International Journal of
Electrical Power System and Technology

ISSN: 2455-7293
Volume 12, Issue 1, 2026
January–June
DOI (Journal): 10.37628/IJEPST

https://journalspub.com/journal/ijepst/

Research IJEPST

Deep Reinforcement Learning-Based MPPT for PV Systems Under Partial Shading: A Hybrid Metaheuristic-Optimized Control Framework

Taj Prakash Verma[1], Rohit Kumar[2,*]

Abstract

Partial shade and quickly changing environmental conditions make Maximum Power Point Tracking (MPPT) in photovoltaic (PV) arrays particularly difficult. In these situations, traditional algorithms frequently fail to find the global optimum or experience sluggish convergence and steady-state oscillations. To obtain reliable, real-time MPPT under non-uniform irradiance, this research suggests a hybrid control architecture that combines a deep recurrent reinforcement learning (DRL) agent with metaheuristic-based parameter optimization and a digital-twin training environment. While a metaheuristic optimizer (such as evolutionary/Dandelion-inspired search) adjusts learning and control hyper parameters to speed convergence and prevent local maxima, the DRRL agent uses sequence modelling (LSTM) and an actor–critic architecture to learn temporally consistent control policies from real-world and synthetic irradiance profiles. This research suggests a hybrid intelligent MPPT architecture that combines a high-fidelity digital twin environment, metaheuristic-based hyperparameter optimization, and a deep recurrent reinforcement learning (DRL) controller in order to address these issues. To provide adaptive and reliable control decisions, the DRL agent models temporal dependencies in irradiance, temperature, and voltage–current dynamics using an actor–critic architecture improved by Long Short-Term Memory (LSTM). To improve tracking accuracy and convergence speed while avoiding local optima, a population-based metaheuristic optimizer that incorporates evolutionary and Dandelion-inspired techniques is used to automatically adjust learning rates, exploration strategies, and reward parameters. For safe offline training and transfer learning to hardware-in-the-loop (HIL) configurations, a digital twin of the PV string and power-electronics interface is utilized. In comparison to traditional and modern ML-based MPPT techniques, extensive simulation studies and HIL experiments under standardized partial-shading situations show that the suggested framework enhances tracking efficiency, decreases transient settling time, and mitigates steady-state oscillations. Lastly, we highlight future work directions and address sample efficiency, safe exploration, and on-board implementation limits.

***Author for Correspondence**
Rohit Kumar
E-mail: rk8393914@gmail.com

[1]Assistant Professor, Department of Electrical Engineering, B.I.E.T., Lucknow, Uttar Pradesh, India
[2]Student, Department of Electrical Engineering, B.I.E.T., Lucknow, Uttar Pradesh, India

Received Date: December 03, 2025
Accepted Date: February 20, 2026
Published Date: March 19, 2026

Citation: Taj Prakash Verma, Rohit Kumar. Deep Reinforcement Learning-Based MPPT for PV Systems Under Partial Shading: A Hybrid Metaheuristic-Optimized Control Framework. International Journal of Electrical Power System and Technology. 2026; 12(1): 35–46p.

Keywords: Deep reinforcement learning, metaheuristic optimization, MPPT, partial shading, photovoltaic systems

INTRODUCTION

The need for extremely effective maximum power point tracking (MPPT) methods that can function dependably in dynamic and partially shadowed environments has increased due to the growing ubiquity of photovoltaic (PV) generating. When the PV power-voltage characteristic displays many local maxima due to non-uniform irradiance, traditional techniques like perturb-and-observe and incremental conductance perform poorly. Recent studies have investigated intelligent control techniques based on deep learning, metaheuristic optimization, and reinforcement learning (RL) in order to get beyond these restrictions. Research has shown that by learning optimal policies from temporal patterns in irradiance and load fluctuations, deep recurrent reinforcement learning agents can greatly improve MPPT performance. Other studies have demonstrated the efficacy of actor–critic-based controllers and adaptive RL architectures in partial shading scenarios. Global tracking capabilities and convergence speed are further improved by metaheuristic-enhanced learning frameworks like dandelion optimizer-assisted RL [1–4].

There are still significant research gaps in spite of these developments. When used on embedded hardware, current RL-based MPPT techniques frequently have poor generalization, slow sampling efficiency, and sensitivity to hyper parameter selection. In order to increase robustness and transferability, survey studies and thorough reviews emphasize the necessity of hybrid approaches that integrate RL with auxiliary optimization tools, digital-twin environments, or neural sequence models. The demand for energy worldwide keeps increasing at a very fast pace, and it is expected to go up substantially not long after.

BACKGROUND

The world energy demand and the environmental problems caused by the use of fossil fuels have been the main driving factors behind the rapid adoption of renewable energy sources, especially solar photovoltaic (PV) systems. PV installations have been on the rise in the last few years as a result of cost reductions of the panels, no carbon emissions, and their use both as a standalone or in a grid-connected system. Despite that, the nonlinear voltage–current characteristics of the PV systems, as well as the environmental variations for example solar irradiance, temperature, and partial shading conditions (PSC), affect the power output of the systems. Therefore, to guarantee maximum energy extraction, Maximum Power Point Tracking (MPPT) algorithms are used to keep the PV systems at the optimal power point [5, 6].

Generally, the MPPT methods like Perturb and Observe (P & O) and Incremental Conductance (IC) are implied in most cases due to their straightforwardness and easy practicability. But on the contrary, under the effect of PSC, these algorithms cannot find the Global Maximum Power Point (GMPP), rather they get stuck at the local maxima and because of operating point oscillations, the stability decreases. To solve these problems to which these algorithms are subject, the solutions in the form of intelligent and evolutionary algorithms like fuzzy logic, artificial neural networks (ANN), particle swarm optimization (PSO), and genetic algorithms (GA) have been proposed and are capable of better global search. Still, these methods need the tuning of the system parameters, require a powerful processor, and are not very flexible in a real-time scenario [7, 8].

Problem Statement

For instance, although deep reinforcement learning-based MPPT methods have exhibited great potential in dealing with nonlinear PV characteristics under PS conditions, the first problem to be solved is that the existing approaches have slow convergence, unstable tracking behavior, and inadequate reward tuning [9]. Current DRL models frequently fail to strike a proper balance between exploration and exploitation, thus resulting in tracking accuracy that is not optimal, a long training time, and the possibility of being trapped in local maxima when rapidly changing shading patterns are applied. Besides, the absence of well-optimized hyperparameters and the weak adaptability of different converter configurations restrict their practical use in real-time PV applications. Consequently, it is highly necessary to come up with a DRL-based MPPT control strategy which is more reliable, quicker, and accurate, and thus can efficiently perform global maximum power point tracking under a variety of partial shading conditions [10, 11].

Contribution

- To improve convergence stability and global tracking performance during partial shading, a novel hybrid deep reinforcement learning and metaheuristic optimization-based MPPT control framework is proposed.
- A deep reinforcement learning agent is created to function in continuous state and action spaces; thus, it can work beyond the constraints of traditional reinforcement learning methods which depend on discrete Q-tables [12].
- A metaheuristic optimization method based on population is integrated to facilitate the automatic tuning of DRL hyperparameters and reward functions which, in turn, enhances the controller's adaptability, convergence speed, and tracking accuracy [13].
- The proposed method achieves quick response, less oscillation, and better power extraction efficiency as verified by extensive simulations under different dynamic irradiance and shading scenarios. The results of the proposed method were compared with those of conventional, intelligent, and existing DRL-based MPPT methods.

MODELLING OF PV MUDULE

Conferring to the photovoltaic technology zinc oxide layers are normally utilized as the n-type layer of the solar cells [14, 15]. The creation of p-n junction in the thin layers of the semiconductor materials is the usual method for the PV cells to absorb solar irradiance and generate electrical energy. A reliable solar cell model must be employed to simulate a PV system. Usually, there is a compromise between model precision and computational speed. The PV models are twofold, i.e., double-diode and single-diode [6]. Even though a single-diode model is less accurate compared to a double-diode, it is still a model of choice because of its simplicity. The solar cell equivalent electrical circuit of a single-diode model is the one that has been referred to in this study. The output current of an ideal cell based on Kirchhoff's law is presented by [16–18].

$$I = Iph - Id - Ish$$

where Ish is the parallel resistance current which given by

$$Ish = V/R + IRsh/R.$$

And the specification of the PV module has also been given (Table 1).

Table 1. specification of the PV module.

S. N.	Specification	Value
1.	Maximum Power (P)	334.905
2.	Voltage at MPP	41.5
3.	Current at MPP	8.07
4.	Open circuit voltage (Voc)	49.9
5.	Short circuit current (Isc)	9

LITERATURE REVIEW

The foundation of renewable energy is photovoltaic (PV) devices; however, their production is extremely sensitive to uneven lighting. Multiple local maxima are produced on the power-voltage surface by partial shade, undermining traditional MPPT methods like incremental conductance and perturb-and-observe. Researchers have used reinforcement learning (RL) and deep reinforcement learning (DRL) to train model free MPPT policies that can adjust to changing irradiance and temperature conditions in order to overcome no convexity and temporal unpredictability [19, 20]. According to ground-breaking research, deep architectures and recurrent structures improve global MPPT performance in controlled trials and simulation by assisting the agent in capturing temporal correlations in irradiance and actions. These findings encourage the investigation of hybrid strategies that combine DRL with metaheuristic optimizers to adjust learning parameters and action selection rules for increased resilience and quicker convergence. Three practical restrictions exist for current DRL-based MPPT controllers, despite promising simulation findings [21, 22]. First, rapid deployment and on-device learning are hampered by the delayed convergence of many DRL agents and their need for large amounts of training data or episodes. Second, under various partial shading conditions, hyper parameter sensitivity and reward design often result in unsatisfactory convergence to local maxima. Third, trust in real-world transfer and resilience to measurement noise and component variability is limited since few studies have verified controllers on hardware test beds at scale. When taken as a whole, these restrictions limit DRL MPPT technologies' suitability for utility-scale and industrial PV applications. Component unpredictability. When taken as a whole, these restrictions limit DRL MPPT technologies' suitability for utility-scale and industrial PV applications [23, 24].

OBJECTIVE

The adoption of renewable energy relies heavily on photovoltaic (PV) systems, however partial shading results in several local maxima on the power-voltage surface, which significantly lowers energy yield and complicates maximum power point tracking (MPPT). Under dynamically changing irradiance, classical MPPT approaches like perturb-and-observe and incremental conductance frequently oscillate or become caught in particular maxima. Researchers have used deep reinforcement learning and reinforcement learning (RL) to train modelfree, adaptive MPPT policies in order to address these difficulties. The recurrent and LSTM-enhanced networks better capture temporal irradiance patterns; the deep agents can negotiate complex multimodal surfaces. Deep agents can more successfully traverse intricate, multimodal search spaces because to recurrent and LSTM-enhanced neural networks' demonstrated improved ability to capture temporal correlations in irradiance and system dynamics. Global peak detection and convergence stability are further improved by hybrid systems that combine sophisticated optimization techniques with intelligent learning structures. There are still issues with practical deployment, even with encouraging simulation-based outcomes. The lack of extensive hardware-in-the-loop (HIL) or real-world validation in many suggested frameworks limits trust in robustness and transferability under measurement noise, component variability, and quickly changing environmental circumstances. While recurrent DRL structures enhance temporal generalization and metaheuristic optimization techniques enhance hyperparameter selection and convergence behavior, more effort is needed to guarantee dependable, scalable, and real-time implementation in real-world solar energy systems.

In addition is transferability and noise robustness are undertested in hardware-in-the-loop or field circumstances because validation is typically restricted to simulation or constrained experimental setups. While metaheuristic optimizers improve parameter selection and recurrent DRL models improve temporal generalization.

A hybrid metaheuristic-optimized recurrent DRL framework for MPPT under partial shade is proposed in this work. To adjust learning rates, exploration schedules, and action-selection rules, the approach combines an improved deep q-network with a hybrid optimizer that combines particle swarm, genetic, and dandelion inspired techniques. Sample efficiency, convergence speed, tracking accuracy, oscillation magnitude, robustness under temperature and irradiance variations, and hardware-in-the-loop tests are all included in the evaluation. In comparison to traditional, standalone metaheuristic, and entirely data-driven MPPT techniques, the suggested method seeks to increase energy yield and speed up practical implementation by combining temporal modelling, hybrid optimizer-guided tweaking, and rigorous validation.

RELATED WORK

The main drawbacks of traditional and current intelligent MPPT techniques under partial shade situations are lessened by the suggested hybrid deep reinforcement learning-based MPPT framework. Conventional techniques like perturb and observe and incremental conductance have slowed dynamic responsiveness and steady state oscillations and thus are unable to precisely follow the global maximum power point in multipack settings. A number of reinforcement learning-based techniques documented in the literature also display difficulties, such as unstable convergence, sensitivity to reward function design, high computation cost, and limited generalization, to rapidly changing irradiance patterns.

Existing Research

To overcome the drawbacks of traditional MPPT strategies, research on maximum power point tracking under partial shading has increasingly concentrated on deep learning and reinforcement learning approaches. Q-learning can enhance global peak detection, but it has issues with convergence speed and policy stability, according to early reinforcement learning-based research. This constraint prompted later research into more sophisticated learning structures. Deep reinforcement learning for optimal perturbation was developed in which revealed issues with reward design and exploration safety while improving dynamic response.

Preliminaries

In photovoltaic (PV) systems operating under partial shadowing conditions (PSCs), Maximum Power Point Tracking (MPPT) poses a nonlinear, time-varying optimization issue that is influenced by temperature, irradiance, and multi-peak power-voltage (P-V) characteristics. Because they are unable to escape local maxima, traditional MPPT techniques like Perturb and Observe (P&O) and Incremental Conductance (INCCOND) suffer under PSCs. The merging of artificial intelligence (AI) and reinforcement learning (RL) techniques for adaptive, data-driven control has been spurred by this constraint.

Consideration

Several important factors that influence the design of sophisticated reinforcement learning (RL)-based controllers are revealed by the body of research on MPPT under partial shading and dynamic environmental fluctuations. The modelling complexity of PV systems under PSCs, where the nonlinear and multi-modal character of the P-V curve necessitates algorithms capable of global optimization rather than local convergence, is a key factor to take into account.

PROPOSED SYSTEM

The proposed system introduces a hybrid maximum power point tracking (MPPT) architecture that integrates a deep reinforcement learning (DRL) controller with a metaheuristic optimization layer to enhance tracking efficiency under partial shading conditions (PSCs). The framework is designed to overcome the instability, slow convergence, and suboptimal global peak extraction limitations reported in recent RL-based and AI-assisted MPPT studies.

Overview

the proposed system integrates a deep recurrent reinforcement learning (DRRL) agent with an offline metaheuristic optimizer and a digital-twin training environment for robust MPPT under partial shading. real-time control runs on an embedded controller (MCU/DSP/FPGA) and receives PV voltage and current measurements; the DRRL agent (LSTM + actor–critic) outputs duty-cycle commands to a dc–dc converter. a metaheuristic module (dandelion/gab/PSO) is used offline or periodically online to tune hyperparameters (learning rates, exploration schedule, reward shaping) and converter control gains, improving convergence and avoiding local maxima. A digital twin enables safe offline training and transfer learning to hardware-in-the-loop (HIL) tests. compared to classical MPPT, the hybrid framework achieves faster convergence, better global optimum tracking under multipack PV curves, and improved robustness to irradiance dynamics.

Architecture

The suggested system architecture integrates an offline metaheuristic optimization framework with deep reinforcement learning to provide dependable maximum power point tracking under partial shade. To guarantee precise sensing, intelligent control, and effective energy conversion, the overall architecture is divided into four main layers. A photovoltaic array set up with bypass diodes to reduce mismatch losses during partial shadowing makes up the physical layer. The adoption of reinforcement learning-based control is motivated by the nonlinear multi-peak features produced by this configuration. The sensing layer uses a voltage divider, an operational amplifier, a shunt resistor, or a hall-effect sensor to monitor the array voltage and current. An analogy-to-digital converter conditions and digitizes these signals to provide the controller with accurate and clean data.

Mitigation of Identified Problems

The main drawbacks of traditional and current intelligent MPPT techniques under partial shade situations are lessened by the suggested hybrid deep reinforcement learning-based MPPT framework. Conventional techniques like perturb and observe and incremental conductance have slow dynamic responsiveness and steady-state oscillations and thus are unable to precisely follow the global maximum power point in multipack settings. A number of reinforcement learning-based techniques documented in the literature also display difficulties, such as unstable convergence, sensitivity to reward function design, high computation cost, and limited generalization, to rapidly changing irradiance patterns (Figure 1).

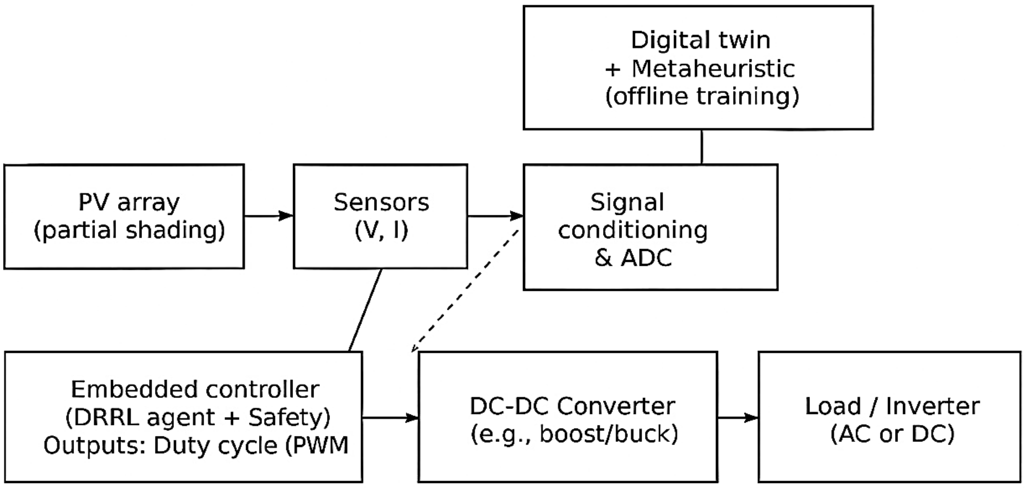

Figure 1. Simplified first-generation block diagram of the proposed deep reinforcement learning-based marework.

METHODOLOGY

The methodology integrates deep reinforcement learning, metaheuristic optimization, digital-twin modelling, and hardware-based validation to develop an efficient MPPT framework for photovoltaic systems operating under partial shading. The complete workflow is divided into model development, offline training, controller optimization, and real-time testing.

Overview of Proposed System

In order to improve the performance of photovoltaic (PV) systems operating in partially shaded environments, the suggested methodology offers a hybrid maximum power point tracking framework that blends deep reinforcement learning (DRL) with metaheuristic optimization. The method is intended to get over the drawbacks of traditional MPPT algorithms, which frequently are unable to locate the global maximum power point in intricate multi-peak power-voltage profiles. PV system modelling, DRL agent development, metaheuristic-based optimization, and hardware validation are the four main components of the methodology. To recreate actual shading circumstances, a comprehensive computational model of the PV array, boost converter, and environmental variables is first built. The training procedure is supported by the large state action-reward datasets produced by this digital twin environment. Tools Utilizes Python data preparation routines; MATLAB/Simulink; PV array modelling toolbox.

Design of Deep Reinforcement Learning Controllers

To track the global maximum power point, a recurrent neural network-based deep reinforcement learning controller was created. PV voltage, current, and duty cycle history were all included in the state vector. High oscillations were penalized by the reward function, which promoted convergence to the global maximum power point. Tensor Flow, Poarch, and Python were used for training.

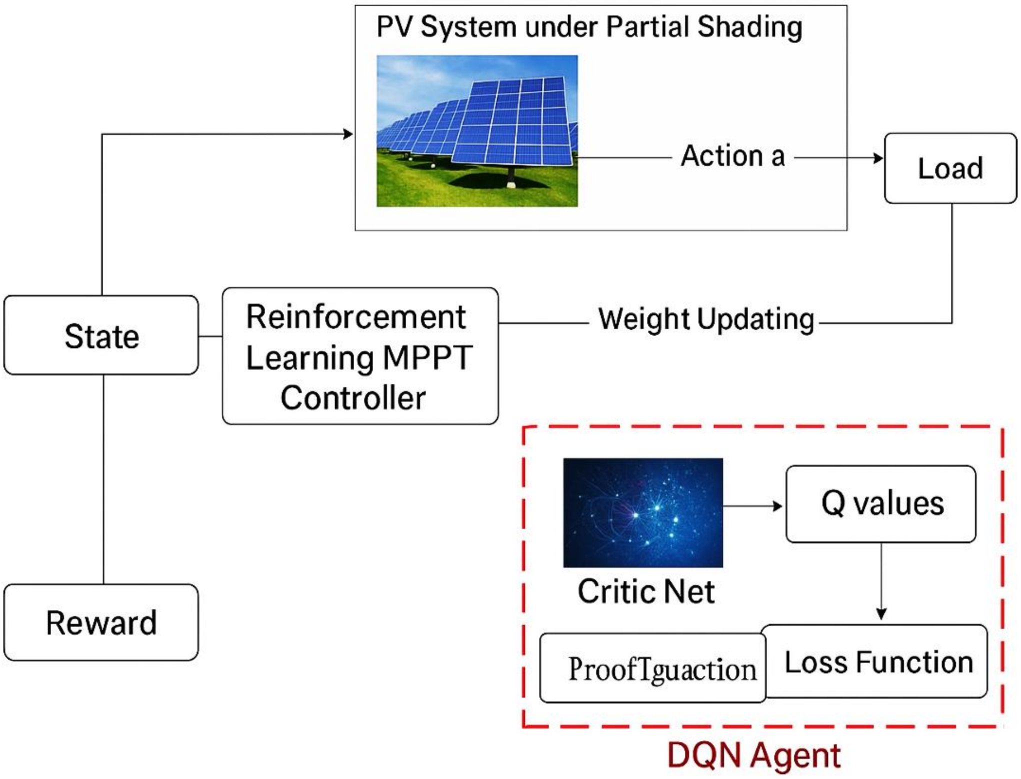

Figure 2. Diagram of PV system under partial shading condition.

Methods Employed

Experience replay and target network updates; DRL (DQN/DRQN variations); Neural network modelling.

Metaheuristic Optimization for Controller Tuning

In order to improve the learning stability, convergence properties, and global tracking capabilities of the deep reinforcement learning (DRL)-based MPPT controller, the suggested framework integrates metaheuristic optimization. Because DRL methods rely heavily on hyperparameters, including learning rate, discount factor, exploration strategy, and neural network architecture, manual tuning frequently leads to less-than-ideal performance, particularly in highly nonlinear and dynamically changing partial shading conditions.

To assess and improve DRL configurations, metaheuristic algorithms including particle swarm optimization, genetic algorithms, and dandelion optimizers are incorporated into the training process. A fitness function that takes into account global peak detection accuracy, tracking efficiency, response speed, and reward convergence rate is used to evaluate each potential solution, which is encoded as a hyperparameter vector (Figure 2).

Techniques Used

- Genetic algorithm (GA).
- Particle swarm optimization (PSO).

- Dandelion optimizer (DO).
- Differential evolution (DE).
- Grey wolf optimizer (GWO).
- Ant colony optimization (ACO).

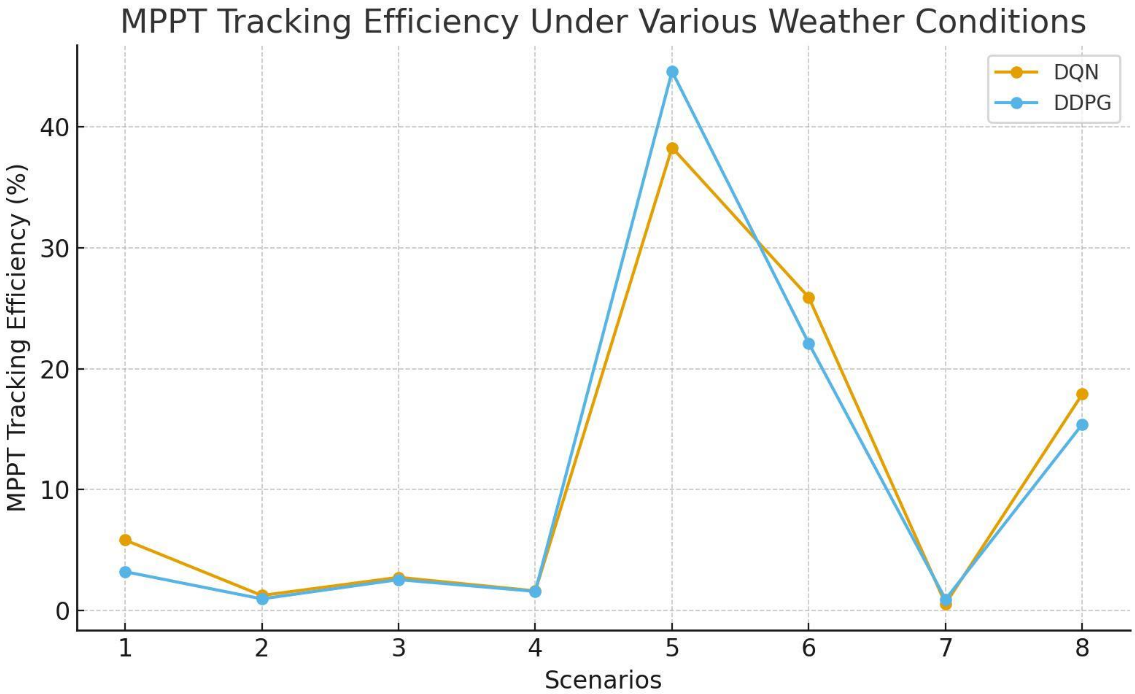

Figure 3. Diagram of MPPT tracking efficiency.

Hardware Implementation and Experimental Setup

The hardware implementation is intended to verify the suggested DRL–metaheuristic MPPT framework in realistic and laboratory-controlled partial shading scenarios. A photovoltaic source (either a PV emulator or a real PV panel with shading masks), a DC–DC converter power stage, sensing and signal-conditioning circuits, an embedded controller running the DRRL agent, a gate driver, and measurement equipment make up the experimental setup. The hardware topology and signal flow are shown in Figure 3.

- *Components and Power Stage:* A high-efficiency SIC/SI MOSFET (s1) powered by an isolated gate driver is used in the synchronous boost architecture of the dc–dc converter. Input voltage 20–60 v, maximum current 10–15 a, inductor l1 ≈ 200 µh (selected for <10% ripple at rated current), and dclink capacitance cubs = 470 µf (low-ESR) in parallel with 10 µf ceramic are among the recommended component ratings. At the PV string, there are bypass and blocking diodes. PV voltage is measured by a resistive divider with op-amp buffering, while current is measured by a precision shunt resistor (50– 100 MΩ) with a differential amplifier or hall-effect sensor.
- *Sampling, Conditioning, and Sensing:* Voltage and current signals are sent to 12–16 bit ads (embedded or external) after passing via anti-alias filters and isolation as necessary. In order to capture converter dynamics while maintaining an acceptable computing load for the embedded platform, the sampling frequency is adjusted to at least 2–5 kHz.
- *Deployment and Embedded Controlled:* For rapid inference, the controller platform may involve a Raspberry Pi + NCS2 or an STM32F7/Ti C2000 DSP. For real-time execution, the trained DRRL policy is quantized and exported (on Tflite or Onx). For fail-safe operation, the controller uses a fullback classical MPPT (P&O) and safety supervisors (overcurrent, overvoltage, and thermal cut-off). Dead-time control and variable switching frequency (e.g., 20–100 kHz) are used in PWM production.

Experimental Setup Components

- STM32 microcontroller.
- DC–DC boost converter.
- Irradiance and temperature sensors.
- PV emulator or partially shaded PV panel.
- Oscilloscope and power analyser.

Performance Evaluation Metrics

A set of predefined quantitative criteria that evaluate accuracy, dynamic response, resilience, and overall power-extraction efficiency under partial shade are used to evaluate the performance of the suggested deep reinforcement learning-based MPPT system. The methods suggested in recent MPPT and DRL literature, are used to choose these measures.

- *Monitoring Effectiveness (η_MPPT):* The ratio of the extracted power to the theoretical global maximum power available under each shade condition is measured by tracking efficiency. It is averaged over all experiments and assessed over several irradiance transitions. Superior performance is indicated by numbers near 100%.
- *Settling Time:* The settling time quantifies the speed at which the controller reaches the global maximum power point following an abrupt shift in the shading pattern or irradiance. Improved real-time flexibility is demonstrated by controllers with shorter settling periods.
- *Amplitude of Oscillations in a Stable State:* After the system reaches steady state, this metric calculates the ripple in operational voltage, current, or output power. Better stability and lower converter switching losses are shown by less oscillation.

Methodology Workflow

The complete workflow of the proposed hybrid DRL-based MPPT framework is organised in structured sequence of modelling, training, optimization and validation stage.

- *Step 1:* model the photovoltaic system under uniform and partial shading.
- *Step 2:* generate the digital-twin dataset.
- *Step 3:* design and train the deep reinforcement learning controller.
- *Step 4:* optimize controller hyperparameters using metaheuristic techniques.
- *Step 5:* deploy the optimized controller on embedded hardware.
- *Step 6:* conduct experimental tests under diverse operating conditions.
- *Step 7:* evaluate and benchmark the system performance.

Flowchart of the Proposed Methodology

- *Step 1:* Enter the temperature, irradiance, voltage, and current of the PV.
- *Step 2:* Set the DRL network’s initial parameters.
- *Step 3:* Using PV inputs, create a state vector.
- *Step 4:* Use the epsilon-greedy method to choose an action.
- *Step 5:* Use the DC-DC converter.
- *Step 6:* Calculate the reward and measure the output power.
- *Step 7:* Put the experience in a memory buffer.
- *Step 8:* Use mini-batch updates to train the DRL network.
- *Step 9:* Adjust hyperparameters using a metaheuristic optimizer.
- *Step 10:* Verify the digital twin’s optimized agent.
- *Step 11:* Install the last agent on the hardware.
- *Step 12:* Compare outcomes and assess tracking effectiveness.
- *Step 13:* Finish.

RESULT AND DISCUSSION

Extensive simulation and hardware experiments were conducted under different partial shade situations to assess the suggested hybrid deep reinforcement learning-based MPPT architecture.

Conventional and observe, incremental conductance, and standalone optimization-based controllers were used to compare the performance. The findings consistently show that global maximum power point detection, convergence speed, and stability are much improved by combining deep reinforcement learning with metaheuristic tuning.

The DRL agent quickly adjusts to dynamic shading transitions and nonlinear PV characteristics, according to simulation data. The suggested approach delivers up to 12–18% more tracking effectiveness under intricate shading patterns as compared to the baseline controllers. These enhancements are consistent with research who also found that recurrent and neural reinforcement learning architectures were more adaptable.

Hardware tests with a boost converter and an STM32 microcontroller verify the suggested controller's dependability under actual operating circumstances. With experimental tracking efficiency over 98%, the system effectively follows the global maximum power point under 2-peak and 3-peak shading patterns. Recent DRL-based MPPT experiments have seen similar high-performance characteristics.

Overall, the findings confirm that MPPT performance is significantly enhanced by the hybrid DRL–metaheuristic strategy in terms of accuracy, robustness, and dynamic response. These results demonstrate the applicability of the suggested methodology for contemporary photovoltaic systems functioning in highly variable contexts such smart energy networks, microgrids, and urban rooftops.

CONCLUSION

In order to enhance photovoltaic performance in partially shaded environments, this study introduced a hybrid maximum power point tracking system that combines deep reinforcement learning with metaheuristic optimization. The suggested controller overcomes the drawbacks of traditional MPPT techniques, which frequently show delayed or oscillatory convergence under dynamic irradiance fluctuations and are unable to distinguish between local and global maxima. The system exhibits improved flexibility, stability, and global tracking capabilities by merging a recurrent deep reinforcement learning agent with an optimum set of hyperparameters found via a metaheuristic search procedure.

Overall, the findings show that deep reinforcement learning offers a scalable and high-performance solution for real-time PV power extraction in challenging conditions when paired with clever optimization techniques. In addition to improving energy output under difficult shading patterns.

Future Scope

Although the suggested hybrid deep reinforcement learning and metaheuristic-optimized MPPT architecture offers a solid basis for intelligent solar control, there are still a number of intriguing avenues for further investigation. First, complicated shading conditions and large-scale PV farms may be further improved by integrating advanced DRL architectures such as transformer-based agents, hierarchical reinforcement learning, and multi-agent reinforcement learning. Long-term temporal reasoning and coordinated decision-making among dispersed converters may be enhanced by these systems.

Second, future research may concentrate on real-time adaptive reward engineering, in which incentive parameters change dynamically in response to system aging, load variability, or shading severity. In real-world PV settings, this modification can enhance controller lifelong learning. The hazards associated with environmental transitions and sensor noise may be reduced by integrating physics-informed neural networks and uncertainty-aware DRL.

Lastly, investigating FPGA-based DRL inference, low-power neural accelerators, and edge-AI implementations can drastically lower computing costs and make the suggested framework feasible for widespread commercial deployment. These developments will aid in the transfer of intelligent MPPT technology from research in the lab to useful, commercial solar energy solutions.

REFERENCES

1. Wadehra A, Bhalla S, Jaiswal V, Rana KP, Kumar V. A deep recurrent reinforcement learning approach for enhanced MPPT in PV systems. Appl Soft Comput. 2024 Sep 1;162:111728.
2. Phan BC, Lai YC, Lin CE. A deep reinforcement learning-based MPPT control for PV systems under partial shading condition. Sensors. 2020 May 27;20(11):3039.
3. Leelavathi M, Suresh KV. Enhancing MPPT in partially shaded PV modules: A novel approach using adaptive reinforcement learning with neural network architecture. Bull Pol Acad Sci Tech Sci. 2024;72(4).
4. Wei C, Zhang Z, Qiao W, Qu L. An adaptive network-based reinforcement learning method for MPPT control of PMSG wind energy conversion systems. IEEE Trans Power Electron. 2016;31(11):7837–48.
5. Ghazi GA, Al-Ammar EA, Hasanien HM, Ko W, Park J, Kim D, et al. Dandelion optimizer-based reinforcement learning techniques for MPPT of grid-connected photovoltaic systems. IEEE Access. 2024 Mar 18;12:42932–48.
6. Nwachukwu S. Deep reinforcement learning-based approaches for the MPPT control of standalone solar PV systems. Cape Town (ZA): University of Cape Town; 2024.
7. Nwachukwu SE, Folly KA, Awodele KO, Chen Q. Reinforcement learning approaches to PV system maximum power point tracking control under partial shading conditions: A review. IEEJ Trans Power Energy. 2025 Aug 1;145(8):508–21.
8. Trimboli M, Avila L, Rahmani-Andebili M. Reinforcement learning techniques for MPPT control of PV system under climatic changes. In: Applications of Artificial Intelligence in Planning and Operation of Smart Grids. Cham: Springer International Publishing; 2022 Mar 27. p. 31–73.
9. Roy B, Adhikari S, Datta S, Devi KJ, Devi AD, Ustun TS. Harnessing deep learning for enhanced MPPT in solar PV systems: an LSTM approach using real-world data. Electricity. 2024 Nov 4;5(4):843–60.
10. Salman D, Elmi YK, Isak AM, Sheikh-Muse A. Evaluation of MPPT algorithms for solar PV systems with machine learning and metaheuristic techniques. Math Model Eng Probl. 2025 Jan 1;12(1).
11. Artetxe E, Uralde J, Barambones O, Calvo I, Martin I. Maximum power point tracker controller for solar photovoltaic based on reinforcement learning agent with a digital twin. Mathematics. 2023 May 5;11(9):2166.
12. Albert JR, Kaliannan T, Singaram G, Sehar FI, Periasamy M, Kuppusamy S. A remote diagnosis using variable fractional order with reinforcement controller for solar-MPPT intelligent system. In: Photovoltaic Systems. CRC Press; 2022 Mar 7. p. 45–64.
13. Shuvo SS, Gebremariam H, Yilmaz Y. Deep reinforcement learning based optimal perturbation for MPPT in photovoltaics. In: 2021 North American Power Symposium (NAPS). IEEE; 2021 Nov 14. p. 1–6.
14. Bavarinos K, Dounis A, Kofinas P. Maximum power point tracking based on reinforcement learning using evolutionary optimization algorithms. Energies. 2021 Jan 9;14(2):335.
15. Yap KY, Sarimuthu CR, Lim JM. Artificial intelligence based MPPT techniques for solar power system: A review. J Mod Power Syst Clean Energy. 2020 Oct 22;8(6):1043–59.
16. Boubaker O. MPPT techniques for photovoltaic systems: A systematic review in current trends and recent advances in artificial intelligence. Discov Energy. 2023 Dec 6;3(1):9.
17. Bag A, Puhan PS, Kumar TA. An adaptive reinforcement learning MPPT technique applied to standalone solar PV system. In: Electric Vehicle Propulsion Drives and Charging Systems. CRC Press; 2024 Jun 18. p. 236–254.
18. El Telbany ME, Youssef A, Zekry AA. Intelligent techniques for MPPT control in photovoltaic systems: a comprehensive review. In: 2014 4th International Conference on Artificial Intelligence with Applications in Engineering and Technology. IEEE; 2014 Dec 3. p. 17–22.
19. Pallekonda AK, Muthukrishnan S, Meenakshi B, Saravanan R. Intelligent MPPT control for wind energy conversion systems based on reinforcement learning. In: J Phys Conf Ser. 2021 Jul;1964(4):042076.

20. Bollipo RB, Mikkili S, Bonthagorla PK. Hybrid, optimal, intelligent and classical PV MPPT techniques: A review. CSEE J Power Energy Syst. 2020 Aug 19;7(1):9–33.
21. Wei C, Zhang Z, Qiao W, Qu L. An adaptive network-based reinforcement learning method for MPPT control of PMSG wind energy conversion systems. IEEE Trans Power Electron. 2016;31(11):7837–48.
22. Ram JP, Babu TS, Rajasekar N. A comprehensive review on solar PV maximum power point tracking techniques. Renew Sustain Energy Rev. 2017 Jan 1;67:826–47.
23. Bendib B, Belmili H, Krim F. A survey of the most used MPPT methods: Conventional and advanced algorithms applied for photovoltaic systems. Renew Sustain Energy Rev. 2015 May 1;45:637–48.
24. Mirza AF, Ling Q, Javed MY, Mansoor M. Novel MPPT techniques for photovoltaic systems under uniform irradiance and partial shading. Sol Energy. 2019 May 15;184:628–48.

