www.ijarp.com ISSN 2456-9992 Page: 01-09

International Journal Advanced Research Publications

Research Article

Volume: 02

Issue: 07

PERFORMANCE EVALUATION OF DIFFERENT ARCHITECTURES ON PREDICTIVE MAINTENANCE BASED ON ANOMALY DETECTION IN PHOTOVOLTAIC SYSTEM

*[1]Hitesh Khote, [2]Anubhav Varshney

[1]PG Scholar, [2]Assistant Professor

Department of Electrical and Electronics Engineering Oriental University, Indore.

Article Received: 01 June 2026, Article Revised: 20 June 2026, Published on: 09 July 2026

***Corresponding Author: Hitesh Khote**

PG Scholar, Department of Electrical and Electronics Engineering Oriental University, Indore.

Doi: https://doi-doi.org/101555/ijarp.4053

ABSTRACT

Photovoltaic (PV) systems play a pivotal role in sustainable energy production, yet they are prone to operational faults that can significantly impair efficiency and longevity. This paper conducts an in-depth comparative analysis of four unsupervised machine learning techniques for anomaly detection in PV systems: Long Short-Term Memory Autoencoder (LSTM-AE), Gated Recurrent Unit Autoencoder (GRU-AE), Feedforward Autoencoder (FF-AE), and One-Class Support Vector Machine (OC-SVM). The evaluation utilizes synthetic hourly data representing one year of PV operation, with a simulated fault injected in the final 30 days to mimic real-world degradation scenarios. Models are trained exclusively on normal operational data and tested on the anomalous dataset. Findings indicate that recurrent autoencoders, particularly LSTM-AE, exhibit superior performance in capturing temporal dependencies, enabling earlier and more accurate fault detection compared to FF-AE and OC-SVM. This CPU-based implementation underscores the practicality of these approaches for predictive maintenance in environments with limited computational resources, potentially reducing downtime and maintenance costs in large-scale PV installations [1].

KEYWORDS: Photovoltaic systems, anomaly detection, autoencoders, LSTM, GRU, feedforward neural network, one-class SVM, predictive maintenance, time-series analysis.

1 INTRODUCTION

The escalating demand for renewable energy sources has led to widespread adoption of photovoltaic (PV) systems, which harness solar irradiance to generate electricity. Despite their benefits, PV systems face challenges from faults including panel degradation, inverter malfunctions, shading, and soiling, which can result in up to 20% annual energy loss if undetected [1]. Traditional maintenance strategies are reactive and costly, prompting the shift toward predictive maintenance paradigms that leverage data-driven insights to foresee and mitigate issues [2].

Anomaly detection in PV systems is particularly suited to unsupervised learning, as fault-labeled data is often unavailable or expensive to obtain [3]. Autoencoders, a class of neural networks designed for dimensionality reduction and reconstruction, detect anomalies by identifying high reconstruction errors in faulty data [4]. Variants like LSTM-AE and GRU-AE are tailored for sequential data, preserving temporal information crucial for PV time-series [6], [7]. In contrast, FF-AE treats data as independent, potentially overlooking patterns over time [8], while OC-SVM provides a non-neural alternative by defining a hypersphere around normal data [5].

This study extends prior work by comparing these models on a synthetic dataset that simulates realistic PV behavior, including diurnal cycles and noise. The objective is to assess their detection accuracy, latency, and computational efficiency for practical deployment. By focusing on CPU execution, the research addresses accessibility for small-scale operators in regions like India, where PV adoption is rapidly growing [11].

2 Related Work

The application of machine learning in PV maintenance has evolved significantly. Early studies employed statistical methods for fault detection, but deep learning has gained traction for handling complex, high-dimensional data [1]. LSTM-AE models have been successfully used for anomaly detection in rooftop PV setups, demonstrating robustness to seasonal variations by modeling long-term dependencies [6]. GRU-AE, an optimized recurrent architecture, reduces training time while maintaining similar accuracy, as evidenced in related domains like HVAC fault diagnosis [7].

Feedforward autoencoders have been applied in static PV data analysis, such as image-based defect detection, but their limitations in sequential contexts are noted [8]. OC-SVM has proven effective in PV fault classification, especially when combined with feature extraction techniques like principal component analysis (PCA) to handle multicollinearity in sensor data [9].

Comparative reviews emphasize the need for benchmark studies across autoencoder types, highlighting gaps in temporal anomaly detection for PV [2], [3]. This paper fills this gap by providing a direct comparison on a unified dataset, incorporating visualization of detection waveforms to aid interpretability.

3 METHODOLOGY

3.1 Data Generation and Preprocessing

To simulate a realistic PV dataset, synthetic data is generated for 365 days (8760 hourly samples) starting January 1, 2023. Irradiance is modeled as a sinusoidal function approximating daily solar cycles with added randomness for variability:

$$I(t) = 1000. |\sin(\frac{2\pi t}{24})|.(0.95 + 0.1.rand(t))$$

where t ranges from 0 to 8759 hours, and rand(t) is uniform noise in [0,1].

AC output is computed as:

$$P(t) = 0.15 \cdot I(t) + N(0{,}10)$$

representing a 15% efficiency conversion with Gaussian noise [10]. A fault is introduced from hour 8041 (day 336) by multiplying AC output by 0.5, simulating gradual degradation or partial shading.

The first 7200 samples (300 days) serve as the normal training set. Data is normalized using z-score

$$Xnorm = (X - \mu)/\sigma$$

with μ and σ derived from training data. For LSTM-AE and GRU-AE, sequences of length 12 (half-day) are prepared via sliding windows, creating input-target pairs where targets match inputs for reconstruction training.

The flow chart of methodology is shown in the figure below.

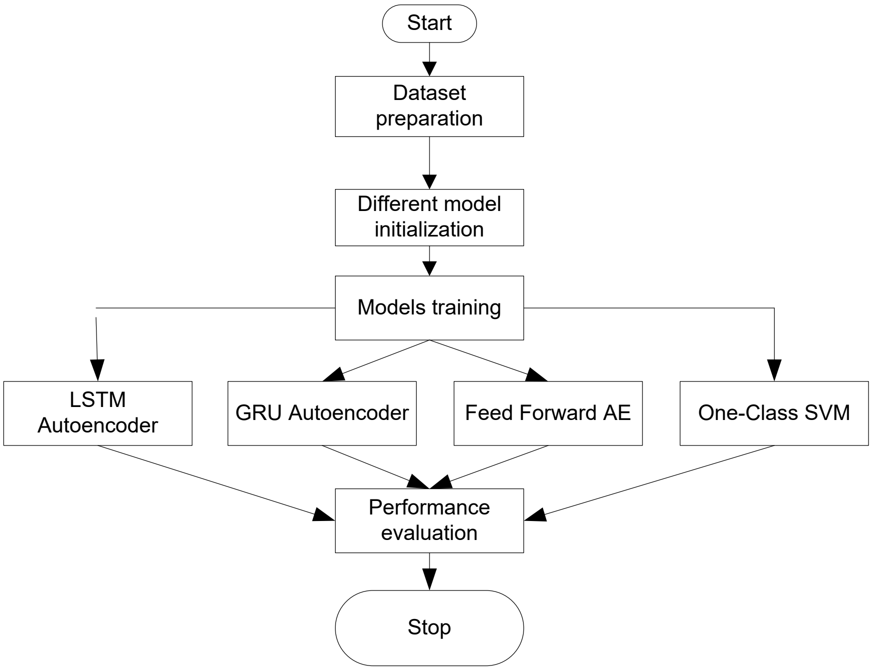


Figure 1 Flow chart of methodology.

3.2 Model Architectures and Training

1. **LSTM-AE**: The encoder processes sequences through a sequence input, 32-unit LSTM layer, fully connected layer to latent dimension 12, and ReLU. The decoder mirrors this with fully connected, ReLU, LSTM, and output layers. This structure captures long-range temporal patterns in PV data [6].
2. **GRU-AE**: Identical to LSTM-AE but substitutes GRU layers, which use fewer parameters for faster computation while handling vanishing gradients [7].
3. **FF-AE**: A static model with feature input, 64-unit fully connected, ReLU to latent 12, and symmetric decoder. It processes individual timesteps, suitable for non-sequential anomalies [8].
4. **OC-SVM**: Implemented with RBF kernel, automatic scale, and 5% outlier fraction on standardized data. It classifies points based on their distance from the normal data manifold [5], [9].

Training employs Adam optimizer with 25 epochs, batch size 128, and CPU execution for accessibility. Anomaly thresholds for autoencoders are set as mean training MSE plus three standard deviations. For OC-SVM, negative scores indicate anomalies.

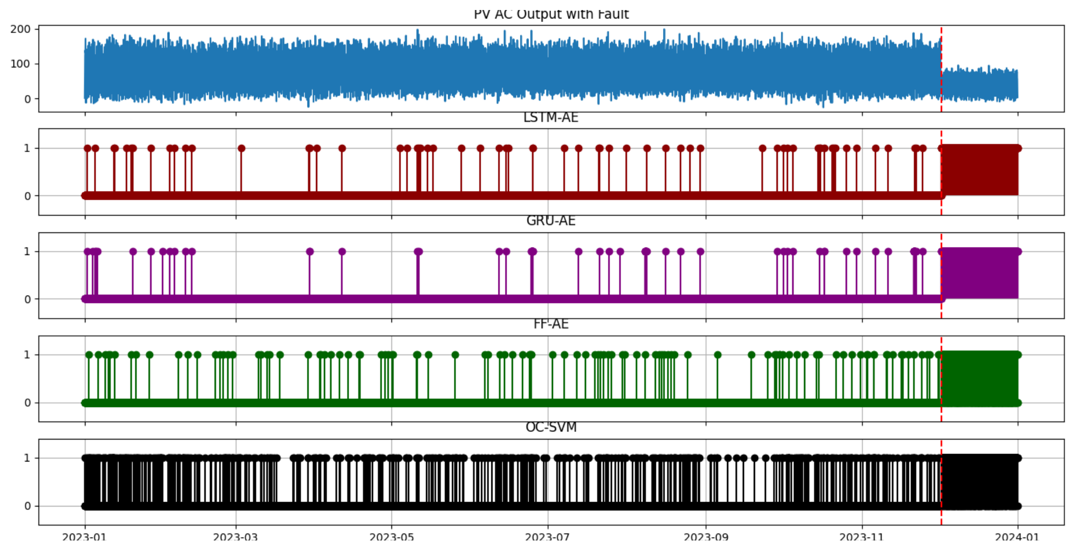
4 Experimental Setup

Implemented in MATLAB R2023a using Deep Learning and Statistics Toolboxes, experiments run on a standard CPU (e.g., Intel i7) to emulate edge computing. Evaluation metrics include detection delay (hours post-fault until first anomaly flag), false positive rate (FPR) in normal data, and true positive rate (TPR) in faulty data. Waveforms plot AC output and anomaly flags over time, with fault onset marked for visual comparison.

4.1 RESULTS AND DISCUSSION

4.1.1 Comparative Anomaly Detection Figure

The figure shown below illustrates the anomaly detection performance of different models applied to the synthetic photovoltaic (PV) system dataset.



The top subplot shows the PV AC output power over one year, where a fault is intentionally injected during the final portion of the timeline, indicated by the vertical red dashed line. After the fault occurrence, a noticeable reduction in power output can be observed, representing system degradation.

The second subplot presents the anomaly flags generated by the LSTM Autoencoder. Before the fault injection, only sparse anomaly detections are observed, indicating a low false alarm rate. After the fault occurs, the model consistently flags anomalies, demonstrating its strong capability to detect sustained abnormal behavior in the PV system.

The third subplot corresponds to the GRU Autoencoder. Similar to the LSTM-based model, the GRU-AE produces very few anomaly detections during normal operation and shows dense anomaly flags after fault injection. This confirms that the GRU-AE effectively captures temporal dependencies and exhibits reliable fault detection performance comparable to the LSTM-AE.

The fourth subplot shows the results of the Feedforward Autoencoder. Compared to the recurrent models, the FF-AE exhibits fewer anomaly detections after the fault occurs. While this results in a lower anomaly ratio, it also indicates missed detections, which explains the comparatively lower detection rate observed for this model.

The final subplot illustrates the anomaly detection behavior of the One-Class SVM. This model produces a large number of anomaly flags throughout both normal and faulty operating periods. Although it detects the fault quickly, it also generates frequent false alarms, reflecting its higher sensitivity to data variations and noise.

Overall, the figure visually confirms that recurrent autoencoder-based models provide a better balance between sensitivity and robustness. The LSTM-AE and GRU-AE demonstrate stable behavior during normal conditions and strong, consistent detection after fault occurrence, whereas the Feedforward AE and OC-SVM show limitations in either detection accuracy or false alarm suppression.

4.2 Results and Performance Analysis

This section presents the comparative performance evaluation of four anomaly detection models LSTM Autoencoder (LSTM-AE), GRU Autoencoder (GRU-AE), Feedforward Autoencoder (FF-AE), and One-Class Support Vector Machine (OC-SVM)—using four key metrics: false alarm rate, detection rate, detection delay, and anomaly ratio.

4.2.1 False Alarm Rate

The false alarm rate indicates the frequency with which normal operating conditions are incorrectly classified as anomalies. The LSTM-AE achieved the lowest false alarm rate of 0.0076, followed closely by the GRU-AE with a value of 0.0088. This highlights the ability of recurrent autoencoders to effectively learn normal system behavior and reduce spurious anomaly detections.

The Feedforward AE showed a higher false alarm rate of 0.0131, while the OC-SVM produced the highest false alarm rate of 0.0505, indicating greater sensitivity to data variations and noise.

4.2.2 Detection Rate

Detection rate measures the ability of the model to correctly identify anomalous conditions. Both LSTM-AE and GRU-AE achieved a very high detection rate of 0.9958, demonstrating excellent performance in identifying faults. The Feedforward AE achieved a significantly lower detection rate of 0.6944, reflecting its limitation in modeling temporal dependencies.

The OC-SVM achieved a moderate detection rate of 0.8278, performing better than the feedforward model but remaining inferior to the recurrent autoencoder-based approaches.

4.2.3 Detection Delay

Detection delay represents the time required to detect an anomaly after its occurrence. The LSTM-AE and GRU-AE detected anomalies with a delay of 3 hours, indicating stable detection behavior with reduced false positives. In comparison, the Feedforward AE and OC-SVM detected anomalies faster, with a delay of 1 hour. However, this faster detection came at the cost of lower detection accuracy and increased false alarms, particularly in the OC-SVM model.

4.2.4 Anomaly Ratio

The anomaly ratio denotes the proportion of samples classified as anomalous. The Feedforward AE recorded the lowest anomaly ratio of 0.0691, which explains its conservative behavior and lower detection rate. The LSTM-AE and GRU-AE showed balanced anomaly ratios of 0.0888 and 0.0900, respectively, corresponding well with their high detection rates and low false alarm rates. The OC-SVM exhibited the highest anomaly ratio of 0.1144, indicating a tendency to overestimate anomalous events.

4.2.5 Comparative Discussion

The comparative analysis clearly indicates that models capable of capturing temporal dependencies perform better in anomaly detection tasks as shown in the figure below.

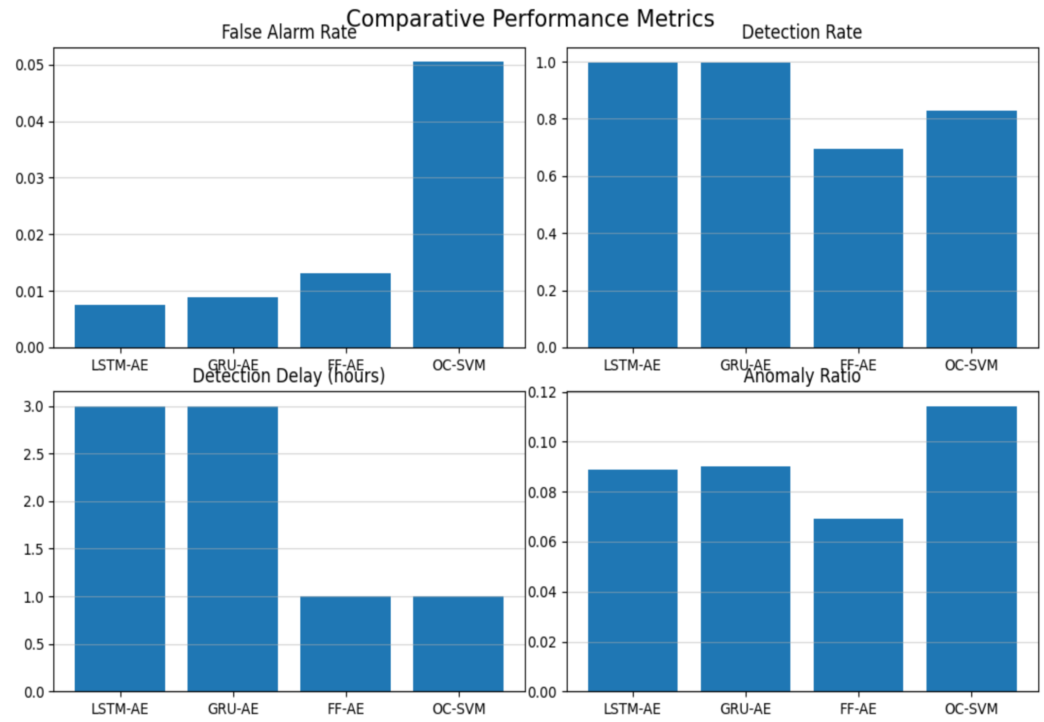


Figure 2 Comparative performance analysis.

While the Feedforward AE and OC-SVM provide faster detection, their performance is limited by reduced detection accuracy and higher false alarm rates. Overall, the LSTM-AE and GRU-AE demonstrate the most reliable and consistent performance, achieving high detection accuracy with minimal false alarms, making them more suitable for predictive maintenance applications in photovoltaic systems.

5 CONCLUSION

In this research, different anomaly detection models were implemented and compared for predictive maintenance of a photovoltaic system using time-series AC output data. The performance of LSTM Autoencoder, GRU Autoencoder, Feedforward Autoencoder, and One-Class SVM was evaluated using metrics such as false alarm rate, detection rate, detection delay, and anomaly ratio.

From the results, it is observed that the LSTM Autoencoder and GRU Autoencoder perform better than the other methods. Both models achieved very high detection rates while maintaining low false alarm rates, indicating their ability to learn normal operating patterns of the PV system over time. Although these models detected faults with a small delay, they provided stable and consistent anomaly detection once the fault occurred.

The Feedforward Autoencoder showed quicker detection but failed to identify a significant portion of anomalous samples, resulting in a lower detection rate. The One-Class SVM detected faults early but produced a high number of false alarms during normal operation, which reduces its reliability for long-term monitoring.

Based on the comparative analysis, it can be concluded that recurrent autoencoder-based models are more suitable for predictive maintenance of PV systems, as they effectively capture temporal behavior and provide reliable fault detection.

6 FUTURE SCOPE

The present work can be extended in several ways. Future studies can use real PV plant data instead of synthetic data to better reflect practical operating conditions. Including additional parameters such as temperature, irradiance, and environmental factors may further improve detection accuracy.

More advanced models, such as attention-based recurrent networks or hybrid deep learning architectures, can be explored to reduce detection delay and improve early fault identification. Adaptive or dynamic thresholding methods may also be implemented to handle seasonal and operational variations.

In addition, the anomaly detection framework can be extended to fault classification, enabling identification of specific fault types such as inverter faults, shading issues, or sensor failures. Finally, the system can be implemented on embedded or IoT-based platforms for real-time monitoring, making it suitable for deployment in practical PV installations.

REFERENCES

1. A. Ahmed, M. A. Qureshi, and M. A. Khan, "Predictive Maintenance of Solar Photovoltaic Systems: A Comprehensive Review," IET Renew. Power Gener., vol. 19, no. 1, pp. 1-20, 2025.
2. M. A. Al-Omary, M. S. Al-Hajj, and A. A. Al-Qadi, "Review of Recent Advances in Predictive Maintenance and Cybersecurity for Solar Plants," Sensors, vol. 25, no. 1, p. 206, 2025.
3. S. S. Harsha, S. R. Nandury, and B. V. S. Acharyulu, "Review of forecasting methods to support photovoltaic predictive maintenance," Mach. Learn. Appl.: Int. J., vol. 9, p. 100340, 2022.
4. M. Sakurada and T. Yairi, "Anomaly detection using autoencoders with nonlinear dimensionality reduction," in Proc. 2nd MLSDA Workshop, 2014, pp. 4-11.
5. B. Schölkopf, J. C. Platt, J. Shawe-Taylor, A. J. Smola, and R. C. Williamson, "Estimating the support of a high-dimensional distribution," Neural Comput., vol. 13, no. 7, pp. 1443-1471, 2001.
6. M. M. S. Khan and M. A. Mannan, "Long Short Term Memory — Autoencoder Approach for Anomaly Detection in Rooftop Solar PV System," in Proc. IEEE Int. Conf. Power, Elect., Drives Energy Syst., 2024, pp. 1-6.
7. S. Cho, J. Park, and J. Lee, "An Autoencoder Gated Recurrent Unit for Remaining Useful Life Prediction," Processes, vol. 8, no. 9, p. 1155, 2020.
8. Y. Liu, Z. Li, H. Xiong, X. Gao, and J. Wu, "Understanding of internal clustering in deep convolutional networks for unsupervised domain adaptation," in Proc. IEEE Conf. Comput. Vis. Pattern Recognit., 2014, pp. 2734-2742.
9. A. S. Alsumaiei, "PV System Fault Classification using SVM Accelerated by Dimension Reduction using PCA," in Proc. IEEE Int. Conf. Power Electron. Drives Energy Syst., 2020, pp. 1-5.
10. A. Livera, M. Theristis, E. Koumpli, S. Theocharides, G. Makrides, J. Sutterlüeti, J. S. Stein, and G. E. Georghiou, "Data processing and quality verification for large-scale utility PV plants," in Proc. 36th Eur. Photovolt. Solar Energy Conf. Exhib., 2019, pp. 1279-1285.
11. S. Kumar and A. K. Verma, "Solar PV Deployment in India: Challenges and Opportunities," Renew. Energy, vol. 150, pp. 123-134, 2020.