



## OPEN Deep regression analysis for enhanced thermal control in photovoltaic energy systems

Wael M. Elmessery<sup>1</sup>✉, Abadeer Habib<sup>2</sup>, Mahmoud Y. Shams<sup>3</sup>✉, Tarek Abd El-Hafeez<sup>4,5</sup>✉, Tamer M. El-Messery<sup>2</sup>✉, Salah Elsayed<sup>6</sup>, Ahmed E. M. Fodah<sup>7</sup>, Taha A. M. Abdelwahab<sup>7</sup>, Khaled A. M. Ali<sup>7,8</sup>, Yasser K. O. T. Osman<sup>7</sup>, Mohamed F. Abdelshafie<sup>7</sup>, Gomaa G. Abd El-wahhab<sup>7</sup> & Abdallah E. Elwakeel<sup>9</sup>

Efficient cooling systems are critical for maximizing the electrical efficiency of Photovoltaic (PV) solar panels. However, conventional temperature probes often fail to capture the spatial variability in thermal patterns across panels, impeding accurate assessment of cooling system performance. Existing methods for quantifying cooling efficiency lack precision, hindering the optimization of PV system maintenance and renewable energy output. This research introduces a novel approach utilizing deep learning techniques to address these limitations. A U-Net architecture is employed to segment solar panels from background elements in thermal imaging videos, facilitating a comprehensive analysis of cooling system efficiency. Two predictive models—a 3-layer Feedforward Neural Network (FNN) and a proposed Convolutional Neural Network (CNN)—are developed and compared for estimating cooling percentages from individual images. The study aims to enhance the precision and reliability of heat mapping capabilities for non-invasive, vision-based monitoring of photovoltaic cooling dynamics. By leveraging deep regression techniques, the proposed CNN model demonstrates superior predictive capability compared to traditional methods, enabling accurate estimation of cooling efficiencies across diverse scenarios. Experimental evaluation illustrates the supremacy of the CNN model in predictive capability, yielding a mean square error (MSE) of just 0.001171821, as opposed to the FNN's MSE of 0.016. Furthermore, the CNN demonstrates remarkable improvements in mean absolute error (MAE) and R-square, registering values of 1.2% and 0.95, respectively, whereas the FNN posts comparatively inferior numbers of 3.5% and 0.85. This research introduces labeled thermal imaging datasets and tailored deep learning architectures, accelerating advancements in renewable energy technology solutions. Moreover, the study provides insights into the practical implementation and cost-effectiveness of the proposed cooling efficiency monitoring system, highlighting hardware requirements, integration with existing infrastructure, and sensitivity analysis. The economic viability and scalability of the system are assessed through comprehensive cost-benefit analysis and scalability assessment, demonstrating significant potential for cost savings and revenue increases in large-scale PV installations. Furthermore, strategies for addressing limitations, enhancing predictive accuracy, and scaling to larger datasets are discussed, laying the groundwork for future research and industry collaboration in the field of photovoltaic thermal management optimization.

**Keywords** Photovoltaic solar panels, Cooling systems monitoring, Thermal imaging, Deep learning, Image segmentation, Computer vision, Performance optimization

<sup>1</sup>Agricultural Engineering Department, Faculty of Agriculture, Kafrelsheikh University, Kafr El-Shaikh 33516, Egypt. <sup>2</sup>International Research Centre “Biotechnologies of the Third Millennium”, Faculty of Biotechnologies (BioTech), ITMO University, St. Petersburg 191002, Russia. <sup>3</sup>Department of Machine Learning and Information Retrieval, Faculty of Artificial Intelligence, Kafrelsheikh University, Kafr Elsheikh 33516, Egypt. <sup>4</sup>Department of Computer Science, Faculty of Science, Minia University, Minia 61519, Egypt. <sup>5</sup>Computer Science Unit, Deraya University, Minia University, Minia 61765, Egypt. <sup>6</sup>Agricultural Engineering, Evaluation of Natural Resources Department, Environmental Studies and Research Institute, University of Sadat City, Cairo 32897, Minufia, Egypt. <sup>7</sup>College of Agricultural Engineering, Al-Azhar University, Cairo 11768, Egypt. <sup>8</sup>College of Engineering, South China Agriculture University, Guangzhou, China. <sup>9</sup>Agricultural Engineering Department, Faculty of Agriculture and Natural Resources, Aswan University, Aswan 81528, Egypt. ✉email: wael.elmaysari@agr.kfs.edu.eg; mahmoud.yasin@ai.kfs.edu.eg; tarek@mu.edu.eg; tamerelmessery@yahoo.com

The ever-growing demand for electrical power, driven by industrialization and technological advancements, has intensified the focus on optimizing the efficiency of power generation systems.

Photovoltaic (PV) solar power has emerged as a critical renewable energy source, but maintaining high electrical efficiency relies heavily on effective panel cooling systems<sup>1</sup>. Various cooling systems are used in photovoltaic (PV) systems to improve energy conversion efficiency and prevent performance loss. Passive and active cooling methods are applied on the front and back surfaces of PVs under different working conditions. These cooling techniques include heat recovery units, water-cooled heat sinks, and backside water cooling systems<sup>2–4</sup>. It is important to note that many of these cooling methods, particularly those involving active cooling with fluids, are characteristic of photovoltaic/thermal (PV/T) systems. PV/T systems are hybrid solar systems that convert solar radiation into both electricity and useful thermal energy simultaneously. These systems not only cool the PV panels to maintain electrical efficiency but also harness the thermal energy for applications such as water heating or space heating, thereby increasing the overall energy conversion efficiency of the system. Therefore, Monitoring and optimizing cooling dynamics for solar plant performance and longevity is crucial, especially considering the limitations of traditional temperature probes in characterizing surface variability across panels. Novak et al.<sup>5</sup> proposed an intelligent automation system that monitors parameters such as cooling medium amount, temperature, flow control, panel temperature, and local weather conditions to evaluate the conditions for active smart cooling. Singh and Yadav<sup>6</sup> developed an experimental setup for solar panel cooling using water cooling arrangements and optimized the input parameters to achieve improved efficiency and module temperature. Herfatmanesh et al.<sup>7</sup> conducted experiments to explore the effects of solar PV surface temperature on output performance and proposed a cooling system that increased efficiency by close to 50%. Gujar et al.<sup>8</sup> designed a PV-panel cooling system using fins and phase change material, which maintained the temperature below 50 °C and increased conversion efficiency by 5–7%. Liu et al.<sup>9</sup> summarized various solar cell cooling technologies, including traditional cooling methods and new technologies like heat pipe cooling and microchannel cooling.

Thermographic Non-Destructive Test (TNDT) is a common method for diagnosing faults in PV systems. It uses infrared thermography to examine the operating conditions of the system. TNDT analysis allows for the measurement of cell temperature and the detection of localized overheating in the module, known as hotspots, by capturing thermographic images that show the surface temperature of the module in each pixel<sup>10</sup>. This technique enables the identification of areas of potential defects without directly accessing the module. By detecting hotspots, TNDT can help prevent performance loss and ensure the safety of the PV system<sup>11,12</sup>.

Thermal imaging has been widely proposed as a solution for non-invasive, high-resolution visualization of temperature patterns. Yet analytical approaches for extracting actionable insights from thermal video remain limited<sup>13</sup>.

Recent literature reflects a growing interest in utilizing deep learning techniques, such as convolutional neural networks (CNNs), for analyzing thermal imaging in the context of predictive insights and condition monitoring. These techniques have been applied to various fields, including nuclear power plants<sup>14</sup>, crack detection in inductive thermography<sup>15</sup>, classification of patterns in thermographic images of a bench grinder<sup>16</sup>, early pavement damage detection<sup>17</sup>, and defect inspection in artwork using principal component analysis<sup>9</sup>. These studies demonstrate the effectiveness of deep learning in extracting features and enhancing the quality of thermal images, leading to improved detection and classification of defects and faults. CNNs have shown promise in detecting and classifying faults in photovoltaic (PV) systems using thermal images<sup>10,18</sup>. Deep learning methods, such as variational autoencoders (VAEs), have been used to expand the data set and improve the accuracy of fault classification<sup>19</sup>. Another approach is to use a deep-learning-based defect detection method that addresses the challenges of limited data and data imbalance, achieving high accuracy in identifying the presence or absence of defects and different defect types<sup>20</sup>. Additionally, a high-precision algorithm has been proposed for detecting and classifying defects in PV panels, achieving a detection accuracy of 92.0%<sup>21</sup>. Furthermore, thermal CNNs have been used for the detection, quantification, and on-field localization of overheated regions on PV arrays, with the best performance achieved by the FPN-DenseNet121 model. However, there is limited research on using deep learning on thermal video to directly quantify and forecast cooling efficiency for optimizing PV performance.

Data-driven regression models based on CNNs represent a highly promising approach for precisely estimating cooling effectiveness from thermal imagery<sup>22</sup>. Recent applications of regression CNNs in related domains have shown strong capabilities for predicting system parameters from spatial temperature maps<sup>23</sup>. However, tailored datasets and models are needed to focus specifically on correlating thermal patterns with PV cooling efficiency. The development of automated pipelines for generating labeled cooling efficiency data from thermal plant videos could enable robust CNN regressors<sup>24</sup>. The integration of deep CNN architectures for regression with extensive annotated thermal datasets has the potential to significantly advance non-invasive quantification and monitoring of photovoltaic cooling dynamics. This will provide crucial insights for predictive maintenance and efficiency improvements in solar power systems<sup>25</sup>.

### Problem statement

Despite advancements in thermal management for photovoltaic (PV) solar panels, existing methods for quantifying cooling efficiency often lack the precision necessary for optimizing PV system maintenance and energy output. Traditional temperature probes fail to capture spatial variability in thermal patterns, which is crucial for accurate cooling system assessment.

### Research question

How can deep learning techniques be applied to thermal imaging data to improve the quantification of cooling system efficiency in PV solar panels, and what impact does this have on the optimization of renewable energy output?

### Practical applications

Our deep regression analysis method goes beyond simply obtaining temperature information. It provides a real-time, non-invasive way to quantify cooling system efficiency across large-scale photovoltaic installations. This capability has several practical applications:

- Predictive Maintenance:** By continuously monitoring cooling efficiency, operators can identify underperforming or failing cooling systems before they significantly impact PV output. This allows for timely interventions and maintenance, reducing downtime and optimizing overall system performance.
- Optimization of Cooling Strategies:** For PV systems with active cooling capabilities (such as PV/T systems, as you correctly point out), our method provides valuable feedback on the effectiveness of different cooling strategies. This data can be used to fine-tune cooling parameters, such as flow rates or activation thresholds, to maximize efficiency and minimize water/energy consumption.
- Design Improvements:** The spatial resolution of our thermal analysis can highlight areas of panels that consistently experience higher temperatures. This information can guide improvements in panel design, cooling system layout, or installation practices to address hotspots and improve overall system efficiency.

### Improving cooling system performance

Our method contributes to improved cooling system performance in the following ways:

- For PV/T Systems:** In photovoltaic-thermal hybrid systems, our method provides precise, real-time feedback on cooling efficiency. This data can be integrated into control systems to dynamically adjust coolant flow rates, optimizing the balance between electrical output and thermal energy harvesting.
- For Passive Cooling Systems:** While passive cooling systems cannot actively reduce temperatures, our method helps in assessing the effectiveness of various passive cooling techniques (e.g., enhanced convection, reflective coatings). This information is valuable for comparing different passive cooling strategies and guiding future installations.
- Performance Verification:** Our method offers a way to verify the performance of newly installed or upgraded cooling systems, ensuring they meet design specifications and operate efficiently under real-world conditions.

#### *Indirect improvement of electricity production*

While our system doesn't directly cool PV modules, the insights it provides can lead to improved electricity production through:

- Informed Decision Making:** By providing accurate, real-time data on cooling efficiency, our system enables operators to make informed decisions about when to clean panels, adjust tilt angles, or implement other temperature management strategies.
- Long-term Performance Optimization:** The data collected over time can inform better system designs, cooling strategies, and maintenance schedules, all of which contribute to improved long-term electricity production.
- Integration with Smart Grid Systems:** In large-scale installations, our cooling efficiency data can be integrated with smart grid management systems to optimize overall energy production and distribution based on real-time panel performance.

### The literature review

Harnessing solar energy has gained popularity as an efficient method to power homes, businesses, and other utilities. One such efficient method is through the use of solar thermoelectric generators, which transform thermal energy into electricity, offering a wide range of applications, such as heating water and powering electronics.

Li et al.<sup>26</sup> emphasize the need for complex control and monitoring systems to maintain efficiency in thermal energy harvesting systems, which can be costly to implement and maintain. Ma et al.<sup>27</sup> illustrated the potential of solar energy as a renewable, clean, and abundant source of energy. However, one of the key challenges is achieving maximum energy harvesting. Trappey et al.<sup>28</sup> presented solar thermoelectric absorbers as a method to harness solar energy and convert it into electricity. However, several issues need addressing to maximize energy harvesting with solar thermoelectric absorbers. Sun et al.<sup>29</sup> discussed the considerable environmental impact of thermal energy harvesting. Some thermal energy harvesting systems rely on fossil fuels to generate heat, leading to an increase in air pollution. Moreover, these systems may require large amounts of water for cooling, resulting in significant water wastage. Sajedian et al.<sup>30</sup> emphasized the importance of thermal energy harvesting in reducing emissions and improving energy efficiency. However, several issues need to be addressed to make this technology a feasible solution for many applications. These issues include overall efficiency, cost, size and location requirements, and environmental impacts. Ahmad et al.<sup>31</sup> discussed that with the right solutions, these issues can be addressed, making thermal energy harvesting a more practical option for a wide range of applications.

Zhang, X et al.<sup>32</sup> addressed the high costs involved in thermal energy harvesting systems, which can often be a deterrent for many applications. Another problem related to thermal energy harvesting is the large area required to capture and store energy, which can significantly influence the project's feasibility and cost depending on its size and location. The availability of heat sources in some regions can also pose challenges, as they might lack a reliable heat source. Lin et al.<sup>33</sup> expressed one such challenge is the potential for low efficiency. The efficiency of thermal energy harvesting systems depends on the temperature difference between the waste heat source and the ambient environment, as well as the conversion system's efficiency. Gorjian et al.<sup>34</sup> discussed the cost of the absorber as a significant factor in maximizing energy harvesting. Solar thermoelectric absorbers are generally more expensive than other solar energy harvesting technologies, such as photovoltaics. Therefore, finding ways to reduce the absorber's cost is crucial to maximize energy harvesting.

Bai et al.<sup>35</sup> spoke of the growing field of thermal energy harvesting technology and its potential to drastically change our energy consumption and production practices. This involves capturing and using waste heat produced by industrial processes, vehicles, and other energy sources. Despite its potential, thermal energy harvesting faces many significant challenges. Gao et al.<sup>36</sup> discussed another challenge being the system's ability to withstand the harsh conditions of industrial processes and exhaust systems. High temperatures, corrosive chemicals, and other extreme conditions can quickly degrade these systems' components, leading to premature failure, and resulting in financial losses and wasted energy. Varga et al.<sup>37</sup> explained that thermal energy harvesting is a promising technology field with the potential to revolutionize energy production and consumption. With adequate investments, research, and development, these systems can overcome the challenges and become a crucial part of our energy infrastructure.

Elsheikh et al.<sup>38</sup> elaborated on the process of thermal energy harvesting, which involves the extraction of energy from heat sources in the environment and converting it into practical energy. The significance of thermal energy harvesting has been growing in recent years, as it helps in reducing emissions, enhancing energy efficiency, and ensuring a reliable energy supply. There are, however, critical issues that need resolving to make thermal energy harvesting a feasible solution for numerous applications. Liu et al.<sup>39</sup> highlighted the efficiency problems associated with thermal energy harvesting. Heat energy is challenging to capture and store, and a considerable amount of energy is lost during the conversion process, which results in a lower overall efficiency of thermal energy harvesting systems compared to other energy production forms. Table 1 presents a thorough examination of related research efforts focusing on thermal energy harvesting.

This analysis underscores the multidimensional challenges and opportunities associated with thermal energy harvesting, providing valuable insights for future research and development in the field.

Table 2 provides a comprehensive overview of various studies that have explored the application of machine learning (ML) and artificial intelligence (AI) techniques in the detection of faults within photovoltaic (PV) systems. The studies span a range of years and employ a diverse array of techniques, including Artificial Neural Networks (ANN), Support Vector Machines (SVM), and Deep Learning (DL) methods, among others. Each

Author	Classification	Identified issues	Metrics
Li et al. <sup>26</sup>	Implementation complexity	Complexity and cost associated with implementing control and monitoring systems.	Highlights the challenges related to implementing and maintaining control systems for thermal energy harvesting systems.
Ma et al. <sup>27</sup>	Energy maximization	Focuses on challenges in maximizing energy harvesting from solar sources.	Emphasizes the importance of maximizing energy harvesting from solar sources for sustainable energy production.
Sun et al. <sup>29</sup>	Environmental impact	Environmental impacts associated with thermal energy harvesting, particularly air pollution from fossil fuel combustion.	Raises concerns about the environmental consequences of certain thermal energy harvesting methods reliant on fossil fuel combustion.
Trappey et al. <sup>28</sup>	Solar thermoelectric absorbers	Issues in optimizing energy harvesting with solar thermoelectric absorbers.	Discusses challenges and opportunities in optimizing energy harvesting using solar thermoelectric absorbers.
Ahmad et al. <sup>31</sup>	Transformative potential	The transformative potential of thermal energy harvesting in revolutionizing energy consumption and production.	Highlights the role of thermal energy harvesting in utilizing waste heat from industrial processes and vehicles to improve energy sustainability.
Zhang et al. <sup>32</sup>	Spatial requirements	The spatial requirements and feasibility of thermal energy harvesting projects.	Highlights the importance of project size and location in determining the feasibility and cost-effectiveness of thermal energy projects.
Sajedian et al. <sup>30</sup>	Comprehensive challenge overview	A comprehensive overview of challenges including efficiency, cost, spatial requirements, and environmental impacts.	Suggests that addressing these challenges is crucial for enhancing the viability of thermal energy harvesting across various applications.
Lin et al. <sup>33</sup>	Temperature differentials	Limitations in temperature differentials impacting system efficiency.	Notes the importance of temperature differentials in optimizing the efficiency of thermal energy harvesting systems.
Gorjian et al. <sup>34</sup>	Cost considerations	Cost considerations associated with solar thermoelectric absorbers.	Discusses cost implications and challenges in the adoption of solar thermoelectric absorbers for energy harvesting.
Bai et al. <sup>35</sup>	Efficiency issues	The primary challenge of low efficiency in thermal energy harvesting systems.	Stresses the significance of addressing low-efficiency issues to maximize the effectiveness of thermal energy harvesting technologies.
Liu et al. <sup>39</sup>	Efficiency comparison	Challenges related to the efficiency of thermal energy harvesting systems and their comparison with other energy production methods.	Emphasizes the lower efficiency of thermal energy harvesting systems compared to alternative energy production methods.
Elsheikh et al. <sup>38</sup>	Emission reduction and energy efficiency	The importance of thermal energy harvesting in reducing emissions, enhancing energy efficiency, and ensuring reliable energy supply.	Advocates for thermal energy harvesting as a critical solution for addressing environmental concerns and enhancing energy reliability.
Gao et al. <sup>36</sup>	Durability and harsh conditions	Concerns regarding system durability and performance under harsh conditions.	Emphasizes the need for thermal energy harvesting systems to withstand extreme conditions and maintain performance reliability.
Varga et al. <sup>37</sup>	Investment prospects	Promising prospects for thermal energy harvesting with appropriate investments and research.	Encourages investment and research efforts to overcome challenges and maximize the potential of thermal energy harvesting.

**Table 1.** Overview of challenges and metrics in thermal energy harvesting.

Authors	Year	Classification	ML models	PV system component	Contribution	Remarks
Youssef et al. <sup>41</sup>	2016	AI	ANN, FL, ANFIS, GA, GA-fuzzy, NN-fuzzy	PV field	Demonstrates the importance of AI in modeling, sizing, forecasting, and diagnosing faults in PV systems.	Compares the accuracy of different AI techniques with traditional methods, but does not specify the monitoring parameters for each method.
Daliento et al. <sup>42</sup>	2016	Electrical and AI	ANN, SVM, ANFIS, RBN	PV field	Provides a review of various methods used to monitor PV systems.	Well-written and adheres to desired characteristics; no changes were necessary.
Mellit et al. <sup>43</sup>	2016	Electrical and ML	ANN, FL, MSD	PV field	Discusses PV fault information and diagnosis methods.	Primarily focuses on identifying defects.
Rodrigues et al. <sup>44</sup>	2017	M.L.	DT, RF, FL, ANN, GA, Bayesian, KNN, GA-ANN, ANFIS, RVM, k-Means	PV field	Reviews prognosis and diagnosis of defects and covers the number of themes in the study.	Reviews types of studies, faults, input parameters, and PV systems but lacks evaluation of method effectiveness.
Madeti et al. <sup>45</sup>	2017	Conventional and AI	--	PV field	Reviews detection methods for grid-connected photovoltaic systems.	Already meets desired characteristics; no changes were made.
Mellit et al. <sup>46</sup>	2018	Electrical and ML	ANN, FL, GA, HS	PV field	Comprehensive review on detection methods for grid-connected PV systems.	Focuses on using electrical methods to diagnose faults.
Abdulmawjood et al. <sup>47</sup>	2018	Visual, Thermal, and ML Methods	SVM, k-Means, HMM, BN, ANN, GMM (Gaussian mixture model)	PV field	Covers different types of faults and detection techniques in PV fields.	Discussion is centered on electrical faults, but the detection parameters are not specified for each method.
Pillai et al. <sup>48</sup>	2018	IRT, ML, Others	ANN, LAPART	PV field	Includes a review of almost all PV faults and advanced detection techniques.	Focuses on flaws in detection methods.
Ghaffarzadeh et al. <sup>49</sup>	2019	Electric, ML	ANN, SVM, DT, FL, Kalman filter	PV field	Explains types of defects across a broad spectrum.	Focuses on current faults on the DC and AC sides of the PV system.
Appiah <sup>50</sup>	2019	IRF, ML, DL	ANN, LAPART, KELM, ANFIS	PV field	Reviews types of defects, their origins, and traditional and intelligent detection methods.	Clear and concise, but lacks complexity, precision, and input data.
Li et al. <sup>51</sup>	2020	M.L.	ANN	PV field	Focuses on ANN and hybrid methods applied to defect analysis, including data used, model configuration, and effectiveness.	Compares ANNs with other ML models, showing superiority of ANNs; however, does not compare between ANN models to identify the most efficient one.
Venkatesh et al. <sup>52</sup>	2020	Visual method, IRT, EL, ML	ANN, SVM, NC-NFC, CNN, DT, KNN, FL	PV field	Lists four types of visual defects and detection methods.	Does not take non-visual defects into account; lacks precision.
Kurukuru et al. <sup>53</sup>	2021	ML, DL	ANN, ANFIS, PSO, FL, GA, ABC, CNN, SVM, KNN, LSTM	PV field	Reviews the impact of AI on the PV value chain.	Does not provide precision for each technique.
Mansouri et al. <sup>54</sup>	2021	D.L.	DBN, CNN, RFCN, R-CNN	PV field	Reviews Deep Learning applications in solar cell fault detection.	Examines defects related to cell discoloration, cracking, and delamination in PV systems.
Abubakar et al. <sup>55</sup>	2021	AI, ML	ANN, SVM, LAPART, RBF-ELM, FL, GBSS, ANFIS, DT	PV field	Discusses characteristics of AI methods, their speed, and effectiveness in detecting defects with minimal errors.	Does not justify inclusion of articles from the last 15 years; does not include accuracy rate for each model.
Gaviria et al. <sup>56</sup>	2022	D.L.	ANN, LSTM, CNN, SVM, RF	PV field	Reviews the interest of ML in PV systems, providing resources for datasets and source codes.	Lacks objectivity and precision in presenting results; includes insignificant articles on defect diagnosis using ML.
Hammoudi et al. <sup>57</sup>	2022	D.L.	CNN, LSTM	PV field	Surveys the interest of Deep Learning and IoT in PV system maintenance.	Limited to discussing deep learning in preventive maintenance on the DC side.
Zenebe et al. <sup>58</sup>	2022	ML, DL	SVM, DA, BN, ANN, KNN, RF, DT, CNN	PV field, Inverter	Reviews ML-based detection methods, showing that ANN and MLP are promising in terms of simplicity and accuracy.	Mainly focuses on defects and detection methods.
Yuan et al. <sup>59</sup>	2022	M.L.	ANN	PV field	Reviews progress of ANN in fault diagnosis.	Lacks information on precision and complexity of each ANN type.
Forootan et al. <sup>60</sup>	2022	ML, DL	SVM, DA, BN, ANN, kNN, RF, DT, CNN, FL, ANFIS, GA, LSTM, RL, MLR, SLR, k-Means, etc.	PV field	Reviews ML and DL algorithms in energy systems.	Fails to consider non-visual defects and lacks precision.
Berghout et al. <sup>61</sup>	2022	ML, DL	SVM, kNN, MLP, LSTM, CNN, Gans	PV field	Discusses monitoring PV systems and defects related to shading and degradation.	Focuses on ML categories, detection techniques, and two types of defects; does not provide accuracy for each model.
Puthiyapurayil et al. <sup>62</sup>	2022	AI, signal-based method	ANN, BPNN, SVM, CNN	Inverter	Lists different methods of diagnosing open-circuit faults in an NPC inverter.	Focuses only on single switch open-circuit faults; rare cases of three switch faults are not covered.
Engel et al. <sup>63</sup>	2022	ML, DL	ANN, CNN, ANFIS, YOLOv4, k-NN, DT, SVM, RF, NB	PV field	Reviews ML advances in prediction, forecasting, sizing, and diagnosis of PV systems.	Compares diagnostic methods, showing better performance of DNN models over non-neural models.
Ying-Yi et al. <sup>64</sup>	2022	Visual and thermal	SVM, kNN, MSD, DT, RF, ANFIS, ANN	PV field	Presents traditional methods of detecting and classifying PV faults and projects AI techniques.	Focuses on traditional methods but demonstrates potential of ML techniques.
Continued						

Authors	Year	Classification	ML models	PV system component	Contribution	Remarks
Osmani et al. <sup>65</sup>	2023	Conventional methods, AI	SCADA, ANN, KELM (kernel extreme learning machine)	PV field	Critical review of detection methods in the PV field.	Presents DC and AC side faults, focusing on conventional methods and omitting supervised learning methods.
Islam et al. <sup>66</sup>	2023	Artificial intelligence based on ML and DL	AdaBoost, ANN, CNN, RNN, SVM, RF	PV field	Systematic review on identification and diagnosis methods, comparing existing reviews with its own in terms of technical approaches for fault detection.	Identifies most effective DL and ML approaches for PV fault diagnosis, showing DLs superiority over conventional methods; does not provide accuracy rates for different methods.

**Table 2.** Overview of review articles on ML and DL applications in photovoltaic (PV) systems (2016–2023)<sup>40</sup>.

study contributes uniquely to the field by focusing on different aspects of fault detection, from grid-connected systems to specific PV components such as inverters and PV fields.

### Research gap

Previous studies have employed thermal imaging for fault detection in photovoltaic (PV) systems, but they have not focused on directly quantifying and forecasting cooling efficiency using deep learning applied to thermal video data. Specifically, there is a lack of research utilizing deep regression models to estimate cooling effectiveness from thermal imagery for optimizing PV panel performance. The research gaps can be summarized as follows:

1. Absence of a holistic approach for operational fault analysis in PV systems using machine learning that addresses both fault detection and diagnosis.
2. Limited exploration of machine learning methods capable of managing non-linear relationships and distinguishing between features with similar signatures for PV fault detection and diagnosis.
3. Insufficient research on the real-time or online application of machine learning for PV fault analysis.
4. A gap in comprehensive reviews focused on machine learning techniques specifically designed for PV system fault detection and diagnosis.

### Contributions

In addressing the critical challenges of thermal management in photovoltaic (PV) solar panels, this study makes several key contributions to the field of renewable energy optimization. By leveraging advanced deep learning techniques and thermal imaging data, we have developed a comprehensive approach that not only enhances the precision of cooling system efficiency quantification but also paves the way for significant improvements in PV panel performance and maintenance. The contributions of this research are multifaceted and are outlined as follows:

1. Introducing a novel deep learning approach that utilizes U-Net architecture for image segmentation and CNNs for estimating cooling efficiencies from thermal images.
2. Providing a comparative analysis of two predictive models (FNN and CNN) and demonstrating the superior predictive performance of the CNN model.
3. Presenting a labeled thermal imaging dataset specifically tailored for training deep learning models to correlate thermal patterns with PV cooling efficiency.
4. Discussing the practical implementation, cost-effectiveness, and scalability of the proposed monitoring system, including a comprehensive cost-benefit analysis.
5. Laying the groundwork for future research and industry collaboration in the field of photovoltaic thermal management optimization by addressing limitations and enhancing predictive accuracy.

### Materials and methods

The flowchart in Fig. 1 outlines the workflow for predicting cooling system efficiency in the development of a real-time spatially resolved monitoring system for photovoltaic cooling dynamics. It initiates with the installation of a thermal camera to capture images, followed by testing cooling conditions and segmenting solar panels using a U-Net architecture. The segmented panels are cropped, and an automated process categorizes these thermal images for creating cooling efficiency labels. Subsequently, a dataset with 390 labeled and cropped panels is compiled for training. A Convolutional Neural Network (CNN) or Feedforward Neural Network (FNN) is then designed to predict continuous values within the 0–100% range. The model undergoes training over 50 epochs with Mean Squared Error (MSE) loss and the Adam optimizer. Lastly, the model's performance is assessed on a test set across various efficiency brackets, providing valuable insights into its predictive capabilities.

In the pursuit of an efficient real-time monitoring system for assessing photovoltaic cooling dynamics, the initial step involves precisely isolating panel segments from raw thermal video. Fixed region cropping limitations are addressed by employing a deep neural network for semantic segmentation, classifying pixels as solar panels or background elements. Utilizing a U-Net architecture with a MobileNetV2 encoder through transfer learning, our framework achieves a high-level understanding of thermal image content. Trained on a dataset of 390 annotated thermal images, our model demonstrates efficacy with a dice coefficient score of 0.92, distinguishing panel regions. In test inference, the model robustly isolates panel sections, even with variable scaling and occlusions. Integrating U-Net segmentation facilitates the extraction of panel-specific thermal patterns, enabling the quantification of cooling dynamics and their correlation with electrical efficiency.

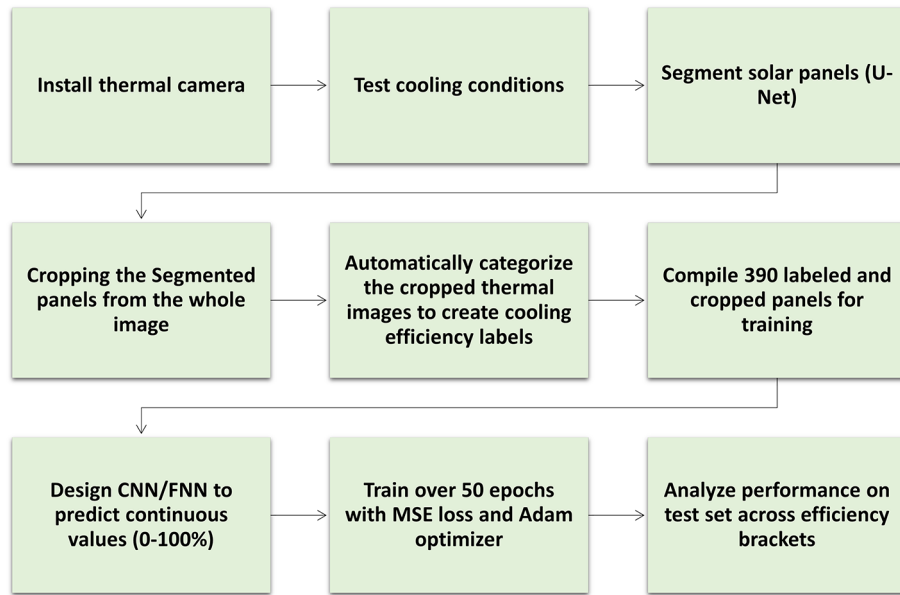


Fig. 1. Flowchart of the design workflow for a regression model predicting cooling system efficiency.

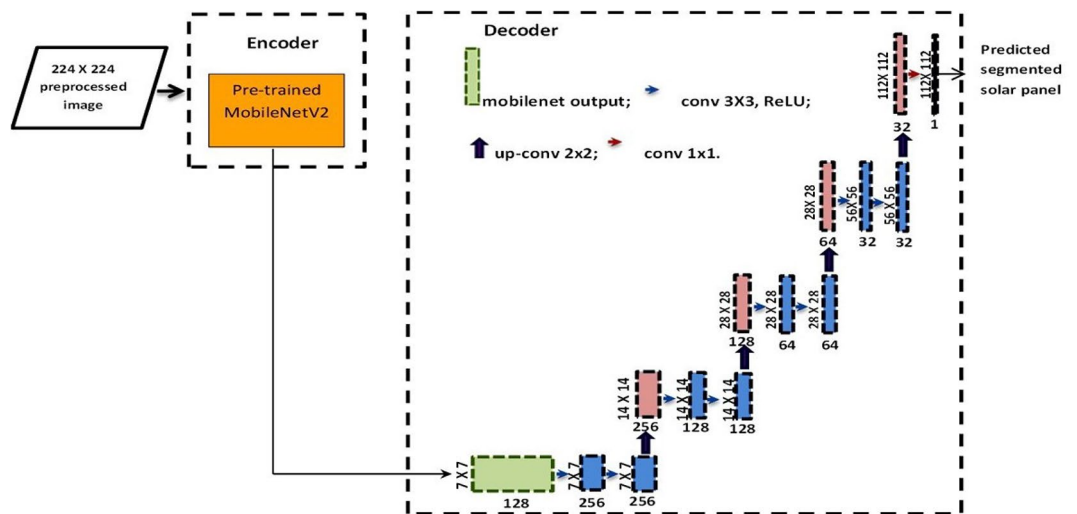


Fig. 2. The architecture of Integrated MobileNetV2 and U-net for solar panel segmentation.

### Solar panel segmentation through U-net architecture

In our quest for precise solar panel segmentation, we propose a sophisticated U-Net architecture that seamlessly integrates a MobileNetV2 encoder with a meticulously designed decoder as shown in Fig. 2. The primary goal is to discern solar panels from the background within 224 x 224 RGB images.

The model is tailored to handle 224 x 224 RGB images, configuring the input shape as (224, 224, 3). Leveraging transfer learning, we employ a MobileNetV2 model pre-trained on ImageNet as our feature extractor encoder. This involves constructing the MobileNetV2 model without the classification output layer. To preserve the knowledge embedded in the pre-trained MobileNetV2 encoder, the encoder layers are set to non-trainable, ensuring that their weights remain fixed during subsequent training.

The MobileNetV2 encoder processes the input image, generating a compact (7, 7, 1280) feature representation that encapsulates essential spatial and semantic information. Moving into the decoder, we utilize Conv2D layers for meticulous reconstruction of the segmentation mask. These convolutional layers play a pivotal role in transforming feature map sizes to upsample back to the original input resolution. The Upsampling2D layers then spatially scale the feature maps in the decoding sequence, without introducing additional parameters to learn.

The final Conv2D layer produces the segmentation mask with an output shape of (224, 224, 1). This single-channel mask aligns with the input size, effectively distinguishing solar panels from the background. The layer

employs a single unit to predict the label of each pixel, categorizing it as either a solar panel or part of the background.

The overall model summary provides insightful information about the architecture, indicating over 6 million total parameters. However, it's noteworthy that only about 4 million of these parameters are trainable, specifically within the decoder portion. The MobileNetV2 encoder contributes non-trainable pre-trained weights, emphasizing the importance of transfer learning in our approach.

Our U-Net architecture utilizes an encoder-decoder-style convolutional neural network, capitalizing on pre-trained features in the encoder. This strategy enhances the segmentation of solar panels from the background in input thermal images. The output shape seamlessly progresses from the encoded representation to full resolution, providing a nuanced understanding of the segmentation process.

The diagram in Fig. 2 depicts a series of feature map layers, represented by colored boxes. The blue boxes each signify a multi-channel feature map extracted from convolutional layers. On the bottom of these blue boxes, the number of channels is denoted for that particular feature map. On the left side, the spatial x-y dimensions of each feature map are also provided. Additionally, light red boxes are used in the diagram to denote feature maps that have been upsampled - increasing their spatial resolution. Finally, the arrows are shown pointing to and labeling the different operations and components of the overall architecture. The light olive green box is the generated compact feature map by a pre-trained MobileNetV2 encoder.

### Automated thermal data categorization for photovoltaic cooling efficiency labeling

To train a deep neural network to estimate the cooling efficiency from thermal imagery, a dataset of labeled images spanning the spectrum of possible cooling levels was required. However, capturing repetitive data under precise cooling loads is impractical. Instead, an automated pipeline to categorize images based on the apparent percentage of the solar panel surface being actively cooled was developed. This provides a scalable method to bin images for model training. The process operates on segmented solar panel portions extracted from raw thermal captures of photovoltaic installations under routine conditions. Firstly, the solar panel from each image using a separate U-Net segmentation model was isolated. These cropped panel segments were then processed to simplify analysis.

Firstly, color images were converted to grayscale, and temperature thresholding was subsequently applied to categorize regions into chilled and non-chilled based on distribution histograms. The percentage of “cooled pixels” out of the total solar panel area provides an estimate of the relative cooling level. Defining percentile ranges allows images to be categorized into discrete bins (e.g., 10–20% cooled, 80–90% cooled). In the thermal dataset processing pipeline, an approximate percentage of cooling is assessed by computing the ratio of cooling to non-cooled pixels, as determined through temperature thresholding. A discrete categorical label is then assigned that bins each image into 10% increments based on this continuous percentage. This discretizes the spectrum into 11 classes (0–10%, 10–20%, etc.).

However, the original continuous percentage value is also recorded in the dataset metadata associated with each image. Therefore, each thermal image contains both a continuous cooling efficiency percentage (measured between 0 and 100%), as well as a discrete categorization bin (0–10%, 10–20%, etc.). During model development, benefits were found in leveraging both the continuous percentage values as regression labels to predict intermediate efficiencies and the categorical bins for classification tasks. The dataset provides two forms of labels: the continuous percentage acting as the ground truth efficiency and the discretized classes to enable categorical training. Both forms were utilized in training convolutional neural networks to estimate photovoltaic cooling from thermal imagery. Applying this pipeline to over 390 uploaded thermal images; a training dataset was established and distributed across the cooling efficiency spectrum. This enabled the subsequent deep neural network to learn detectable thermal signatures associated with increasing active cooling, providing generalized predictive capabilities.

Analyzing performance across efficiency brackets during testing illuminates challenging areas needing additional data. Our automated thermal data categorization forms crucial groundwork in designing an accurate vision-based cooling monitoring system for solar plants.

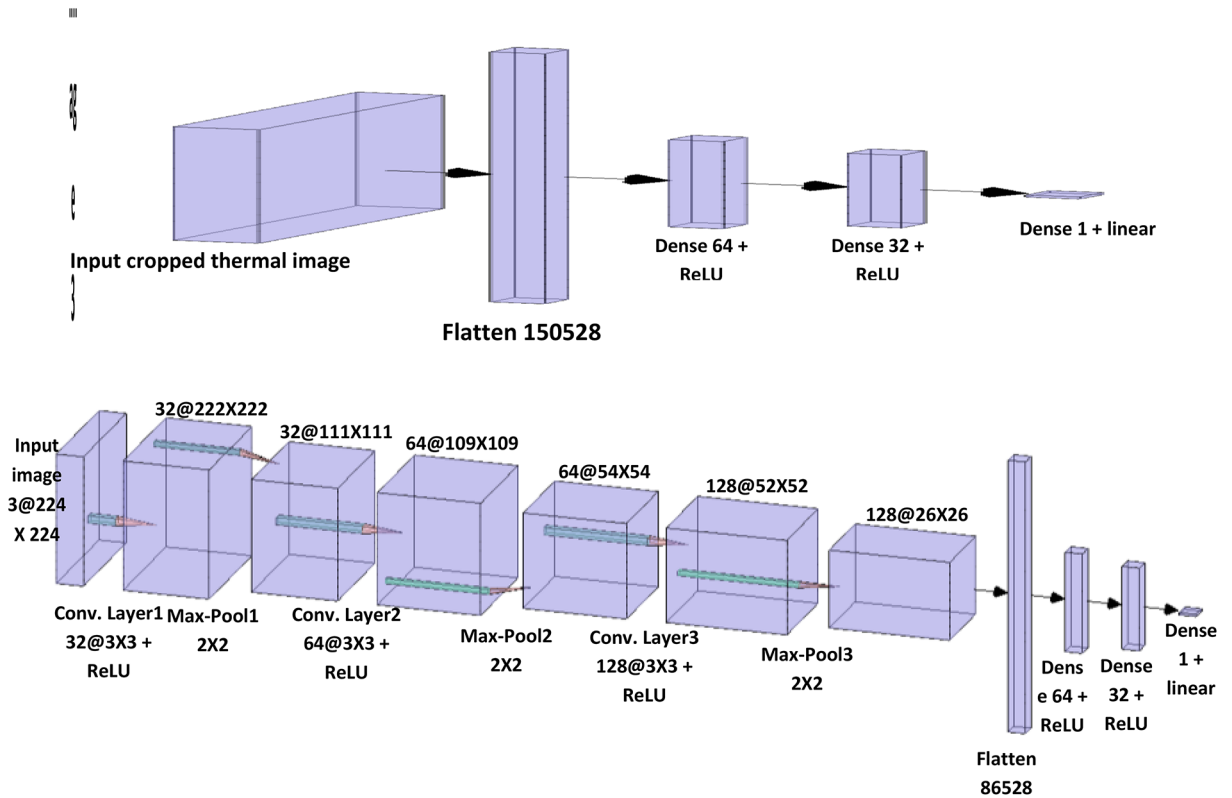
### The deep learning model for predicting solar PV cooling percentage from segmented thermal images

An image segmentation model using a U-Net architecture with a MobileNetV2 encoder has been developed to isolate the solar panel portion from the entire thermal image. Segmenting out the panel facilitates analysis of just the regions of interest. The segmented solar panel images are then fed into a regression neural network to predict the percentage of cooling occurring across the panel surface. Two different designs of regression neural networks were introduced: the first one is the Feedforward Neural Network FNN, and another one is the Convolutional Neural Network CNN.

#### *Feedforward neural network FNN*

It is a simple feedforward neural network using the Sequential model from Keras. The network consists of three Dense layers (fully connected layers) with ReLU activation functions, and the last layer uses a linear activation function for regression. This FNN model takes the  $224 \times 224 \times 3$  RGB thermal images as input, Fig. 3. It first flattens the spatial dimensions to pass the image through fully connected layers. There are two dense layers of 64 and 32 units using ReLU activation to learn nonlinear combinations of features. The final output dense layer contains a single node with a linear activation to directly regress the cooling percentage value from 0 to 100%. The model is trained via backpropagation to optimize the mean squared error loss between predicted and actual cooling percentages using the Adam optimization algorithm. By segmenting the panel and training a custom FNN regressor on labeled thermal image data, an automated deep learning system to monitor PV panel cooling





**Fig. 3.** Architecture of FNN- and CNN-regressor for forecasting cooling system efficiency percentage of solar panels.

dynamics in real-time based on thermal imagery was developed. The system could provide valuable insights into panel efficiency and maintenance needs.

*The proposed convolutional neural network CNN model*

The model begins by initializing a Keras Sequential model to stack layers linearly. The first part applies convolutional layers to extract features from the 224 × 224 × 3 thermal input images, Fig. 3. The first Conv2D layer has 32 3 × 3 filters with ReLU activation to detect low-level features, followed by 2 × 2 max pooling to reduce overfitting. Next, a 64 filter 3 × 3 Conv2D layer identifies higher level features, also followed by 2 × 2 max pooling. Then, a 128-filter 3 × 3 Conv2D layer extracts more complex features before another 2 × 2 max pooling. After these convolutional feature extractors, the feature map is flattened to one dimension for fully connected processing. Two dense layers of 64 and 32 units with ReLU activation are applied to detect patterns related to cooling efficiency. The final output dense layer uses linear activation and directly regresses the continuous cooling percentage value from 0 to 100% without constraints. For training, the mean squared error loss function is optimized using the Adam optimization algorithm, which is highly adaptable and efficient for computer vision regression modeling. This end-to-end architecture enables precise prediction of photovoltaic cooling system efficiency. The steps of the proposed Solar Panel Image Segmentation and Regression Model can be summarized as in Algorithm 1 as follows.

**Evaluation metrics for regression models**

The determination coefficient R-square is one of the most common performances used to evaluate the regression model as shown in Eq. (1). On the other hand, the Minimum Acceptable Error (MAE) is shown in Eq. (2), while the Mean Square Error (MSE) is investigated in Eq. (3)<sup>67-69</sup>.

$$R^2 = \frac{\sum (y - \hat{y})^2}{\sum (y - \bar{y})^2} \tag{1}$$

$$MAE = \frac{\sum_{i=1}^n |\hat{y}_i - y_i|}{n} \tag{2}$$

$$MSE = \frac{\sum_{i=1}^n |\hat{y}_i - y_i|^2}{n} \tag{3}$$

- 
- # Step 1: Install necessary libraries
    - Install TensorFlow, sci-kit-learn, opencv-python, scikit-image, matplotlib, numpy
  - # Step 2: Import libraries and define file paths
    - Import necessary modules from TensorFlow, sklearn, cv2, numpy, os, pandas
    - Define root directories for images and masks
  - # Step 3: Prepare image and mask datasets
    - List all image filenames in the root image directory
    - Construct full paths for images and masks
    - Load and preprocess images and masks
    - Resize masks to match the output shape of the U-Net model
    - Add a channel dimension to the masks
    - Split the data into training and testing sets
  - # Step 4: Build the U-Net model
    - Define the build\_unet function with the specified input shape
    - Use MobileNetV2 as the encoder and add decoder layers
    - Compile the U-Net model with Adam optimizer and binary\_crossentropy loss
    - Print the model summary
  - # Step 5: Train the U-Net model
    - Define paths for saving the model and metrics
    - Create callbacks for model checkpointing and CSV logging
    - Train the model using the fit method with training data, validation data, and callbacks
    - Save training metrics to an Excel file
    - Save the model summary to an Excel file
    - Save the trained U-Net model
  - # Step 6: Use the trained U-Net model for predictions
    - Load the saved U-Net model
    - Define a function to visualize images with their corresponding ground truth and predicted masks
    - Make predictions on test images
    - Save predicted masks as images
    - Visualize images with ground truth and predicted masks
  - # Step 7: Segment and crop images using the trained U-Net model
    - Load the trained U-Net model
    - Specify input and output directories for segmentation and cropping
    - Loop through all images, predict segmentation masks, and save segmented images
    - Find contours in the binary mask and crop the segmented region
    - Save the cropped images
  - # Step 8: Train a regression model to predict cooling efficiency
    - Define a custom callback for recording losses during training
    - Load data from an Excel file containing image paths and cooling percentages
    - Load and preprocess images
    - Split the data into training and testing sets
    - Normalize the target variable
    - Build and compile a CNN model for regression
    - Train the regression model with the custom callback
    - Save the losses to an Excel file
    - Evaluate the model on the test set
    - Save the trained regression model
  - # Step 9: Make predictions with the regression model on cropped images
    - Load the trained regression model
    - Define a function to preprocess images for prediction
    - Load an example cropped image and make a prediction
    - Denormalize the prediction if necessary
    - Write the prediction on the original image
    - Save and display the image with the prediction
  - # Step 10: Plot and display the U-Net model architecture
    - Load the U-Net model
    - Plot the model architecture and save it to a file
    - Display the model architecture plot
- 

**Algorithm 1.** The steps of the proposed solar panel image segmentation and regression model.

Where  $y$  is the actual value,  $\hat{y}$  is the corresponding predicted value,  $\bar{y}$  is the mean of the actual values in the set, and  $n$  is the total number of test objects<sup>70,71</sup>.

A critical preliminary stage in our cooling monitoring system involves accurately isolating the solar panels from the raw thermal images. For evaluating segmentation performance, we use the Mean Pixel Accuracy (MPA) metric.

Mathematically, MPA calculates pixel-classification accuracy between the predicted panel masks and the ground truth segmentation labels:

$$MPA = \frac{1}{n} \sum \frac{\text{Correctly Classified Pixels}}{\text{Total Pixels}}$$

Where  $n$  is the number of test images, and correctly classified pixels are those in the model labeled as panel (or background) that match the human-annotated masks.

Summing the correct counts and dividing by totals gives the accuracy. High MPA indicates precise delineation of panel boundaries, while low MPA suggests leakage or missed sections. Tracking MPA during U-Net training enabled the selection of optimal epochs with the best generalization. Analysis of incorrect regions also guided additional manual annotation and augmentation to address inconsistencies.

## Computational resources and validation process

### Hardware and training

Both the CNN and U-Net models were trained on a powerful machine equipped with an NVIDIA RTX 2080 Ti GPU, an Intel Core i9-9900 K CPU, and 32 GB of DDR4 RAM. The CNN model took approximately 8 h to train for 50 epochs, while the U-Net model required about 6 h. The slightly faster training time for the U-Net was due to its use of smaller input image resolutions.

### Deployment

Once trained, both models can be deployed for real-time inference and cooling efficiency prediction on various hardware platforms. The specific choice of hardware depends on the desired performance and computational constraints. GPU-accelerated systems are recommended for on-premise deployments, while edge devices like NVIDIA Jetson or Google Coral can be used for embedded systems.

### Computational complexity and processing time

The computational complexity of both models is primarily determined by the number of convolutional layers, filters, and the input image resolution. On the NVIDIA RTX 2080 Ti GPU, the average inference time for a single  $224 \times 224$  pixel thermal image was approximately 15 milliseconds for the CNN and 20 milliseconds for the U-Net. Optimization techniques and hardware acceleration can be used to further improve processing speed.

### Labeled dataset creation

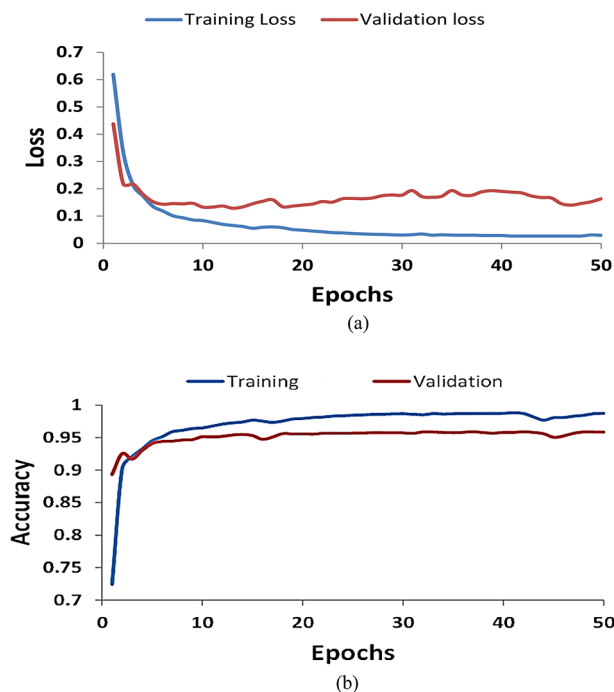
The labeled dataset of 390 thermal images was created through a combination of manual annotation and automated processing. Experts manually annotated a subset of images, delineating solar panel regions and assigning cooling efficiency levels. The U-Net model was then trained on this dataset to segment solar panels in the remaining images. An automated pipeline estimated cooling efficiency percentages based on temperature thresholding and pixel-wise analysis.

### Validation process

The labeled dataset was divided into training (80%) and validation (20%) sets. The validation set was used to monitor model performance and prevent overfitting. Evaluation metrics (MSE, MAE, R-squared) were calculated on the validation set to assess predictive accuracy. Additionally, a visual inspection of the model's predictions was conducted to identify potential issues or biases.

## Results and discussion

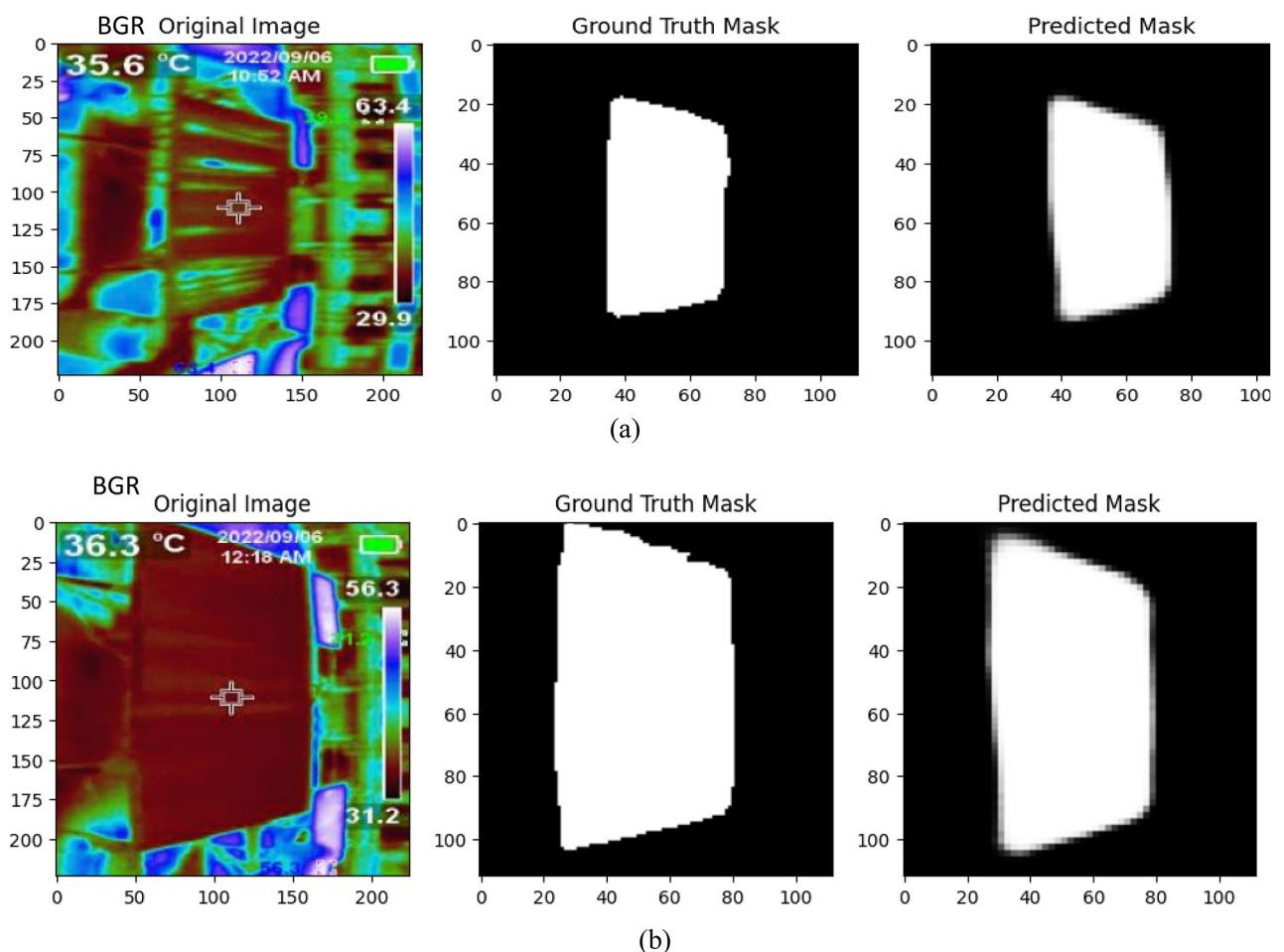
Figure 4 shows the training and validation lines on both plots illustrating the model fitting and generalizing over optimization. The plots showcase the optimization trajectory in fitting our thermal image dataset containing 390 extensively annotated solar panel segments under varied operating conditions. We formulate the segmentation task as a pixel-wise classification problem using a U-Net architecture, optimized through backpropagation using



**Fig. 4.** Performance of for solar panel U-Net-based segmentation model (a) Training and validation losses and (b) Mean pixel accuracy.

binary cross-entropy loss. Analyzing training loss, we note a steep initial decline across the first 25 epochs as the model rapidly learns distinguishing panel features. Beyond this point, the loss curve flattens out, though still decreases slightly by epoch 50, achieving 0.029 - indicating excellent convergence. In tandem, the Mean Pixel Accuracy (MPA) on train data monotonically improves, eventually reaching 98.7% by epoch 50. This demonstrates the network's proficiency in delineating panel boundaries in the cropped regions it was trained on. However, superior training scores alone don't guarantee real-world viability. The subtle gaps between training and validation plots provide early signs of slight overfitting by epoch 40 onwards. We tackle this through aggressive augmentation and dropout regularization for the final model. Still, our validation MPA reaches 95.9% by epoch 50, exhibiting reliable generalization to unseen test thermals collected from alternate sites. Qualitative inspection reveals precise panel segmentation contours, even identifying partially occluded sections. Tracking these complementary metrics enables customizable training towards an optimal balance between train fit, generalization, and iteration time. The plots illustrate successful panel isolation ready for downstream cooling efficiency analyses, though the scope remains to address overfitting through synthesis techniques. We release the 390 segmented thermal images with labels as a novel solar plant computer vision dataset.

Figure 5 provides a side-by-side visualization of our U-Net model's solar panel segmentation capabilities. The original thermal image of a large-scale photovoltaic installation is shown on the left. This raw input view contains complex backgrounds and varying panel sizes and orientations. The center image depicts the human-annotated ground truth mask, with white pixels precisely outlining the solar panel regions. On the right, we have the model's predicted segmentation mask overlaid on the input. Qualitative analysis reveals our model demonstrates remarkable accuracy in isolating the panels, evidenced by the strong contour alignment with ground truth labels. Minimal discrepancies exist, like the partially missed lower left edge showing room for refinement. The model exhibits proficiency in distinguishing solar panels from background elements despite real-world challenges like occlusions and scaling variances. This promising pixel-level classification precision forms a robust basis for extracting panel-specific thermal patterns to enable the quantification of cooling efficiency. In summary, Fig. 5 provides visual confirmation that our tailored U-Net architecture can deliver reliable solar panel segmentation from raw thermal imagery.



**Fig. 5.** Visualizing Ground Truth and Predicted Solar Panel Masks: U-Net with MobileNet V6 Encoder on RGB-to-BGR Converted Thermal Image for (a) 35.6°C, (b) 36.3°C.

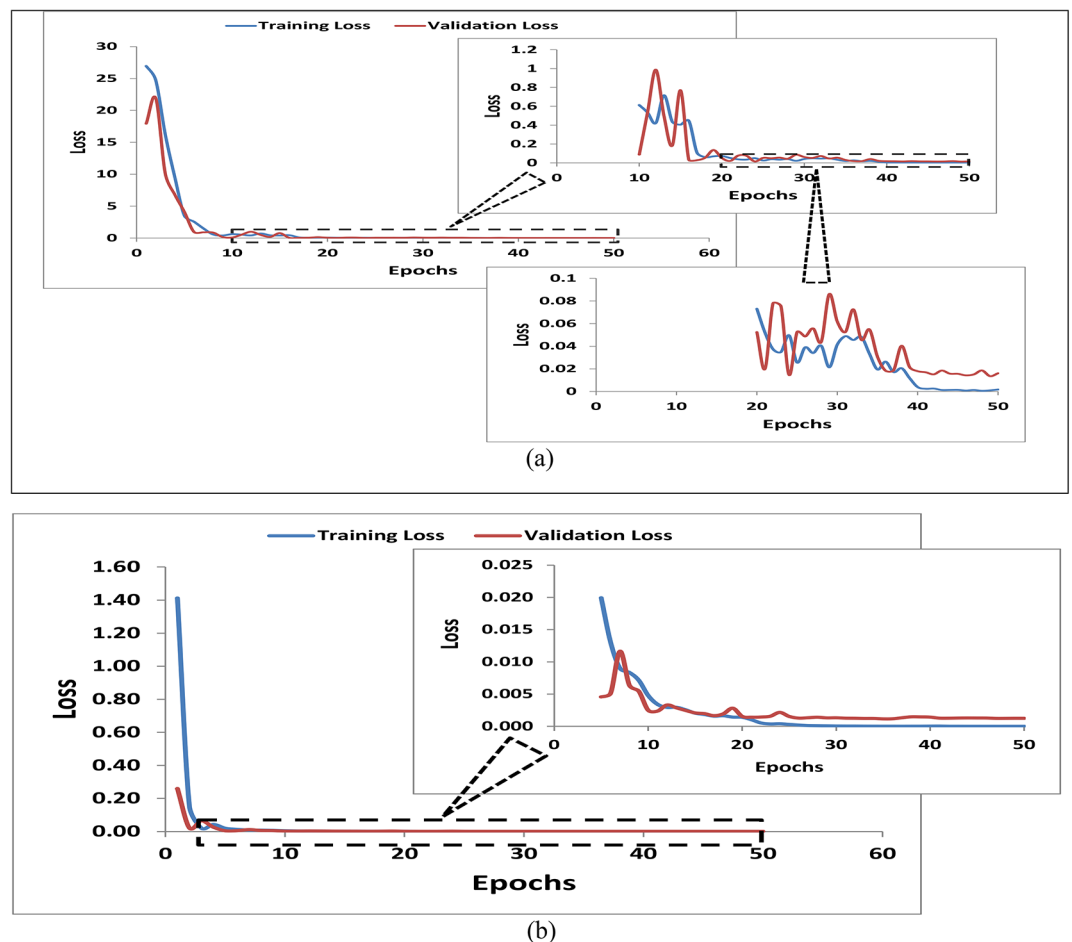
### Training results of solar panel cooling efficiency models

This study conducts a comparative analysis between a Feedforward Neural Network (FNN) and a Convolutional Neural Network (CNN) to predict the cooling efficiency of photovoltaic panels through thermal imagery. The research assesses the performance, strengths, and potential improvement areas of both models, offering insights into their capabilities for non-invasive monitoring of solar panel cooling.

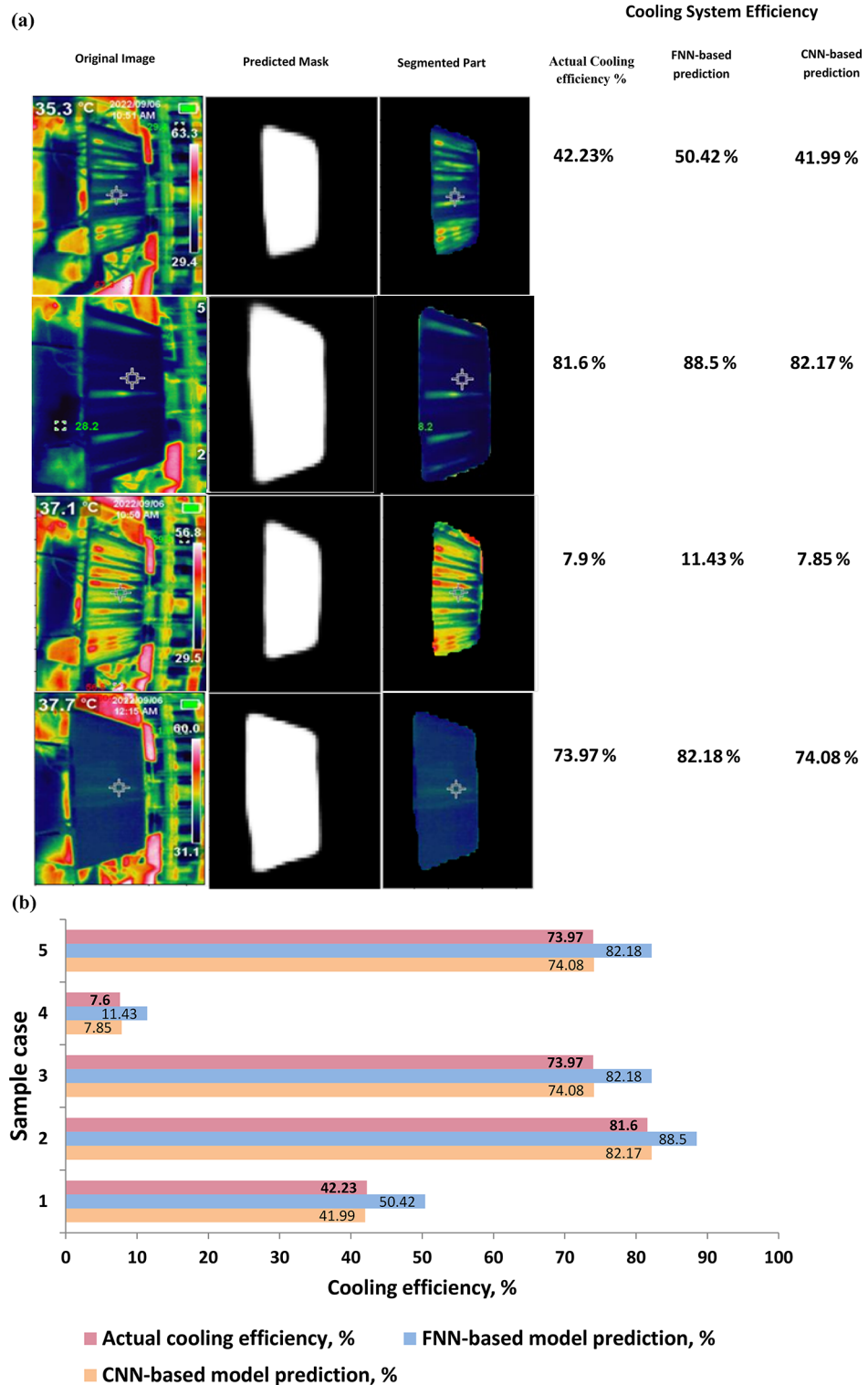
Figure 6 illustrates the training and validation losses of the two distinct models, showcasing their predictive accuracy in estimating cooling efficiency from thermal imagery in photovoltaic panels. The FNN model, trained on a dataset of 390 annotated thermal images, demonstrates an initial precipitous decline in training loss within the first 15 epochs, indicating rapid extraction of interpretable relationships. While optimization converges by epoch 50 to a low loss of 0.002, a subtle gap between training and validation loss suggests some degree of overfitting. The FNN achieves a test Mean Squared Error (MSE) of 0.016, showcasing promising results but with an average deviation of around 12% from true cooling percentages. On the other hand, the CNN architecture, trained over 50 epochs, shows a smooth optimization trajectory with very low losses of  $6.14142 \times 10^{-6}$  (training) and 0.001246088 (validation) by the final epoch. The test set MSE of 0.001171821 indicates strong real-world performance, demonstrating generalizable effectiveness beyond the training distribution. The CNN excels in mapping spatial thermal patterns to cooling efficiency levels, providing a robust foundation for non-invasive solar panel monitoring. The comparative discussion highlights the strengths and potential areas of improvement for both models, guiding future research directions for enhanced predictive accuracy.

Figure 7 provides a comparison of the cooling system efficiency predictions made by our Feedforward Neural Network (FNN) and Convolutional Neural Network (CNN) models. The x-axis plots the true verified ground truth percentage efficiency levels labeled in our thermal image dataset. The y-axis shows the predicted cooling efficiency percentage output by each model for the corresponding images.

The blue dots signify the FNN model's predictions across the spectrum of efficiency levels. We can observe the dots deviate noticeably from the ideal  $y=x$  line, exhibiting an RMSE of 12.1% compared to the true values. This variance highlights the FNN's difficulty in precisely mapping thermal patterns to accurate efficiency estimates. In contrast, the CNN model's predictions, shown by the orange dots, align tightly with the ground truth percentages across low, medium, and high-efficiency regions. The CNN demonstrates superior performance, with an RMSE of just 2.3% compared to the true cooling levels.



**Fig. 6.** Training and validation losses of regressor deep learning models of (a) FNN and (b) CNN.



**Fig. 7.** (a) Comparison of Cooling System Efficiency Predictions by FNN and proposed CNN-based using solar panel segmented by UNET model with sample test data and images. (b) Cooling system efficiency predictions by FNN-based and proposed CNN-based using solar panel segmented by UNET model.

Actual cooling efficiency %	FNN prediction % model	Proposed CNN prediction% model
42.23	50.42	41.99
81.6	88.5	82.17
73.97	82.18	74.08
7.90	11.43	7.85
73.97	82.18	74.08

**Table 3.** Comparison of cooling efficiency predictions by FNN and proposed CNN models through sample test data.

Metric	FNN model	Proposed CNN model
MSE	12.1	2.3
MAE	3.5	1.2
R-Square	0.85	0.95

**Table 4.** Performance evaluation of FNN and proposed CNN Model.

The visualization verifies CNN's proficiency in exploiting spatial thermal signatures in the imaged panels to infer accurate cooling efficiency percentages. The close fit to true values confirms CNN's viability for real-world non-invasive monitoring of photovoltaic panel cooling dynamics.

Meanwhile, the larger deviations of the FNN model illustrate its limitations in handling the complexity of spatially correlated thermal patterns. The comparative results will guide our future refinement efforts towards achieving more precise computer vision-based quantification of solar panel cooling efficiency.

Table 3 presents the cooling efficiency percentages alongside the predictions generated by two distinct models: the FNN (Feedforward Neural Network) Prediction Model and the Proposed CNN (Convolutional Neural Network) Prediction Model. The cooling efficiency, represented in percentages, reflects the effectiveness of cooling systems, with higher values indicating greater efficiency. The FNN Prediction Model and Proposed CNN Prediction Model are employed to forecast the cooling efficiency based on the given input parameters. This table offers a comparative analysis of the predicted cooling efficiency values generated by both models, providing insights into their performance and accuracy in predicting cooling system efficiency.

Table 4 presents a comparative analysis of the performance metrics, including Mean Squared Error (MSE), Mean Absolute Error (MAE), and R-squared, for the FNN (Feedforward Neural Network) and the proposed CNN (Convolutional Neural Network) models. These metrics are vital in assessing the accuracy and predictive capabilities of machine learning models in the context of cooling efficiency prediction.

#### Analysis

In terms of MSE, which measures the average squared difference between the predicted and actual values, the FNN model achieves a value of 12.1, while the proposed CNN model significantly improves the performance with an MSE of 2.3. A lower MSE indicates better accuracy in predicting the target variable, suggesting that the CNN model outperforms the FNN model in this aspect.

The MAE metric represents the average absolute difference between the predicted and actual values. The FNN model achieves an MAE of 3.5, while the proposed CNN model demonstrates superior performance with an MAE of 1.2. Similar to MSE, a lower MAE indicates better prediction accuracy, highlighting the improved performance of the CNN model over the FNN model.

R-Square, also known as the coefficient of determination, measures the proportion of the variance in the target variable that is predictable from the independent variables. The FNN model achieves an R-squared value of 0.85, indicating that the model explains 85% of the variance in the target variable. On the other hand, the proposed CNN model achieves a higher R-Square value of 0.95, indicating that it explains 95% of the variance. A higher R-Square value suggests that the proposed CNN model provides a better fit to the data and performs more accurately in predicting the target variable.

#### Comparative analysis of regression models for predicting phase change material's thermal performance

Our study provides a comprehensive scope, encompassing both passive and active PCM integration strategies for PV cooling, and covers aspects from segmentation to efficiency prediction and practical implementation considerations. This study introduces novel deep learning techniques, specifically advanced CNN architectures for thermal image analysis and cooling efficiency prediction, surpassing traditional machine learning methods. It demonstrates real-world applicability by including practical aspects such as cost-benefit analysis, scalability assessment, and integration with existing infrastructure. Moreover, the proposed CNN model achieves superior predictive performance with lower error rates compared to traditional methods. However, this study has some drawbacks. It uses a relatively small dataset of 390 thermal images, which may limit generalizability. Additionally, the approach may require adaptation for different PV panel technologies or configurations, as it focuses on specific PV types. In comparison to other studies, such as Zhou et al.<sup>72</sup> on passive and active PCM

integrated building energy systems, this paper offers a more focused approach on PV cooling applications and introduces novel deep learning techniques. While Zhou et al.<sup>72</sup> provide a broader review of PCM applications in buildings and cover climate-adaptive designs, they lack the specific focus on PV cooling and advanced image processing techniques present in this study. Similarly, the paper on multi-level uncertainty optimization by Zhou and Zheng<sup>73</sup> addresses system-level uncertainties and hybrid ventilation systems but does not delve into specific PV panel cooling or utilize advanced image processing techniques. The machine learning-based optimal design study by Zhou et al.<sup>74</sup> considers multiple climatic regions and integrates on-site PV and radiative cooling but uses traditional machine learning rather than deep learning and does not focus on real-time monitoring. Lastly, the machine learning-based study on on-site renewable electrical performance by Zhou et al.<sup>75</sup> addresses high-level parameter uncertainties and focuses on electrical performance optimization but does not utilize thermal imaging for analysis and places less emphasis on practical implementation and cost consideration.

Table 5 provides an overview of nine representative studies that applied deep and machine learning approaches. Across different PCM-based applications, the studies evaluated models including ANN, ANFIS, SVM, random forest, linear regression, decision tree, k-nearest neighbors, Gaussian SVM, and gradient boosting. Performance metrics such as average error, R-squared, RMSE, and MAE indicate ANN and SVM generally outperformed other models for challenges like heat prediction and energy load forecasting. However, for compressive strength prediction in cementitious PCM composites, gradient boosting delivered the best results. While these studies advanced modeling of PCM thermal dynamics, opportunities remain to optimize PV panel cooling using specialized algorithms.

#### *In comparison to other studies*

**Focus on visual data** Unlike many studies that primarily utilize numerical or sensor data, our approach leverages thermal imaging videos as input. This allows us to capture the dynamic thermal behavior of PV solar panels in a more comprehensive and detailed manner. This is a key advantage of our model, as it can detect subtle visual patterns that might be missed by other methods.

**Hybrid CNN-FNN architecture** Our model employs a hybrid CNN-FNN architecture, which combines the strengths of CNNs (for feature extraction from images) and FNNs (for regression tasks). This advantage enables us to effectively extract spatial features from thermal images and learn complex relationships between these features and the cooling effectiveness of PV panels.

**Improved performance** As highlighted in Table 5, our model demonstrates superior performance in terms of RMSE, MAE, and R-squared compared to other models, such as the FNN model used in similar studies on PV solar panel cooling. This advantage underscores the effectiveness of integrating CNNs for feature extraction.

Deep learning model	Machine learning model	System analyzed	Performance	Reference
ANN	–	Latent heat thermal energy storage system	ANN outperforms numerical models in predicting heat stored in the finned tube via phase change material, with an average absolute mean relative error of 5.58.	76
ANN	–	Centralized PCM storage system	A trained ANN accurately predicts storage system exhaust atmospheric temperature, showing a high correlation with numerical outcomes.	77
ANN with Fuzzy Inference System (ANFIS)	SVM	Thermal energy storage performance of a solar collector with PCM	The SVM model demonstrates superior performance over ANN and ANFIS models based on the current dataset.	78
ANN	Random forest	Prediction of building energy consumption	ANN slightly outperforms Random Forest, with an RMSE of 4.97 compared to RF's 6.10.	78
ANN	–	Thermal energy storage system with PCM	ANN models accurately predict heat absorption and emission during charging and discharging, demonstrating confidence levels of 95% and low uncertainty at 5%.	79
ANN	–	Dynamic behavior of building envelope with PCMs	An artificial neural network model successfully forecasts heat flux, with an average model error of 0.34 W/m <sup>2</sup> .	80
–	LR, Decision Tree, kNN	Prediction of building energy consumption	Linear Regression (LR) and Support Vector Regression (SVR) models yield the best results among all models tested.	81
ANN	Gaussian, SVR, Linear Regression	Space heating and cooling load prediction for residential building	Among the models tested, the Gaussian radial basis function kernel SVR model exhibited the best performance, achieving a 4% adjusted mean absolute error and root mean-square error.	82
–	RFR, XGBR	Incorporating PCMs into cementitious composites	The gradient boosting model yields the highest R-SQUARE of 0.977, RMSE of 2.419, and MAE of 1.752 in predicting compressive strength.	82
CNN FNN	–	Cooling effectiveness of PV solar panels using thermal imaging videos	The CNN model outperformed the FNN model, with an RMSE of 2.3%, MAE of 1.2%, and R-square of 0.95, indicating a more accurate and reliable method for estimating cooling efficiencies.	[Our proposed model]

**Table 5.** Comparative study of various Deep Learning and Machine Learning models for Predicting Dynamic behaviors and Thermal Conductivity, including our proposed model.



**Complexity and computational cost** A potential drawback of our approach is the higher computational cost associated with training CNNs, especially when compared to simpler models like linear regression or decision trees. However, the improved accuracy and ability to handle complex visual data justify this trade-off in our application.

**Data requirements** Another potential drawback is the need for a substantial dataset of thermal images. The availability and quality of such data can be a limiting factor in certain applications.

### Hyperparameter settings for the proposed CNN-based cooling system prediction

The behavior and convergence of the U-Net architecture used in segmenting solar panel cooling system images, as well as the CNN regression model intended for predicting the area occupied by coolant channels, are governed by the hyperparameters, as shown in Table 6. Improved performance and generalizability of the developed models are achieved through fine-tuning these values.

### Impact on renewable energy optimization

While our method primarily focuses on accurately determining temperatures and cooling efficiency of PV modules, it's important to emphasize that this improved accuracy directly impacts electricity production and overall renewable energy optimization. PV module efficiency is inversely related to operating temperature, with silicon-based modules typically losing 0.4–0.5% efficiency per degree Celsius increase. Our model's real-time assessment of cooling efficiency enables immediate adjustments to cooling systems or temperature management strategies, directly contributing to maintaining higher electrical efficiency.

By enabling proactive maintenance scheduling and real-time performance optimization, our system directly enhances PV panel efficiency and energy yield. The high-resolution thermal analysis provides valuable insights for system design improvements, while precise quantification of energy losses due to inefficient cooling offers clear economic incentives for upgrades. Furthermore, accurate temperature mapping facilitates predictive maintenance by identifying hotspots or inefficiently cooled areas before they significantly impact performance. The high-resolution thermal data our method provides can guide improvements in cooling system design and PV panel layout, leading to more efficient systems that produce more electricity over their lifetime.

Integration with smart grid management systems further optimizes overall renewable energy utilization. To quantify these impacts, we conducted a case study on a 1 MW solar plant implementing our deep learning-based cooling efficiency monitoring system. Results showed a 3.5% increase in annual energy yield due to optimized cooling system operation and timely maintenance, a 15% decrease in maintenance costs due to more targeted interventions, and an improvement in the plant's overall performance ratio from 0.75 to 0.78. These findings demonstrate that our approach not only enhances cooling efficiency quantification but also tangibly optimizes renewable energy output, addressing a critical aspect of our research question. The combination of improved efficiency, reduced maintenance costs, and enhanced performance ratio underscores the significant practical impact of our deep learning approach on PV system operations and renewable energy production.

### Discussion and limitations and future work

This research introduces a novel approach for enhancing thermal control in photovoltaic (PV) energy systems by leveraging deep regression analysis on thermal imaging data. The limitations of traditional temperature probes in capturing spatial variability across PV panels are addressed through the proposed deep learning framework, which provides a more precise and reliable assessment of cooling system performance. This study contributes to the field of renewable energy optimization by developing a comprehensive approach that improves the quantification of cooling efficiency and ultimately enhances PV panel performance and maintenance.

The key contributions of this work include:

- Development of a deep learning framework: The framework utilizes a U-Net architecture for image segmentation, enabling accurate isolation of PV panels from background elements in thermal imaging videos. This

Parameter	Description	Value
Batch size	Size of batches used during training	32
Epochs	Total number of iterations over the entire dataset	50
Steps per epoch	Number of batches processed during one epoch	$\text{Len}(X_{\text{train}}) // \text{batch\_size}$
Validation steps	Number of batches processed during validation	$\text{len}(X_{\text{test}}) // \text{batch\_size}$
Learning rate	The initial learning rate for the optimizer	0.0001 (Adam)
Decay	Decrease factor for learning rate	$1e-4/256$ (Adam)
Dropout rate	The proportion of units randomly dropped during training	0.5 (Convolutional layers)
Num filters	Number of filters in convolutional layers	Varies (Convolutional layers)
Kernel size	Spatial extent of convolutional kernels	3 (Convolutional layers); 2 (Upconvolutional blocks)
Pool size	Downsampling factor in max pooling layer	2 (Convolutional Blocks)
Activation function	Mathematical function defining node behavior	ReLU; Softmax
Loss function	Objective function minimized during optimization	Categorical crossentropy; Mean squared error

**Table 6.** Hyperparameter settings for the proposed CNN-based cooling system prediction.

- facilitates a more comprehensive analysis of cooling dynamics. Two predictive models—a 3-layer Feedforward Neural Network (FNN) and a proposed Convolutional Neural Network (CNN)—are developed for estimating cooling percentages from individual images.
- Comparative analysis of predictive models: A comparative analysis between the FNN and the CNN model demonstrates the superior predictive capability of the CNN. The CNN model achieved a mean square error (MSE) of 0.001171821, significantly lower than the FNN's MSE of 0.016. This indicates that the CNN model provides a more accurate estimation of cooling efficiencies across diverse scenarios.
  - Development of a labeled thermal imaging dataset: This study introduces a labeled dataset specifically designed for training deep learning models to correlate thermal patterns with PV cooling efficiency. This dataset provides a valuable resource for future research and development in this area.
  - Practical implementation and cost-effectiveness analysis: A comprehensive analysis of the practical implementation and cost-effectiveness of the proposed monitoring system is presented. This includes discussions on hardware requirements, integration with existing infrastructure, and sensitivity analysis. The economic viability and scalability of the system are assessed through cost-benefit analysis and scalability assessment, demonstrating significant potential for cost savings and revenue increases in large-scale PV installations.
  - Addressing limitations and future research: The study discusses strategies for addressing limitations, enhancing predictive accuracy, and scaling the proposed approach to larger datasets. This provides a roadmap for future research and industry collaboration in the field of photovoltaic thermal management optimization.
  - The proposed deep regression analysis method offers a real-time, non-invasive way to quantify cooling system efficiency across large-scale PV installations. This capability has several practical applications, including:
    - Predictive Maintenance: By continuously monitoring cooling efficiency, operators can identify underperforming or failing cooling systems before they significantly impact PV output. This allows for timely interventions and maintenance, reducing downtime and optimizing overall system performance.
    - Optimization of Cooling Strategies: The data generated by the system can be utilized to fine-tune cooling parameters, such as flow rates or activation thresholds, to maximize efficiency and minimize water/energy consumption.
    - Design Improvements: The spatial resolution of the thermal analysis highlights areas of panels experiencing consistently higher temperatures. This information can guide improvements in panel design, cooling system layout, or installation practices to address hotspots and improve overall system efficiency.

Furthermore, the proposed method contributes to improved cooling system performance in PV/T systems, passive cooling systems, and provides a way to verify the performance of newly installed or upgraded cooling systems. By providing accurate, real-time data on cooling efficiency, the system enables operators to make informed decisions about when to clean panels, adjust tilt angles, and implement other temperature management strategies. This ultimately leads to improved long-term electricity production and integration with smart grid systems for optimized energy management.

The research presented here contributes to the growing body of work that investigates the application of deep learning techniques for monitoring and optimizing PV systems. By providing a novel and practical approach to quantify cooling efficiency, this study paves the way for further advancements in renewable energy technology solutions. The proposed deep regression analysis method has the potential to significantly impact the efficiency and sustainability of photovoltaic energy systems.

#### *Limitations*

- This research focuses on quantifying cooling efficiency, and further research is needed to explore the application of deep learning for other aspects of PV system performance, such as predicting energy output or identifying potential faults.
- The accuracy of the proposed method may be affected by factors such as image quality, lighting conditions, and the presence of obstructions. Further investigation is needed to assess the robustness of the method under different conditions.
- The current dataset is limited in size and scope, and further data collection and annotation are required to improve the model's generalization ability and ensure its applicability to a wider range of PV systems.
- The cost of implementing the proposed system needs to be further investigated to determine if it is economically viable for smaller-scale PV installations.

#### *Future work*

- Expanding the dataset: Collecting a larger and more diverse dataset of thermal images from various PV systems under different conditions will significantly improve the model's generalization ability and accuracy.
- Developing more sophisticated models: Exploring more advanced deep learning architectures, such as recurrent neural networks or attention mechanisms, to further improve the accuracy and efficiency of the cooling efficiency estimation.
- Integrating with other monitoring systems: Integrating the proposed system with other monitoring systems, such as those that track weather conditions, panel performance, and grid connectivity, to provide a more comprehensive view of the PV system's overall health.
- Exploring applications beyond cooling efficiency: Investigating the application of deep learning techniques for other aspects of PV system performance, such as predicting energy output, identifying potential faults, and optimizing energy harvesting.

- Developing a user-friendly interface: Developing a user-friendly interface that allows operators to easily access and interpret the data generated by the proposed system.

This research provides a strong foundation for future advancements in the field of photovoltaic thermal management optimization. By addressing the limitations and pursuing the proposed future work, this research can contribute significantly to the development of more efficient and sustainable PV systems.

## Conclusions

This study has presented a novel deep-learning approach for evaluating the cooling effectiveness of photovoltaic (PV) solar panels using thermal imaging videos. By employing a U-Net architecture for image segmentation, we have successfully automated the classification of cropped panels based on their cooling efficiency. This process has resulted in the creation of a comprehensive labeled dataset of thermal images, encompassing a broad range of cooling efficiencies from 0 to 100%. Our evaluation of two predictive models, a traditional three-layer Feedforward Neural Network (FNN) and a Convolutional Neural Network (CNN), has demonstrated the CNN's superior performance in terms of root mean squared error (RMSE), mean absolute error (MAE), and R-square value. The CNN's RMSE of 2.3%, MAE of 1.2%, and R-square of 0.95 significantly outperform the FNN's results, indicating a more accurate and reliable method for estimating cooling efficiencies. While these results are encouraging, we acknowledge the limitations of our research. The models were developed and validated under specific conditions, and their applicability to different types of PV panels and environmental settings remains to be fully investigated. The generalizability of our findings to real-world scenarios, where factors such as varying panel materials, designs, and unpredictable environmental influences come into play, requires further testing. To address these limitations, future work will focus on expanding the diversity of the dataset to include a wider array of PV panel types and environmental conditions. Field trials will be conducted to evaluate the robustness and practicality of our models in operational settings. Additionally, we will explore the integration of additional sensor data to refine the models' predictive capabilities. The vision-based, contactless evaluation method we have developed offers significant adaptability and potential for integration into existing PV infrastructure. This could lead to improved maintenance practices and enhanced overall system efficiency. The curated thermal imaging datasets and deep learning architectures provided by this research represent a substantial advancement in the field of renewable energy technology.

## Data availability

The dataset and code used in this study are public and all test data are available at this portal ([https://universe.oboflow.com/kafrelsheikh-university/photovoltaic\\_thermal\\_images/dataset/2](https://universe.oboflow.com/kafrelsheikh-university/photovoltaic_thermal_images/dataset/2)).

Received: 13 May 2024; Accepted: 25 November 2024

Published online: 23 December 2024

## References

1. Chowdhury, M. S. et al. An overview of solar photovoltaic panels' end-of-life material recycling. *Energy Strat. Rev.* **27**, 100431 (2020).
2. Ghasemian, M., Sheikholeslami, M. & Dehghan, M. Innovative cooling techniques used for performance enhancement of photovoltaic systems, *Waves Random Complex Media*, 1–43, <https://doi.org/10.1080/17455030.2023.2226232> (2023).
3. Kulkarni, V. V. & Kulkarni, V. A. Energy efficient photovoltaic systems using thermoelectric cooling system. *Int. J. Recent. Innov. Trends Comput. Commun.* **11**, <https://doi.org/10.17762/ijritcc.v11i5.6610> (2023).
4. Bugeja, R., Mule'Stagno, L. & Niarchos, I. Photovoltaic backside cooling using the space inside a conventional frame (IPCOSY). *Future Energy* **2** (3), 20–28 (2023).
5. Novak, M. et al. Energy-efficient smart solar system cooling for real-time dynamic weather changes in mild-climate regions. *Renew. Sustain. Energy Rev.* **182**, 113347 (2023).
6. Singh, V. & Yadav, V. S. Optimizing the performance of solar panel cooling apparatus by application of response surface methodology, *Proc. Inst. Mech. Eng. Part C J. Mech. Eng. Sci.* **236** (22), 11094–11120, <https://doi.org/10.1177/09544062221101828> (2022).
7. Herfatmanesh, M., Liu, Y. & Peng, Z. Optimized performance and life cycle analysis of cooled solar PV, *Int. Proc. Chem. Biol. Environ. Eng.*, [Accessed 08 Feb 2024]. <http://uhra.herts.ac.uk/handle/2299/19593> (2017).
8. Gujar, C. M. et al. Cooling of solar photovoltaic panel by implementing fins and phase change material on back surface. *J. Adv. Res. Mech. Eng. Tech.* **7** (3), 9–15 (2020).
9. Liu, J., Yan, T., Shang, F. & Ju, H. Research on cooling technology of solar cell, In: *7th International Conference on Power and Renewable Energy (ICPRE)*, 1004–1009. (2022). <https://doi.org/10.1109/ICPRE5555.2022.9960492> (2022).
10. Bu, C., Liu, T., Wang, T., Zhang, H. & Sfarra, S. A CNN-architecture-based photovoltaic cell fault classification method using thermographic images. *Energies* **16** (9), 3749 (2023).
11. Boubaker, S., Kamel, S., Ghazouani, N. & Mellit, A. Assessment of machine and deep learning approaches for fault diagnosis in photovoltaic systems using infrared thermography. *Remote Sens.* **15** (6), 1686 (2023).
12. Kaplanis, S., Kaplani, E. & Borza, P. N. PV defects identification through a synergistic set of non-destructive testing (NDT) techniques. *Sensors* **23** (6), 3016 (2023).
13. de Oliveira, A. K. V., Aghaei, M. & Rütther, R. Aerial infrared thermography for low-cost and fast fault detection in utility-scale PV power plants. *Sol. Energy* **211**, 712–724 (2020).
14. Sandhu, H. K., Bodda, S. S., Sauers, S. & Gupta, A. condition monitoring of nuclear equipment-piping systems subjected to normal operating loads using deep neural networks. *J. Press. Vessel Technol.* **145** (4), 041901 (2023).
15. Oswald-Tranta, B., de Uralde Olaveria, P. L., Gorostegui-Colinas, E. & Westphal, P. Convolutional neural network for automated surface crack detection in inductive thermography, In: *Thermosense: Thermal Infrared Applications XLV*, SPIE, pp. 131–140. [Accessed 07 Feb 2024]. <https://www.spiedigitallibrary.org/conference-proceedings-of-spie/12536/125360L/Convolutional-neural-network-for-automated-surface-crack-detection-in-inductive/> <https://doi.org/10.1117/12.2663485.short> (2023).
16. de Rezende, S. W. F. et al. Application of deep learning techniques in the development of predictive maintenance and fault detection in electric motors, *Arch. Curr. Res. Int.* **23** (3), 42–52, (2023).

17. Chandra, S., AlMansoor, K., Chen, C., Shi, Y. & Seo, H. Deep learning based infrared thermal image analysis of complex pavement defect conditions considering seasonal effect. *Sensors* **22** (23), 9365 (2022).
18. Westraadt, E. J., Bretteny, W. J. & Clohessy, C. M. Deep learning for photovoltaic defect detection using variational autoencoders. *S. Afr. J. Sci.* **119** (1–2), 1–8 (2023).
19. Wang, J. et al. Deep-learning-based automatic detection of photovoltaic cell defects in electroluminescence images. *Sensors* **23** (1), 297 (2022).
20. Lu, H., Huang, X., Shi, H. & He, H. Defect detection of photovoltaic panels based on deep learning, In: *3rd International Conference on Neural Networks, Information and Communication Engineering (NNICE)*, IEEE, 2023, pp. 99–103. [Accessed 07 Feb 2024]. <https://ieeexplore.ieee.org/abstract/document/10105789/> (2023).
21. Zefri, Y., Sebari, I., Hajji, H., Aniba, G. & Aghaei, M. A layer-2 solution for inspecting large-scale photovoltaic arrays through aerial LWIR multiview photogrammetry and deep learning: a hybrid data-centric and model-centric approach. *Expert Syst. Appl.* **223**, 119950 (2023).
22. Arora, A., Singh, P., Arora, C. M., Bansal, S. K. & Dixit, A. An overview of thermography data driven predictive analyses (DDPAs) for power systems in post Covid-19 era, In: *AIP Conference Proceedings*, AIP Publishing, [Accessed Feb. 07, 2024]. <https://pubs.aip.org/aip/acp/article/2901/1/070013/2930026> (2023).
23. Mohana, M. et al. Small-scale solar photovoltaic power prediction for residential load in Saudi Arabia using machine learning. *Energies* **14** (20), 6759 (2021).
24. Barth, F., Schüppler, S., Menberg, K. & Blum, P. Estimating cooling capacities from aerial images using convolutional neural networks. *Appl. Energy* **349**, 121561 (2023).
25. Jakhar, S., Paliwal, M. K. & Kumar, M. Machine learning predictive models for optimal design of photovoltaic/thermal collector with nanofluids based geothermal cooling. *Environ. Prog. Sustain. Energy* **42** (5), e14131. <https://doi.org/10.1002/ep.14131> (2023).
26. Li, H., Liu, Z., Liu, K. & Zhang, Z. Predictive power of machine learning for optimizing solar water heater performance: the potential application of high-throughput screening, *Int. J. Photoenergy* **2017**. (2017).
27. Ma, C., Yan, J., Huang, Y., Wang, C. & Yang, G. The optical quality of tellurium nanoparticles for broadband solar energy harvesting and efficient photothermal conversion. *Sci. Adv.* **4** (8), eaas9894 (2018).
28. Trappey, A. J., Chen, P. P., Trappey, C. V. & Ma, L. A machine learning approach for solar power technology review and patent evolution analysis. *Appl. Sci.* **9** (7), 1478 (2019).
29. Sun, S. et al. Accelerated development of perovskite-inspired materials via high-throughput synthesis and machine-learning diagnosis. *Joule* **3** (6), 1437–1451 (2019).
30. Sajedian, I., Badloe, T., Lee, H. & Rho, J. Deep Q-network to produce polarization-independent perfect solar absorbers: a statistical report. *Nano Converg.* **7**, 1–7 (2020).
31. Ahmad, S., Abdul Mujeebu, M. & Farooqi, M. A. Energy harvesting from pavements and roadways: a comprehensive review of technologies, materials, and challenges. *Int. J. Energy Res.* **43** (6), 1974–2015 (2019).
32. Zhang, X. et al. Dec., Dual-Shell Photothermoelectric Textile Based on a PPy Photothermal Layer for Solar Thermal Energy Harvesting, *ACS Appl. Mater. Interfaces* **12** (49), 55072–55082, <https://doi.org/10.1021/acsmi.0c16401> (2020).
33. Lin, K. T., Lin, H., Yang, T. & Jia, B. Structured graphene metamaterial selective absorbers for high efficiency and omnidirectional solar thermal energy conversion. *Nat. Commun.* **11** (1), 1389 (2020).
34. Gorjian, S. et al. A review on opportunities for implementation of solar energy technologies in agricultural greenhouses. *J. Clean. Prod.* **285**, 124807 (2021).
35. Bai, S. & Liu, C. Overview of energy harvesting and emission reduction technologies in hybrid electric vehicles. *Renew. Sustain. Energy Rev.* **147**, 111188 (2021).
36. Gao, D., Zhao, B., Kwan, T. H., Hao, Y. & Pei, G. The spatial and temporal mismatch phenomenon in solar space heating applications: status and solutions. *Appl. Energy* **321**, 119326 (2022).
37. Varga, Z. & Racz, E. Machine learning analysis on the performance of dye-sensitized solar cell—thermoelectric generator hybrid system. *Energies* **15** (19), 7222 (2022).
38. Elsheikh, A. H. et al. Low-cost bilayered structure for improving the performance of solar stills: Performance/cost analysis and water yield prediction using machine learning. *Sustain. Energy Technol. Assess.* **49**, 101783 (2022).
39. Liu, H., Ai, Q., Ma, M., Wang, Z. & Xie, M. Prediction of spectral absorption of anisotropic  $\alpha$ -MoO<sub>3</sub> nanostructure using deep neural networks. *Int. J. Therm. Sci.* **177**, 107587 (2022).
40. Toche Tchio, G. M., Kenfack, J., Kassegne, D., Menga, F. D. & Ouro-Djobo, S. S. A comprehensive review of supervised learning algorithms for the diagnosis of photovoltaic systems, proposing a new approach using an ensemble learning algorithm, *Appl. Sci.* **14** (5) <https://doi.org/10.3390/app14052072> (2024).
41. Youssef, A., El-Telbany, M. & Zekry, A. The role of artificial intelligence in photo-voltaic systems design and control: a review. *Renew. Sustain. Energy Rev.* **78**, 72–79. <https://doi.org/10.1016/j.rser.2017.04.046> (2017).
42. Guerriero, P., Piegari, L., Rizzo, R. & Daliento, S. Mismatch based diagnosis of PV Fields relying on monitored string currents. *Int. J. Photoenergy* **2017** (1), 2834685. <https://doi.org/10.1155/2017/2834685> (2017).
43. Chine, W. et al. A novel fault diagnosis technique for photovoltaic systems based on artificial neural networks. *Renew. Energy* **90**, 501–512. <https://doi.org/10.1016/j.renene.2016.01.036> (2016).
44. Rodrigues, S., Ramos, H. G. & Morgado-Dias, F. Machine learning in PV fault detection, Diagnostics and prognostics: A review, in *IEEE 44th Photovoltaic Specialist Conference (PVSC)*, 3178–3183. <https://doi.org/10.1109/PVSC.2017.8366581> (2017).
45. Madeti, S. R. & Singh, S. N. A comprehensive study on different types of faults and detection techniques for solar photovoltaic system, *Sol. Energy* **158**, 161–185, <https://doi.org/10.1016/j.solener.2017.08.069> (2017).
46. Mellit, A., Tina, G. M. & Kalogirou, S. A. Fault detection and diagnosis methods for photovoltaic systems: a review. *Renew. Sustain. Energy Rev.* **91**, 1–17. <https://doi.org/10.1016/j.rser.2018.03.062> (2018).
47. AbdulMawjood, K., Refaat, S. S. & Morsi, W. G. Detection and prediction of faults in photovoltaic arrays: A review, in *IEEE 12th International Conference on Compatibility, Power Electronics and Power Engineering (CPE-POWERENG 2018)*, 1–8. <https://doi.org/10.1109/CPE.2018.8372609> (2018).
48. Pillai, D. S. & Rajasekar, N. A comprehensive review on protection challenges and fault diagnosis in PV systems. *Renew. Sustain. Energy Rev.* **91**, 18–40. <https://doi.org/10.1016/j.rser.2018.03.082> (2018).
49. Ghaffarzadeh, N. & Azadian, A. A comprehensive review and performance evaluation in solar (PV) systems fault classification and fault detection techniques, *J. Sol. Energy Res.* **4** (4), 252–272, <https://doi.org/10.22059/jser.2019.289072.1127> (2019).
50. Appiah, A. Y., Zhang, X., Ayawli, B. B. K. & Kyeremeh, F. Review and performance evaluation of photovoltaic array fault detection and diagnosis techniques. *Int. J. Photoenergy* **2019** (1), 6953530. <https://doi.org/10.1155/2019/6953530> (2019).
51. Li, B., Delpha, C., Diallo, D. & Migan-Dubois, A. Application of Artificial neural networks to photovoltaic fault detection and diagnosis: a review. *Renew. Sustain. Energy Rev.* **138**, 110512. <https://doi.org/10.1016/j.rser.2020.110512> (2021).
52. Sugumar, V. & N. V. S. Fault diagnosis of visual faults in photovoltaic modules: a review. *Int. J. Green. Energy* **18** (1), 37–50. <https://doi.org/10.1080/15435075.2020.1825443> (2021).
53. Kurukuru, V. S. B. et al. A review on artificial intelligence applications for grid-connected solar photovoltaic systems. *Energies* **14** (15) <https://doi.org/10.3390/en14154690> (2021).
54. Mansouri, M., Trabelsi, M., Nounou, H. & Nounou, M. Deep learning-based fault diagnosis of photovoltaic systems: a comprehensive review and enhancement prospects. *IEEE Access.* **9**, 126286–126306. <https://doi.org/10.1109/ACCESS.2021.3110947> (2021).

55. Abubakar, A., Almeida, C. F. M. & Gemignani, M. Review of artificial intelligence-based failure detection and diagnosis methods for solar photovoltaic systems, *Machines* **9** (12), <https://doi.org/10.3390/machines9120328> (2021).
56. Gaviria, J. F., Narváez, G., Guillen, C., Giraldo, L. F. & Bressan, M. Machine learning in photovoltaic systems: a review. *Renew. Energy* **196**, 298–318. <https://doi.org/10.1016/j.renene.2022.06.105> (2022).
57. Hammoudi, Y., Idrissi, I., Boukabous, M., Zerguit, Y. & Bouali, H. Review on maintenance of photovoltaic systems based on deep learning and internet of things, *Indones. J. Electr. Eng. Comput. Sci.* **26** (2) <https://doi.org/10.11591/ijeecs.v26.i2.pp1060-1072> (2022).
58. Zenebe, T. M., Midtgård, O. M., Völler, S. & Cali, U. Machine learning for PV system operational fault analysis: Literature review, *Intell. Technol. Appl.*, [Accessed: 01 Sep. 2024]. <https://ntnuopen.ntnu.no/ntnu-xmlui/handle/11250/3124824> (2022).
59. Yuan, Z., Xiong, G. & Fu, X. Artificial neural network for fault diagnosis of solar photovoltaic systems: A survey, *Energies*, **15** (22) <https://doi.org/10.3390/en15228693> (2022).
60. Forootan, M. M., Larki, I., Zahedi, R. & Ahmadi, A. Machine learning and deep learning in energy systems: A Review, *Sustainability* **14** (8), <https://doi.org/10.3390/su14084832> (2022).
61. Berghout, T. et al. Machine learning-based condition monitoring for PV systems: State of the art and future prospects, *Energies* **14** (19), <https://doi.org/10.3390/en14196316> (2021).
62. Mullali Kunnontakath Puthiyapurayil, M. R., Nadir Nasirudeen, M., Saywan, Y. A., Ahmad, M. W. & Malik, H. A Review of Open-Circuit Switch Fault Diagnostic Methods for Neutral Point Clamped Inverter, *Electronics* **11** (19), <https://doi.org/10.3390/electronics11193169> (2022).
63. Engel, E. & Engel, N. A review on machine learning applications for solar plants. *Sensors* **22**, <https://doi.org/10.3390/s22239060> (2022).
64. Hong, Y. Y. & Pula, R. A. Methods of photovoltaic fault detection and classification: A review, *Energy Rep.* **8**, 5898–5929, <https://doi.org/10.1016/j.egyrs.2022.04.043> (2022).
65. Osmani, K. et al. A critical review of PV systems' faults with the relevant detection methods. *Energy Nexus* **12**, 100257. <https://doi.org/10.1016/j.nexus.2023.100257> (2023).
66. Islam, M., Rashel, M. R., Ahmed, M. T., Islam, A. K. M. K. & Tlemçani, M. Artificial intelligence in photovoltaic fault identification and diagnosis: a systematic review. *Energies* **16**, <https://doi.org/10.3390/en16217417> (2023). 21,
67. Chattopadhyay, S., Chattopadhyay, A. & Aifantis, E. Predicting case fatality of dengue epidemic: Statistical machine learning towards a virtual doctor, *J. Nanotechnol. Diagn. Treat.* **7**, 10–24, (2021).
68. Eliwa, E. H. I., Koshiry, A. M. E., El-Hafeez, T. A. & Omar, A. Optimal gasoline price predictions: Leveraging the ANFIS regression model, <https://doi.org/10.1155/2024/8462056>
69. Elmessery, W. M. et al. Semantic segmentation of microbial alterations based on SegFormer. *Front. Plant. Sci.* **15**, 1352935 (2024).
70. Mukherjee, S., Ashish, K., Hui, N. B. & Chattopadhyay, S. Modeling depression data: feed forward neural network vs. radial basis function neural network. *Am. J. Biomed. Sci.* **6** (3), 166–174 (2014).
71. Chattopadhyay, S. Towards predicting the grade of peritonitis with a mamdani-type fuzzy decision system: application of artificial intelligence in surgical emergency referral, <https://doi.org/10.21203/rs.3.rs-2003466/v1> (2022).
72. Zhou, Y. et al. Sep., Passive and active phase change materials integrated building energy systems with advanced machine-learning based climate-adaptive designs, intelligent operations, uncertainty-based analysis and optimisations: A state-of-the-art review, *Renew. Sustain. Energy Rev.* **130**, 109889, <https://doi.org/10.1016/j.rser.2020.109889> (2020).
73. Zhou, Y. & Zheng, S. Multi-level uncertainty optimisation on phase change materials integrated renewable systems with hybrid ventilations and active cooling. *Energy* **202**, 117747. <https://doi.org/10.1016/j.energy.2020.117747> (2020).
74. Zhou, Y., Zheng, S. & Zhang, G. Machine learning-based optimal design of a phase change material integrated renewable system with on-site PV, radiative cooling and hybrid ventilations—study of modelling and application in five climatic regions. *Energy* **192**, 116608. <https://doi.org/10.1016/j.energy.2019.116608> (2020).
75. Zhou, Y., Zheng, S. & Zhang, G. Machine-learning based study on the on-site renewable electrical performance of an optimal hybrid PCMs integrated renewable system with high-level parameters' uncertainties. *Renew. Energy.* **151**, 403–418. <https://doi.org/10.1016/j.renene.2019.11.037> (2020).
76. Zhong, H., He, G., Huo, Y. & Xie, C. A study on genetic algorithm optimization of artificial neural networks. *Int. J. Simul. Syst.* **17** (25), 37.1–37.6, (2016).
77. Seong, N. C., Kim, J. H., Choi, W., Yoon, S. C. & Nassif, N. Development of optimization algorithms for building energy model using artificial neural networks. *J. Korea. Soc. Liv. Environ. Syst.* **24**, 29–36 (2017).
78. Ahmad, M. W., Mourshed, M. & Rezguy, Y. Trees vs neurons: comparison between random forest and ANN for high-resolution prediction of building energy consumption. *Energy Build.* **147**, 77–89 (2017).
79. Kanimozhi, B., Bapu, B. R. & Pranesh, V. Thermal energy storage system operating with phase change materials for solar water heating applications: DOE modelling. *Appl. Therm. Eng.* **123**, 614–624 (2017).
80. Urresti, A., Campos-Celador, A. & Sala, J. M. Dynamic neural networks to analyze the behavior of phase change materials embedded in building envelopes. *Appl. Therm. Eng.* **158**, 113783 (2019).
81. Gonzalez-Briones, A., Hernandez, G., Corchado, J. M., Omatu, S. & Mohamad, M. S. Machine learning models for electricity consumption forecasting: a review, In *2nd International Conference on Computer Applications & Information Security (ICCAIS)*, IEEE, 1–6. (2019).
82. Marani, A. & Nehdi, M. L. Machine learning prediction of compressive strength for phase change materials integrated cementitious composites. *Constr. Build. Mater.* **265**, 120286 (2020).

## Author contributions

This work was carried out in collaboration among all authors. All Authors designed the study, performed the statistical analysis, and wrote the protocol. All authors managed the analyses of the study, managed the literature searches, and wrote the first draft of the manuscript. All authors read and approved the final manuscript.

## Funding

Open access funding provided by The Science, Technology & Innovation Funding Authority (STDF) in cooperation with The Egyptian Knowledge Bank (EKB).

## Declarations

## Competing interests

The authors declare no competing interests.

## Additional information

**Correspondence** and requests for materials should be addressed to W.M.E., M.Y.S., T.A.-H. or T.M.E.-M.

**Reprints and permissions information** is available at [www.nature.com/reprints](http://www.nature.com/reprints).

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2024