

Article

Anomaly Identification for Photovoltaic Power Stations Using a Dual Classification System and Gramian Angular Field Visualization

Zihan Wang ¹ , Qiushi Cui ^{1,*} , Zhuowei Gong ¹ , Lixian Shi ¹ , Jie Gao ¹ and Jiayong Zhong ²¹ Chongqing University, Chongqing 400044, China; wang5zh@mail.uc.edu (Z.W.); gongze@mail.uc.edu (Z.G.); shilixian1996@163.com (L.S.); gaojie0268@126.com (J.G.)² State Grid Chongqing Electric Power Company Electric Power Science Research Institute, Chongqing 401123, China; liaozhong10000@163.com

* Correspondence: qcui@cqu.edu.cn

Abstract: With the increasing scale of photovoltaic (PV) power stations, timely anomaly detection through analyzing the PV output power curve is crucial. However, overlooking the impact of external factors on the expected power output would lead to inaccurate identification of PV station anomalies. This study focuses on the discrepancy between measured and expected PV power generation values, using a dual classification system. The system leverages two-dimensional Gramian angular field (GAF) data and curve features extracted from one-dimensional time series, along with attention weights from a CNN network. This approach effectively classifies anomalies, including normal operation, aging pollution, and arc faults, achieving an overall classification accuracy of 95.83%.

Keywords: PV power station; anomaly detection; time series data; Gramian angular field; CNN; attention matrix



Citation: Wang, Z.; Cui, Q.; Gong, Z.; Shi, L.; Gao, J.; Zhong, J. Anomaly Identification for Photovoltaic Power Stations Using a Dual Classification System and Gramian Angular Field Visualization. *Processes* **2024**, *12*, 690. <https://doi.org/10.3390/pr12040690>

Academic Editor: Sara Pescetelli

Received: 1 March 2024

Revised: 22 March 2024

Accepted: 26 March 2024

Published: 29 March 2024



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

1. Introduction

The exponential growth in the installed capacity of renewable power generation systems stands as a pivotal driver in transitioning towards sustainable electric power systems. Among the promising power sources, PV power generation is witnessing a substantial surge, with an increasing number of PV systems being deployed as decentralized power generation units in medium- and low-voltage transmission networks [1]. The variability and stochastic fluctuations in the output power of PV technology pose a considerable challenge as large-scale grid connections of PV power generation could significantly impact the power grid [2–4].

The power time-series data within solar power station datasets constitutes a fundamental aspect, as the magnitude of output power directly impacts the economic feasibility of power generation [5,6]. Additionally, the morphology of the power time curve serves as a crucial indicator of the operational health of the station [7]. Extensive research has been conducted to forecast the power generation of PV stations, with the aim of ensuring power system stability amidst the increasing integration of PV systems and proactively implementing adaptive measures [8]. Different patterns in power curves can reflect corresponding fault scenarios [9]. For instance, localized power declines or irregular changes on the curve may occur in PV panels near hot spots, contrasting with the behavior of normally operating panels. Various failures have been observed in practical PV systems, including issues such as series resistance disconnection, current leakage to ground, and module branch disconnection. Typically, these failures manifest when PV modules exhibit discoloration, cracking, snail tracks, or degradation of the antireflection coating [10]. The gradual degradation of components due to aging, pollution, and connection issues leads to a decline in power curve performance. Irregular fluctuations in the power time curve may indicate

arc faults, partial shielding, or local faults, while sudden declines or interruptions may suggest anomalies in equipment line faults.

Accurate photovoltaic generation forecasting is an important feature that can assist utilities and plant operators in the direction of energy management and dispatchability planning [11]. Approaches combining the weather data, PV station's physical parameters, historical generation data, and other conditions would be prioritized, then conditional factors, other than the factors of the photovoltaic power station, can be reflected in the prediction timing curve [12,13]. Digital twin technology utilizes interactive simulations among physical entities, sensors, and historical databases to establish a high-fidelity virtual mapping of real-world equipment. This enables the synchronous evolution of the digital twin and physical equipment throughout their entire life cycle [14–17]. Integrating the digital twin model into PV power forecasting enhances the system by combining environmental parameters, prediction data, historical data, and reference data, thereby boosting the efficiency and decision-making speed of the scheduling center [18–20]. This holistic approach ensures comprehensive and reliable forecasting, crucial for effective energy management.

The hybrid photovoltaic–thermoelectric generation system (PVTEG) efficiently utilizes solar energy by converting solar irradiance and thermal energy into electricity, while enhancing PV efficiency by reducing module surface temperature. Ref. [21] proposes a machine learning-based approach with efficient fault detection methods to achieve fast real-time global maximum power point tracking for addressing the low efficiency of PV and thermoelectric generation devices in hybrid PVTEG systems. This demonstrates the potential of machine learning and efficient fault detection methods to significantly enhance system performance. By identifying and scrutinizing instabilities in the power generation performance, the monitoring of abnormalities within power stations becomes possible, thus minimizing potential losses [22]. Deep learning (DL) has emerged as an effective tool for power system stability prediction due to its exceptional representational learning capabilities [23,24]. However, the application of DL methods has also received wide acceptance in stability assessment [25–28], which is not directly suitable for processing power system time-series data. DL models in stability analysis often suffer from overfitting due to redundant information, limiting robustness. Handling massive transient data increases storage demands, while diverse raw curve dynamics hinder transfer learning across different power grids.

To address this gap, advanced time-series methods such as the recurrence plot [29,30] present promising alternatives. The Markov Transition Field (MTF) captures transition frequencies between time stamps, and the Gramian angular field (GAF) is a valuable tool for converting time-series data into a format amenable to image-based analysis, making it compatible with CNNs and enabling effective feature extraction and pattern recognition in time-series datasets. Ref. [31] encodes a univariate time series while ref. [32] extends it to apply DIM identification to a multivariate time-series transformation with the help of the RGB channel. The proposed framework in [32] integrates curve filtering rules and a dual-channel VGGNet, enabling the simultaneous processing of rotor angle and voltage images. DL methodologies in [33] have enabled an accurate prediction by leveraging GAF representations and variable-order Markov models. Using daily time-series data as input, refs. [34–37] propose methods employed for classification. Ref. [38] proposes a spectral clustering-based anomaly detection method for mountainous PV power plants, which sorts current time-series data to partially eliminate the influence of dip angle and azimuth, and distinguishes anomalies in strings through spectral clustering. The model described in [34] utilizes statistical analysis, machine learning, or computer vision methods. These approaches for data classification have significant dimensional and computational requirements. Ref. [35] proposes an ensemble scheme for time-series data classification based on constructing classifiers on different data representations. The scheme employs standard baseline algorithms used in time-series classification research, including the 1-NN algorithm and methods using Euclidean distance and/or dynamic time warping. Ref. [36] proposes a method to identify anomalies in power quality (PQ) time-series data, addressing

the increasing volume of PQ data and the need for automated analysis. It focuses on short-term PQ characteristics to detect deviations from typical behavior caused by changes in customer or network behavior. An ensemble model based on the multi-channel deep CNN (MDCNN) mentioned in [37] describes a method based on time-series classification that utilizes a CNN and an ensemble model based on the PyTorch package's MDCNN to classify the installation status of PV systems within a 24 h period. The above approaches offer significant improvements over traditional statistical learning methods, enhancing the restoration of statistical characteristics in generated time-series data.

This study specifically focuses on using PV output active power data as input data, providing advantages such as an enhanced reflection of power station health through power timing curve morphology and reduced data parameters for faster calculations. It aligns with the broader context of power system stability research. The model outputs are categorized as normal, irregular fluctuations, and evident variance, addressing issues such as component aging, pollution, connection problems, and equipment or circuit faults leading to an abrupt decline in the system. This progression narrows down to the main research objective, excluding external factors like weather, and focuses on the differences between predicted and measured curves, utilizing Devi-preprocessed GAF (DP-GAF) encoding images with an adaptive matrix for the selective enhancement of image features. The main contributions of this paper are as follows:

- We propose an integrated analysis methodology that combines expected and actual daily output active power data through meticulous deviation calculations. This approach provides a detailed view of power station performance, enhancing anomaly interpretability, which is superior to directly detecting output active power curve data.
- Due to the difficulty of classifying images generated using traditional GAF encoding under different categories of PV power station conditions, we introduce a pre-processing method tailored to PV systems before GAF transformation. This method enhances the effectiveness of GAF encoding and leads to a better capture of curve features for more accurate anomaly detection.
- A convolutional neural network (CNN) architecture with attention from identified curve features is proposed, seamlessly combining DP-GAF-encoded images with curve evaluation metrics. This integrated approach significantly advances anomaly detection for PV power stations, promising heightened accuracy in identifying anomalies.

The remainder of this paper is organized as follows. Section 2 presents the overall structure of the proposed framework for anomaly identification, and the implementation process of the three modules is introduced in detail in this part. Section 3 presents how to train the CNN network and some performance metrics to evaluate the model. Experiments are conducted on two different power systems in Section 4. Conclusions are drawn in Section 5.

2. Overall Structure of the Proposed Framework for Anomaly Detection

The anomaly detection model framework proposed in this paper consists of three main components: the Devi-preprocessed Gramian angular field transformation, curve features calculation, and the integration of the CNN-based classification with the attention mechanism model. These components are intricately interconnected within the framework, as illustrated in Figure 1. We will explain the corresponding process in Figure 1 more in detail as we introduce the methodology in this section. Together, these components constitute a robust framework that seamlessly integrates advanced time-series analysis and machine learning techniques to enhance the discernment of anomalies in the intricate dynamics of photovoltaic power generation.

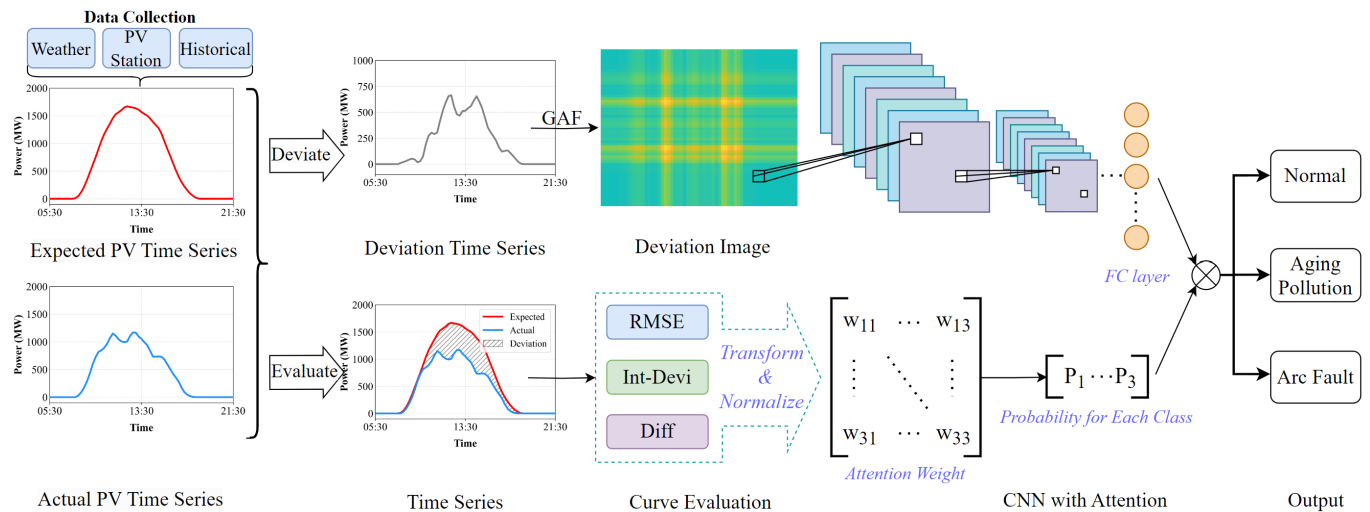


Figure 1. Overall framework of the CNN model with deviation series and GAF for anomaly identification.

2.1. Encoding Power Deviation Series by Gramian Angular Field

The first component of the proposed framework involves the utilization of the Gramian angular field to transform the preprocessed deviation time-series data. Inspired by the widespread use of CNN in computer vision, encoding time-series data as images and then using CNN for classification have also shown the potential of anomaly identification. The time-series information in the transient process is rich, and it is crucial to extract it reasonably to help with anomaly detection. The proposed encoding method can, on the one hand, encode time series as images for convenient and efficient use of machine learning models, and on the other hand, provide a dynamic feature extraction method.

2.1.1. Data Collection

Using predicted solar irradiance and weather data, along with empirical formulas based on photovoltaic power station parameters, we employed an indirect prediction method. This method was complemented by a statistical learning model that considers historical power generation. The goal was to forecast the daily PV output active power curve one day in advance. The principle block diagram of the prediction model of generated power is shown in Figure 2. This forecast serves as the day-ahead forecast value for the time-series data of the PV output.

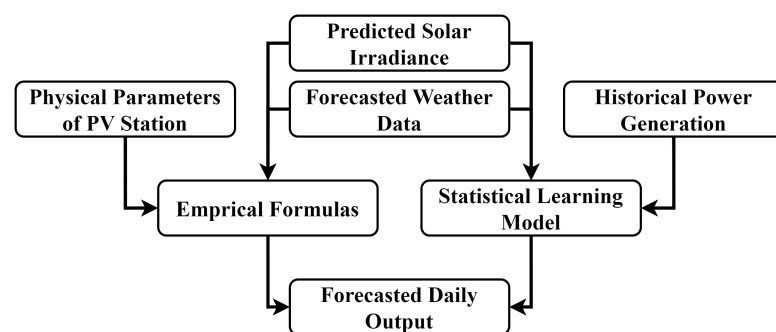


Figure 2. Principle block diagram of prediction model of generated power.

Therefore, as illustrated in the leftmost part of Figure 1, the expected PV time series can be obtained from data collection, along with the actual PV time series, leading to the next section on preprocessing and curve evaluation.

2.1.2. Preprocessing for Photovoltaic Characteristics

Before encoding with GAF, it is imperative to preprocess each set of deviation data to align with the daily output characteristics of photovoltaic systems. This refers to the deviation time series obtained through the “Deviate” step in Figure 1. Given the limited duration of sunlight each day, the output active power before sunrise and after sunset invariably registers as zero. Moreover, the timing of sunrise and sunset shifts with the seasons, resulting in variations in the time span of the non-zero photovoltaic output active power. Concurrently, seasonal changes may cause the output curve to contract or expand horizontally. To mitigate the impact of seasonal variations on zero-point changes in the power active output and to filter out such invalid information, we first redefined the range of each set of deviation data: the time corresponding to the peak of observed predicted power time-series data is denoted as t_m ; both preceding and following t_m , the predicted power diminishes until it reaches a certain proportion of the maximum predicted output power. The times corresponding to these points before and after t_m are taken as the initial and final values of the deviation data range, respectively. This preprocessing allows the deviation data to focus more precisely on the power station’s output scenario during operation, the effect of which is shown in Figure 3.

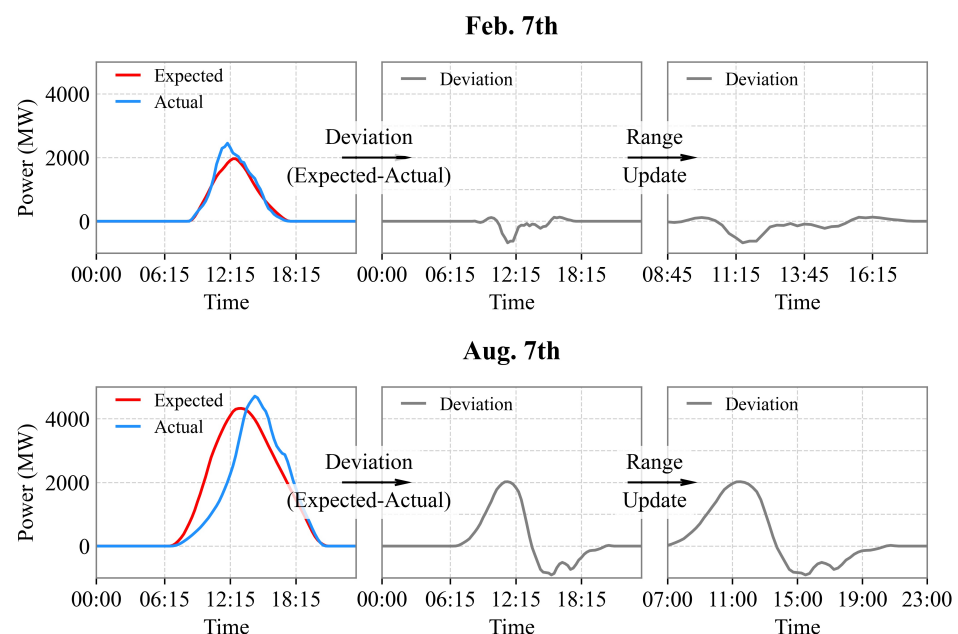


Figure 3. The effects after optimizing the time range.

Additionally, prior to the GAF transformation, data normalization was implemented to ensure comparability within a defined range. However, this normalization process may inadvertently have altered the data distribution, particularly near extreme values, leading to unwarranted fluctuations or anomalies post-normalization. To mitigate these fluctuations, an adaptive extreme value was introduced at both ends of the dataset. Initially, the maximum predicted value in each set of photovoltaic data was identified, and then its $a\%$ was computed. Subsequently, the entire deviation dataset was augmented by b , with the first element of the deviation set to 0, and the last element increased by b . This strategic approach effectively diminished the influence of extreme values during normalization, thus ensuring the stability and reliability of the data.

The above preprocessing was designed to improve the traditional encoding method, which exhibited reduced interference and enhanced the capturing of curve features during the generation of GAF images, providing more informative inputs for subsequent data analysis and classification.

2.1.3. GAF Encoding

Next, we proceeded to the step in Figure 1 where the deviation image was obtained through the GAF encoding with the PV-characterized preprocessing process. The processed power deviation time-series data $P = \{p_1, p_2, \dots, p_n\}$ of n values was encoded into an image G . Firstly, P was scaled to be $\tilde{P} = \{\tilde{p}_1, \tilde{p}_2, \dots, \tilde{p}_n\}$ with all values falling in the interval $[-1, 1]$:

$$\tilde{p} = \frac{[p_i - \max(p)] + [p_i - \min(p)]}{\max(p) - \min(p)}, \quad i = 1, 2, \dots, n, \quad (1)$$

where p is the original power deviation time-series data with n values, and p_i represents the i th value in the original power deviation time series. $\tilde{p}$ represents scaled power deviation time-series data, ensuring all values can be effectively encoded into an image representation.

The GAF is used to represent the inner product relationships within a set of vectors. For a given set of vectors $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n$, the element G_{ij} in the Gram Matrix represents the inner product between vectors $\mathbf{v}_i$ and $\mathbf{v}_j$. The calculation method for the Gram Matrix can be expressed by the following formula:

$$G_{GAF}(i, j) = \langle \tilde{\mathbf{p}}_i, \tilde{\mathbf{p}}_j \rangle = \tilde{\mathbf{p}}_i^T \cdot \tilde{\mathbf{p}}_j, \quad (2)$$

Here, $\langle \cdot, \cdot \rangle$ denotes the inner product operation, and $\tilde{\mathbf{p}}_i^T$ represents the transpose of vector $\tilde{\mathbf{p}}_i$. The element G_{ij} in the Gram Matrix is the result of the inner product between vectors $\tilde{\mathbf{p}}_i$ and $\tilde{\mathbf{p}}_j$ at corresponding positions.

The scaled value $\tilde{p}_i$ is then converted to polar coordinates (ϕ_i, r_i) for $1 \leq i \leq n$, where $\phi_i = \cos^{-1}(\tilde{p}_i)$, $p_i = \frac{r_i}{N}$, and N is a constant factor for regularizing all polar coordinates' span. The Gramian summation angular fields (GASFs) and Gramian difference angular fields (GADF) are defined as follows [3]:

$$\begin{cases} G_{GASF}(i, j) = [\cos(\phi_i + \phi_j)] \\ G_{GADF}(i, j) = [\sin(\phi_i - \phi_j)] \end{cases}, \quad i, j = 1, 2, \dots, n, \quad (3)$$

where $G_{GASF}(i, j)$ and $G_{GADF}(i, j)$ represent the elements in the GASF and GADF, respectively. They are computed as the cosine of the sum of angles and sine of the difference between angles, respectively.

While these three functions are all designed to compute Gramian angle fields, they employ different formulas when calculating the Gram Matrix, resulting in subtle differences in the obtained GAF. The choice of which to use depends on the modeling requirements for time-series features and the specific application context.

Overall, in this section, the Devi-preprocessed GAF, noted as DP-GAF, first calculated the deviation from the actual and predicted photovoltaic output active power data, then performed preprocessing based on the photovoltaic characteristics. After that, we used DP-GAF encoding to transform the data into images.

2.2. Extra Curve Features Calculation

Following the encoding transformation, identified curve features were calculated to encapsulate various aspects of the power generation process. These extra features serve as a holistic measure, incorporating factors such as root mean square error (RMSE), integral differences, and differential analysis. This calculated metric acts as a pivotal input for the subsequent phases of the framework. This corresponds to the "Curve Evaluation" part of Figure 1.

2.2.1. RMSE Computation

The RMSE is calculated as a measure of the differences between the forecasted and measured power values:

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (measure[i] - forecast[i])^2}, \quad (4)$$

where n is the number of data points in the measured and forecasted values, and $measure[i]$ and $forecast[i]$ represent the i th data point of the measured and forecasted values, respectively.

This component of the composite indicator gauges the precision of the forecasting model, providing an indication of how well the predicted values align with the actual measurements. RMSE contributes a nuanced understanding of the forecasting accuracy to the overall composite indicator.

2.2.2. Integral Differences Analysis

Integral differences, noted as Δ , are computed by integrating the absolute differences between the measured and forecasted power values.

$$\Delta = \sum_{i=1}^n |(measure[i] - forecast[i])|. \quad (5)$$

This formula calculates the cumulative sum of absolute differences between the measured and forecasted values at each time point, reflecting the significance of deviations across the entire time series. The integration provides a cumulative measure of dissimilarity between the two sets of data, emphasizing the importance of deviations over the entire time series. The resulting integral differences parameter enhances the composite indicator by providing a holistic perspective on the cumulative impact of forecast inaccuracies.

2.2.3. Differential Analysis

Integral differentials, noted as δ , are computed by integrating the absolute differences between the value of the curve at the previous time step and the value at the subsequent time step. So, we obtain the following formula:

$$\delta = \sum_{i=1}^n \left| \frac{measure(i + \delta t) - measure(i)}{\delta t} \right|, \quad (6)$$

which calculates the integral differential δ as the sum of the absolute differences between the values of the curve at consecutive time steps, normalized by the time interval δt . This parameter enables the measurement of fluctuations in the observed curve, thereby aiding in the classification of curve characteristics.

To ensure the consistent scaling of these metrics during subsequent neural network training, we performed normalization. Each metric was scaled to a predefined range, typically between 0 and 1, using the minimum–maximum normalization method.

Finally, these four normalized features were incorporated into the fully connected layer of the neural network. This integration aimed to provide the neural network with a comprehensive set of features, including dynamic characteristics from time-series analysis, enhancing its ability to identify anomalies. This holistic feature representation enabled the neural network to adapt to the complex dynamics of performance fluctuations in photovoltaic power stations.

This methodology contributes to a more nuanced understanding of deviations in power generation and enhances the interpretability and robustness of anomaly detection in the photovoltaic domain.

2.3. CNN-Based Classification with Attention Mechanism

The integration of the CNN-based classification with the attention mechanism model is central to our anomaly identification framework, corresponding to the “CNN With Attention” part of Figure 1. It is specifically designed to exploit intricate patterns encoded within the data transformed with DP-GAF. This section focuses on the process of integrating the CNN

with the attention mechanism based on three distinct features. The integration involves the generation of a training weight matrix and the subsequent application of attention during the classification phase.

The process begins with the generation of a pre-training weight matrix, characterized by the steps illustrated in Figure 4. This process involves the following:

1. The model takes m normalized features, calculated in section B, as input. Each of the m features undergoes a logarithmic transformation with base e .
2. The transformed feature values are normalized using min–max normalization, with min and max representing the minimum and maximum values within each respective feature.
3. The normalized features are convolved with a weight matrix.
4. The features are classified, and the weight matrix is subsequently updated.

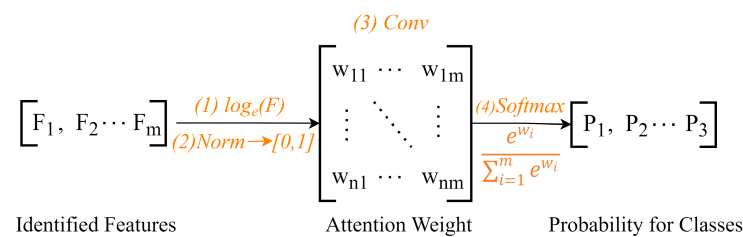


Figure 4. Pre-training and testing matrix weights.

During the classification phase, the weight matrix is applied as shown in Figure 5. The process includes the following steps:

1. Feature values are inputted into the system.
2. The input feature values undergo logarithmic transformation and are normalized using min–max normalization.
3. The transformed features are multiplied with the pre-trained weight matrix, yielding three values representing the transformation into probabilities.
4. Probabilities are computed, resulting in an output probability matrix.
5. Considering the potential inaccuracies in pre-trained models, where predicting the correct class with the lowest probability is highly improbable, a mapping is applied to ensure that the predicted class with the lowest probability is assigned a relatively low value. This effectively prevents the misidentification of images as the class with the lowest predicted probability.

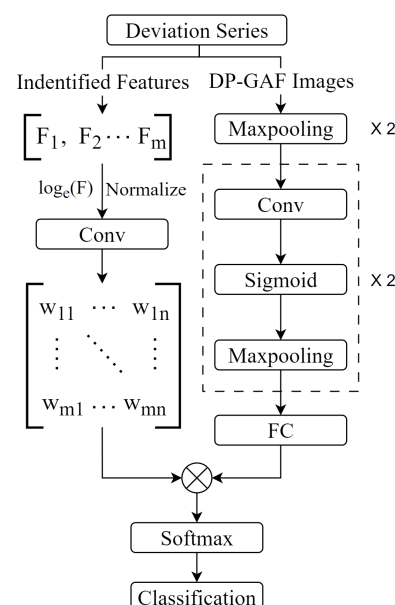


Figure 5. The integration of the CNN-based classification with the attention mechanism model.

In summary, the CNN with attention model integration optimizes classification performance by incorporating attention mechanisms based on feature metrics, pre-training weight matrices, and probability mapping strategies.

3. Model Training and Performance

Generating samples is the foundation of deep learning. In data analysis, the uncertainty fluctuations caused by hourly anomaly detection in the curve lead to a high probability of false alarms. Therefore, a specific power plant's daily active power generation time-series data can be chosen as the detection and analysis database. Each day's active power generation is predicted based on factors such as weather and power plant parameters, resulting in day-ahead forecast power generation data. The actual monitored active power generation each day serves as the real-time upscaled measurement data. By observing the active power generation waveforms containing two sets of time-series data each day, data with significant differences (concave, spiky fluctuations, or large deviations) between the measured and predicted curves can be classified as anomalous data for the power plant. Data with power generation curves close to the baseline for both curves can be classified as normal for the power plant. Following this classification rule, dates corresponding to anomalous and normal power generation can be obtained. Consequently, the deviation curves' GAF images for these corresponding dates are classified, resulting in the final classification dataset.

Our proposed model is built with the integration of an attention mechanism tailored for GAF data. The model was trained using a custom dataset specifically designed for GAF data. The dataset includes instances from normal, fluctuating, and evident variance conditions. The training process involved optimizing the model parameters to minimize the average cross-entropy loss, which is calculated as follows:

$$\mathcal{L}_{\text{class}} = -\frac{1}{N} \sum_{i=1}^N (y_i \cdot \log(p_i) + (1 - y_i) \cdot \log(1 - p_i)), \quad (7)$$

where N is the number of instances, y_i is the ground truth label, and p_i is the predicted probability.

To assess the performance of the attention mechanism, we computed the confusion matrix, denoted as $C_{\text{Attention}}$. The confusion matrix is a table used to evaluate the performance of a classification model. It displays the model's predictions for each class, including true positives, false positives, true negatives, and false negatives. By inputting the true labels of the test data and the model's predictions into a confusion matrix, the calculation function can be obtained.

Then, we calculated the accuracy, noted as A , for each class using the confusion matrix $C_{\text{Attention}}$. The accuracy for class i was computed as the ratio of the correct predictions to the total predictions for that class:

$$A_{\text{Class}}(i) = \frac{C_{\text{Attention}}(i, i)}{\sum_j C_{\text{Attention}}(i, j)}, \quad (8)$$

where $C_{\text{Attention}}(i, i)$ represents the element in the i -th row and i -th column of the matrix $C_{\text{Attention}}$, indicating the number of samples of class i that were correctly classified. $\sum C_{\text{Attention}}(i, :)$ denotes the total number of samples in the i -th row.

The overall accuracy, considering all classes, was computed as the ratio of the sum of correct predictions for each class (diagonal elements) to the total number of predictions:

$$A_{\text{Total}} = \frac{\sum \text{diag}(C_{\text{Attention}})}{\sum_{i,j} C_{\text{Attention}}(i, j)}, \quad (9)$$

where $\sum \text{diag}(C_{\text{Attention}})$ represents the sum of the diagonal elements of the matrix $C_{\text{Attention}}$, which corresponds to the total number of samples correctly classified. $\sum C_{\text{Attention}}(:)$ de-

notes the sum of all elements in the matrix $C_{\text{Attention}}$, representing the total number of samples in the dataset.

These calculations provide a comprehensive assessment of the model's performance, delving into the classification accuracy for each class and the overall accuracy. Such results are crucial for understanding the model's classification efficacy and the impact of the attention mechanism.

The integrated CNN model was trained using a stochastic gradient descent with momentum (SGDM) optimizer. The training process involved multiple epochs with a mini-batch size of 16, and the learning rate was initialized at 0.0002. We incorporated a validation set, consisting of both image and extra features, to monitor the model's performance during training.

The training loop updated the model parameters based on the gradient of the composite loss function with respect to the model parameters θ :

$$\theta \leftarrow \theta - \eta \cdot \nabla_{\theta} \mathcal{L}_{\text{class}}, \quad (10)$$

where η is the learning rate of the optimizer. $\nabla_{\theta} \mathcal{L}_{\text{class}}$ is the gradient of the loss function $\mathcal{L}_{\text{class}}$ with respect to the model parameters θ , indicating the rate of change of the loss function at the current parameter values. The evaluation of the model involves assessing its performance on a separate test dataset.

4. Experimental Evaluation

To validate the effectiveness of the proposed anomaly detection framework, multiple case studies were conducted. This section focuses on analyzing the active power generation curve of a photovoltaic (PV) power station in Belgium and the anomalies it reflects. The primary dataset was sourced from <https://www.elia.be/en/grid-data/power-generation/solar-pv-power-generation-data>, accessed on 20 March 2024. Specifically, we selected the solar power generation data for Belgium recorded from 2018 to 2023, when the system had a monitored capacity of 8.887 megawatts at its peak. For example, the measured and the day-ahead forecast active power output time series are presented as follows in Figure 6.

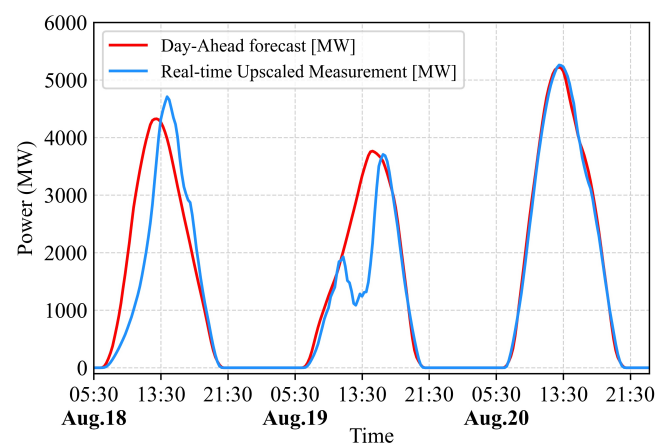


Figure 6. Solar PV measured and forecast (data source: [39]).

As observed, despite the preliminary consideration of factors such as illumination, weather conditions, power plant parameters, and historical power generation, there still exists a deviation between the day-ahead forecast curve of PV generation for the next day and the measured curve of the power plant. This phenomenon is evident in the contrast curves on 18 August in Figure 6, where a general shift is observed. On 19 August, it can be noted that there is a fault-induced abrupt decline in actual active power generation around noon, suggesting the possibility of sudden weather changes or other factors affecting the stability of the power plant on that day. However, on 20 August, the predicted curve closely

aligns with the measured curve, indicating normal operation of the power plant on that day without any apparent anomalies.

The deviation could be seen more evidently in the deviation curve in Figure 7. It is evident that the deviation data exhibits significantly larger deviations from zero on 18 and 19 August, whereas the curve for 20 August, when the output power is normal, shows only minor fluctuations near zero, indicating a more stable performance.

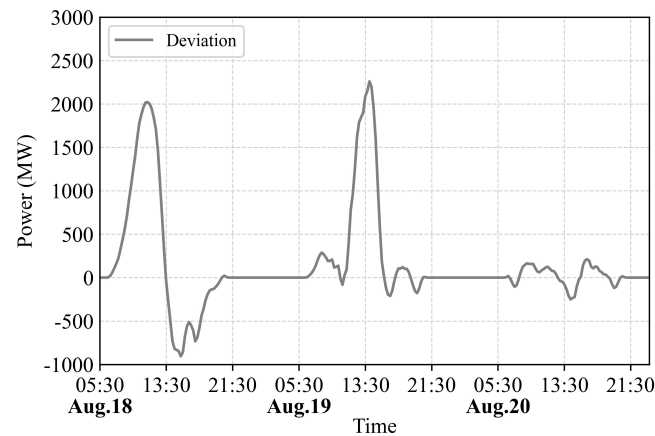


Figure 7. Deviation PV output curve.

Through these three steps in Section 2, we encoded the two-dimensional image data corresponding to the three different operational states of the photovoltaic power station with the features of each state's image as shown in Figure 8. Under normal conditions at the photovoltaic power station, the images we encoded were almost solid-colored. However, in cases of aging or pollution, irregular fluctuations appeared in the output power curve, resulting in a grid-like pattern in the images. In the event of an arc fault, which causes a sudden and significant drop in the curve, the images exhibited a distinct highlight.

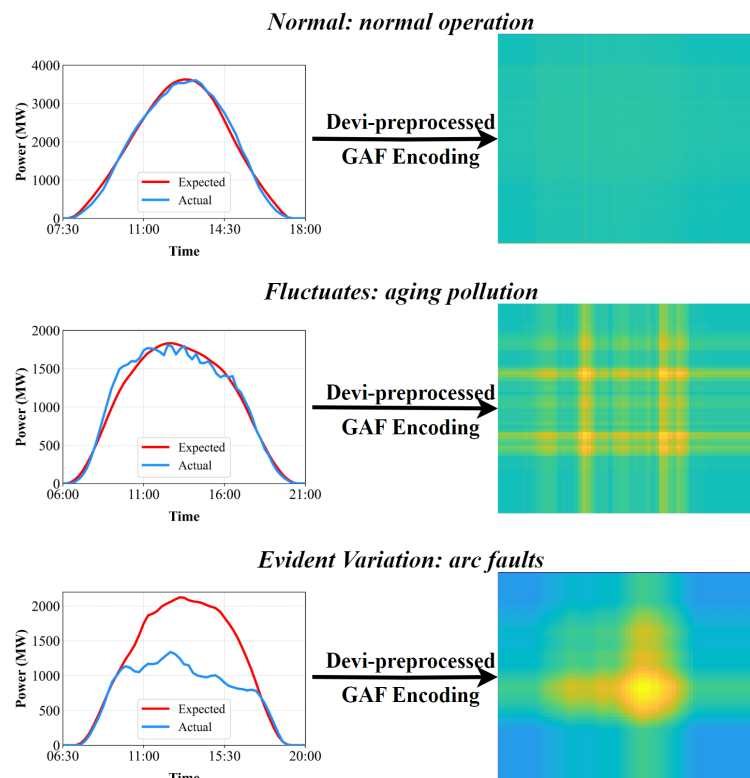


Figure 8. Devi-preprocessed GAF-encoded images of 3 states.

Subsequently, we further processed the deviation time-series data and encoded them into images containing information about the deviations using the Gramian angular field, as illustrated in Figure 9. The visual differences between images obtained from GAF encoding without and with preprocessing for photovoltaic characteristics were significant. In the former, the images exhibited a grid-like pattern under both normal and irregular fluctuations in the PV curve, with localized brightness observed in cases of evident variation. However, the images obtained with preprocessing for PV characteristics showed more distinct features in all three scenarios, facilitating easier differentiation: the image was a nearly solid color under normal conditions, displayed a grid-like pattern during fluctuations, and exhibited significant brightness under evident variation.

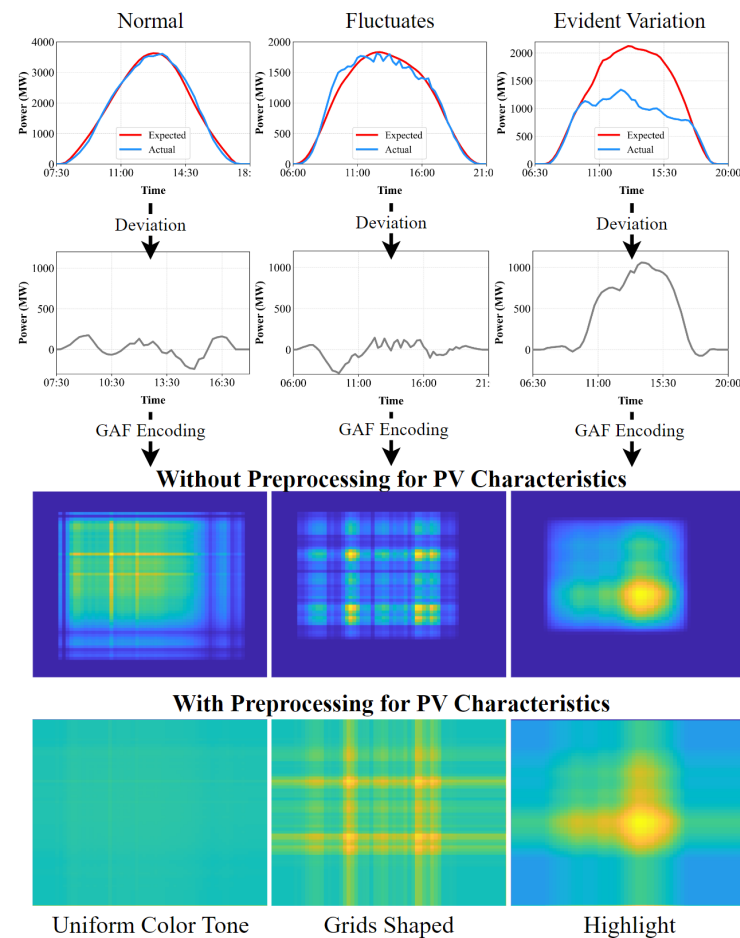


Figure 9. Comparison of GAF encoding with and without preprocessing for PV characteristics.

4.1. Dataset Description

The experimental evaluation was conducted on a dataset comprising images encoded using the DP-GAF method, representing time-series data. The original dataset was derived from daily PV output time-series data spanning each month from 2018 to 2023, totaling 2191 data points. After initial cleaning, 2000 both actual and expected data points were retained. This dataset includes three categories based on anomalies: normal (21%), aging pollution (47%), and arc faults (32%). These categories were processed using the GAF encoding with preprocessing for PV characteristics method mentioned in Section 2, resulting in an image dataset with corresponding labels. Subsequently, the dataset was divided into training and testing sets, with 80% of the data used for training and the remaining 20% for testing. This division ensured an ample amount of data for model training while enabling a robust evaluation of unseen samples.

4.2. Results

Upon completion of training, the model was evaluated on the test set, and the confusion matrix was computed to assess the classification performance. Additionally, class-specific accuracies provided insights into the model's effectiveness in identifying anomalies across different categories.

4.2.1. Classification Results of Different Inputs

To demonstrate that the deviation series more comprehensively incorporates factors related to the operation of the photovoltaic power station itself, and to enhance anomaly detection by excluding redundant information, both the traditionally measured time-series data and the deviation time-series data were employed as inputs to the network. Similarly, features without encoding images were utilized as input in order to show the effect of the deviation series encoded images. The classification accuracy results can be compared in Table 1.

Table 1. Test results of the model with different input series.

Input Series	$\mathcal{A}_{\text{Total}} \%$	$\mathcal{A}_{\text{Normal}} \%$	$\mathcal{A}_{\text{Fluct}} \%$	$\mathcal{A}_{\text{Evi-vari}} \%$
Deviation Series	95.83	99.0	98.21	89.74
Original Series	47.93	19.45	63.24	35.29
Features Only	71.89	72.80	44.13	61.68

In Table 1, the test results of the model with different input series are presented. The “Input Series” column indicates the type of input data used for classification, while the subsequent columns represent the classification accuracy (%) for different categories: “Total”, “Normal”, “Fluctuation”, and “Evident-Variability”.

The deviation series, incorporating deviation time-series data, demonstrates the highest overall classification accuracy of 95.83%. It exhibits remarkable performance in identifying normal instances (99.0%) and fluctuations (98.21%). However, it shows relatively lower accuracy in detecting instances of evident variability (89.74%).

Utilizing traditionally measured time-series data as input (original series) results in significantly lower classification accuracies compared to the deviation series, with an overall accuracy of only 47.93%. Particularly low accuracies are observed in identifying normal instances (19.45%) and evident variability (35.29%). Meanwhile, using features without encoding images (features only) yields higher accuracies than the original series but lower than the deviation series, with an overall accuracy of 71.89%. It demonstrates relatively high accuracies in identifying normal instances (72.80%) and evident variability (61.68%), but lower accuracy in detecting fluctuations (44.13%).

These results underscore the effectiveness of utilizing deviation time-series data, particularly when encoded into images, for anomaly detection in photovoltaic power stations. The deviation series outperforms both the original series and the features-only approach, indicating its ability to capture essential operational information and enhance anomaly detection capability.

4.2.2. Classification Results of Different Encoding Methods (with Attention)

The methods for encoding time series into images are not limited to the traditional Gramian angular field. Other methods include the traditional GADF, GASF, as well as MTF, and so on. In this study, an improved normalization approach was applied to adaptively enhance the encoding matrices. The resulting images from the enhanced encoding methods were more conducive to the classification task performed by the neural network in this study. To demonstrate the superiority of the proposed enhancement and to investigate which improved encoding method is optimal, Table 2 presents a comparison of classification accuracies with and without preprocessing for PV characteristics before GAF encoding, as well as those obtained using different encoding methods. As depicted in

Table 2, the GAF encoding method exhibits outstanding performance with preprocessing, achieving an impressive overall accuracy of 95.83%. It demonstrates high accuracy in distinguishing between normal, fluctuation, and evident variance scenarios, with accuracies of 99.0%, 98.21%, and 89.74%, respectively. Conversely, without preprocessing, the overall accuracy of the GAF encoding method drops to 88.33%, with significantly lower accuracies in normal and fluctuation scenarios. These findings underscore the critical role of preprocessing in enhancing the model's performance, particularly in the context of PV generation data analysis.

Table 2. Test results of the model with attention.

Approach	$\mathcal{A}_{\text{Total}}$ %	$\mathcal{A}_{\text{Normal}}$ %	$\mathcal{A}_{\text{Fluct}}$ %	$\mathcal{A}_{\text{Evi-vari}}$ %
With Preprocessing for PV Characteristics				
DP-GAF+CNN	95.83	99.0	98.21	89.74
DP-GASF+CNN	95.0	99.0	94.64	92.31
DP-GADF+CNN	95.0	95.0	96.43	92.31
DP-MTF+CNN	87.5	60.0	94.64	94.87
Without Preprocessing for PV Characteristics				
GAF+CNN	88.33	72.0	96.43	87.18
GASF+CNN	89.17	60.0	98.21	94.87
GADF+CNN	90.0	68.0	98.21	92.31
MTF+CNN	87.5	60.0	98.21	89.74

4.2.3. Classification Results of Different Encoding Methods (without Attention)

The CNN-based classification with attention mechanism serves as the core component of the model. The attention during image classification was optimized based on pre-trained weight matrices obtained from custom curve features. To demonstrate the enhancement in model detection results due to the optimized attention mechanism, Table 3 presents the classification accuracy results for various image encoding methods with and without attention optimization.

Table 3. Test results of the model without attention.

Approach	$\mathcal{A}_{\text{Total}}$ %	$\mathcal{A}_{\text{Normal}}$ %	$\mathcal{A}_{\text{Fluct}}$ %	$\mathcal{A}_{\text{Evi-vari}}$ %
With Attention Matrix				
DP-GAF+CNN	95.83	99.0	98.21	89.74
DP-GASF+CNN	95.0	99.0	94.64	92.31
DP-GADF+CNN	95.0	95.0	96.43	92.31
DP-MTF+CNN	87.5	60.0	94.64	94.87
Without Attention Matrix				
DP-GAF+CNN	92.5	88.0	96.43	89.74
DP-GASF+CNN	90.0	96.0	85.71	92.31
DP-GADF+CNN	82.5	48.0	91.07	92.31
DP-MTF+CNN	76.67	20.0	91.07	92.31

With the attention matrix, the GAF encoding method achieves an outstanding overall accuracy of 95.83%. It also demonstrates high accuracies in identifying normal (99.0%), fluctuation (98.21%), and evident variance (89.74%) scenarios. In contrast, without the attention matrix, the overall accuracy of the GAF encoding method drops to 92.5%, with lower accuracies in normal (88.0%) and evident variance (89.74%) scenarios. These results underscore the effectiveness of the attention matrix in improving the model's performance, particularly in distinguishing between different scenarios in PV generation data analysis.

4.2.4. Comparison with Other Time-Series Classification Methods

The classification accuracy of our model was compared with the existing methods proposed in [35–37] using daily time-series data for anomaly classification. Table 4 presents the comparison results. Our model, based on Devi-preprocessed GAF encoding and CNN with attention, achieves a peak accuracy of 95.83%, demonstrating its effectiveness in classifying daily output power time-series data from photovoltaic generation. When compared with existing methods such as the Collective of Transformation-Based Ensembles (COTE) approach [35] with an accuracy of 93.00%, and the approach in [36] using the Locally Weighted Scatterplot Smoothing (LOWESS) and Threshold algorithm based on interquartile range (IQR) for power quality anomaly classification with an accuracy of 94.51%, our model shows superior performance. Additionally, the multi-channel deep CNN (MCD-CNN) approach [37], focusing on the ensemble model based on the multi-channel deep CNN for anomaly classification in the net consumption of nodes, achieves an accuracy of 95.10%, which is comparable to our model's performance. These results highlight the effectiveness of our proposed DP-GAF+CNN approach for anomaly classification in photovoltaic power generation.

Table 4. Comparison of classification accuracy.

Approach	Time-Series Data	Algorithm	Accuracy
DP-GAF+CNN	Active power	DP-GAF and CNN with attention	95.83%
COTE [35]	Active power	Collective of Transformation-Based Ensembles	93.00%
[36]	Power quality	LOWESS & Threshold algorithm based on IQR	94.51%
MCD-CNN [37]	Net consumption of nodes	Ensemble model based on the multi-channel deep CNN	95.10%

Overall, these results demonstrate the superior performance of our proposed DP-GAF+CNN approach in classifying anomalies in PV power generation, showcasing its potential for enhancing anomaly detection and monitoring systems in the field of renewable energy.

4.3. Discussion

In this section, we discuss the observed results, highlighting any trends, patterns, or noteworthy findings. We also consider the implications of the classification and anomaly detection metrics, providing insights into the model's performance and potential areas for improvement.

The results of the proposed approach demonstrate the effectiveness in identifying anomalies in PV power plant output. The use of modified Gramian transition images and a convolutional neural network with attention has enabled the accurate classification of deviations in output power. This method has the potential to significantly improve the reliability and efficiency of anomaly detection systems in PV systems.

One key finding is the importance of preprocessing and feature extraction in anomaly detection. By transforming deviation time-series data into informative images, the model can capture subtle operational deviations that may not be apparent in raw data. This highlights the value of advanced data processing techniques in enhancing anomaly detection accuracy.

However, there are limitations to consider. The model's performance may be influenced by the quality and completeness of the input data. Additionally, the model's interpretability could be further enhanced. Future research could focus on developing more interpretable models or incorporating additional data sources to improve model performance. Further research and development in this area could lead to significant advancements in the field of renewable energy monitoring and management.

5. Conclusions

To address the adverse effects caused by overlooking the impact of external factors on the expected power output in detecting anomalies in PV time-series data, this paper highlights the key contributions of utilizing deviation GAF images and curve features for

reliable anomaly detection and classification in PV systems operations. The approach's effectiveness was demonstrated through its ability to identify anomalies such as normal, aging pollution, and arc faults with an overall classification accuracy of 95.83%.

A robust approach for daily anomaly detection in large-scale photovoltaic power stations was presented, leveraging daily output power and deviation time-series data with Gramian angular field encoding. The methodology combines expected and actual daily output power through deviation calculations, enhancing anomaly interpretability. Notably, GAF encoding with preprocessing for PV characteristics refines anomaly detection by transforming deviation time-series data into informative images with an adaptive normalization processing, facilitating enhanced feature extraction for subtle operational deviations. The integrated CNN architecture focuses on training attention matrices through the combination of deviation GAF images and curve features, proving to be highly effective in identifying anomalies such as normal, aging pollution, and arc faults. In summary, this study offers a comprehensive solution emphasizing the utilization of deviation GAF images and curve features for reliable anomaly detection and classification for PV systems operations.

For future work, enhancing the approach with advanced techniques like ensemble learning or deep reinforcement learning could improve its accuracy and robustness. Integrating real-time monitoring and adaptive learning mechanisms is also a promising direction. Additionally, expanding the dataset to include more diverse anomalies and data from different PV systems could further enhance the approach. Investigating the integration of other data types and applying the approach to other renewable energy systems could also extend its applicability.

Author Contributions: Conceptualization, Q.C. and Z.W.; methodology, Z.W. and Z.G.; software, Z.W. and Z.G.; validation, Z.W. and Z.G.; formal analysis, Z.W.; investigation, Z.W. and Z.G.; resources, Q.C. and Z.W.; data curation, Z.W.; writing—original draft preparation, Z.W.; writing—review and editing, Z.W., Q.C. and L.S.; supervision, Q.C.; project administration, J.G. and J.Z. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Data Availability Statement: Data are contained within the article.

Conflicts of Interest: Author Jiayong Zhong was employed by the State Grid Chongqing Electric Power Company. The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

References

1. Jafari, M.R.; Parniani, M.; Ravanji, M.H. Decentralized Control of OLTC and PV Inverters for Voltage Regulation in Radial Distribution Networks with High PV Penetration. *IEEE Trans. Power Deliv.* **2022**, *37*, 4827–4837. [\[CrossRef\]](#)
2. Bae, Y.; Vu, T.K.; Kim, R.Y. Implemental Control Strategy for Grid Stabilization of Grid-Connected PV System Based on German Grid Code in Symmetrical Low-to-Medium Voltage Network. *IEEE Trans. Energy Convers.* **2013**, *28*, 619–631. [\[CrossRef\]](#)
3. Pramanik, A.S.; Sepasi, S. Transient Behavior Analysis of Microgrids in Grid-Connected and Islanded Modes: A Comparative Study of LVRT and HVRT Capabilities. *Clean Technol.* **2023**, *5*, 1287–1303. [\[CrossRef\]](#)
4. Yuan, W.; Yuan, X.; Xu, L.; Zhang, C.; Ma, X. Harmonic Loss Analysis of Low-Voltage Distribution Network Integrated with Distributed Photovoltaic. *Sustainability* **2023**, *15*, 4334. [\[CrossRef\]](#)
5. Hu, X.; Ge, Y.; Liu, G.; Li, H.; Tan, J.; Zhao, F. Analysis on Time Series Coupling Characteristics of New Energy with Load in the Southeast Coast of China. In Proceedings of the 2021 IEEE Sustainable Power and Energy Conference (iSPEC), Nanjing, China, 23–25 December 2021; pp. 25–32. [\[CrossRef\]](#)
6. Wang, C.; Zhang, G.; Zhang, X. Experimental and Photothermal Performance Evaluation of Multi-Wall Carbon-Nanotube-Enhanced Microencapsulation Phase Change Slurry for Efficient Photothermal Conversion and Storage. *Energies* **2022**, *15*, 7627. [\[CrossRef\]](#)
7. Ren, G.; Liu, J.; Wan, J.; Wang, W.; Fang, F.; Hong, F.; Yu, D. Investigating the Complementarity Characteristics of Wind and Solar Power for Load Matching Based on the Typical Load Demand in China. *IEEE Trans. Sustain. Energy* **2022**, *13*, 778–790. [\[CrossRef\]](#)
8. Wu, Q.H.; Bose, A.; Singh, C.; Chow, J.H.; Mu, G.; Sun, Y.; Liu, Z.; Li, Z.; Liu, Y. Control and Stability of Large-scale Power System with Highly Distributed Renewable Energy Generation: Viewpoints from Six Aspects. *CSEE J. Power Energy Syst.* **2023**, *9*, 8–14. [\[CrossRef\]](#)

9. Palma-Behnke, R.; Valencia, F.; Vega-Herrera, J.; Núñez-Mata, O. Synthetic Time Series Generation Model for Analysis of Power System Operation and Expansion with High Renewable Energy Penetration. *J. Mod. Power Syst. Clean Energy* **2021**, *9*, 849–858. [\[CrossRef\]](#)
10. Ma, W.; Ma, M.; Wang, H.; Zhang, Z.; Zhang, R.; Wang, J. Shading Fault Detection Method for Household Photovoltaic Power Stations Based on Inherent Characteristics of Monthly String Current Data Mapping. *CSEE J. Power Energy Syst.* **2023**, *9*, 1370–1382. [\[CrossRef\]](#)
11. Cui, H.; Hu, A.; Sang, B.; Yang, B.; Li, G.; Tao, Y. Research on Method of Photovoltaic Power Station Assisting Power Grid Black Start Based on Energy Storage. In Proceedings of the 2018 2nd IEEE Conference on Energy Internet and Energy System Integration (EI2), Beijing, China, 20–22 October 2018; pp. 1–6. [\[CrossRef\]](#)
12. Barhmi, K.; Heynen, C.; Golroodbari, S.; van Sark, W. A Review of Solar Forecasting Techniques and the Role of Artificial Intelligence. *Solar* **2024**, *4*, 99–135. [\[CrossRef\]](#)
13. Leva, S.; Mussetta, M.; Ogliari, E. PV Module Fault Diagnosis Based on Microconverters and Day-Ahead Forecast. *IEEE Trans. Ind. Electron.* **2019**, *66*, 3928–3937. [\[CrossRef\]](#)
14. Lee, D.S.; Son, S.Y. PV Forecasting Model Development and Impact Assessment via Imputation of Missing PV Power Data. *IEEE Access* **2024**, *12*, 12843–12852. [\[CrossRef\]](#)
15. Mihai, S.; Yaqoob, M.; Hung, D.V.; Davis, W.; Towakel, P.; Raza, M.; Karamanoglu, M.; Barn, B.; Shetve, D.; Prasad, R.V.; et al. Digital Twins: A Survey on Enabling Technologies, Challenges, Trends and Future Prospects. *IEEE Commun. Surv. Tutor.* **2022**, *24*, 2255–2291. [\[CrossRef\]](#)
16. Viceconti, M.; De Vos, M.; Mellone, S.; Geris, L. Position Paper From the Digital Twins in Healthcare to the Virtual Human Twin: A Moon-Shot Project for Digital Health Research. *IEEE J. Biomed. Health Inform.* **2024**, *28*, 491–501. [\[CrossRef\]](#)
17. Zhang, B.; Zhang, M.; Dong, T.; Lu, M.; Li, H. Design of Digital Twin System for DC Contactor Condition Monitoring. *IEEE Trans. Ind. Appl.* **2023**, *59*, 3904–3909. [\[CrossRef\]](#)
18. Huang, Y.; Chen, S.; Tan, X.; Hu, M.; Zhang, C. Power Prediction Method of Distributed Photovoltaic Digital Twin System Based on GA-BP. In Proceedings of the 2022 4th International Conference on Electrical Engineering and Control Technologies (CEEET), Shanghai, China, 16–18 December 2022; pp. 241–245. [\[CrossRef\]](#)
19. Yuan, J.; Tian, Z.; Ma, J.; Man, K.L.; Li, B. A Digital Twin Approach for Modeling Electrical Characteristics of Bifacial Solar Panels. In Proceedings of the 2022 International Conference on Industrial IoT, Big Data and Supply Chain (IIoTBDSC), Beijing, China, 23–25 September 2022; pp. 317–321. [\[CrossRef\]](#)
20. Yonce, J.; Walters, M.; Venayagamoorthy, G.K. Short-Term Prediction of Solar Photovoltaic Power Generation Using a Digital Twin. In Proceedings of the 2023 North American Power Symposium (NAPS), Asheville, NC, USA, 15–17 October 2023; pp. 1–6. [\[CrossRef\]](#)
21. Khan, K.; Rashid, S.; Mansoor, M.; Khan, A.; Raza, H.; Zafar, M.H.; Akhtar, N. Data-driven green energy extraction: Machine learning-based MPPT control with efficient fault detection method for the hybrid PV-TEG system. *Energy Rep.* **2023**, *9*, 3604–3623. [\[CrossRef\]](#)
22. Yeganejou, M.; Ryan, T.; Reshadi, M.; Dick, S.; Lipsett, M.G. Condition Monitoring of Underground Power Cables Via Power-Line Modems and Anomaly Detection. *IEEE Trans. Power Deliv.* **2024**, *39*, 3–12. [\[CrossRef\]](#)
23. Mestav, K.R.; Wang, X.; Tong, L. A Deep Learning Approach to Anomaly Sequence Detection for High-Resolution Monitoring of Power Systems. *IEEE Trans. Power Syst.* **2023**, *38*, 4–13. [\[CrossRef\]](#)
24. Massaoudi, M.S.; Abu-Rub, H.; Ghrayeb, A. Navigating the Landscape of Deep Reinforcement Learning for Power System Stability Control: A Review. *IEEE Access* **2023**, *11*, 134298–134317. [\[CrossRef\]](#)
25. Zhao, T.; Wang, J.; Lu, X.; Du, Y. Neural Lyapunov Control for Power System Transient Stability: A Deep Learning-Based Approach. *IEEE Trans. Power Syst.* **2022**, *37*, 955–966. [\[CrossRef\]](#)
26. Hossain, R.R.; Huang, Q.; Huang, R. Graph Convolutional Network-Based Topology Embedded Deep Reinforcement Learning for Voltage Stability Control. *IEEE Trans. Power Syst.* **2021**, *36*, 4848–4851. [\[CrossRef\]](#)
27. Shi, Z.; Yao, W.; Tang, Y.; Ai, X.; Wen, J.; Cheng, S. Bidirectional Active Transfer Learning for Adaptive Power System Stability Assessment and Dominant Instability Mode Identification. *IEEE Trans. Power Syst.* **2023**, *38*, 5128–5142. [\[CrossRef\]](#)
28. Guanoluisa, R.; Arcos-Aviles, D.; Flores-Calero, M.; Martinez, W.; Guinjoan, F. Photovoltaic Power Forecast Using Deep Learning Techniques with Hyperparameters Based on Bayesian Optimization: A Case Study in the Galapagos Islands. *Sustainability* **2023**, *15*, 12151. [\[CrossRef\]](#)
29. Amiri, A.; Samet, H.; Ghanbari, T. Recurrence Plots Based Method for Detecting Series Arc Faults in Photovoltaic Systems. *IEEE Trans. Ind. Electron.* **2022**, *69*, 6308–6315. [\[CrossRef\]](#)
30. Nath, A.G.; Udmale, S.S.; Raghuwanshi, D.; Singh, S.K. Improved Structural Rotor Fault Diagnosis Using Multi-Sensor Fuzzy Recurrence Plots and Classifier Fusion. *IEEE Sens. J.* **2021**, *21*, 21705–21717. [\[CrossRef\]](#)
31. Li, Y.L.; Jiang, J.R. Anomaly Detection for Non-Stationary and Non-Periodic Univariate Time Series. In Proceedings of the 2020 IEEE Eurasia Conference on IOT, Communication and Engineering (ECICE), Yunlin, Taiwan, 23–25 October 2020; pp. 177–179. [\[CrossRef\]](#)
32. Zhang, R.; Yao, W.; Shi, Z.; Ai, X.; Tang, Y.; Wen, J. Encoding Time Series as Images: A Robust and Transferable Framework for Power System DIM Identification Combining Rules and VGGNet. *IEEE Trans. Power Syst.* **2023**, *38*, 5781–5793. [\[CrossRef\]](#)

33. Guo, L.; Zhang, L.; Huang, Q. Bearing Fault Diagnosis Based on Multi-Channel GAF-MTF and Res2Net. In Proceedings of the 2023 35th Chinese Control and Decision Conference (CCDC), Yichang, China, 20–22 May 2023; pp. 5321–5327. [[CrossRef](#)]
34. Bagnall, A.; Lines, J.; Bostrom, A.; Large, J.; Keogh, E. The great time series classification bake off: A review and experimental evaluation of recent algorithmic advances. *Data Min. Knowl. Discov.* **2017**, *31*, 606–660. [[CrossRef](#)] [[PubMed](#)]
35. Bagnall, A.; Lines, J.; Hills, J.; Bostrom, A. Time-series classification with COTE: The collective of transformation-based ensembles. In Proceedings of the 2016 IEEE 32nd International Conference on Data Engineering (ICDE), Helsinki, Finland, 16–20 May 2016; pp. 1548–1549. [[CrossRef](#)]
36. Zyabkina, O.; Domagk, M.; Meyer, J.; Schegner, P. Classification and identification of anomalies in time series of power quality measurements. In Proceedings of the 2016 IEEE PES Innovative Smart Grid Technologies Conference Europe (ISGT-Europe), Ljubljana, Slovenia, 9–12 October 2016; pp. 1–6. [[CrossRef](#)]
37. Petrushev, A.; Bauer, R.; Rigo-Mariani, R.; Debusschere, V.; Reignier, P.; Hadjsaid, N. Comparing Time Series Classification And Forecasting To Automatically Detect Distributed Generation. In Proceedings of the 2021 IEEE Madrid PowerTech, Madrid, Spain, 28 June–2 July 2021; pp. 1–6. [[CrossRef](#)]
38. Ma, W.; Ma, M.; Zhang, Z.; Ma, J.; Zhang, R.; Wang, J. Anomaly Detection of Mountain Photovoltaic Power Plant Based on Spectral Clustering. *IEEE J. Photovolt.* **2023**, *13*, 621–631. [[CrossRef](#)]
39. Solar-PV Power Generation Data. Available online: <https://www.elia.be/en/grid-data/power-generation/solar-pv-power-generation-data> (accessed on 15 February 2024).

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