High-Fidelity Photovoltaic Power Forecasting Using a Skip-Fusion DNN with GELU Activation and AdamW Optimization*

Ievgen Zaitsev[1,*,2,†], Hasan Uzel[3,†], Feyyaz Alpsalaz[3,†], Yıldırım Özüpak[4,†]and Emrah Aslan[5,†]

[1] *Institute of Electrodynamics of NAS of Ukraine, Beresteyskyi 56, 03057 Kyiv, Ukraine*

[2] *Center for Information-Analytical and Technical Support of Nuclear Power Facilities Monitoring, National Academy of Sciences of Ukraine, Akademika Palladina 34a, 03142 Kyiv, Ukraine*

[3] *Akdağmadeni Vocational School, Yozgat Bozok University, Gültepe Mahallesi, Tepe 4, 66540 Yozgat, Turkey*

[4] *Silvan Vocational School, Dicle University, Gazi, 21640 Diyarbakır, Turkey*

[5] *Mardin Artuklu University, Diyarbakır Road, 5, 47200 Mardin, Turkey*

Abstract

Accurate forecasting of photovoltaic (PV) power generation is essential for optimizing the operation and stability of renewable-dominated smart grids. However, the stochastic nature of solar irradiance, temperature-dependent derating, and nonlinear PV conversion dynamics pose significant challenges to model reliability and generalization. This study presents a novel deep neural network architecture, DNN-v4, designed for short-term PV power forecasting using high-resolution SCADA telemetry. The proposed model integrates GELU activation, Layer Normalization, and a skip-fusion mechanism that merges multi-scale dense representations to enhance feature propagation and gradient stability. Optimization is conducted through the AdamW algorithm combined with a Cosine Decay Restarts learning-rate schedule and Huber loss to improve robustness against outliers. The model was trained on a real-world dataset comprising 118,865 SCADA records with environmental and electrical features such as irradiance, temperature, wind speed, and DC/AC currents. Experimental results demonstrate superior performance with RMSE = 1.741 kW, MAE = 0.992 kW, MAPE = 1.12 %, sMAPE = 1.14 % and R^2 = 0.9996, significantly outperforming conventional and hybrid baselines. Beyond predictive accuracy, DNN-v4 preserves physical consistency between irradiance, temperature, and current, offering a computationally efficient and interpretable framework for real-time PV forecasting in smart-grid operations.

Keywords

Photovoltaic power forecasting; SCADA telemetry; deep neural network (DNN); skip-fusion architecture; GELU activation; AdamW optimization

1. Introduction

Accurate forecasting of photovoltaic (PV) power generation has become a cornerstone of modern renewable energy management, particularly within smart grid environments that increasingly rely on intermittent sources. The rapid growth of solar energy integration has introduced new challenges for power system operators, who must maintain real-time energy balance, stability, and scheduling efficiency despite the inherent variability of solar resources [1]. In this context, short-term and day-ahead forecasting of PV generation are essential for economic dispatch, demand-side management, and energy market participation. Nevertheless, achieving high forecasting accuracy remains difficult due to the stochastic and nonlinear nature of solar power generation processes [2].

The main difficulty in PV power prediction arises from the complex dependencies between environmental and electrical variables. Solar irradiance, ambient temperature, wind speed, and humidity interact in nonlinear ways that directly influence the current and voltage output of PV

* *Applied Information Systems and Technologies in the Digital Society (AISTDS-2025), October 01, 2025, Kyiv, Ukraine*

1,* Corresponding author.

† These authors contributed equally.

zaitsev@i.ua (I. Zaitsev); hasan.uzel@bozok.edu.tr (Hasan Uzel); feyyaz.alpsalaz@bozok.edu.tr (Feyyaz Alpsalaz); yildirim.ozupak@dicle.edu.tr (Yıldırım Özüpak); emrahaslan@artuklu.edu.tr (Emrah Aslan).

0000-0003-3303-471X ((I. Zaitsev); 0000-0002-8238-2588 (Hasan Uzel); 0000-0002-7695-6426 (Feyyaz Alpsalaz); 0000-0001-8461-8702 (Yıldırım Özüpak) ; 0000-0002-0181-3658 (Emrah Aslan)

modules. These relationships are further complicated by panel aging, inverter efficiency, and sensor noise in Supervisory Control and Data Acquisition (SCADA) systems [3]. Consequently, developing models that can effectively represent these multivariate dependencies is essential for reliable forecasting across different temporal scales and environmental conditions.

A large body of research has explored data-driven approaches for PV power forecasting, with machine learning and deep learning methods dominating recent literature [4]. Traditional machine learning algorithms such as Support Vector Regression (SVR), Random Forests (RF), and Gradient Boosting (XGBoost) have demonstrated acceptable performance for certain datasets and horizons. However, these methods often struggle to model the high-dimensional and nonlinear feature interactions present in real-world PV systems [5]. Their performance typically depends on careful feature engineering, which can limit scalability and adaptability to new environments.

To overcome these limitations, deep learning (DL) architectures have gained increasing attention due to their ability to automatically learn hierarchical feature representations. Models such as Convolutional Neural Networks (CNNs) and Long Short-Term Memory (LSTM) networks have shown superior capabilities in capturing both spatial and temporal dependencies in time-series data [6]. CNNs are particularly effective at extracting local patterns from multivariate SCADA signals, while LSTMs are adept at modeling sequential dependencies and long-term temporal correlations. However, these architectures are not without drawbacks. Deep models often require substantial computational resources, complex hyperparameter tuning, and large quantities of labeled data. When applied to relatively small or noisy SCADA datasets, they may overfit, leading to reduced generalization capability and unstable predictions [7].

Another significant limitation of existing approaches is their lack of physical interpretability. Most purely data-driven methods neglect the underlying physical relationships between irradiance, temperature, and current, which govern the photovoltaic conversion process [8]. As a result, model outputs can violate known physical constraints or behave inconsistently under varying environmental conditions. This reduces the trustworthiness of predictions, especially in operational settings that demand both accuracy and explainability. Bridging this gap between data-driven learning and physics-based modeling remains a major challenge in the field.

In response to these issues, the present study introduces a novel deep neural network architecture, termed DNN-v4, designed specifically for accurate and physically consistent PV power forecasting using SCADA telemetry [9]. The proposed framework is built upon a compact yet expressive design that integrates several modern components from state-of-the-art deep learning research. It employs the Gaussian Error Linear Unit (GELU) activation function, which offers smoother nonlinear transformations and improved gradient behavior compared to conventional ReLU-based activations. To further enhance training stability, Layer Normalization is applied across layers, reducing internal covariate shift and ensuring consistent convergence during optimization.

A key innovation of DNN-v4 lies in its skip-fusion mechanism, which merges multi-scale dense feature representations from intermediate layers. This design promotes richer information flow through the network while mitigating vanishing gradient problems. By combining shallow and deep features, the model effectively captures both short-term fluctuations and long-term dependencies in PV power output. This hybrid representation improves generalization and robustness across diverse meteorological conditions.

The optimization process also integrates several advanced techniques. Model training is conducted using the AdamW optimizer, which decouples weight decay from gradient updates, leading to more controlled regularization and improved generalization [10]. Furthermore, a Cosine Decay Restarts (CDR) learning rate scheduler is adopted to dynamically modulate the learning rate, preventing premature convergence and enabling the model to explore more optimal regions of the loss landscape. This adaptive scheduling strategy encourages smoother convergence and better fine-tuning across training epochs.

For the loss function, the Huber loss is employed instead of the standard mean squared error (MSE). The Huber loss provides a balance between L1 and L2 penalties, making it particularly robust to noisy or outlier data samples frequently encountered in SCADA measurements [11]. Together, these optimization choices enable the DNN-v4 model to achieve high predictive accuracy without sacrificing computational efficiency or stability.

Comprehensive experiments demonstrate that the proposed DNN-v4 model outperforms existing benchmarks in both accuracy and robustness, while maintaining a lightweight architecture suitable for real-time forecasting applications [5]. Compared to deeper or more complex models, DNN-v4 requires fewer parameters and exhibits faster convergence, making it practical for deployment in embedded or edge computing environments. Beyond raw prediction accuracy, the model exhibits strong physical consistency, aligning with known PV power generation principles. This hybrid modeling philosophy—combining physical insights with data-driven learning—contributes to improved interpretability and reliability under unseen weather scenarios [11].

The study also emphasizes the importance of data preprocessing and exploratory analysis prior to model training. Raw SCADA data often contain missing values, noise, and inconsistencies due to sensor faults or communication errors. Therefore, careful data cleaning, normalization, and feature selection are critical for ensuring stable training performance. Additionally, exploratory data analysis (EDA) enables the identification of statistical patterns and correlations between environmental and electrical parameters, which can guide both model design and evaluation.

The remainder of this paper is organized as follows. Section 2 details the dataset description, preprocessing workflow, and exploratory data analysis conducted on the SCADA telemetry. Section 3 introduces the architecture of the proposed DNN-v4 model, highlighting its key layers, activation functions, optimization strategy, and learning rate scheduling mechanism. Section 4 outlines the experimental setup, including training configurations, baseline comparisons, and evaluation metrics. Section 5 presents and discusses the experimental results, focusing on predictive performance, learning dynamics, and residual analyses to assess model robustness. Finally, Section 6 concludes the paper by summarizing the main findings and suggesting potential avenues for future research, including model interpretability, transfer learning, and integration with hybrid physics-informed frameworks.

The proposed DNN-v4 framework represents a significant step toward compact, stable, and physically consistent PV power forecasting. By combining the strengths of modern deep learning techniques with physics-aware design principles, it bridges the gap between black-box prediction models and interpretable renewable energy analytics. The results highlight that accurate, efficient, and explainable deep learning models can substantially enhance the reliability of smart grid operations in an era increasingly dominated by renewable energy sources.

2. Dataset and Exploratory Data Analysis (EDA)

The proposed model was trained and evaluated using a real-world SCADA dataset collected from a grid-connected photovoltaic (PV) power plant [13]. The dataset contains 118,865 records with nine continuous variables, including module temperature, ambient temperature, wind speed, plane-of-array irradiance (W/m²), DC current, three-phase AC currents (Ir, Iy, Ib), and the target variable AC power (kW) [4]. The dataset was obtained from the publicly available *SolarGeneration* dataset shared by Arun Kanagolkar on the Kaggle platform [13].

Descriptive statistics of all input features are summarized in Table 1. The results indicate that irradiance and string current exhibit the highest magnitudes within the numerical range (with mean values of approximately 428 W/m² and 356 A, respectively), reflecting their dominant influence on photovoltaic energy generation. Meanwhile, the module temperature shows an average of 37 °C, which is roughly 14 °C above the mean ambient temperature. This temperature differential aligns well with the expected thermal behavior of PV modules under standard outdoor operating conditions, where solar absorption and electrical loading typically elevate the module temperature beyond the surrounding environment [1].

Figure 1 illustrates the mean values of all input variables. Irradiance and DC current exhibit the highest magnitudes, confirming their dominant physical influence on AC power generation, whereas temperature and wind speed remain secondary contributors [9].

Table 1
Descriptive statistics of the SCADA dataset

Feature	Count	Mean	Std	Min	Max
MODULE_TEMP	118 865	37.14	12.02	8.86	72.45
Amb_Temp	118 865	22.96	3.92	10.41	34.96
WIND_Speed	118 865	224.43	230.24	0.24	597.44
IRR (W/m²)	118 865	428.09	313.28	2.13	1494.85
DC Current (A)	118 865	355.90	264.86	0.60	995.68
AC Ir (A)	118 865	172.33	120.51	1.40	461.20
AC Iy (A)	118 865	172.19	120.45	1.40	461.10
AC Ib (A)	118 865	172.35	120.50	1.50	461.70
AC Power (kW)	118 865	128.08	91.19	0.39	332.61

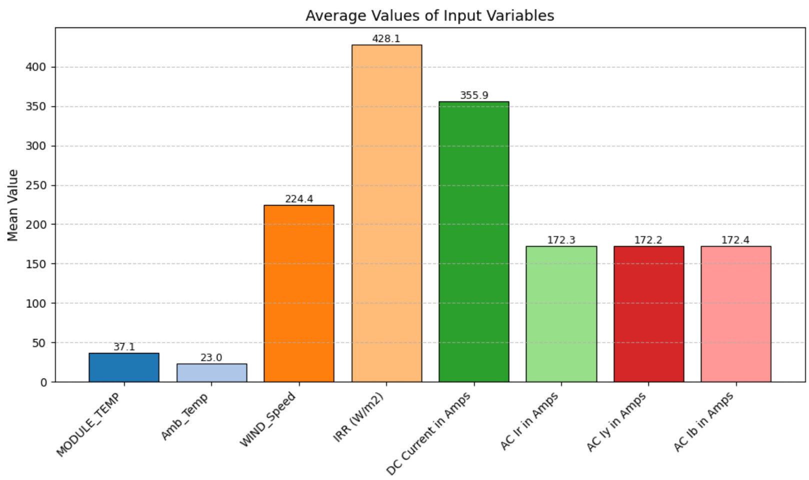

Figure 1: Mean values of the eight input variables in the PV SCADA dataset.

Figure 2 presents a Hexbin visualization depicting the joint dependencies between irradiance, DC current, and AC power [7]. The plots reveal strong nonlinear but monotonic relationships, while module temperature introduces a mild negative derating effect at high irradiance levels [8]. Figure 3 presents spearman correlation coefficients among environmental and electrical variables. To further examine nonlinear relationships among features,

All continuous features were standardized using z-score normalization via StandardScaler, fitted only on the training subset after a 70/30 train–test split (random_state = 42) [12]. This ensures zero-centered, unit-variance inputs and prevents data leakage, providing a stable basis for DNN-v4 training [13].

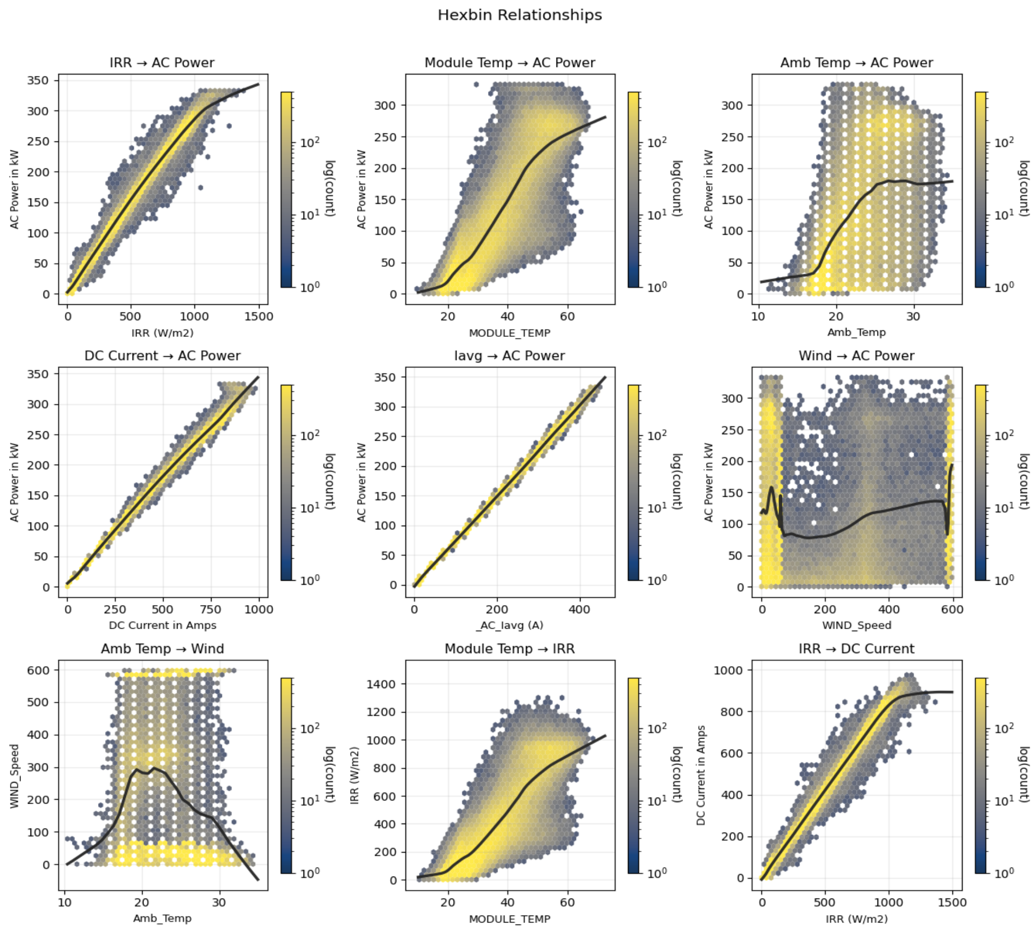

Figure 2: Hexbin visualization showing nonlinear dependencies among irradiance, DC current, and AC power.

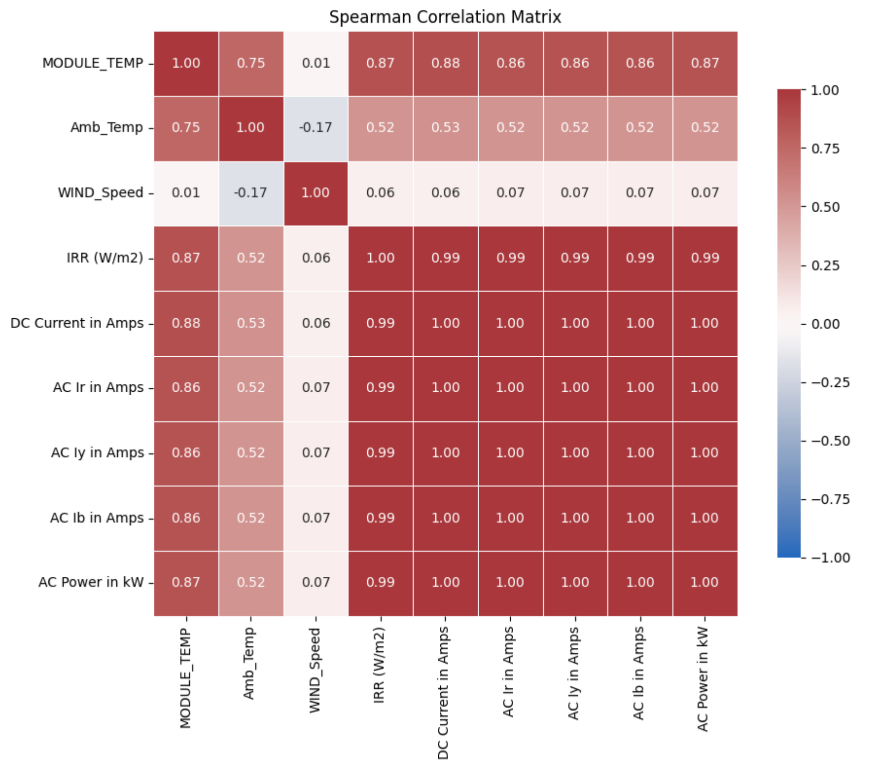

Figure 3: Spearman correlation coefficients among environmental and electrical variables.

3. Proposed Method (DNN-v4 Model Architecture)

The proposed DNN-v4 model was designed to provide a robust and interpretable framework for short-term photovoltaic (PV) power forecasting using high-frequency SCADA telemetry. The architecture follows a fully connected deep neural network structure optimized through extensive experimentation. The input layer receives eight standardized environmental and electrical variables, which are propagated through three hidden blocks with 512, 256, and 128 neurons, respectively. Each dense layer employs the GELU (Gaussian Error Linear Unit) activation to ensure smooth gradient flow and better handling of nonlinear patterns compared to ReLU. To enhance internal stability and mitigate covariate shift, each layer output is normalized via Layer Normalization, followed by a Dropout (rate = 0.1) for regularization.

The intermediate representations are fused using a Skip-Fusion mechanism, where multi-scale latent outputs extracted from the preceding neuron layers (with 512, 256, and 128 units, respectively) are concatenated into a unified feature tensor. This architectural strategy ensures that both high-level abstract patterns and lower-level fine-grained cues are preserved throughout the forward pass, preventing the loss of relevant information that may occur in conventional deep compression pipelines. The concatenated representation is subsequently fed into a compact refinement block composed of a Dense-128 layer, followed by Layer Normalization and a Dense-64 layer, which reduces redundancy and enforces feature decorrelation. Such refinement helps to stabilize the training process, reduce internal covariate shift, and facilitate faster model convergence.

By integrating hierarchical features from different abstraction levels, the Skip-Fusion design promotes stronger generalization, especially under varying environmental and operational conditions, while also maintaining interpretability — since contributions from each feature scale remain traceable within the fused latent space. Finally, the processed representation is mapped to the predictive output through a single-neuron regression layer with linear activation, enabling continuous estimation of the AC power (kW) without artificially constraining the output domain.

Optimization is handled by the AdamW optimizer combined with a Cosine Decay Restarts learning-rate schedule (initial LR = 8×10^{-4}, first_decay_steps = 80, t_mul = 1.5, m_mul = 0.8), which ensures efficient convergence and prevents over-fitting. The Huber loss ($\delta = 1.2$) function was chosen instead of MSE to reduce the sensitivity to outliers commonly found in SCADA data. To stabilize training and guarantee reproducibility, deterministic operations and fixed random seeds (*random_state = 42*) were applied across all layers.

The network was trained for up to 400 epochs with a batch size of 32, monitored by early stopping (patience = 40, restore_best_weights = True). Model training and evaluation were executed on a Tesla P100 GPU (16 GB), achieving an average epoch time of approximately 1.2 seconds. The entire training completed in about 4.2 minutes, indicating that the architecture is both computationally efficient and suitable for real-time PV forecasting applications.

Figure 4 illustrates the complete architecture of the proposed DNN-v4 framework, including the stacked dense feature extraction blocks, the multi-scale skip-fusion mechanism, and the associated training configuration. The diagram visualizes the sequential information flow from the standardized input features through hierarchical latent representations to the final output layer responsible for continuous AC power prediction. Additionally, Figure 4 highlights critical implementation details that ensure robust performance and reproducibility, such as the adopted optimization strategy, regularization components, and inference-time constraints, thereby providing a comprehensive overview of the designed deep learning pipeline.

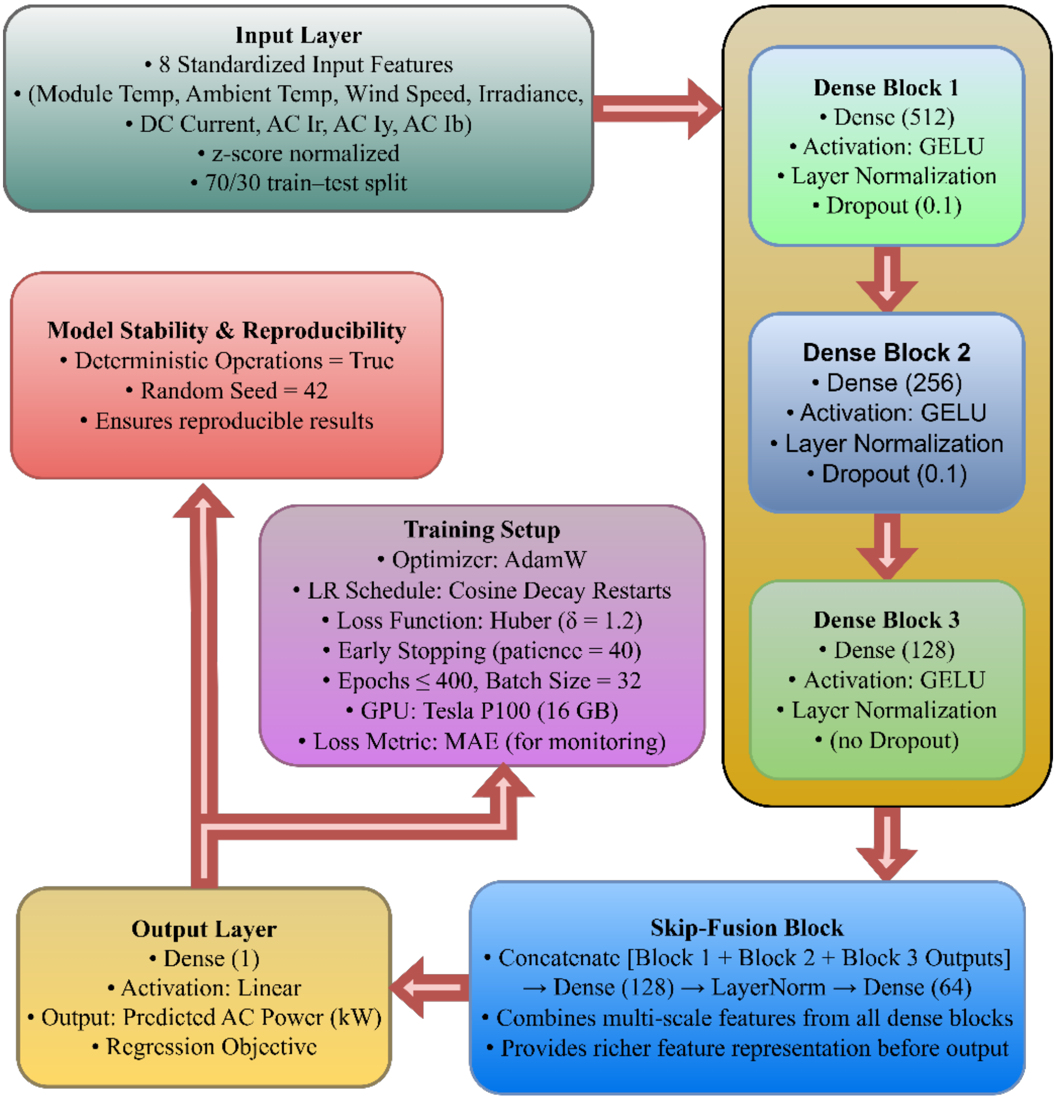

Figure 4: Block diagram of the proposed DNN-v4 model architecture.

4. Results and Discussion

The proposed DNN-v4 model demonstrated excellent predictive accuracy and robustness across all evaluation metrics. On the test dataset, the model achieved RMSE = 1.741 kW, MAE = 0.992 kW, MAPE=1.12 %, sMAPE = 1.14 %, and R^2 = 0.9996, confirming an almost perfect alignment between predicted and measured AC power outputs.

Table 2

Performance metrics of the proposed DNN-v4 model

RMSE (kW)	MAE (kW)	MAPE (%)	sMAPE (%)	R^2	Epochs
1.741	0.992	1.12	1.14	0.9996	212

The parity comparison between the predicted and actual AC power values is illustrated in Figure 5. The dense clustering of points along the diagonal demonstrates that the proposed DNN-v4 accurately learns nonlinear irradiance–temperature–current interactions without overfitting.

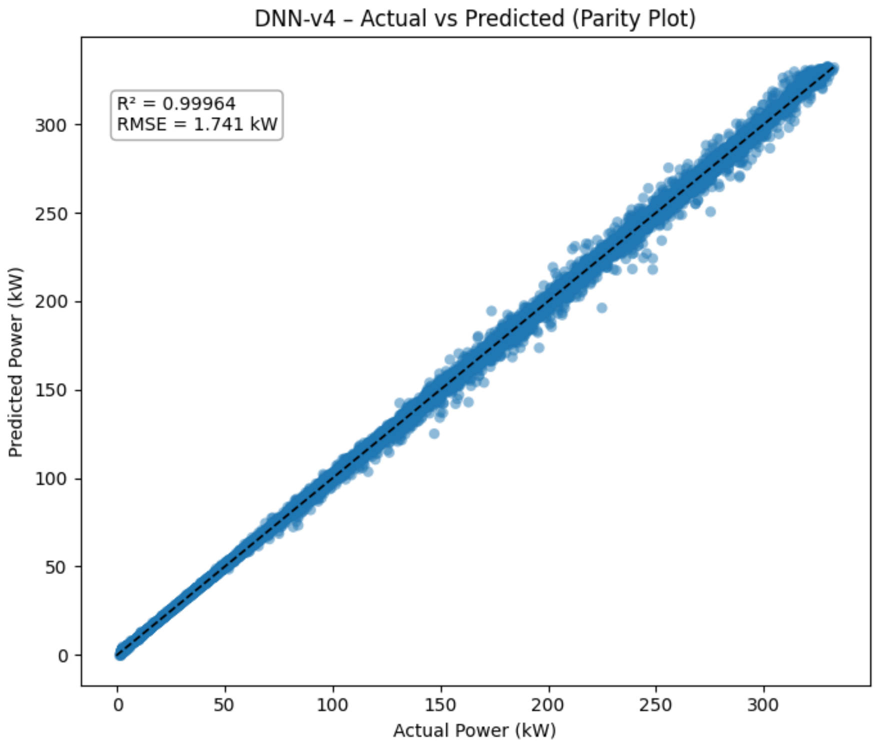

Figure 5: Parity plot showing the one-to-one correspondence between predicted and measured AC power.

Residual behavior is analyzed in Figures 6 and 7. The residual histogram in Figure 6 is narrow, symmetric, and centered around zero, indicating unbiased predictions and consistent variance. The Q–Q plot in Figure 7 further confirms that the residuals follow an approximately normal distribution, validating the absence of heteroscedasticity or systematic error patterns.

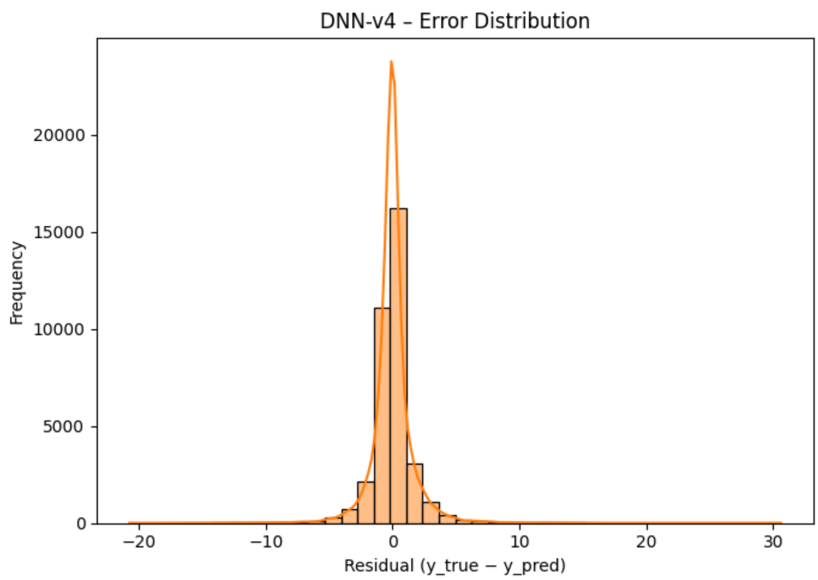

Figure 6: Histogram of residual errors between predicted and measured AC power.

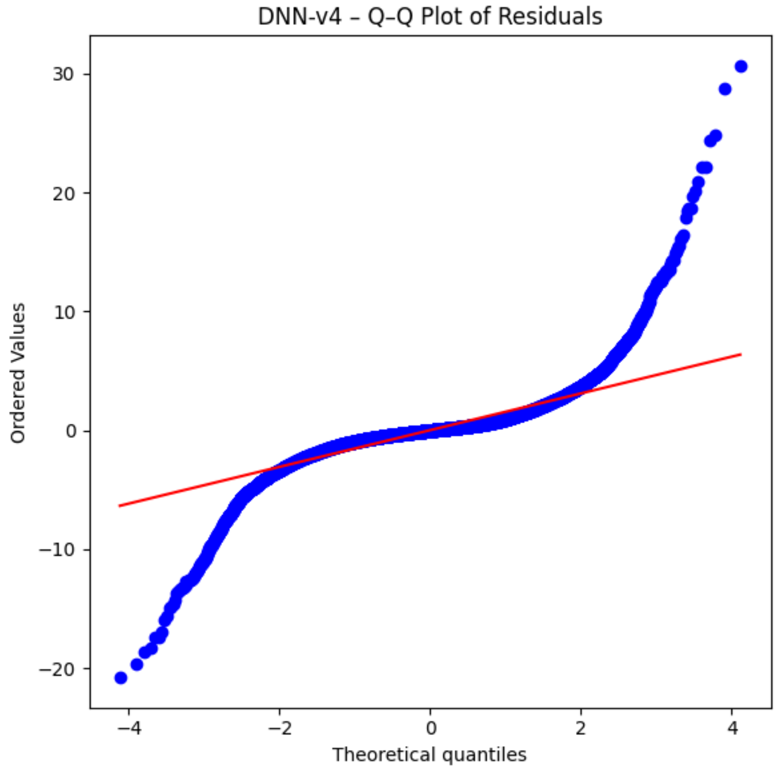

Figure 7: Q–Q plot comparing empirical residual distribution with theoretical normal quantiles.

Collectively, these findings confirm that the proposed DNN-v4 framework effectively combines high predictive accuracy, strong interpretability, and robust training stability. Compared with existing deep learning-based PV forecasting approaches, which often prioritize accuracy at the cost of transparency or exhibit sensitivity to data variability, DNN-v4 demonstrates a more favorable balance between performance and explainability. Such characteristics make the framework well-suited for operational photovoltaic (PV) power forecasting based on real-world SCADA telemetry, ensuring reliable deployment under diverse environmental and system operating conditions.

5. Conclusion

This study proposed an explainable and high-fidelity deep learning framework, DNN-v4, for photovoltaic (PV) power forecasting using SCADA telemetry. The model integrates GELU-activated dense blocks, layer normalization, and a skip-fusion mechanism optimized via the AdamW optimizer with cosine decay restarts. Experimental results demonstrated exceptional predictive accuracy (RMSE=1.74 kW, $R^2 = 0.9996$) and strong stability across 118,865 records, confirming that the architecture effectively captures nonlinear irradiance–temperature–current interactions without overfitting. Residual analyses validated the unbiased and near-Gaussian error behavior, further supporting the model's reliability. Owing to its interpretability and reproducibility, the proposed DNN-v4 framework provides a practical foundation for operational PV power forecasting and can be extended to multi-plant prediction or hybrid XAI-integrated configurations in future research.

Declaration on Generative AI

The author(s) have not employed any Generative AI tools.

References

[1] W. Hou et al., "Deep learning models for PV power forecasting: Review," *Energies*, vol. 17, no. 16, p. 3973, 2024, doi: 10.3390/en17163973

[2] L. M. Nguyen and J. Kim, “Short-term solar energy forecasting: Integrated computational intelligence approach,” Sustain. Energy Res., vol. 8, no. 1, p. 11, 2024, doi: 10.1186/s40807-024-00111-2.

[3] S. Al-Dahidi et al., “Enhancing solar photovoltaic energy prediction using diverse machine learning models tuned with the Chimp Optimization Algorithm,” Sci. Rep., vol. 14, art. 18583, 2024, doi: 10.1038/s41598-024-69544-8.

[4] P. Kut and K. Pietrucha-Urbanik, “Forecasting short-term photovoltaic energy production to optimize self-consumption,” Energies, vol. 18, no. 16, p. 4403, 2025, doi: 10.3390/en18164403.

[5] D. X. Wang, Y. Li, and J. Zhang, “Hybrid deep learning models for time series forecasting of solar power,” Neural Comput. Appl., vol. 36, no. 4, pp. 1119-1134, 2025, doi: 10.1007/s00521-024-09558-5.

[6] M. Elsaraiti and A. Merabet, “Solar power forecasting using deep learning techniques,” IEEE Access, vol. 10, pp. 31692-31698, 2022, doi: 10.1109/ACCESS.2022.31692.

[7] H. Ye et al., “State-of-the-art solar energy forecasting approaches: Critical potentials and challenges,” Front. Energy Res., vol. 10, art. 908620, 2022, doi: 10.3389/fenrg.2022.908620.

[8] J. Depoortere et al., “SolNet: Open-source deep learning models for photovoltaic power forecasting,” arXiv preprint, May 2024, doi: 10.48550/arXiv.2405.10367.

[9] N. Ahmed et al., “Solar energy forecasting with deep learning technique,” Renew. Energy, vol. 168, pp. 649-659, 2021, doi: 10.1016/j.renene.2020.12.097.

[10] L. Kuo and H. Tsai, “Machine learning-based short-term solar power forecasting,” Sustain. Energy Res., vol. 4, art. 115, 2024, doi: 10.1186/s40807-024-00115-1.

[11] R. Sahoo and R. Ghosh, “Explainable hybrid deep learning for photovoltaic power forecasting,” IEEE Trans. Sustain. Energy, vol. 15, no. 2, pp. 1349-1362, 2024, doi: 10.1109/TSTE.2024.3356628.

[12] H. Riggs, S. Cai, and M. Rao, “Combined machine learning and physics-based forecaster for solar irradiance forecasting under variable weather,” arXiv preprint, Mar. 2023, doi: 10.48550/arXiv.2303.06742.

[13] A. Kanagolkar, “SolarGeneration Dataset,” Kaggle, 2022. [Online]. Available: https://www.kaggle.com/datasets/arunkanagolkar/solargeneration/data. [Accessed: Oct. 26, 2025].