed on deep learning-based probabilistic forecasting models. For solar energy time series data, the probabilistic forecasting model can be used to objectively quantify the level of uncertainty. Hence, it is essential to the operation and administration of electric power and energy systems daily [178].

3. Unified Predictive Model: Solar and renewable energy data have a variety of deep features in different seasons and meteorological and topographic situations. Accordingly, the forecasting model varies based on the situation. However, it is difficult to determine whether a forecasting model could work effectively with different types of data because of the unique solar energy datasets used in existing research.

4. Features Extraction Method: The use of feature extraction is important for effective solar energy data forecasting; therefore, it is always recommended to develop a novel future extraction method for deep forecasting data that will better serve this purpose. Currently, there is only one deep learning strategy that is used for feature extraction in the current deep learning forecasting model. Thus, it is recommended to develop feature extraction for deep forecasting features by combining different deep learning algorithms for future work.

5. Hybrid physical forecasting models: In general, it is normal to include numerical weather forecasting data in short-term forecasting models to improve solar energy forecasting performance. However, it is challenging to determine an effective way to include the relationships between many ground measures in a deep learning-based forecasting model. Thus, this topic is considered one of the most important future directions of study.

7. Conclusions

A comprehensive review of numerous deep learning-based models used to forecast solar irradiance and accessible in the literature is provided in this paper. These models' sophisticated architectures make it easier to extract high-level, non-linear features from the solar data. This study elaborates on several popular deep learning models, such as LTSM, DBN, CNN, RNN, RGU, GAN, and an Attention Mechanism based on solar irradiance forecasting, and improves the training of models by increasing the training data and avoiding overfitting models. These models also include their associated architecture, strengths, and limitations. In addition, the influential factors of forecasting model accuracy, such as forecasting horizons, weather classifications, input feature optimization, and assessment measures, are presented. Based on the literature review, the existing research demonstrates the superiority of deep learning-based solar irradiance forecasting models compared to conventional machine learning approaches. In most cases, the forecasting accuracy performance is further enhanced by incorporating the hybrid of deep learning models.

Thus, the hybrid model, which extracts both temporal and spatial data features, enhances accuracy even further and should be preferred. Additionally, extending the data are an approach that improves the performance of deep learning, particularly deep networks. It involves augmenting the training data by adding new data generated from existing data or creating modified copies of old data. Furthermore, conducting correlation analysis between features and irradiation and removing redundant information in the feature set leads to reduced model complexity and improved forecasting. Based on the literature provided, hybrid LSTM and CNN models demonstrated superior performances compared to those of other models considered. The main findings of this study can be summarized as follows:

1. The forecasting horizon is considered the key factor for improving accuracy in forecasting models. The performance of a solar irradiance forecasting model typically

degrades as the forecasting horizon lengthens. Therefore, forecasting models should be considered based on the selected forecasting horizon.

2. The use of a single deep learning model is limited. Our findings show that the implementation of hybrid models and using different models can improve accuracy and are preferred over basic deep learning models. The training duration should be considered during performance evaluation, since the hybrid model requires the extraction of both temporal and spatial features of data; therefore, this model has a longer training time compared to other models.

3. Model complexity is also a key factor. While hybrid models perform better than individual models, their computational time is generally higher. There is a trade-off between the accuracy of the model and its computational complexity and time. Therefore, using high-performance processors and large data frameworks may be difficult.

4. Changes in the weather influence how effectively a forecasting model performs. For example, some forecasting models generate lower accuracy for sky circumstances, while others have substantially better performance. Therefore, it is important to prioritize weather classification to ensure accuracy in forecasting.

5. The implementation of deep learning algorithms in solar irradiance forecasting has been widely explored. Some of the most used algorithms include LSTM, CNN, and DBN, while rarely used algorithms include GRU, ESN, and RNN. These models, however, have trade-offs because they can be highly precise while also being computationally expensive. In conclusion, deep learning techniques are still in their infancy, and their potential to resolve challenging time-series forecasting challenges should be further investigated.

The recently discussed comparative analysis based on existing deep learning models introduced in this study will further serve as a reference for future academics, planners, and forecasting experts who use solar energy systems to determine the most appropriate deep learning model required to achieve an excellent forecasting model.

References

1. Duffy, A.; Rogers, M.; Ayompe, L. *Renewable Energy and Energy Efficiency: Assessment of Projects and Policies*; John Wiley & Sons: Hoboken, NJ, USA, 2015.
2. Das, U.K.; Tey, K.S.; Seyedmahmoudian, M.; Mekhilef, S.; Idris, M.Y.I.; Van Deventer, W.; Horan, B.; Stojcevski, A. Forecasting of photovoltaic power generation and model optimization: A review. *Renew. Sustain. Energy Rev.* **2018**, *81*, 912–928. [CrossRef]
3. Dincer, I. Energy and environmental impacts: Present and future perspectives. *Energy Sour.* **1998**, *20*, 427–453. [CrossRef]

4.   Campbell-Lendrum, D.; Prüss-Ustün, A. Climate change, air pollution and noncommunicable diseases. *Bull. World Health Organ.* **2019**, *97*, 160. [CrossRef]

5.   van der Hoeven, M. *Technology Roadmap-Solar Photovoltaic Energy*; International Energy Agency: Paris, France, 2014.

6.   Zhang, Y.; Beaudin, M.; Taheri, R.; Zareipour, H.; Wood, D. Day-ahead power output forecasting for small-scale solar photovoltaic electricity generators. *IEEE Trans. Smart Grid* **2015**, *6*, 2253–2262. [CrossRef]

7.   Yang, C.; Thatte, A.A.; Xie, L. Multitime-scale data-driven spatio-temporal forecast of photovoltaic generation. *IEEE Trans. Sustain. Energy* **2014**, *6*, 104–112. [CrossRef]

8.   Haque, A.U.; Nehrir, M.H.; Mandal, P. Solar PV power generation forecast using a hybrid intelligent approach. In Proceedings of the 2013 IEEE Power & Energy Society General Meeting, Vancouver, BC, Canada, 21–25 July 2013; pp. 1–5.

9.   RG Al-Shakarchi, M.; Ghulaim, M.M. Short-term load forecasting for baghdad electricity region. *Electr. Mach. Power Syst.* **2000**, *28*, 355–371. [CrossRef]

10.   Keyno, H.S.; Ghaderi, F.; Azade, A.; Razmi, J. Forecasting electricity consumption by clustering data in order to decline the periodic variable's affects and simplification the pattern. *Energy Convers. Manag.* **2009**, *50*, 829–836. [CrossRef]

11.   Doorga, J.R.S.; Dhurmea, K.R.; Rughooputh, S.; Boojhawon, R. Forecasting mesoscale distribution of surface solar irradiation using a proposed hybrid approach combining satellite remote sensing and time series models. *Renew. Sustain. Energy Rev.* **2019**, *104*, 69–85. [CrossRef]

12.   Zhang, X.; Li, Y.; Lu, S.; Hamann, H.F.; Hodge, B.-M.; Lehman, B. A solar time based analog ensemble method for regional solar power forecasting. *IEEE Trans. Sustain. Energy* **2018**, *10*, 268–279. [CrossRef]

13.   Kumari, P.; Toshniwal, D. Deep learning models for solar irradiance forecasting: A comprehensive review. *J. Clean. Prod.* **2021**, *318*, 128566. [CrossRef]

14.   Duvenhage, D.F. Sustainable Future CSP Fleet Deployment in South Africa: A Hydrological Approach to Strategic Management. Ph.D. Thesis, Stellenbosch University, Stellenbosch, South Africa, 2019.

15.   Alfares, H.K.; Nazeeruddin, M. Electric load forecasting: Literature survey and classification of methods. *Int. J. Syst. Sci.* **2002**, *33*, 23–34. [CrossRef]

16.   Pedregal, D.J.; Trapero, J.R. Mid-term hourly electricity forecasting based on a multi-rate approach. *Energy Convers. Manag.* **2010**, *51*, 105–111. [CrossRef]

17.   Weron, R. *Modeling and Forecasting Electricity Loads and Prices: A Statistical Approach*; John Wiley & Sons: Hoboken, NJ, USA, 2007.

18.   Lappalainen, K.; Wang, G.C.; Kleissl, J. Estimation of the largest expected photovoltaic power ramp rates. *Appl. Energy* **2020**, *278*, 115636. [CrossRef]

19.   Husein, M.; Chung, I.-Y. Day-ahead solar irradiance forecasting for microgrids using a long short-term memory recurrent neural network: A deep learning approach. *Energies* **2019**, *12*, 1856. [CrossRef]

20.   Hammer, A.; Heinemann, D.; Lorenz, E.; Lückehe, B. Short-term forecasting of solar radiation: A statistical approach using satellite data. *Sol. Energy* **1999**, *67*, 139–150. [CrossRef]

21.   Nonnenmacher, L.; Coimbra, C.F. Streamline-based method for intra-day solar forecasting through remote sensing. *Sol. Energy* **2014**, *108*, 447–459. [CrossRef]

22.   Murata, A.; Ohtake, H.; Oozeki, T. Modeling of uncertainty of solar irradiance forecasts on numerical weather predictions with the estimation of multiple confidence intervals. *Renew. Energy* **2018**, *117*, 193–201. [CrossRef]

23.   Perez, R.; Lorenz, E.; Pelland, S.; Beauharnois, M.; Van Knowe, G.; Hemker, K., Jr.; Heinemann, D.; Remund, J.; Müller, S.C.; Traunmüller, W. Comparison of numerical weather prediction solar irradiance forecasts in the US, Canada and Europe. *Sol. Energy* **2013**, *94*, 305–326. [CrossRef]

24.   Hao, Y.; Tian, C. A novel two-stage forecasting model based on error factor and ensemble method for multi-step wind power forecasting. *Appl. Energy* **2019**, *238*, 368–383. [CrossRef]

25.   Singh, S.; Mohapatra, A. Repeated wavelet transform based ARIMA model for very short-term wind speed forecasting. *Renew. Energy* **2019**, *136*, 758–768.

26.   Huang, J.; Korolkiewicz, M.; Agrawal, M.; Boland, J. Forecasting solar radiation on an hourly time scale using a Coupled AutoRegressive and Dynamical System (CARDS) model. *Sol. Energy* **2013**, *87*, 136–149. [CrossRef]

27.   Reikard, G. Predicting solar radiation at high resolutions: A comparison of time series forecasts. *Sol. Energy* **2009**, *83*, 342–349. [CrossRef]

28.   Kumari, P.; Toshniwal, D. Real-time estimation of COVID-19 cases using machine learning and mathematical models—The case of India. In Proceedings of the 2020 IEEE 15th International Conference on Industrial and Information Systems (ICIIS), Rupnagar, India, 26–28 November 2020; pp. 369–374.

29.   Lee, J.; Wang, W.; Harrou, F.; Sun, Y. Reliable solar irradiance prediction using ensemble learning-based models: A comparative study. *Energy Convers. Manag.* **2020**, *208*, 112582. [CrossRef]

30.   Bouzgou, H.; Gueymard, C.A. Minimum redundancy–maximum relevance with extreme learning machines for global solar radiation forecasting: Toward an optimized dimensionality reduction for solar time series. *Sol. Energy* **2017**, *158*, 595–609. [CrossRef]

31.   Hou, K.; Shao, G.; Wang, H.; Zheng, L.; Zhang, Q.; Wu, S.; Hu, W. Research on practical power system stability analysis algorithm based on modified SVM. *Prot. Control Mod. Power Syst.* **2018**, *3*, 11. [CrossRef]

32. Zhang, T.; Lv, C.; Ma, F.; Zhao, K.; Wang, H.; O'Hare, G.M. A photovoltaic power forecasting model based on dendritic neuron networks with the aid of wavelet transform. *Neurocomputing* **2020**, *397*, 438–446. [CrossRef]

33. Wang, F.; Zhang, Z.; Liu, C.; Yu, Y.; Pang, S.; Duić, N.; Shafie-Khah, M.; Catalão, J.P. Generative adversarial networks and convolutional neural networks based weather classification model for day ahead short-term photovoltaic power forecasting. *Energy Convers. Manag.* **2019**, *181*, 443–462. [CrossRef]

34. McCandless, T.; Haupt, S.; Young, G. A regime-dependent artificial neural network technique for short-range solar irradiance forecasting. *Renew. Energy* **2016**, *89*, 351–359. [CrossRef]

35. AlKandari, M.; Ahmad, I. Solar power generation forecasting using ensemble approach based on deep learning and statistical methods. *Appl. Comput. Inform.* **2020**, *2020*, 1–20. [CrossRef]

36. Ramli, M.A.; Twaha, S.; Al-Turki, Y.A. Investigating the performance of support vector machine and artificial neural networks in predicting solar radiation on a tilted surface: Saudi Arabia case study. *Energy Convers. Manag.* **2015**, *105*, 442–452. [CrossRef]

37. Zang, H.; Cheng, L.; Ding, T.; Cheung, K.W.; Wei, Z.; Sun, G. Day-ahead photovoltaic power forecasting approach based on deep convolutional neural networks and meta learning. *Int. J. Electr. Power Energy Syst.* **2020**, *118*, 105790. [CrossRef]

38. Singh, P.; Singh, N.K.; Singh, A.K. Solar Photovoltaic Energy Forecasting Using Machine Learning and Deep Learning Technique. In Proceedings of the 2022 IEEE 9th Uttar Pradesh Section International Conference on Electrical, Electronics and Computer Engineering (UPCON), Prayagraj, India, 2–4 December 2022; pp. 1–6.

39. Wang, Z.; Hong, T.; Piette, M.A. Building thermal load prediction through shallow machine learning and deep learning. *Appl. Energy* **2020**, *263*, 114683. [CrossRef]

40. Li, L.; Yuan, Z.; Gao, Y. Maximization of energy absorption for a wave energy converter using the deep machine learning. *Energy* **2018**, *165*, 340–349. [CrossRef]

41. Hu, Q.; Zhang, R.; Zhou, Y. Transfer learning for short-term wind speed prediction with deep neural networks. *Renew. Energy* **2016**, *85*, 83–95. [CrossRef]

42. Mishra, M.; Dash, P.B.; Nayak, J.; Naik, B.; Swain, S.K. Deep learning and wavelet transform integrated approach for short-term solar PV power prediction. *Measurement* **2020**, *166*, 108250. [CrossRef]

43. Qing, X.; Niu, Y. Hourly day-ahead solar irradiance prediction using weather forecasts by LSTM. *Energy* **2018**, *148*, 461–468. [CrossRef]

44. Kuremoto, T.; Kimura, S.; Kobayashi, K.; Obayashi, M. Time series forecasting using a deep belief network with restricted Boltzmann machines. *Neurocomputing* **2014**, *137*, 47–56. [CrossRef]

45. Kaba, K.; Sarıgül, M.; Avcı, M.; Kandırmaz, H.M. Estimation of daily global solar radiation using deep learning model. *Energy* **2018**, *162*, 126–135. [CrossRef]

46. Yadav, A.K.; Chandel, S. Solar radiation prediction using Artificial Neural Network techniques: A review. *Renew. Sustain. Energy Rev.* **2014**, *33*, 772–781. [CrossRef]

47. Pazikadin, A.R.; Rifai, D.; Ali, K.; Malik, M.Z.; Abdalla, A.N.; Faraj, M.A. Solar irradiance measurement instrumentation and power solar generation forecasting based on Artificial Neural Networks (ANN): A review of five years research trend. *Sci. Total Environ.* **2020**, *715*, 136848. [CrossRef]

48. Voyant, C.; Notton, G.; Kalogirou, S.; Nivet, M.-L.; Paoli, C.; Motte, F.; Fouilloy, A. Machine learning methods for solar radiation forecasting: A review. *Renew. Energy* **2017**, *105*, 569–582. [CrossRef]

49. Zang, H.; Liu, L.; Sun, L.; Cheng, L.; Wei, Z.; Sun, G. Short-term global horizontal irradiance forecasting based on a hybrid CNN-LSTM model with spatiotemporal correlations. *Renew. Energy* **2020**, *160*, 26–41. [CrossRef]

50. Antonanzas-Torres, F.; Urraca, R.; Polo, J.; Perpiñán-Lamigueiro, O.; Escobar, R. Clear sky solar irradiance models: A review of seventy models. *Renew. Sustain. Energy Rev.* **2019**, *107*, 374–387. [CrossRef]

51. 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]

52. Engerer, N. Minute resolution estimates of the diffuse fraction of global irradiance for southeastern Australia. *Sol. Energy* **2015**, *116*, 215–237. [CrossRef]

53. Olatomiwa, L.; Mekhilef, S.; Shamshirband, S.; Mohammadi, K.; Petković, D.; Sudheer, C. A support vector machine–firefly algorithm-based model for global solar radiation prediction. *Sol. Energy* **2015**, *115*, 632–644. [CrossRef]

54. Heng, J.; Wang, J.; Xiao, L.; Lu, H. Research and application of a combined model based on frequent pattern growth algorithm and multi-objective optimization for solar radiation forecasting. *Appl. Energy* **2017**, *208*, 845–866. [CrossRef]

55. Kaushika, N.; Tomar, R.; Kaushik, S. Artificial neural network model based on interrelationship of direct, diffuse and global solar radiations. *Sol. Energy* **2014**, *103*, 327–342. [CrossRef]

56. Jiang, Y. Prediction of monthly mean daily diffuse solar radiation using artificial neural networks and comparison with other empirical models. *Energy Policy* **2008**, *36*, 3833–3837. [CrossRef]

57. Hong, T.; Wilson, J.; Xie, J. Long term probabilistic load forecasting and normalization with hourly information. *IEEE Trans. Smart Grid* **2013**, *5*, 456–462. [CrossRef]

58. Perez, R.; Kivalov, S.; Schlemmer, J.; Hemker, K., Jr.; Renné, D.; Hoff, T.E. Validation of short and medium term operational solar radiation forecasts in the US. *Sol. Energy* **2010**, *84*, 2161–2172. [CrossRef]

59. Li, J.; Ward, J.K.; Tong, J.; Collins, L.; Platt, G. Machine learning for solar irradiance forecasting of photovoltaic system. *Renew. Energy* **2016**, *90*, 542–553. [CrossRef]

60. Wang, Z.; Tian, C.; Zhu, Q.; Huang, M. Hourly solar radiation forecasting using a volterra-least squares support vector machine model combined with signal decomposition. *Energies* **2018**, *11*, 68. [CrossRef]

61. Dong, Z.; Yang, D.; Reindl, T.; Walsh, W.M. A novel hybrid approach based on self-organizing maps, support vector regression and particle swarm optimization to forecast solar irradiance. *Energy* **2015**, *82*, 570–577. [CrossRef]

62. Royer, J.C.; Wilhelm, V.E.; Junior, L.A.T.; Franco, E.M.C. Short-term solar radiation forecasting by using an iterative combination of wavelet artificial neural networks. *Indep. J. Manag. Prod.* **2016**, *7*, 271–288. [CrossRef]

63. Bae, K.Y.; Jang, H.S.; Sung, D.K. Hourly solar irradiance prediction based on support vector machine and its error analysis. *IEEE Trans. Power Syst.* **2016**, *32*, 935–945. [CrossRef]

64. Kwon, Y.; Kwasinski, A.; Kwasinski, A. Solar irradiance forecast using naïve Bayes classifier based on publicly available weather forecasting variables. *Energies* **2019**, *12*, 1529. [CrossRef]

65. Yu, Y.; Cao, J.; Zhu, J. An LSTM short-term solar irradiance forecasting under complicated weather conditions. *IEEE Access* **2019**, *7*, 145651–145666. [CrossRef]

66. Wang, F.; Zhen, Z.; Mi, Z.; Sun, H.; Su, S.; Yang, G. Solar irradiance feature extraction and support vector machines based weather status pattern recognition model for short-term photovoltaic power forecasting. *Energy Build.* **2015**, *86*, 427–438. [CrossRef]

67. Wang, F.; Yu, Y.; Zhang, Z.; Li, J.; Zhen, Z.; Li, K. Wavelet decomposition and convolutional LSTM networks based improved deep learning model for solar irradiance forecasting. *Appl. Sci.* **2018**, *8*, 1286. [CrossRef]

68. Akarslan, E.; Hocaoglu, F.O.; Edizkan, R. Novel short term solar irradiance forecasting models. *Renew. Energy* **2018**, *123*, 58–66. [CrossRef]

69. Wang, W.; Zhen, Z.; Li, K.; Lv, K.; Wang, F. An ultra-short-term forecasting model for high-resolution solar irradiance based on SOM and deep learning algorithm. In Proceedings of the 2019 IEEE Sustainable Power and Energy Conference (iSPEC), Beijing, China, 21–23 November 2019; pp. 1090–1095.

70. Lima, F.J.; Martins, F.R.; Pereira, E.B.; Lorenz, E.; Heinemann, D. Forecast for surface solar irradiance at the Brazilian Northeastern region using NWP model and artificial neural networks. *Renew. Energy* **2016**, *87*, 807–818. [CrossRef]

71. Stehman, S.V. Selecting and interpreting measures of thematic classification accuracy. *Remote Sens. Environ.* **1997**, *62*, 77–89. [CrossRef]

72. Wang, F.; Zhen, Z.; Liu, C.; Mi, Z.; Shafie-khah, M.; Catalão, J.P. Time-section fusion pattern classification based day-ahead solar irradiance ensemble forecasting model using mutual iterative optimization. *Energies* **2018**, *11*, 184. [CrossRef]

73. Nann, S.; Riordan, C. Solar spectral irradiance under clear and cloudy skies: Measurements and a semiempirical model. *J. Appl. Meteorol. Climatol.* **1991**, *30*, 447–462. [CrossRef]

74. Aguiar, L.M.; Pereira, B.; Lauret, P.; Díaz, F.; David, M. Combining solar irradiance measurements, satellite-derived data and a numerical weather prediction model to improve intra-day solar forecasting. *Renew. Energy* **2016**, *97*, 599–610. [CrossRef]

75. Kumari, P.; Wadhvani, R. Wind power prediction using klms algorithm. In Proceedings of the 2018 International Conference on Inventive Research in Computing Applications (ICIRCA), Coimbatore, India, 11–12 July 2018; pp. 154–161.

76. Gutierrez-Corea, F.-V.; Manso-Callejo, M.-A.; Moreno-Regidor, M.-P.; Manrique-Sancho, M.-T. Forecasting short-term solar irradiance based on artificial neural networks and data from neighboring meteorological stations. *Sol. Energy* **2016**, *134*, 119–131. [CrossRef]

77. Ozgoren, M.; Bilgili, M.; Sahin, B. Estimation of global solar radiation using ANN over Turkey. *Expert Syst. Appl.* **2012**, *39*, 5043–5051. [CrossRef]

78. Shaddel, M.; Javan, D.S.; Baghernia, P. Estimation of hourly global solar irradiation on tilted absorbers from horizontal one using Artificial Neural Network for case study of Mashhad. *Renew. Sustain. Energy Rev.* **2016**, *53*, 59–67. [CrossRef]

79. Yagli, G.M.; Yang, D.; Srinivasan, D. Automatic hourly solar forecasting using machine learning models. *Renew. Sustain. Energy Rev.* **2019**, *105*, 487–498. [CrossRef]

80. Bashir, Z.; El-Hawary, M. Applying wavelets to short-term load forecasting using PSO-based neural networks. *IEEE Trans. Power Syst.* **2009**, *24*, 20–27. [CrossRef]

81. Bao, Y.; Liu, Z. A fast grid search method in support vector regression forecasting time series. In Proceedings of the International Conference on Intelligent Data Engineering and Automated Learning, Burgos, Spain, 20–23 September 2006; pp. 504–511.

82. Li, H.; Guo, S.; Zhao, H.; Su, C.; Wang, B. Annual electric load forecasting by a least squares support vector machine with a fruit fly optimization algorithm. *Energies* **2012**, *5*, 4430–4445. [CrossRef]

83. Niu, D.; Wang, Y.; Wu, D.D. Power load forecasting using support vector machine and ant colony optimization. *Expert Syst. Appl.* **2010**, *37*, 2531–2539. [CrossRef]

84. Hong, W.-C. Application of chaotic ant swarm optimization in electric load forecasting. *Energy Policy* **2010**, *38*, 5830–5839. [CrossRef]

85. Hong, W.-C. Electric load forecasting by seasonal recurrent SVR (support vector regression) with chaotic artificial bee colony algorithm. *Energy* **2011**, *36*, 5568–5578. [CrossRef]

86. Hong, W.-C. Electric load forecasting by support vector model. *Appl. Math. Model.* **2009**, *33*, 2444–2454. [CrossRef]

87. Tao, Y.; Chen, Y. Distributed PV power forecasting using genetic algorithm based neural network approach. In Proceedings of the 2014 International Conference on Advanced Mechatronic Systems, Kumamoto, Japan, 10–12 August 2014; pp. 557–560.

88. Pedro, H.T.; Coimbra, C.F. Assessment of forecasting techniques for solar power production with no exogenous inputs. *Sol. Energy* **2012**, *86*, 2017–2028. [CrossRef]

89. Zhen, Z.; Pang, S.; Wang, F.; Li, K.; Li, Z.; Ren, H.; Shafie-khah, M.; Catalão, J.P. Pattern classification and PSO optimal weights based sky images cloud motion speed calculation method for solar PV power forecasting. *IEEE Trans. Ind. Appl.* **2019**, *55*, 3331–3342. [CrossRef]

90. Wang, F.; Zhou, L.; Ren, H.; Liu, X. Search improvement process-chaotic optimization-particle swarm optimization-elite retention strategy and improved combined cooling-heating-power strategy based two-time scale multi-objective optimization model for stand-alone microgrid operation. *Energies* **2017**, *10*, 1936. [CrossRef]

91. Wang, F.; Zhou, L.; Wang, B.; Wang, Z.; Shafie-Khah, M.; Catalão, J.P. Modified chaos particle swarm optimization-based optimized operation model for stand-alone CCHP microgrid. *Appl. Sci.* **2017**, *7*, 754. [CrossRef]

92. Mellit, A.; Kalogirou, S.A. Artificial intelligence techniques for photovoltaic applications: A review. *Prog. Energy Combust. Sci.* **2008**, *34*, 574–632. [CrossRef]

93. Lee, J.-T.; Kim, H.-G.; Kang, Y.-H.; Yun, C.-Y.; Kim, C.K.; Kim, B.-Y.; Kim, J.-Y.; Park, Y.Y.; Kim, T.H.; Jo, H.N. Trend Review of Solar Energy Forecasting Technique. *J. Korean Sol. Energy Soc.* **2019**, *39*, 41–54.

94. Rajagukguk, R.A.; Ramadhan, R.A.; Lee, H.-J. A review on deep learning models for forecasting time series data of solar irradiance and photovoltaic power. *Energies* **2020**, *13*, 6623. [CrossRef]

95. Hameed, W.I.; Sawadi, B.A.; Al-Kamil, S.J.; Al-Radhi, M.S.; Al-Yasir, Y.I.; Saleh, A.L.; Abd-Alhameed, R.A. Prediction of solar irradiance based on artificial neural networks. *Inventions* **2019**, *4*, 45. [CrossRef]

96. Paoli, C.; Voyant, C.; Muselli, M.; Nivet, M.-L. Forecasting of preprocessed daily solar radiation time series using neural networks. *Sol. Energy* **2010**, *84*, 2146–2160. [CrossRef]

97. Kehl, C.; Varbanescu, A.L. Towards Distributed, Semi-Automatic Content-Based Visual Information Retrieval (CBVIR) of Massive Media Archives. *Adv. Neural Inf. Process. Syst.* **2012**, *2012*, 1097–1105.

98. Yamashita, R.; Nishio, M.; Do, R.K.G.; Togashi, K. Convolutional neural networks: An overview and application in radiology. *Insights Into Imaging* **2018**, *9*, 611–629. [CrossRef]

99. Li, G.; Wu, S.X.; Zhang, S.; Li, Q. Detect insider attacks using CNN in decentralized optimization. In Proceedings of the ICASSP 2020—2020 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), Barcelona, Spain, 4–8 May 2020; pp. 8758–8762.

100. Gu, J.; Wang, Z.; Kuen, J.; Ma, L.; Shahroudy, A.; Shuai, B.; Liu, T.; Wang, X.; Wang, G.; Cai, J. Recent advances in convolutional neural networks. *Pattern Recognit.* **2018**, *77*, 354–377. [CrossRef]

101. Zhou, D.-X. Universality of deep convolutional neural networks. *Appl. Comput. Harmon. Anal.* **2020**, *48*, 787–794. [CrossRef]

102. Kumari, P.; Toshniwal, D. Long short term memory–convolutional neural network based deep hybrid approach for solar irradiance forecasting. *Appl. Energy* **2021**, *295*, 117061. [CrossRef]

103. Hüsken, M.; Stagge, P. Recurrent neural networks for time series classification. *Neurocomputing* **2003**, *50*, 223–235. [CrossRef]

104. Hochreiter, S.; Schmidhuber, J. Long short-term memory. *Neural Comput.* **1997**, *9*, 1735–1780. [CrossRef] [PubMed]

105. Bandara, K.; Bergmeir, C.; Smyl, S. Forecasting across time series databases using recurrent neural networks on groups of similar series: A clustering approach. *Expert Syst. Appl.* **2020**, *140*, 112896. [CrossRef]

106. Liu, Y.; Guan, L.; Hou, C.; Han, H.; Liu, Z.; Sun, Y.; Zheng, M. Wind power short-term prediction based on LSTM and discrete wavelet transform. *Appl. Sci.* **2019**, *9*, 1108. [CrossRef]

107. Srivastava, S.; Lessmann, S. A comparative study of LSTM neural networks in forecasting day-ahead global horizontal irradiance with satellite data. *Sol. Energy* **2018**, *162*, 232–247. [CrossRef]

108. Cho, K.; Van Merriënboer, B.; Gulcehre, C.; Bahdanau, D.; Bougares, F.; Schwenk, H.; Bengio, Y. Learning phrase representations using RNN encoder-decoder for statistical machine translation. *arXiv* **2014**, arXiv:1406.1078.

109. Schuster, M.; Paliwal, K.K. Bidirectional recurrent neural networks. *IEEE Trans. Signal Process.* **1997**, *45*, 2673–2681. [CrossRef]

110. Ogawa, A.; Hori, T. Error detection and accuracy estimation in automatic speech recognition using deep bidirectional recurrent neural networks. *Speech Commun.* **2017**, *89*, 70–83. [CrossRef]

111. Hinton, G.E.; Osindero, S.; Teh, Y.-W. A fast learning algorithm for deep belief nets. *Neural Comput.* **2006**, *18*, 1527–1554. [CrossRef]

112. Chen, Y.; Zhao, X.; Jia, X. Spectral–spatial classification of hyperspectral data based on deep belief network. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2015**, *8*, 2381–2392. [CrossRef]

113. Lee, H.; Ekanadham, C.; Ng, A. Sparse deep belief net model for visual area V2. *Adv. Neural Inf. Process. Syst.* **2007**, *20*, 1–8.

114. Qiao, J.; Wang, G.; Li, W.; Li, X. A deep belief network with PLSR for nonlinear system modeling. *Neural Netw.* **2018**, *104*, 68–79. [CrossRef] [PubMed]

115. Bahdanau, D.; Cho, K.; Bengio, Y. Neural machine translation by jointly learning to align and translate. *arXiv* **2014**, arXiv:1409.0473.

116. Sutskever, I.; Vinyals, O.; Le, Q.V. Sequence to sequence learning with neural networks. *Adv. Neural Inf. Process. Syst.* **2014**, *27*, 777–780.

117. Choi, H.; Cho, K.; Bengio, Y. Fine-grained attention mechanism for neural machine translation. *Neurocomputing* **2018**, *284*, 171–176. [CrossRef]

118. Song, K.; Yao, T.; Ling, Q.; Mei, T. Boosting image sentiment analysis with visual attention. *Neurocomputing* **2018**, *312*, 218–228. [CrossRef]

119. Li, W.; Guo, D.; Fang, X. Multimodal architecture for video captioning with memory networks and an attention mechanism. *Pattern Recognit. Lett.* **2018**, *105*, 23–29. [CrossRef]

120. Goodfellow, I.; Pouget-Abadie, J.; Mirza, M.; Xu, B.; Warde-Farley, D.; Ozair, S.; Courville, A.; Bengio, Y. Generative adversarial networks. *Commun. ACM* **2020**, *63*, 139–144. [CrossRef]

121. Radford, A.; Metz, L.; Chintala, S. Unsupervised representation learning with deep convolutional generative adversarial networks. *arXiv* **2015**, arXiv:1511.06434.

122. Wang, Y.; Li, H. A novel intelligent modeling framework integrating convolutional neural network with an adaptive time-series window and its application to industrial process operational optimization. *Chemom. Intell. Lab. Syst.* **2018**, *179*, 64–72. [CrossRef]

123. Creswell, A.; White, T.; Dumoulin, V.; Arulkumaran, K.; Sengupta, B.; Bharath, A.A. Generative adversarial networks: An overview. *IEEE Signal Process. Mag.* **2018**, *35*, 53–65. [CrossRef]

124. Mirza, M.; Osindero, S. Conditional generative adversarial nets. *arXiv* **2014**, arXiv:1411.1784.

125. Muhammad, A.; Lee, J.M.; Hong, S.W.; Lee, S.J.; Lee, E.H. Deep learning application in power system with a case study on solar irradiation forecasting. In Proceedings of the 2019 International Conference on Artificial Intelligence in Information and Communication (ICAIIC), Okinawa, Japan, 11–13 February 2019; pp. 275–279.

126. Chandola, D.; Gupta, H.; Tikkiwal, V.A.; Bohra, M.K. Multi-step ahead forecasting of global solar radiation for arid zones using deep learning. *Procedia Comput. Sci.* **2020**, *167*, 626–635. [CrossRef]

127. Jeon, B.-k.; Kim, E.-J. Next-day prediction of hourly solar irradiance using local weather forecasts and LSTM trained with non-local data. *Energies* **2020**, *13*, 5258. [CrossRef]

128. Sorkun, M.C.; Paoli, C.; Incel, Ö.D. Time series forecasting on solar irradiation using deep learning. In Proceedings of the 2017 10th International Conference on Electrical and Electronics Engineering (ELECO), Bursa, Turkey, 30 November–2 December 2017; pp. 151–155.

129. Alzahrani, A.; Shamsi, P.; Dagli, C.; Ferdowsi, M. Solar irradiance forecasting using deep neural networks. *Procedia Comput. Sci.* **2017**, *114*, 304–313. [CrossRef]

130. Alzahrani, A.; Shamsi, P.; Ferdowsi, M.; Dagli, C. Solar irradiance forecasting using deep recurrent neural networks. In Proceedings of the 2017 IEEE 6th International Conference on Renewable Energy Research and Applications (ICRERA), San Diego, CA, USA, 5–8 November 2017; pp. 988–994.

131. Obiora, C.N.; Ali, A.; Hasan, A.N. Forecasting Hourly Solar Irradiance Using Long Short-Term Memory (LSTM) Network. In Proceedings of the 2020 11th International Renewable Energy Congress (IREC), Hammamet, Tunisia, 29–31 October 2020; pp. 1–6.

132. Chu, T.-P.; Jhou, J.-H.; Leu, Y.-G. Image-based Solar Irradiance Forecasting Using Recurrent Neural Networks. In Proceedings of the 2020 International Conference on System Science and Engineering (ICSSE), Kagawa, Japan, 31 August–3 September 2020; pp. 1–4.

133. Mukherjee, A.; Ain, A.; Dasgupta, P. Solar irradiance prediction from historical trends using deep neural networks. In Proceedings of the 2018 IEEE International Conference on Smart Energy Grid Engineering (SEGE), Ottawa, ON, Canada, 12–15 August 2018; pp. 356–361.

134. Ashfaq, Q.; Ulasyar, A.; Zad, H.S.; Khattak, A.; Imran, K. Hour-ahead global horizontal irradiance forecasting using long short term memory network. In Proceedings of the 2020 IEEE 23rd International Multitopic Conference (INMIC), Bahawalpur, Pakistan, 5–7 November 2020; pp. 1–6.

135. Justin, D.; Concepcion, R.S.; Calinao, H.A.; Alejandrino, J.; Dadios, E.P.; Sybingco, E. Using stacked long short term memory with principal component analysis for short term prediction of solar irradiance based on weather patterns. In Proceedings of the 2020 IEEE Region 10 Conference (TENCON), Osaka, Japan, 16–19 November 2020; pp. 946–951.

136. Fernando, W.; Jayalath, W.; Kanagasundaram, A.; Valluvan, R. *Solar Irradiance Forecasting Using Deep Learning Approaches*; Research Repository; University of Jaffna: Kokuvil East, Sri Lanka, 2019.

137. Mishra, S.; Palanisamy, P. An integrated multi-time-scale modeling for solar irradiance forecasting using deep learning. *arXiv* **2019**, arXiv:1905.02616.

138. Hossain, M.S.; Mahmood, H. Short-term photovoltaic power forecasting using an LSTM neural network and synthetic weather forecast. *IEEE Access* **2020**, *8*, 172524–172533. [CrossRef]

139. Zhao, X.; Wei, H.; Wang, H.; Zhu, T.; Zhang, K. 3D-CNN-based feature extraction of ground-based cloud images for direct normal irradiance prediction. *Sol. Energy* **2019**, *181*, 510–518. [CrossRef]

140. Feng, C.; Zhang, J. SolarNet: A sky image-based deep convolutional neural network for intra-hour solar forecasting. *Sol. Energy* **2020**, *204*, 71–78. [CrossRef]

141. Wojtkiewicz, J.; Hosseini, M.; Gottumukkala, R.; Chambers, T.L. Hour-ahead solar irradiance forecasting using multivariate gated recurrent units. *Energies* **2019**, *12*, 4055. [CrossRef]

142. Yan, K.; Shen, H.; Wang, L.; Zhou, H.; Xu, M.; Mo, Y. Short-term solar irradiance forecasting based on a hybrid deep learning methodology. *Information* **2020**, *11*, 32. [CrossRef]

143. Mukhoty, B.P.; Maurya, V.; Shukla, S.K. Sequence to sequence deep learning models for solar irradiation forecasting. In Proceedings of the 2019 IEEE Milan PowerTech, Milano, Italy, 23–27 June 2019; pp. 1–6.

144. Li, Q.; Wu, Z.; Ling, R.; Feng, L.; Liu, K. Multi-reservoir echo state computing for solar irradiance prediction: A fast yet efficient deep learning approach. *Appl. Soft Comput.* **2020**, *95*, 106481. [CrossRef]

145. Pang, Z.; Niu, F.; O'Neill, Z. Solar radiation prediction using recurrent neural network and artificial neural network: A case study with comparisons. *Renew. Energy* **2020**, *156*, 279–289. [CrossRef]

146. Zang, H.; Cheng, L.; Ding, T.; Cheung, K.W.; Wang, M.; Wei, Z.; Sun, G. Application of functional deep belief network for estimating daily global solar radiation: A case study in China. *Energy* **2020**, *191*, 116502. [CrossRef]

147. Li, Z.; Wang, K.; Li, C.; Zhao, M.; Cao, J. Multimodal deep learning for solar irradiance prediction. In Proceedings of the 2019 International Conference on Internet of Things (iThings) and IEEE Green Computing and Communications (GreenCom) and IEEE Cyber, Physical and Social Computing (CPSCom) and IEEE Smart Data (SmartData), Atlanta, GA, USA, 14–17 July 2019; pp. 784–792.

148. Lago, J.; De Brabandere, K.; De Ridder, F.; De Schutter, B. Short-term forecasting of solar irradiance without local telemetry: A generalized model using satellite data. *Sol. Energy* **2018**, *173*, 566–577. [CrossRef]

149. Peng, T.; Zhang, C.; Zhou, J.; Nazir, M.S. An integrated framework of Bi-directional long-short term memory (BiLSTM) based on sine cosine algorithm for hourly solar radiation forecasting. *Energy* **2021**, *221*, 119887. [CrossRef]

150. Brahma, B.; Wadhvani, R. Solar irradiance forecasting based on deep learning methodologies and multi-site data. *Symmetry* **2020**, *12*, 1830. [CrossRef]

151. Li, Q.; Wu, Z.; Zhang, H. Spatio-temporal modeling with enhanced flexibility and robustness of solar irradiance prediction: A chain-structure echo state network approach. *J. Clean. Prod.* **2020**, *261*, 121151. [CrossRef]

152. Andrianakos, G.; Tsourounis, D.; Oikonomou, S.; Kastaniotis, D.; Economou, G.; Kazantzidis, A. Sky Image forecasting with Generative Adversarial Networks for cloud coverage prediction. In Proceedings of the 2019 10th International Conference on Information, Intelligence, Systems and Applications (IISA), Patras, Greece, 15–17 July 2019; pp. 1–7.

153. Prado-Rujas, I.-I.; García-Dopico, A.; Serrano, E.; Pérez, M.S. A flexible and robust deep learning-based system for solar irradiance forecasting. *IEEE Access* **2021**, *9*, 12348–12361. [CrossRef]

154. Yeom, J.-M.; Deo, R.C.; Adamowski, J.F.; Park, S.; Lee, C.-S. Spatial mapping of short-term solar radiation prediction incorporating geostationary satellite images coupled with deep convolutional LSTM networks for South Korea. *Environ. Res. Lett.* **2020**, *15*, 094025. [CrossRef]

155. Ziyabari, S.; Du, L.; Biswas, S. A spatio-temporal hybrid deep learning architecture for short-term solar irradiance forecasting. In Proceedings of the 2020 47th IEEE Photovoltaic Specialists Conference (PVSC), Calgary, AB, Canada, 15 June–21 August 2020; pp. 0833–0838.

156. Ghimire, S.; Deo, R.C.; Raj, N.; Mi, J. Deep solar radiation forecasting with convolutional neural network and long short-term memory network algorithms. *Appl. Energy* **2019**, *253*, 113541. [CrossRef]

157. Huang, X.; Zhang, C.; Li, Q.; Tai, Y.; Gao, B.; Shi, J. A comparison of hour-ahead solar irradiance forecasting models based on LSTM network. *Math. Probl. Eng.* **2020**, *2020*, 4251517. [CrossRef]

158. Luong, M.-T.; Pham, H.; Manning, C.D. Effective approaches to attention-based neural machine translation. *arXiv* **2015**, arXiv:1508.04025.

159. He, H.; Lu, N.; Jie, Y.; Chen, B.; Jiao, R. Probabilistic solar irradiance forecasting via a deep learning-based hybrid approach. *IEEJ Trans. Electr. Electron. Eng.* **2020**, *15*, 1604–1612. [CrossRef]

160. Brahma, B.; Wadhvani, R.; Shukla, S. Attention mechanism for developing wind speed and solar irradiance forecasting models. *Wind Eng.* **2021**, *45*, 1422–1432. [CrossRef]

161. Siddiqui, T.A.; Bharadwaj, S.; Kalyanaraman, S. A deep learning approach to solar-irradiance forecasting in sky-videos. In Proceedings of the 2019 IEEE Winter Conference on Applications of Computer Vision (WACV), Waikoloa Village, HI, USA, 7–11 January 2019; pp. 2166–2174.

162. Kumari, P.; Toshniwal, D. Extreme gradient boosting and deep neural network based ensemble learning approach to forecast hourly solar irradiance. *J. Clean. Prod.* **2021**, *279*, 123285. [CrossRef]

163. Rai, A.; Shrivastava, A.; Jana, K.C. A CNN-BiLSTM based deep learning model for mid-term solar radiation prediction. *Int. Trans. Electr. Energy Syst.* **2021**, *31*, e12664. [CrossRef]

164. Chen, Y.; Shi, J.; Cheng, X.; Ma, X. Hybrid Models Based on LSTM and CNN Architecture with Bayesian Optimization for ShortTerm Photovoltaic Power Forecasting. In Proceedings of the 2021 IEEE/IAS Industrial and Commercial Power System Asia (I&CPS Asia), Chengdu, China, 18–21 July 2021; pp. 1415–1422.

165. Ahmad, R.; Kumar, R. Very Short-Term Photovoltaic (PV) Power Forecasting Using Deep Learning (LSTMs). In Proceedings of the 2021 International Conference on Intelligent Technologies (CONIT), Grimstad, Norway, 25–27 June 2021; pp. 1–6.

166. Li, Y.; Ye, F.; Liu, Z.; Wang, Z.; Mao, Y. A Short-Term Photovoltaic Power Generation Forecast Method Based on LSTM. *Math. Probl. Eng.* **2021**, *2021*, 6613123. [CrossRef]

167. Jebli, I.; Belouadha, F.-Z.; Kabbaj, M.I.; Tilioua, A. Deep learning based models for solar energy prediction. *Adv. Sci* **2021**, *6*, 349–355. [CrossRef]

168. Jalali, S.M.J.; Ahmadian, S.; Kavousi-Fard, A.; Khosravi, A.; Nahavandi, S. Automated deep CNN-LSTM architecture design for solar irradiance forecasting. *IEEE Trans. Syst. Man Cybern. Syst.* **2021**, *52*, 54–65. [CrossRef]

169. Wang, K.; Qi, X.; Liu, H. Photovoltaic power forecasting based LSTM-Convolutional Network. *Energy* **2019**, *189*, 116225. [CrossRef]

170. Boubaker, S.; Benghanem, M.; Mellit, A.; Lefza, A.; Kahouli, O.; Kolsi, L. Deep neural networks for predicting solar radiation at Hail Region, Saudi Arabia. *IEEE Access* **2021**, *9*, 36719–36729. [CrossRef]

171. Kolsi, L.; Al-Dahidi, S.; Kamel, S.; Aich, W.; Boubaker, S.; Ben Khedher, N. Prediction of Solar Energy Yield Based on Artificial Intelligence Techniques for the Ha'il Region, Saudi Arabia. *Sustainability* **2022**, *15*, 774. [CrossRef]

172. Ghimire, S.; Deo, R.C.; Casillas-Pérez, D.; Salcedo-Sanz, S.; Sharma, E.; Ali, M. Deep learning CNN-LSTM-MLP hybrid fusion model for feature optimizations and daily solar radiation prediction. *Measurement* **2022**, *202*, 111759. [CrossRef]

173. Peng, Y.; Wang, S.; Chen, W.; Ma, J.; Wang, C.; Chen, J. LightGBM-Integrated PV Power Prediction Based on Multi-Resolution Similarity. *Processes* **2023**, *11*, 1141. [CrossRef]

174. Ghimire, S.; Nguyen-Huy, T.; Prasad, R.; Deo, R.C.; Casillas-Perez, D.; Salcedo-Sanz, S.; Bhandari, B. Hybrid convolutional neural network-multilayer perceptron model for solar radiation prediction. *Cogn. Comput.* **2023**, *15*, 645–671. [CrossRef]

175. Mishra, S.; Palanisamy, P. Multi-time-horizon solar forecasting using recurrent neural network. In Proceedings of the 2018 IEEE Energy Conversion Congress and Exposition (ECCE), Portland, OR, USA, 23–27 September 2018; pp. 18–24.

176. Bendali, W.; Saber, I.; Bourachdi, B.; Boussetta, M.; Mourad, Y. Deep learning using genetic algorithm optimization for short term solar irradiance forecasting. In Proceedings of the 2020 Fourth International Conference on Intelligent Computing in Data Sciences (ICDS), Fez, Morocco, 21–23 October 2020; pp. 1–8.

177. Zhou, H.; Zhang, Y.; Yang, L.; Liu, Q.; Yan, K.; Du, Y. Short-term photovoltaic power forecasting based on long short term memory neural network and attention mechanism. *IEEE Access* **2019**, *7*, 78063–78074. [CrossRef]

178. Wang, H.; Lei, Z.; Zhang, X.; Zhou, B.; Peng, J. A review of deep learning for renewable energy forecasting. *Energy Convers. Manag.* **2019**, *198*, 111799. [CrossRef]