



ARTICLE

Edge-Intelligent Photovoltaic Fault Localization via NAS-Optimized Feature-Space Sub-Pixel Matching

Hongjiang Wang¹, Jian Yu², Tian Zhang³, Na Ren⁴, Nan Zhang² and Zhenyu Liu^{1,*}

¹School of Information Science and Engineering, Shenyang University of Technology, Shenyang, China

²School of Computer Science and Technology, Shenyang Institute of Engineering, Shenyang, China

³College of Software, Northeastern University, Shenyang, China

⁴School of Artificial Intelligence, Shenyang University of Technology, Shenyang, China

*Corresponding Author: Zhenyu Liu. Email: liuzhenyu_1973@163.com

Received: 22 December 2025; Accepted: 26 January 2026; Published: 09 April 2026

ABSTRACT: The rapid deployment of Industrial Internet of Things (IIoT) systems, such as large-scale photovoltaic (PV) power stations in modern power grids, has created a strong demand for edge-intelligent fault localization methods that can operate reliably under strict computational and memory constraints. In this work, we propose an edge-intelligent photovoltaic fault localization framework that integrates intelligent computation with classical sub-pixel optimization. The framework adopts a modular, edge-oriented design in which a radial basis function (RBF) network is first employed as a lightweight screening module to enable conditional execution, thereby reducing unnecessary computation for non-faulty samples. For suspicious samples, a compact convolutional feature extractor is activated to generate discriminative representations. The architecture of this feature extractor is automatically optimized using neural architecture search (NAS) in an offline design stage, explicitly balancing localization accuracy and computational efficiency for industrial edge hardware. Sub-pixel displacement estimation and recursive partitioning are then performed in the learned feature space using a sum of squared differences-based, preserving the mathematical transparency of classical sub-pixel matching while significantly improving robustness to thermal noise and background interference. Unlike large end-to-end detection models, the proposed framework combines intelligent feature representation with interpretable localization mechanisms, resulting in a flexible and resource-efficient solution for edge deployment. Experimental results on a photovoltaic infrared fault image dataset demonstrate that the proposed NAS-optimized feature-space sub-pixel matching framework achieves more stable fault localization than other baselines, with only marginal additional computational overhead.

KEYWORDS: Edge intelligence; neural architecture search; sub-pixel localization; feature-based matching; photovoltaic fault localization; industrial internet of things

1 Introduction

With the rapid deployment of large-scale photovoltaic (PV) power stations, intelligent inspection and fault diagnosis have become essential components of modern power grids and industrial internet of things (IIoT) systems [1]. PV modules are often exposed to harsh outdoor environments, where defects such as hotspots, cracks, and degradation can significantly reduce energy efficiency and pose safety risks. To ensure reliable operation, unmanned aerial vehicles (UAVs) equipped with infrared (IR) sensors are increasingly used to perform large-area PV inspections, generating massive amounts of thermal images that require efficient and accurate fault analysis at the edge [2].

The rapid deployment of IIoT systems, such as large-scale PV power stations in modern power grids, has created a strong demand for edge-intelligent fault localization methods that can operate reliably under strict computational and memory constraints. In this work, we propose an edge-intelligent PV fault localization framework that integrates intelligent computation with classical sub-pixel optimization. The framework adopts a modular, edge-oriented design in which a radial basis function (RBF) network is first employed as a lightweight screening module to enable conditional execution, thereby reducing unnecessary computation for non-faulty samples. For suspicious samples, a compact convolutional feature extractor is activated to generate discriminative representations. The architecture of this feature extractor is automatically optimized using neural architecture search (NAS) in an offline design stage, explicitly balancing localization accuracy and computational efficiency for industrial edge hardware. Sub-pixel displacement estimation and recursive partitioning are then performed in the learned feature space using a sum of squared differences (SSD)-based, preserving the mathematical transparency of classical sub-pixel matching while significantly improving robustness to thermal noise and background interference. Unlike large end-to-end detection models, the proposed framework combines intelligent feature representation with interpretable localization mechanisms, resulting in a flexible and resource-efficient solution for edge deployment. Experimental results on a PV-IR fault image dataset demonstrate that the proposed NAS-optimized feature-space sub-pixel matching framework achieves more stable fault localization than other baselines, with only marginal additional computational overhead.

Among various inspection tasks, accurate fault localization plays a critical role in subsequent maintenance and decision-making. Compared with coarse fault detection or classification, precise localization provides actionable information for identifying defective modules and estimating fault severity. In particular, sub-pixel localization is highly desirable in PV inspection scenarios, as IR images often suffer from low spatial resolution, weak contrast, and significant thermal noise [3]. Traditional pixel-domain sub-pixel matching techniques, such as interpolation-based and SSD-based methods, offer mathematically interpretable formulations but are highly sensitive to noise and background interference, limiting their robustness in real-world PV inspection environments [4]. In recent years, deep learning-based methods have been widely applied to PV fault analysis, including classification, object detection, and segmentation approaches [5,6]. These methods leverage convolutional neural networks to learn discriminative features and have demonstrated strong performance in controlled settings. However, most existing deep models focus on pixel-level or bounding-box-level predictions and do not explicitly address sub-pixel localization accuracy. Moreover, end-to-end deep learning models typically require substantial computational resources and memory, making them difficult to deploy on resource-constrained edge devices commonly used in IIoT scenarios, such as embedded controllers and onboard UAV processors [7]. The lack of interpretability and the high deployment cost further limit their adoption in safety-critical industrial applications.

To bridge the gap between classical sub-pixel localization methods and modern learning-based approaches, feature-space matching has emerged as a promising direction. By performing matching operations on learned feature representations rather than raw pixel intensities, feature-space methods can significantly improve robustness to noise, illumination changes, and background clutter [8,9]. Nevertheless, most existing feature-based localization frameworks rely on manually designed network architectures or generic deep feature extractors, which are not optimized for the specific characteristics of PV-IR images or the stringent resource constraints of edge devices. As a result, these methods often suffer from suboptimal accuracy, efficiency trade-offs in industrial deployment.

Intelligent computation techniques, particularly NAS, provide a principled way to address this challenge [10,11]. NAS aims to automatically discover task-specific and resource-aware neural architectures,

reducing reliance on manual design and enabling adaptive optimization for different hardware constraints [12]. While NAS has been extensively studied in image classification, detection, and segmentation tasks, its application to fine-grained localization problems, especially sub-pixel fault localization in industrial inspection, remains largely unexplored. Furthermore, existing NAS-based approaches are rarely integrated with classical optimization frameworks, limiting their interpretability and practical usability in engineering systems.

In this paper, we propose an intelligent computation-enabled, edge-oriented framework for PV fault localization based on NAS-optimized feature-space sub-pixel matching. Instead of replacing classical localization pipelines with black-box end-to-end models, the proposed method preserves the well-established SSD-based sub-pixel refinement formulation and performs it in a learned feature space tailored by NAS. This design combines the interpretability and precision of classical sub-pixel methods with the robustness of learned representations. To further enhance edge suitability, an RBF-based screening module is introduced to enable conditional execution, allowing computationally intensive localization to be applied only to suspicious samples and thereby reducing average latency and runtime variability on edge devices. The main contributions of this work include:

- (1) **Problem perspective:** Different from conventional pixel-domain sub-pixel matching approaches, the proposed method introduces a lightweight feature-space representation to enhance robustness under resource-constrained industrial environments. The overall diagnostic pipeline is specifically designed for edge deployment in IIoT scenarios, such as PV power stations in modern power grids.
- (2) **Methodological perspective:** A feature-space sub-pixel localization framework is proposed by extending classical SSD-based matching from the pixel domain to learned representations. NAS is further incorporated to automatically derive a lightweight and task-specific feature extractor, enabling robust and efficient sub-pixel localization while preserving the interpretability of classical optimization-based methods.
- (3) **Experimental perspective:** Extensive experiments and a strict ablation study on PV-IR inspection data demonstrate the effectiveness of the proposed framework. The results show improved localization accuracy, robustness, and edge efficiency compared with representative baseline methods, validating the practical applicability of the proposed approach in real-world PV inspection.

The remainder of this paper is organized as follows. [Section 2](#) reviews related work on PV fault detection and localization, sub-pixel matching, and intelligent computation for edge intelligence. [Section 3](#) describes the proposed framework in detail. [Section 4](#) presents experimental results and ablation studies. Finally, [Section 5](#) concludes the paper.

2 Related Work

2.1 Photovoltaic Fault Detection and Localization

PV fault detection and localization are essential tasks in modern power grids and IIoT systems, as undetected or inaccurately localized faults may lead to power loss, accelerated module degradation, and potential safety risks [13]. IR thermography has become a widely used inspection modality, as it enables non-contact identification of thermal anomalies associated with defects such as hotspots, cracks, and degradation [14]. Compared with visible imaging, IR inspection is less sensitive to illumination variations and is therefore more suitable for large-scale outdoor PV monitoring under complex environmental conditions.

Early studies primarily focused on fault detection and classification using handcrafted features and conventional classifiers [15,16]. These methods typically extract temperature statistics or texture descriptors

from IR images and apply thresholding or shallow learning models for decision making. Although computationally efficient, such approaches are highly sensitive to noise, environmental variations, and parameter tuning, which limits their robustness in real-world inspection scenarios. More recent works adopt deep learning models for PV inspection, including classification, object detection, and segmentation [17,18]. While deep models can effectively identify faulty modules, most methods provide coarse localization results in the form of bounding boxes or pixel-level masks, which are insufficient for applications requiring precise fault positioning, such as targeted maintenance or quantitative fault assessment. Several studies attempt to refine localization results through post-processing or pixel-domain sub-pixel techniques [19,20]. However, pixel-domain sub-pixel localization is highly sensitive to thermal noise, low contrast, and background interference commonly observed in outdoor IR images. Moreover, many deep learning-based approaches rely on computationally intensive models with high memory and energy consumption, limiting their deployment on resource-constrained edge devices in IIoT scenarios [21,22]. These limitations become particularly critical when real-time processing and stable latency are required in large-scale PV inspection systems. In parallel, recent PV fault diagnosis studies indicate that optimization-enhanced machine-learning pipelines remain effective for real-time monitoring. In particular, a hybrid LightGBM-based detector optimized via Bayesian hyperparameter optimization has been reported for real-time identification of partial shading and PV module faults [23].

In summary, existing PV inspection methods either focus on coarse fault detection or rely on pixel-domain localization and heavy deep models, leaving the problem of robust and efficient sub-pixel fault localization on edge devices insufficiently addressed.

2.2 Sub-Pixel Localization and Feature-Based Matching Methods

Sub-pixel localization has long been studied in computer vision and image processing, where the goal is to estimate target positions with accuracy beyond pixel resolution [24]. Classical sub-pixel localization methods typically rely on interpolation-based refinement, correlation analysis, or optimization of similarity measures such as SSD and mutual information [25]. These methods are mathematically well-founded and computationally efficient, making them attractive for engineering applications. However, when applied directly in the pixel domain, their performance is highly sensitive to noise, illumination variation, and local intensity distortion. To improve robustness, feature-based matching strategies have been introduced, where matching is performed on extracted descriptors rather than raw pixel intensities. Early feature-based methods employ handcrafted descriptors, such as SIFT or SURF, to establish correspondences with improved invariance properties [26]. While effective in certain scenarios, these handcrafted features are not specifically designed for IR imagery and often fail to capture subtle thermal patterns in PV inspection tasks.

With the advancement of deep learning, learned feature representations have been increasingly adopted for image matching and localization. Deep neural networks can automatically learn discriminative features that are more robust to noise and appearance variations [27,28]. Several studies perform dense matching or correspondence estimation in feature space, demonstrating improved stability compared with pixel-domain methods. Nevertheless, most existing deep feature-based localization approaches rely on manually designed network architectures and are primarily developed for generic vision tasks, such as optical flow or image registration, rather than industrial fault localization [29]. Despite their improved robustness, deep feature-based methods introduce new challenges. End-to-end deep localization frameworks often involve heavy models with high computational and memory costs, which are difficult to deploy on resource-constrained edge devices [30]. Moreover, many deep approaches replace classical optimization formulations with black-box regression, sacrificing interpretability and controllability, properties that are often required in safety-critical industrial applications.

In summary, existing sub-pixel localization methods face a trade-off between robustness and efficiency. Pixel-domain methods are efficient but fragile, while deep feature-based methods are robust but computationally expensive and less interpretable.

2.3 Intelligent Computation for Edge-Oriented Visual Inspection

With the increasing adoption of edge computing in IIoT systems, visual inspection models are increasingly required to operate under strict constraints on computation, memory, and energy. In industrial scenarios such as smart grids and PV inspection, visual algorithms are often deployed on embedded processors or onboard UAV platforms, where conventional deep learning models designed for cloud environments are difficult to apply directly [31,32]. To improve edge suitability, extensive research has focused on lightweight neural network design and model compression, including pruning, quantization, and knowledge distillation [33]. These approaches can effectively reduce model complexity, but they typically rely on manually designed backbone architectures and post hoc optimization, which may lead to suboptimal performance for task-specific industrial inspection problems. In addition to architectural design, automated hyperparameter optimization has been increasingly adopted to improve diagnostic reliability in power-system fault analysis. Recent work reports an Optuna-optimized hybrid model with explicit handling of class imbalance for fault detection and classification, illustrating the broader trend toward automated optimization pipelines in practical IIoT-oriented inspection tasks [34].

NAS has emerged as a representative intelligent computation technique that automates network design by optimizing architectural structures under predefined objectives and constraints [35,36]. Resource-aware NAS frameworks further incorporate latency, memory footprint, and hardware characteristics, enabling the generation of architectures tailored for edge devices [37]. NAS has shown strong performance in image classification, object detection, and semantic segmentation, particularly in mobile and edge computing scenarios. However, the application of NAS to fine-grained localization and industrial visual inspection remains limited. Most existing NAS-based methods focus on high-level recognition tasks and evaluate performance primarily in terms of classification or detection accuracy. The integration of NAS with classical optimization-based localization pipelines, such as sub-pixel matching, has received little attention [38]. Moreover, many NAS-driven inspection systems adopt end-to-end deep architectures, which may reduce interpretability and controllability, properties that are often required in safety-critical industrial applications. In addition to model efficiency, runtime stability and predictable latency are important considerations for edge-oriented inspection systems. Adaptive or conditional computation strategies have been proposed to dynamically adjust inference cost, but such mechanisms are rarely combined with precise localization tasks, especially at the sub-pixel level [39].

In summary, while intelligent computation techniques such as NAS provide powerful tools for designing efficient edge models, existing studies primarily target high-level recognition tasks and overlook fine-grained localization under industrial constraints. These limitations motivate the development of edge-oriented inspection frameworks that integrate intelligent computation with classical, interpretable localization principles.

3 Our Method

3.1 Problem Formulation and Edge-Oriented Framework Overview

3.1.1 Problem Definition

PV fault diagnosis in IIoT environments aims to automatically identify abnormal components and precisely localize fault regions based on sensing data acquired in the field. In practical inspection scenarios,

IR or visible-light images of PV panels are collected by distributed sensing devices deployed in PV power stations and processed locally or near the data source. Given an input inspection image can be formulated as follows:

$$\mathbf{I} \in \mathbb{R}^{H \times W} \quad (1)$$

where H and W denote the image height and width, respectively.

The task considered in this work is formulated as a two-stage problem. First, the system determines whether the inspected PV component is faulty, yielding a fault indicator $y \in \{0, 1\}$. $y = 1$ denotes a faulty component and $y = 0$ indicates normal operation. Second, for samples identified as faulty, the system estimates the spatial location of the fault region with sub-pixel accuracy. This localization task is formulated as the estimation of a continuous displacement vector (dx, dy) , which represents the sub-pixel offset of the fault region relative to the image coordinate system.

Unlike coarse-grained fault detection tasks that only require image-level classification or pixel-level segmentation, PV inspection often demands fine-grained localization precision. Small thermal anomalies or micro-defects may correspond to early-stage faults that significantly affect system safety and energy efficiency. Therefore, the objective of this work is to achieve accurate and stable sub-pixel fault localization while maintaining low computational complexity suitable for deployment in industrial edge environments.

3.1.2 Edge-Oriented Constraints

In real-world IIoT-based PV inspection systems, fault diagnosis algorithms are commonly deployed on edge devices, such as embedded processors, industrial controllers, or lightweight computing units integrated with sensing equipment. These devices typically operate under strict constraints in terms of computational power, memory capacity, and energy consumption. Consequently, deploying computationally intensive end-to-end deep learning models is often impractical in such industrial settings. In addition to hardware limitations, PV inspection systems are subject to real-time processing requirements. Inspection results must be produced with low latency to support timely maintenance decisions and ensure operational safety. This necessitates diagnostic pipelines that avoid unnecessary computation and are able to allocate resources adaptively according to the inspection outcome.

Furthermore, industrial inspection data are frequently collected under harsh and dynamic environmental conditions. Variations in temperature, illumination, sensor noise, and background clutter introduce significant uncertainty into the acquired images. Combined with the fact that labeled industrial datasets are often small in scale, these factors pose substantial challenges for robust fault localization. As a result, algorithms designed for edge-oriented PV inspection must simultaneously address resource constraints, real-time requirements, and robustness to noise and data scarcity.

3.1.3 Framework Overview

To address the above challenges, we propose an NAS-optimized feature-space sub-pixel matching (NAS-FSPM) framework that integrates intelligent computation with classical sub-pixel localization techniques. As shown in Fig. 1, the overall architecture of the proposed framework is illustrated, which highlights a compact, modular, and resource-aware diagnostic pipeline designed for industrial edge deployment.

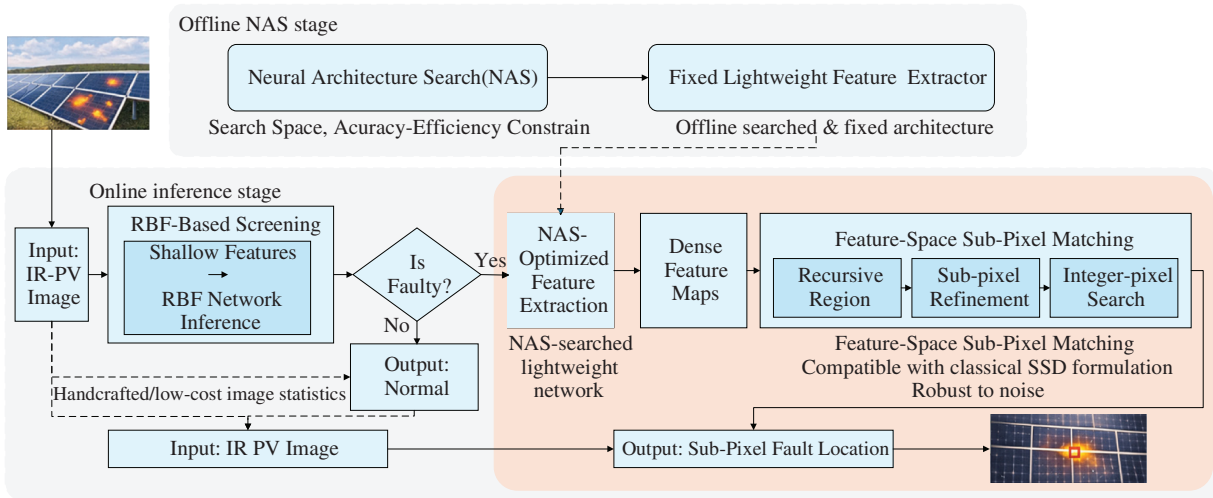


Figure 1: The architecture of NAS-FSPM framework.

First, an RBF neural network is employed as a lightweight fault identification module to enable conditional execution. This stage acts as an efficient front-end screening mechanism that rapidly distinguishes suspicious samples from normal ones with minimal computational overhead. By filtering out non-faulty samples at an early stage, the framework significantly reduces unnecessary downstream processing, thereby conserving computational resources and improving runtime stability on edge devices.

Second, for samples identified as suspicious, a feature extraction stage is activated to generate compact and discriminative representations of the input IR images. Instead of relying on manually designed architectures, a lightweight convolutional feature extractor is automatically optimized using NAS in an offline design phase. The NAS process explicitly considers accuracy-efficiency trade-offs, resulting in a task-specific and resource-aware feature extractor that can be directly deployed on industrial edge hardware. This design reflects the role of intelligent computation in enabling adaptive model customization without introducing excessive runtime complexity.

Finally, precise fault localization is achieved through sub-pixel matching and recursive partitioning performed in the learned feature space. Rather than adopting large end-to-end detection models or heavy post-processing pipelines, the framework preserves classical SSD-based sub-pixel refinement and extends it to operate on learned representations. This hybrid design combines the robustness of learned features with the interpretability and mathematical transparency of classical localization methods. Moreover, the modular structure of the framework allows individual components to be adjusted or replaced according to deployment requirements, further enhancing its flexibility and suitability for edge-oriented industrial inspection applications.

3.1.4 Design Rationale

The design choices of the proposed framework are guided by the characteristics of PV fault diagnosis tasks and the constraints of edge intelligence in IIoT environments.

First, sub-pixel matching is not performed directly in the pixel domain. Pixel-domain representations are highly sensitive to noise, thermal fluctuation, and background interference, which are common in industrial inspection scenarios. These sensitivities often lead to unstable matching costs and unreliable localization results, particularly when image quality varies across operating conditions. Second, feature-space

representation is adopted to improve robustness. By transforming raw image intensities into a compact and discriminative feature space, irrelevant variations caused by noise and environmental factors can be effectively suppressed, while fault-related structural information is preserved. This provides a more stable basis for sub-pixel matching and recursive localization, without altering the underlying mathematical formulation of sub-pixel displacement estimation. Third, NAS is employed to optimize the feature extraction network instead of relying on manually designed architectures. Given the diversity of edge hardware platforms and the stringent resource constraints in industrial environments, manually designing a universally optimal feature extractor is difficult. NAS provides an intelligent computation mechanism that automatically balances localization accuracy and computational efficiency, resulting in lightweight architectures well suited for edge deployment.

Overall, the proposed framework leverages intelligent computation to enhance feature representation quality, while preserving the interpretability and precision of classical sub-pixel localization methods. This design achieves an effective balance between performance, robustness, and computational efficiency, making it particularly suitable for edge-intelligent PV fault diagnosis in IIoT systems.

3.2 RBF-Based Lightweight Fault Identification

In practical PV inspection scenarios, performing high-precision sub-pixel localization for all inspection images is computationally inefficient and unnecessary, especially under normal operating conditions. Therefore, a lightweight fault identification module is introduced as a front-end stage to rapidly determine whether an inspection sample contains potential faults. This module serves as a decision gate that selectively activates subsequent intelligent computation and fine-grained localization processes only when necessary.

An RBF neural network is adopted for this task due to its simple structure, low inference latency, and strong capability to approximate nonlinear decision boundaries with limited training data. The RBF network performs binary fault identification, classifying inspection samples into normal and faulty categories.

3.2.1 Input Representation and RBF Network Structure

Let $\mathbf{x} \in \mathbb{R}^d$ denote a shallow and computationally inexpensive feature vector extracted directly from a PV inspection image. These features are designed to capture coarse fault-related characteristics, such as basic intensity statistics, temperature distribution descriptors, or hotspot area ratios, and can be computed efficiently on edge devices without invoking deep feature extraction. The input features are designed to be low-dimensional and computationally inexpensive, and may include basic statistical descriptors, thermal distribution characteristics, or other shallow image features that can be efficiently computed on edge devices. The RBF network consists of an input layer, a hidden layer composed of radial basis neurons, and a linear output layer. Each radial basis neuron computes a Gaussian activation function can be formulated as follows:

$$\phi_i(\mathbf{x}) = \exp\left(-\frac{\|\mathbf{x} - \boldsymbol{\mu}_i\|^2}{2\sigma_i^2}\right) \quad (2)$$

where $\boldsymbol{\mu}_i \in \mathbb{R}^d$ represents the center of the i -th RBF and σ_i controls its width.

3.2.2 Fault Identification Output and Training Objective

The output of the RBF network is obtained by a linear combination of the hidden-layer can be formulated as follows:

$$y = \sum_{i=1}^M w_i \phi_i(\mathbf{x}) \quad (3)$$

where w_i represents the output weight associated with the i -th hidden unit, and M is the number of RBF.

For binary fault identification, the network output y is compared against a predefined threshold τ . If $y \geq \tau$, the inspection sample is classified as faulty; otherwise, it is regarded as normal. During training, the network parameters are optimized by minimizing the mean squared error loss can be formulated as follows:

$$E = \frac{1}{2} (y - t)^2 \quad (4)$$

where $t \in \{0, 1\}$ denotes the ground-truth label indicating the fault status. The centers μ_i , widths σ_i , and weights w_i can be determined using standard training procedures, such as clustering-based initialization followed by least-squares or gradient-based optimization. Since the RBF network is only required to perform coarse fault identification, the training process remains stable and efficient even when only limited labeled industrial data are available.

3.2.3 Edge-Oriented Role in the Overall Framework

From an edge-oriented perspective, the RBF-based fault identification module is intentionally designed to be lightweight and computationally efficient. By acting as a front-end decision gate, it conditionally triggers the subsequent NAS-optimized feature extraction and feature-space sub-pixel localization stages only for samples identified as faulty. This conditional execution strategy avoids unnecessary high-cost computation for normal operating conditions and significantly reduces overall resource consumption. As a result, the proposed framework achieves an effective balance between detection reliability and computational efficiency, making it suitable for deployment in IIoT environments with strict resource constraints. It is important to note that the RBF-based module operates independently of the NAS-optimized feature extractor. Only inspection samples classified as faulty are forwarded to the subsequent NAS-based feature extraction stage, ensuring that expensive feature computation is conditionally activated.

3.3 NAS-Optimized Lightweight Feature Extraction

Following the RBF-based lightweight fault identification stage, only inspection samples identified as faulty are processed by the NAS-optimized feature extraction module. Given such an inspection image \mathbf{I} , the feature extraction network produces a dense feature map $\mathbf{F} = f_{\theta}(\mathbf{I})$, which serves as the input for subsequent sub-pixel localization.

To clarify how intelligent computation is integrated into the proposed edge-oriented diagnostic framework, Algorithm 1 presents the procedural workflow of the NAS-optimized lightweight feature extraction module. The algorithm summarizes the system-level execution logic rather than the internal optimization details of NAS, highlighting how automated model customization is incorporated under accuracy-efficiency constraints for edge deployment. The following subsections describe the motivation, design principles, and integration of the NAS-optimized feature extraction module in detail.

Algorithm 1: NAS-optimized lightweight feature extraction

Input: \mathbf{I} is inspection image; \mathcal{A} is search space of lightweight feature extraction architectures;
 c is accuracy-efficiency constraints.

Output: Feature map \mathbf{F} .

```

1  $\mathcal{A} \leftarrow$  Initialize candidate architecture set;
2 for each candidate architecture  $a \in \mathcal{A}$  do
3   Evaluate  $a$  under constraints  $c$ ;
4 end for
5 Select optimal architecture  $a^*$  satisfying the accuracy-efficiency trade-off;
6 Deploy the optimized feature extraction network  $f_\theta$  corresponding to  $a^*$ ;
7 if the inspection sample is identified as faulty then
8   Extract feature map  $\mathbf{F} = f_\theta(\mathbf{I})$ ;

10 Skip feature extraction and localization;
11 end if
12 Return feature map  $\mathbf{F}$ 

```

3.3.1 Motivation for NAS-Based Feature Optimization

Manually designing a feature extraction network that simultaneously satisfies robustness, accuracy, and computational efficiency is challenging in IIoT environments. Edge devices deployed in PV power stations differ significantly in terms of processing capability, memory availability, and energy consumption. A manually engineered architecture that performs well on one platform may not generalize well to others, especially under strict real-time constraints.

NAS provides an intelligent computation mechanism that enables automated model customization under predefined constraints. Rather than searching for deep or complex architectures, the goal of NAS in this work is to identify a lightweight feature extractor that offers sufficient discriminative capability while maintaining low computational overhead. By leveraging NAS, the feature extraction module can be adapted to the edge-oriented requirements of PV inspection without relying on extensive manual tuning.

3.3.2 Feature Extraction Network Design

Let the input inspection image be denoted as $\mathbf{I} \in \mathbb{R}^{H \times W}$. For clarity, the output feature map preserves a spatial resolution of $H \times W$, i.e., $\mathbf{F} \in \mathbb{R}^{H \times W \times C}$, where H and W denote the feature-map height and width, respectively. The feature extraction network is defined as a mapping can be formulated as follows:

$$\mathbf{F} = f_\theta(\mathbf{I}), \mathbf{F} \in \mathbb{R}^{H \times W \times C} \quad (5)$$

where $f_\theta(\cdot)$ represents the NAS-optimized feature extraction function parameterized by θ , and C denotes the number of feature channels. Importantly, the spatial resolution of the feature map is preserved, ensuring compatibility with subsequent sub-pixel matching operations. Accordingly, the spatial index in Eq. (5) is defined on this $H \times W$ feature grid.

In this work, the NAS search space \mathcal{A} for the lightweight feature extraction module is intentionally constrained to compact, resolution-preserving convolutional architectures to meet edge-intelligence requirements. Specifically, an architecture $a \in \mathcal{A}$ is composed of a shallow stack of B lightweight blocks, and each block selects one operator from a predefined candidate set consisting of commonly used efficient building blocks (e.g., 1×1 pointwise convolution, small-kernel convolution such as 3×3 and 5×5 ,

depthwise-separable convolution with small kernels, lightweight pooling operators, and identity/skip). The macro-topology is kept feed-forward (blocks are connected sequentially), while local skip/residual connections are allowed within a block (identity shortcut when dimensions match, or a 1×1 projection otherwise) to facilitate stable feature learning without introducing heavy computation. To ensure compatibility with subsequent feature-space sub-pixel matching, all candidate operators are restricted to stride 1 (i.e., no spatial downsampling), so that the feature map spatial resolution is preserved as $H \times W$. In addition, the channel width of intermediate features is limited to a small discrete set (and the output channel number is fixed to C) to avoid over-parameterized designs. During NAS, candidate architectures are searched under explicit accuracy-efficiency constraints (e.g., model complexity and inference latency budgets), and the final architecture a^* is selected as the best accuracy–efficiency trade-off within \mathcal{A} .

3.3.3 Resource-Aware Optimization via Intelligent Computation

The NAS process is conducted under explicit accuracy-efficiency constraints, reflecting the practical requirements of edge-intelligent PV inspection. During the search process, candidate architectures are evaluated based on their ability to produce discriminative feature representations while satisfying computational budget limitations, such as inference latency and model complexity. It is worth noting that NAS in this framework is not intended to replace the overall diagnostic pipeline with an end-to-end deep learning model. Instead, it serves as an intelligent computation tool to optimize a specific component within a hybrid system. By automatically identifying an appropriate feature extraction architecture, NAS enables the proposed framework to balance robustness and efficiency without increasing system complexity.

The resulting feature extractor is lightweight and task-specific, making it well suited for deployment on industrial edge devices. Moreover, the automated nature of NAS reduces reliance on domain-specific heuristics and manual architecture design, enhancing the adaptability of the framework to different inspection scenarios and hardware platforms.

3.4 Feature-Space Sub-Pixel Matching and Recursive Localization

Based on the NAS-optimized feature representation, precise PV fault localization is achieved through sub-pixel matching performed in the learned feature space. Different from conventional pixel-domain approaches, the proposed method computes the matching cost on discriminative feature maps while preserving the classical formulation of integer-pixel search and sub-pixel refinement. This design enables improved robustness against noise and environmental variations without altering the underlying sub-pixel estimation mechanism. To clearly present the execution logic of the proposed localization strategy, Algorithm 2 summarizes the feature-space sub-pixel matching and recursive localization procedure adopted in this work.

Algorithm 2: Feature-space sub-pixel matching with recursive localization

Input: F is feature map; Ω_0 is initial region of interest; L is maximum recursion depth.

Output: Sub-pixel fault location (u^*, v^*) .

```

1 Initialize region  $\Omega \leftarrow \Omega_0$ ;
2 for level  $l = 1$  to  $L$  do
3   Perform integer-pixel matching within  $\Omega$ ;
4   Compute feature-space SSD cost  $J(u, v)$ ;
5   Estimate sub-pixel displacement  $(dx, dy)$ ;
6   Update location  $(u, v) \leftarrow (u_0 + dx, v_0 + dy)$ ;

```

(Continued)

Algorithm 2 (continued)

```

7 Refine region  $\Omega$  centered at  $(u, v)$ ;
8 end for
9 Return final location  $(u^*, v^*)$ 

```

At each recursion level, the region of interest is re-centered at the current localization estimate and its spatial extent is reduced by a fixed ratio. The recursion terminates when the maximum depth L is reached or when the displacement update becomes negligible.

3.4.1 Feature-Space SSD Matching Formulation

For a given reference feature block F_r and a target feature block F_t , the similarity between the two blocks under a candidate displacement (u, v) is evaluated using the SSD criterion. Specifically, the feature-space SSD matching cost is defined as:

$$J(u, v) = \sum_{(x, y) \in \Omega} \|F_r(x, y) - F_t(x + u, y + v)\|_2^2 \quad (6)$$

where $F_r(x, y)$ and $F_t(x, y)$ denote the C -dimensional feature vectors at spatial location (x, y) in the reference and target feature maps, respectively. Ω represents a square matching window of size $W \times W$, centered at the reference location. The operator $\|\cdot\|_2$ represents the Euclidean norm.

Compared with pixel-domain SSD, Eq. (5) aggregates discrepancies across multiple feature channels rather than relying on single-channel intensity differences. This multi-dimensional aggregation suppresses the influence of local noise, illumination variation, and background clutter, leading to a smoother and more discriminative matching cost surface. Importantly, the SSD criterion retains its original interpretation: smaller values of $J(u, v)$ indicate higher similarity, which is consistent with classical sub-pixel matching theory. From an implementation perspective, the use of SSD in feature space preserves computational simplicity. The cost evaluation involves only basic arithmetic operations and can be efficiently executed on edge devices, making it suitable for real-time PV inspection under resource constraints.

3.4.2 Integer and Sub-Pixel Displacement Decomposition

To estimate the relative displacement between the reference and target feature blocks, the displacement vector is decomposed into an integer-pixel component and a sub-pixel offset, expressed as:

$$u = u_0 + dx, \quad v = v_0 + dy \quad (7)$$

where $(u_0, v_0) \in \mathbb{Z}^2$ represents the integer-pixel displacement and $dx, dy \in (-0.5, 0.5)$ denote the sub-pixel offsets along the horizontal and vertical directions, respectively.

The integer-pixel displacement is first obtained through discrete search within a predefined region S can be formulated as follows, and S is typically centered around an initial estimate and constrained to limit computational cost.

$$(u_0, v_0) = \operatorname{argmin}_{(u, v) \in S} J(u, v) \quad (8)$$

where $S = \{(u, v) \mid |u| \leq r, |v| \leq r\}$ denotes a bounded integer-pixel search region with radius r , which is set according to the expected displacement magnitude and computational constraints.

This two-stage decomposition strategy separates coarse alignment from fine-grained refinement. Integer-pixel matching provides a reliable initial estimate that reduces the risk of local minima, while

sub-pixel refinement enables precise localization beyond pixel resolution. Such a separation is particularly important in industrial inspection scenarios, where feature responses may vary due to noise and environmental factors, and direct continuous optimization may be unstable.

3.4.3 Sub-Pixel Displacement Estimation via Quadratic Interpolation

After determining the integer displacement (u_0, v_0) , sub-pixel refinement is performed by locally approximating the SSD cost function around this position. Along the horizontal direction, the SSD values at three neighboring integer locations are defined as:

$$J_- = J(u_0 - 1, v_0), J_0 = J(u_0, v_0), J_+ = J(u_0 + 1, v_0) \quad (9)$$

Assuming a quadratic approximation of the cost surface, the sub-pixel offset dx is obtained as:

$$dx = \frac{J_- - J_+}{2(J_- - 2J_0 + J_+)} \quad (10)$$

Similarly, along the vertical direction can be formulated and compute as follows:

$$K_- = J(u_0, v_0 - 1), K_0 = J(u_0, v_0), K_+ = J(u_0, v_0 + 1) \quad (11)$$

$$dy = \frac{K_- - K_+}{2(K_- - 2K_0 + K_+)} \quad (12)$$

The refined displacement is given by (u^*, v^*) and is used for fault localization. The final sub-pixel localization result is then given as follows:

$$(u^*, v^*) = (u_0 + dx, v_0 + dy) \quad (13)$$

Quadratic interpolation provides a closed-form solution with low computational complexity, making it well suited for edge-oriented applications. For numerical stability, when the denominator in Eqs. (10) and (11) approaches zero, the corresponding sub-pixel offset is set to zero or clipped to the valid range $[-0.5, 0.5]$. This safeguard ensures robust operation under noisy industrial conditions.

4 Experimental Results and Analysis

4.1 Experimental Setup

4.1.1 Datasets

The experimental evaluation is conducted on two IR PV inspection datasets, covering both fine-grained fault localization benchmarking and scale-and-diversity validation under challenging field conditions.

In-house PV-IR inspection dataset. The primary benchmark is collected from a real-world PV power station using a UAV-mounted IR imaging system, reflecting typical IIoT-based inspection scenarios in modern power grids. The raw data consist of IR images capturing the operational thermal distribution of PV modules. For consistency, all images are resized to a unified resolution and normalized before further processing. To improve viewpoint diversity and mitigate limited-data effects commonly encountered in industrial inspection, standard data augmentation is applied, including rotation, scaling, flipping, and translation. PV module images are annotated using Labellmg and organized into two categories: normal samples G_1 and faulty samples G_2 . Specifically, G_1 contains IR images collected under normal operating conditions, whereas G_2 includes images exhibiting evident thermal anomalies. Each subset contains 1000

samples, yielding a total of 2000 annotated inspection images for experimental evaluation. For fine-grained localization benchmarking, we additionally provide bounding-box annotations for hotspot regions in G_2 . Representative examples are illustrated in Fig. 2, where PV fault patterns often exhibit weak contrast, blurred boundaries, and interference from background thermal variations.

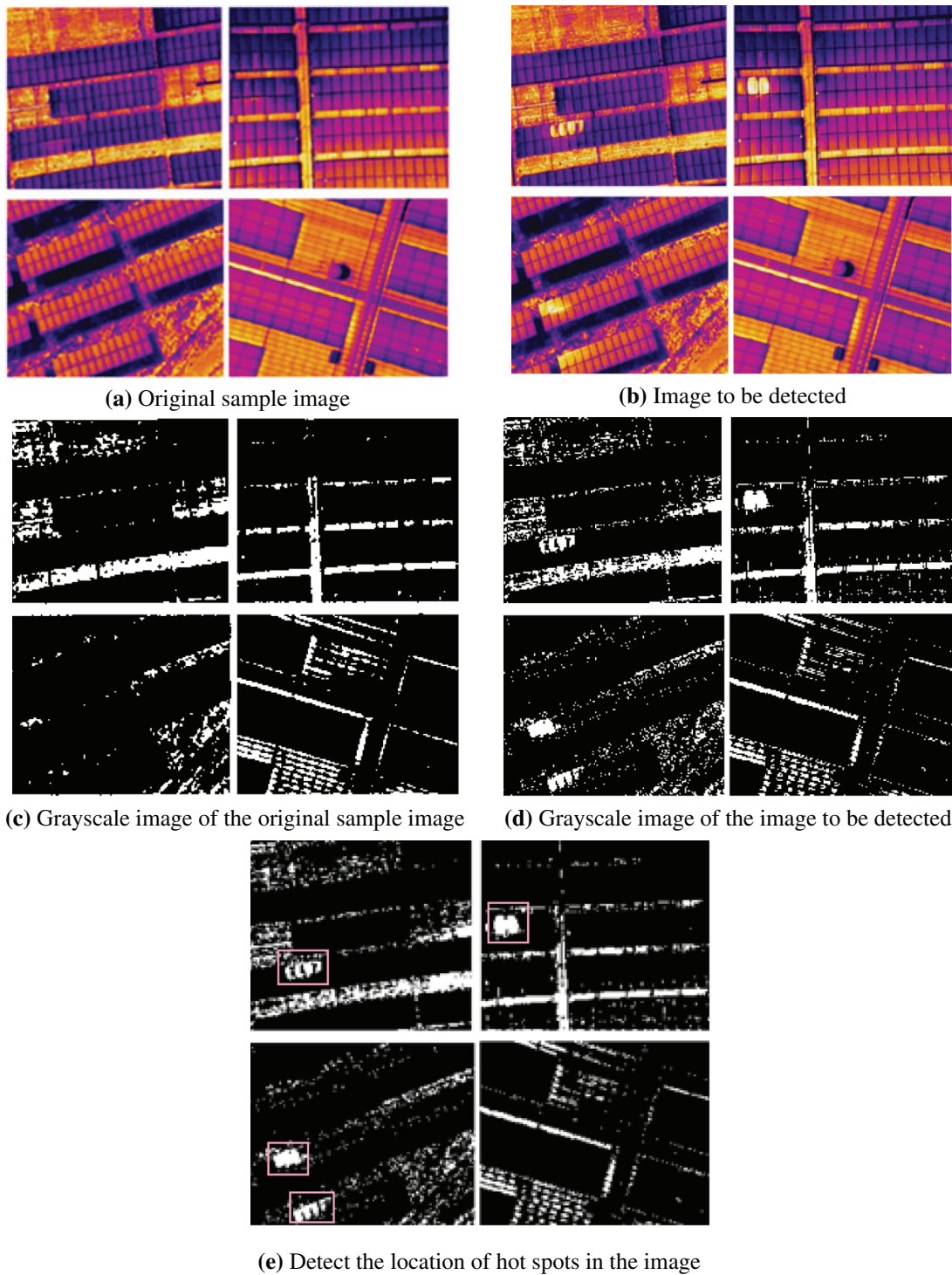


Figure 2: Five examples of detection images.

RaptorMaps dataset for scale-and-diversity validation. We additionally adopt the public RaptorMaps InfraredSolarModules dataset [40]. This dataset contains 20,000 low-resolution IR module images (24×40) and includes challenging field conditions such as soiling (e.g., dust/debris accumulation) and shadowing (partial shading). Since RaptorMaps is primarily designed for module-level inspection and does not provide fine-grained localization annotations required for sub-pixel displacement evaluation, it is used as an auxiliary benchmark for scale-and-diversity validation by formulating an abnormality screening task (Normal vs. Abnormal). Here, Normal corresponds to the normal category, and Abnormal merges all non-normal categories. In addition to overall binary screening performance, we further report subset-wise results on soiling and shadowing samples to explicitly evaluate robustness under dust accumulation and partial shading conditions.

4.1.2 Implementation Details

The overall framework consists of three sequential stages: lightweight fault identification, NAS-optimized feature extraction, and feature-space sub-pixel localization. For the front-end fault identification stage, shallow and computationally inexpensive features are extracted directly from the inspection images, including basic statistical descriptors and coarse thermal distribution characteristics. These features are used as inputs to the RBF-based classifier, which performs binary fault identification and serves as a decision gate. The RBF network is trained using labeled normal and faulty samples from the training set, and only samples classified as faulty are forwarded to the subsequent localization stages.

The NAS-based feature extractor is optimized offline under accuracy-efficiency constraints, and the searched lightweight architecture is deployed for inference without further architectural adaptation. During inference, the feature extractor produces a dense feature map $\mathbf{F} \in \mathbb{R}^{H \times W \times C}$, where C denotes the feature channel dimension. This feature representation is used exclusively for sub-pixel localization and is not involved in the RBF-based decision stage, ensuring clear separation between fault screening and fine-grained localization.

For feature-space sub-pixel localization, the SSD cost is computed within a local matching window of size $W \times W$. Integer-pixel displacement is first estimated by searching within a bounded region of radius r , followed by sub-pixel refinement using quadratic interpolation along horizontal and vertical directions. To improve localization precision, a recursive refinement strategy is applied, where the region of interest is re-centered and progressively reduced at each level. The recursion depth is set to a fixed value L , and the refinement process terminates once the maximum depth is reached or the displacement update becomes negligible. All hyperparameters, including the matching window size W , integer search radius r , feature channel dimension C , and recursion depth L , are selected to balance localization accuracy and computational efficiency. The implementation is designed to reflect practical edge deployment scenarios, where computational resources and latency are constrained. Unless otherwise specified, the same parameter settings are used across all experiments to ensure fair comparisons.

Implementation on the RaptorMaps benchmark. For the additional RaptorMaps dataset, we perform binary abnormality screening (Normal vs. Abnormal) using a lightweight classifier built upon the deployed feature extractor, and train it under a consistent supervised learning protocol. The low-resolution IR inputs are resized using the same preprocessing pipeline, and the evaluation is conducted on a held-out test split with standard binary classification metrics. To explicitly assess the impact of challenging field conditions, subset-wise evaluation is further carried out on samples annotated as soiling and shadowing.

We additionally conduct power/thermal benchmarking on an NVIDIA Jetson Orin Nano 8GB Developer Kit. The device runs JetPack 5.1.2 with CUDA 11.4, and the models are executed using PyTorch 2.0.0

in inference mode. All measurements are performed with batch size 1 under the same input resolution and preprocessing as the screening experiments. For each model, we run continuous inference after a warm-up period while keeping a fixed power mode and clock setting. Device power and temperature are logged using tegrastats at 1 Hz, and we report the average/peak power consumption and steady-state operating temperature.

4.1.3 Evaluation Metrics

To comprehensively evaluate the proposed intelligent computation-enabled edge-oriented PV fault localization framework, multiple evaluation metrics are adopted to assess fault identification performance, localization accuracy, and computational efficiency.

For the front-end fault identification stage, classification accuracy is used to evaluate the performance of the RBF-based lightweight decision module. Let N denote the total number of inspection samples, and let N_{correct} denote the number of samples correctly classified as normal or faulty. The fault identification accuracy is defined as follow:

$$\text{Accuracy} = \frac{N_{\text{correct}}}{N} \quad (14)$$

In addition to localization error metrics, the identification accuracy of the RBF-based screening module is evaluated separately to assess the reliability of the conditional execution mechanism. For fault localization performance, the sub-pixel localization error is measured as the Euclidean distance between the estimated fault position and the ground-truth annotation. Let (u_i^*, v_i^*) denote the estimated sub-pixel location for the i -th test sample, and let (u_i, v_i) denote the corresponding ground-truth location. The localization error for the i -th sample is computed as follow:

$$e_i = \sqrt{(u_i^* - u_i)^2 + (v_i^* - v_i)^2} \quad (15)$$

The overall localization accuracy is evaluated using the average localization error:

$$\bar{e} = \frac{1}{N_f} \sum_{i=1}^{N_f} e_i \quad (16)$$

where N_f denotes the number of faulty samples in the test set. In addition, the median localization error is reported to reduce the influence of outliers and provide a robust estimate of typical localization performance.

To assess the robustness and stability of sub-pixel localization under industrial conditions, the variance of the localization error is also evaluated. The variance is defined as follow:

$$\text{Var}(e) = \frac{1}{N_f} \sum_{i=1}^{N_f} (e_i - \bar{e})^2 \quad (17)$$

A lower variance indicates that the localization results are less sensitive to noise, background variations, and environmental disturbances, which is critical for practical PV inspection scenarios.

Computational efficiency is evaluated by measuring the average runtime per inspection image. Let denote the processing time for the i -th sample. The average runtime is defined as follow:

$$\bar{T} = \frac{1}{N} \sum_{i=1}^N T_i \quad (18)$$

To quantify the benefit of the RBF-based conditional execution strategy, the average runtime is compared between two execution modes: (i) always performing feature extraction and sub-pixel localization for all samples, and (ii) conditionally activating these stages only for samples classified as faulty. This comparison directly reflects the efficiency gains achieved by the proposed edge-oriented design.

For RaptorMaps abnormality screening, Abnormal is treated as the positive class, and standard binary classification metrics are reported, including Accuracy, Precision, Recall, and F1-score. In addition, to explicitly evaluate robustness under challenging environmental conditions highlighted by the reviewers, we further report subset-wise performance on samples annotated as soiling and shadowing, using their recall (TPR) and the corresponding false negative rate (FNR).

4.2 Baseline Methods

We compare the proposed method with the following representative baseline approaches for PV fault localization:

- (1) **Adaptive Pixel-Domain Sub-Pixel Matching PD-SPM** [41] follows recent advances in digital image correlation and adaptive correlation-based refinement. Sub-pixel localization is performed directly on image intensity values using SSD-based matching, integer-pixel search, and adaptive quadratic interpolation.
- (2) **CNN-Based Deep Feature Matching (CNN-DFM)** [42] employs learned feature representations for image matching and localization. Following recent deep feature matching frameworks, sub-pixel localization is achieved by matching deep features and refining correspondences, without explicit neural architecture optimization.
- (3) **Lightweight CNN-Based Feature Extraction (LCNN)** [43] uses a manually designed mobile-oriented convolutional neural network to extract feature maps for localization. The network architecture follows widely adopted lightweight design principles for edge vision without NAS.
- (4) **End-to-End Deep Learning-Based Photovoltaic Fault Detection (E2E-DL)** [44] directly predicts fault locations from IR images using deep neural networks. This baseline follows recent PV inspection studies that employ deep learning for defect detection under unconstrained computational settings.
- (5) **ViT-ANN** [45] is a transformer-based baseline that leverages a Vision Transformer (ViT) artificial neural network (ANN) head to model global contextual dependencies in IR thermography, enabling robust fault detection and multi-class fault classification for photovoltaic modules.
- (6) **RepAlexSolarNet** [46] is a reparameterization-based AlexNet-style CNN baseline for imbalanced solar panel fault classification, which trains with multi-branch convolutional blocks and imbalance-aware optimization to improve minority-class recognition, and then reparameterizes these blocks into a single-branch structure for efficient deployment-time inference.

All baseline methods are evaluated under the same dataset split, annotation protocol, and preprocessing pipeline. For localization-oriented baselines involving sub-pixel matching, identical search window sizes, search radii, and interpolation strategies are used. Fault screening baselines (e.g., ViT-ANN and RepAlexSolarNet) are trained and evaluated under the same binary screening protocol and metrics.

4.3 Experimental Results

4.3.1 Evaluation of Sub-Pixel Localization Accuracy

Table 1 reports the quantitative fault localization performance of all compared methods. The average localization error (Avg. Error (px)), median error (Median Error (px)), error variance (Error Variance),

and maximum error (Max Error (px)) are provided to jointly evaluate accuracy, robustness, and worst-case behavior.

Table 1: Comparison of sub-pixel fault localization performance.

Method	Avg. Error (px)	Median Error (px)	Error Variance	Max Error (px)
PD-SPM	1.42	1.31	0.58	2.94
CNN-DFM	1.18	1.05	0.44	2.36
LCNN	1.02	0.91	0.37	2.01
E2E-DL	0.89	0.83	0.35	1.96
NAS-FSPM (Ours)	0.71	0.65	0.21	1.34

As shown in [Table 1](#), compared with PD-SPM, NAS-FSPM reduces the average localization error from 1.42 px to 0.71 px, corresponding to nearly a 50% improvement. Similar trends are observed for the median error, which decreases from 1.31 px to 0.65 px. These results indicate that replacing raw pixel intensities with learned feature representations substantially improves sub-pixel localization accuracy. DFM and LCNN further reduce localization error compared with PD-SPM, confirming the benefit of feature-space matching. However, NAS-FSPM achieves consistently lower errors than these methods, suggesting that generic deep features and manually designed lightweight networks may not be fully adequate for robust sub-pixel localization in this setting. Although E2E-DL achieves competitive average accuracy, it exhibits higher variance and larger worst-case errors than NAS-FSPM, indicating less stable localization behavior in challenging samples. In addition, NAS-FSPM exhibits lower error variance and a smaller maximum error, indicating improved stability under thermal noise and background interference. Overall, the results in [Table 1](#) suggest that combining NAS-optimized feature representations with classical sub-pixel refinement can provide more accurate and stable PV fault localization than the compared pixel-domain and deep baselines in our evaluation.

4.3.2 Computational Efficiency

To evaluate the suitability of different methods for edge deployment, [Table 2](#) summarizes the computational efficiency in terms of average runtime per image (e.g., Avg. Runtime (ms/image) and Runtime Std. (ms)). For NAS-FSPM, results are reported with and without the RBF-based conditional execution mechanism.

Table 2: Comparison of computational efficiency.

Method	Avg. Runtime (ms/Image)	Runtime Std. (ms)
PD-SPM	18.6	3.1
CNN-DFM	32.4	4.7
LCNN	28.7	4.2
E2E-DL	74.3	9.6
NAS-FSPM (w/o RBF gate)	34.1	5.0
NAS-FSPM	22.9	3.4

As shown in [Table 2](#), the pixel-domain baseline PD-SPM achieves the lowest runtime (e.g., 18.6 ms/image) due to its simple computations, but its localization accuracy is limited. DFM and LCNN

require 32.4 ms/image and 28.7 ms/image, respectively, while E2E-DL is the most computationally expensive at 74.3 ms/image, making it unsuitable for resource-constrained edge deployment. Without conditional execution, NAS-FSPM requires 34.1 ms/image, which is comparable to DFM and slightly higher than LCNN, reflecting the cost of feature extraction. When the RBF-based gate is enabled, the average runtime of NAS-FSPM is reduced to 22.9 ms/image, corresponding to a reduction of approximately 33%. This brings the runtime close to that of PD-SPM while maintaining substantially higher localization accuracy and stability. Across all experiments, the RBF-based screening module achieves an identification accuracy of 96.2%, indicating that the conditional execution strategy does not compromise the reliability of subsequent sub-pixel localization.

Overall, NAS-FSPM with conditional execution provides a favorable trade-off between computational efficiency and localization performance, outperforming learning-based baselines in edge suitability and remaining significantly more efficient than end-to-end deep learning approaches.

4.3.3 Hard-Case Analysis for Hotspot Localization

The UAV-based IR inspection often encounters “hard” hotspot instances where the thermal pattern is weak, small, or embedded in cluttered backgrounds. To better characterize such challenging cases and to understand error patterns beyond aggregate metrics, we conduct a hard-case analysis on the in-house dataset by grouping hotspot instances according to three common factors: low contrast, small hotspot, and high background variation. Each factor is derived directly from the hotspot bounding-box annotations, and hard subsets are constructed using dataset-driven quantiles to avoid arbitrary thresholding. Specifically, low contrast corresponds to the bottom 25% of the contrast score (computed as the mean intensity difference between the bounding-box region and its surrounding ring region), small hotspot corresponds to the bottom 25% of bounding-box area, and high background variation corresponds to the top 25% of the background standard deviation measured in the surrounding ring region. This results in three subsets with sizes $N = 248$, $N = 252$, and $N = 245$, respectively.

Table 3 summarizes the localization performance on these hard subsets using the same metrics as in Table 1 (e.g., Avg. Error, Median Error, Error Variance, and Max Error), together with an additional Failure (%) statistic that quantifies the proportion of samples whose localization error exceeds a fixed tolerance. Across all three hard conditions, NAS-FSPM consistently achieves the best accuracy and stability, outperforming both E2E-DL and the classical pixel-domain baseline PD-SPM.

Across all three hard conditions, NAS-FSPM shows consistently lower errors and improved stability compared with E2E-DL and PD-SPM. For low-contrast hotspots, where weak thermal separation flattens the matching landscape, NAS-FSPM achieves 0.98/0.92 px (Avg./Median) and markedly improves worst-case behavior (Max 1.88 px, Var 0.18, Failure 33.8%) over E2E-DL (1.23/1.16 px, Max 2.35 px) and PD-SPM (1.92/1.81 px, Max 3.12 px). For small hotspots, which provide limited structural cues, NAS-FSPM remains robust (0.91/0.85 px, Var 0.16, Max 1.74 px, Failure 28.4%), outperforming E2E-DL (1.14/1.07 px) and PD-SPM (1.79/1.68 px). Under high background variation, where clutter introduces competing extrema and yields the highest failure rates overall, NAS-FSPM still maintains the most favorable trade-off (1.05/0.99 px, Var 0.21, Max 2.02 px, Failure 39.6%), indicating that its feature representation better suppresses background-induced distractors while preserving hotspot-relevant cues.

Table 3: Hard-case analysis for hotspot localization on the in-house dataset.

Method		Low Contrast	Small Hotspot	High Background
		(N = 248)	(N = 252)	(N = 245)
PD-SPM	Avg. Error (px)	1.92	1.79	1.97
	Median Error (px)	1.81	1.68	1.86
	Error Variance	0.44	0.39	0.47
	Max Error (px)	3.12	2.98	3.24
	Failure (%)	72.6	64.8	75.1
E2E-DL	Avg. Error (px)	1.23	1.14	1.28
	Median Error (px)	1.16	1.07	1.21
	Error Variance	0.26	0.23	0.29
	Max Error (px)	2.35	2.18	2.47
	Failure (%)	45.7	38.9	49.3
NAS-FSPM	Avg. Error (px)	0.98	0.91	1.05
	Median Error (px)	0.92	0.85	0.99
	Error Variance	0.18	0.16	0.21
	Max Error (px)	1.88	1.74	2.02
	Failure (%)	33.8	28.4	39.6

4.3.4 RaptorMaps Screening with on-Device Profiling

As shown in Table 4, we further evaluate the proposed framework on the large-scale public RaptorMaps dataset under a binary screening protocol and profile the on-device power/thermal footprint on an NVIDIA Jetson Orin Nano. This setting enables a practical assessment of the accuracy-efficiency trade-off that is central to edge-oriented PV inspection.

Table 4: RaptorMaps screening results and Jetson Orin Nano power/thermal profiling.

Method	RaptorMaps Screening				Jetson Orin Nano Power/Thermal		
	Acc (%)	Pre (%)	Recall (%)	F1 (%)	Avg Power (W)	Peak Power (W)	Steady Temp (°C)
ViT-ANN	94.12	93.05	94.46	93.75	10.62	16.41	71.3
RepAlexSolarNet	91.38	89.62	92.14	90.86	8.95	13.78	66.1
NAS-FSPM	92.67	91.84	93.02	92.43	6.47	9.83	58.7

In terms of screening performance, ViT-ANN achieves the highest overall accuracy on RaptorMaps (Acc. 94.12%, F1 93.75%). Notably, NAS-FSPM remains competitive rather than being disadvantaged: it attains 92.67% accuracy and 92.43% F1-score, ranking between the two recent deep baselines and achieves higher scores than RepAlexSolarNet on all four screening metrics. These results suggest that the proposed screening design preserves robust discriminative capability on a larger and more diverse benchmark, rather than relying on dataset-specific patterns.

More importantly, NAS-FSPM is explicitly optimized for edge deployment, where energy consumption and thermal dissipation are key bottlenecks for long-duration field operation. The on-device profiling results

provide results show: NAS-FSPM yields the lowest average power (6.47 W), lowest peak power (9.83 W), and lowest steady-state temperature (58.7°C) among all compared methods. Compared with ViT-ANN, NAS-FSPM reduces average power by 39.1%, peak power by 40.1%, and steady-state temperature by 12.6°C, while maintaining strong screening accuracy.

Table 5 provides a condition-specific robustness evaluation on RaptorMaps by isolating two representative environmental factors, shadowing ($N = 1056$) and soiling ($N = 204$), both of which are known to degrade the discriminability of IR module patterns. The results indicate that performance varies across conditions, with soiling generally exhibiting a higher false-negative tendency than shadowing, reflecting the stronger visual ambiguity introduced by surface debris and dust.

Table 5: Subset-wise screening robustness under shadowing and soiling on RaptorMaps.

Method	Shadowing N	Shadowing TPR (%)	Shadowing FNR (%)	Soiling N	Soiling TPR (%)	Soiling FNR (%)	Avg. FNR (%)
ViT-ANN	1056	93.40	6.60	204	91.18	8.82	7.71
RepAlexSolarNet	1056	90.95	9.05	204	87.75	12.25	10.65
NAS-FSPM	1056	92.10	7.90	204	89.46	10.54	9.22

Overall, ViT-ANN delivers the strongest robustness on both subsets, achieving TPRs of 93.40% (shadowing) and 91.18% (soiling), and the lowest average FNR of 7.71%. Importantly, NAS-FSPM remains consistently competitive under these adverse conditions, with TPRs of 92.10% and 89.46% on shadowing and soiling, respectively. Compared with RepAlexSolarNet, NAS-FSPM reduces the shadowing FNR from 9.05% to 7.90% and the soiling FNR from 12.25% to 10.54%, yielding a lower average FNR (9.22% vs. 10.65%). These subset-wise results suggest that the proposed screening design maintains stable detection capability when faced with partial shading and dust accumulation, which underscores its applicability to field PV inspection scenarios where such environmental effects are prevalent.

To further examine how the screening branch responds to IR inputs, we visualize Gradient-weighted Class Activation Maps (Grad-CAM) on representative RaptorMaps samples. The resulting heat maps are aligned with the corresponding grayscale images, highlighting the regions that contribute most to the abnormal prediction in the classifier head. As illustrated in Fig. 3 (10 abnormal and 2 normal cases), abnormal modules typically exhibit concentrated activations around thermally salient patterns (e.g., hotspot-like blobs or elongated abnormal streaks), whereas normal samples yield comparatively weak and diffuse responses. This visualization provides an intuitive complement to the quantitative results, indicating that the decision-making process is primarily driven by localized anomaly-related cues rather than spurious global statistics.

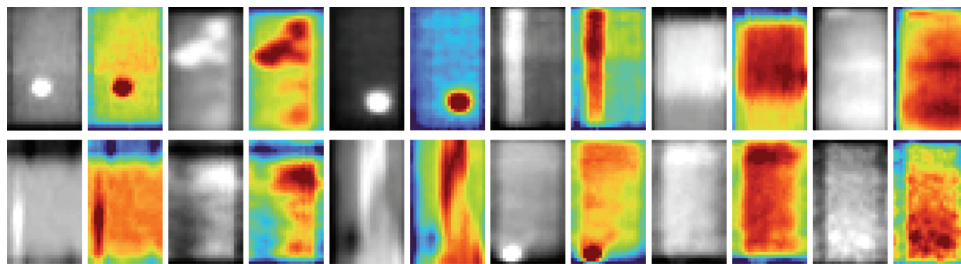


Figure 3: Heat map visualization of the NAS-FSPM for representative RaptorMaps IR samples.

4.4 Ablation Study

4.4.1 Impact of Individual Components

This section presents a strict ablation study to analyze the contribution of individual components within the proposed NAS-FSPM framework. All ablation variants are constructed by modifying or disabling specific modules while keeping the rest of the pipeline unchanged. There variants are considered:

(1) NAS-FSPM (Pixel-Space), where the learned feature representation is removed and SSD-based sub-pixel matching is performed directly in the pixel domain. (2) NAS-FSPM (Fixed-Feature), where the NAS-optimized feature extractor is replaced by a manually designed lightweight CNN with comparable parameter size. (3) NAS-FSPM (w/o RBF Gate), where the RBF-based screening module is disabled and localization is executed on all samples. (4) **The complete NAS-FSPM framework.**

Comparing NAS-FSPM (Pixel-Space) with the full model reveals the impact of moving sub-pixel matching from the pixel domain to the feature space. As shown in [Table 6](#), operating in the pixel domain results in a substantially higher average localization error (1.45 px) and larger error variance (0.60), together with an increased maximum error (2.98 px). These results confirm that raw pixel intensities are highly sensitive to thermal noise and background interference in PV-IR images. In contrast, feature-space matching significantly improves both accuracy and robustness, forming the primary source of performance gains in the proposed method.

Table 6: Ablation study on different NAS-FSPM variants.

Evaluation Metrics	NAS-FSPM (Pixel-Space)	NAS-FSPM (Fixed-Feature)	NAS-FSPM (w/o RBF Gate)	NAS-FSPM
Avg. Error (px)	1.45	1.03	0.71	0.71
Median Error (px)	1.33	0.92	0.65	0.65
Error Variance	0.60	0.36	0.21	0.21
Max Error (px)	2.98	2.12	1.34	1.34
Avg. Runtime (ms/image)	19.1	29.4	34.1	22.9
Runtime Std. (ms)	3.2	4.3	5.0	3.4

To isolate the benefit of NAS, the fixed-feature variant is compared with the full NAS-FSPM. Although both variants perform sub-pixel matching in the feature space, replacing the NAS-optimized extractor with a manually designed lightweight CNN leads to higher average and median errors (1.03 px and 0.92 px, respectively) and a larger maximum error (2.12 px). This performance degradation indicates that manually designed lightweight architectures are suboptimal for this task, even when their parameter size is constrained. By contrast, NAS automatically discovers task-specific and resource-aware feature extractors that better balance representational capacity and robustness, resulting in improved localization accuracy and stability.

Disabling the RBF-based screening module does not affect localization accuracy, as evidenced by identical average, median, variance, and maximum errors for NAS-FSPM (w/o RBF Gate) and the full model. However, the computational cost increases significantly when the gate is removed. As reported in [Table 6](#), the average runtime rises from 22.9 to 34.1 ms/image, and the runtime standard deviation increases from 3.4 to 5.0 ms. This demonstrates that conditional execution not only reduces average latency but also mitigates runtime fluctuations, which is important for stable real-time processing on edge devices.

Overall, the ablation results demonstrate that each component of NAS-FSPM plays a distinct and complementary role: (1) Feature-space representation is the dominant factor for improving localization

accuracy and robustness. (2) NAS-based optimization further enhances stability by tailoring the feature extractor to the task and resource constraints. (3) And RBF-based conditional execution significantly improves computational efficiency while maintaining reliable fault identification. These findings validate the design choices of NAS-FSPM as a coherent and effective edge-oriented framework for PV fault localization.

4.4.2 Sensitivity Analysis of τ and L

To further characterize the robustness of NAS-FSPM under realistic industrial IR noise, we conduct a sensitivity analysis on two key hyperparameters: the RBF decision threshold τ and the recursion depth L used in feature-space sub-pixel refinement. Since IR inspection is often affected by stochastic sensor noise and environmental disturbances, we emulate different noise conditions by injecting additive noise with controlled signal-to-noise ratio (SNR). Here, SNR is used as a standard noise-control index, where a lower SNR indicates a noisier acquisition condition. In the following experiments, we vary τ and L over practical ranges and report their effects on both localization performance and conditional-execution behavior.

Fig. 4 presents sensitivity heatmaps on the τ - L plane under different SNR levels. Each heatmap visualizes the localization error obtained by the full pipeline when scanning τ and L . Two consistent trends can be observed. First, the low-error region is generally not a symmetric smooth bowl but an irregular band-like area, indicating that the interaction between the gate strictness and recursive refinement depth is nontrivial in noisy IR imagery. Second, as SNR decreases, the low-error region becomes narrower and shifts toward more conservative settings, implying that the feasible hyperparameter space that maintains stable sub-pixel matching shrinks under severe noise. Overall, Fig. 4 provides a global view that supports the existence of a stable operating zone rather than a single fragile optimum.

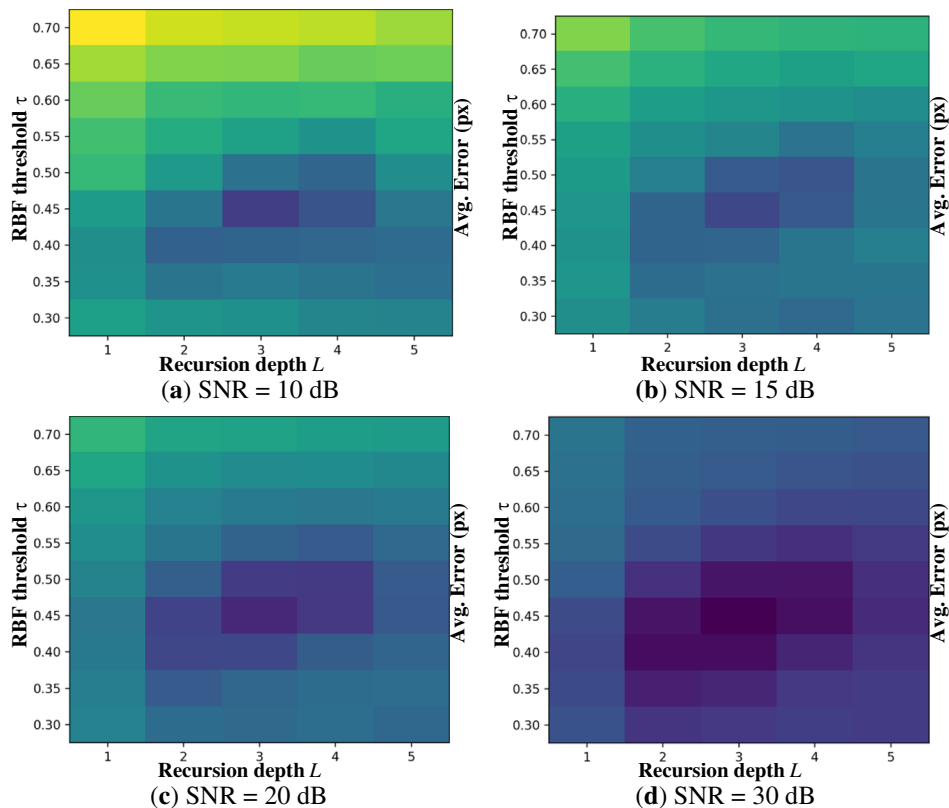


Figure 4: Sensitivity heatmaps of τ and L under varying SNR.

Fig. 5 analyzes τ by fixing L and plotting two complementary curves under varying SNR: fault recall and gate trigger ratio as functions of τ . The fault recall curve reflects whether the decision gate tends to miss faulty samples when the threshold becomes strict, while the trigger ratio indicates how frequently the downstream localization branch is activated. As expected, increasing τ reduces the trigger ratio, which improves average efficiency by skipping more samples, but an overly large τ may cause a noticeable recall drop, particularly under low SNR where the RBF output distribution becomes less separable. In contrast, smaller τ values typically preserve high recall but may increase unnecessary triggers, leading to higher average computation. The results in Fig. 5 thus reveal a clear reliability-efficiency trade-off governed by τ , and suggest that using a moderate threshold yields stable recall while keeping conditional execution effective.

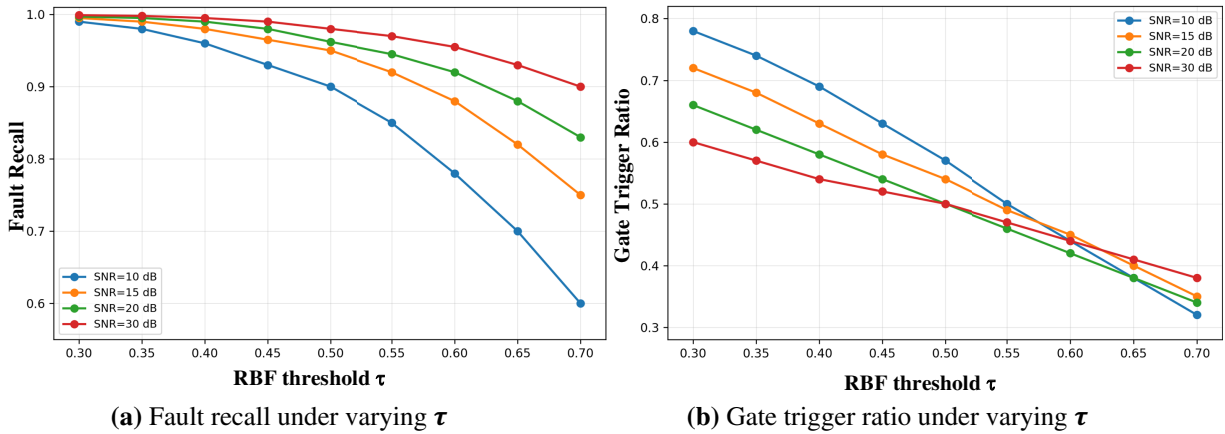


Figure 5: Sensitivity of the RBF threshold τ under varying SNR.

Fig. 6 evaluates the recursion depth by fixing τ and plotting (i) average localization error under varying SNR and (ii) runtime overhead as L increases. Fig. 6a shows that deeper recursion generally reduces localization error because the matching region is progressively refined around the current estimate; however, the improvement exhibits diminishing returns, error reduction becomes marginal, especially under higher SNR where the matching landscape is already stable. Fig. 6b shows that runtime increases monotonically with L due to the additional refinement iterations, providing an explicit cost curve for selecting L under edge constraints. Together, Fig. 6 indicates that a moderate recursion depth can achieve most of the attainable accuracy gain while keeping overhead controllable, which aligns with the edge-oriented design objective of NAS-FSPM.

4.5 Discussion

Although the proposed NAS-FSPM framework achieves strong localization accuracy and favorable edge-oriented efficiency across both the in-house PV-IR dataset and the public RaptorMaps benchmark, several practical considerations merit discussion. First, the NAS-optimized lightweight feature extraction module is searched and validated under the data distributions and sensing conditions represented in our training/evaluation sets. In real PV inspection deployments, however, the input distribution may shift due to previously unseen PV panel types (e.g., different cell layouts, encapsulation materials, or module textures), sensor variations (e.g., camera response, calibration, optics), and environmental conditions not covered by the training data (e.g., extreme irradiance/temperature regimes, snow/water films, and complex background thermal interference). While feature-space sub-pixel matching provides a degree of robustness by reducing sensitivity to raw-pixel noise, the overall performance can still be affected when the learned feature

representation becomes less discriminative under such domain shifts. This suggests that broader cross-site validation and domain-robust training strategies remain important for further improving generalizability.

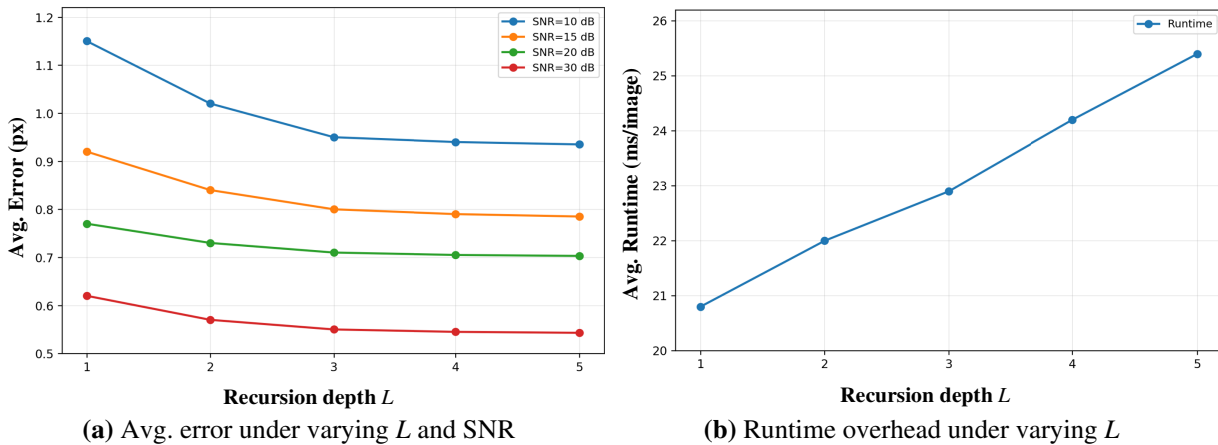


Figure 6: Sensitivity of the recursion depth L under varying SNR.

Second, NAS-FSPM relies on an RBF-based lightweight fault identification module to enable conditional execution, thereby saving computation by only activating the more expensive feature extraction and localization stages for suspicious samples. The RBF component is intentionally designed to be shallow and data-efficient, which is advantageous under limited labeled industrial data. Nevertheless, limited supervision may increase the risk of reduced reliability under long-term deployment, especially when the inspection data distribution drifts over time (e.g., camera aging, maintenance cycles, changing operating conditions). In practice, this motivates lightweight reliability measures such as periodic re-calibration when new labeled samples become available, incremental updating with conservative regularization, and confidence-aware screening (e.g., forwarding uncertain samples to the localization stage by default) to avoid missed faults.

Finally, while our on-device profiling demonstrates the feasibility of edge deployment on representative embedded hardware, deployment constraints can vary across platforms and operating budgets. Future work will therefore include expanding evaluation to a wider range of PV types and field conditions, as well as investigating more robust and maintainable updating mechanisms for the RBF-based screening module under limited annotation and distribution drift.

5 Conclusion

This paper presented NAS-FSPM, an edge-oriented photovoltaic fault localization framework that integrates a NAS-optimized lightweight feature extraction module with feature-space sub-pixel matching to achieve accurate and efficient defect localization from PV infrared imagery. By coupling a data-efficient RBF-based lightweight fault identification module with a conditional execution strategy, NAS-FSPM avoids unnecessary computation on non-suspicious samples and provides a practical balance between localization accuracy and on-device efficiency. Extensive experiments on both our in-house PV-IR dataset and the public RaptorMaps benchmark demonstrate consistent improvements over representative baselines, while the embedded profiling results further confirm the feasibility of deployment on resource-constrained platforms.

Future work will focus on validating the generalizability of NAS-FSPM under broader real-world conditions, including previously unseen PV panel types and environmental scenarios not covered by the current training data. We will also investigate more robust long-term maintenance strategies for the

RBF-based screening module under limited labeled industrial data, such as periodic re-calibration, incremental updating, and confidence-aware mechanisms to better handle distribution drift during continuous field operation.

Acknowledgement: We thank all the members who have contributed to this work with us.

Funding Statement: This work was supported by the Key R & D Projects of Liaoning Provincial Department of Science and Technology: Research on Fault Monitoring and Catastrophe Prediction Technologies for New Energy Power Stations Oriented to Wind-Solar-Storage Complementary Systems (2024JH2/102500074).

Author Contributions: Conceptualization: Hongjiang Wang; methodology: Hongjiang Wang, Jian Yu, Tian Zhang; formal analysis: Hongjiang Wang, Na Ren, Nan Zhang; writing-original draft preparation: Hongjiang Wang; writing review and editing: Hongjiang Wang, Zhenyu Liu. All authors reviewed and approved the final version of the manuscript.

Availability of Data and Materials: The data used to support the findings of this study are available from the corresponding author upon request.

Ethics Approval: Not applicable.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Islam M, Rashel MR, Ahmed MT, Islam AKMK, Tlemçani M. Artificial intelligence in photovoltaic fault identification and diagnosis: a systematic review. *Energies*. 2023;16(21):7417. doi:10.3390/en16217417.
2. Thakfan A, Bin Salamah Y. Artificial-intelligence-based detection of defects and faults in photovoltaic systems: a survey. *Energies*. 2024;17(19):4807. doi:10.3390/en17194807.
3. Wang D, Jiang Y, Wang W, Wang Y. Bias reduction in sub-pixel image registration based on the anti-symmetric feature. *Meas Sci Technol*. 2016;27(3):035206. doi:10.1088/0957-0233/27/3/035206.
4. Wu S, Zeng W, Chen H. A sub-pixel image registration algorithm based on SURF and M-estimator sample consensus. *Pattern Recognit Lett*. 2020;140:261–6. doi:10.1016/j.patrec.2020.09.031.
5. Haidari P, Hajiahmad A, Jafari A, Nasiri A. Deep learning-based model for fault classification in solar modules using infrared images. *Sustain Energy Technol Assess*. 2022;52:102110. doi:10.1016/j.seta.2022.102110.
6. Wang Y, Shen L, Li M, Sun Q, Li X. PV-YOLO: lightweight YOLO for photovoltaic panel fault detection. *IEEE Access*. 2023;11:10966–76. doi:10.1109/access.2023.3240894.
7. Chen S, Wen H, Wu J, Lei W, Hou W, Liu W, et al. Internet of things based smart grids supported by intelligent edge computing. *IEEE Access*. 2019;7:74089–102. doi:10.1109/access.2019.2920488.
8. DeTone D, Malisiewicz T, Rabinovich A. SuperPoint: self-supervised interest point detection and description. In: *Proceedings of the 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)*; 2018 Jun 18–22; Salt Lake City, UT, USA. p. 224–36. doi:10.1109/CVPRW.2018.00060.
9. Truong P, Danelljan M, Timofte R. GLU-Net: global-local universal network for dense flow and correspondences. In: *Proceedings of the 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*; 2020 Jun 13–19; Seattle, WA, USA. p. 6257–67. doi:10.1109/cvpr42600.2020.00629.
10. Ma L, Li N, Yu G, Geng X, Cheng S, Wang X, et al. Pareto-wise ranking classifier for multiobjective evolutionary neural architecture search. *IEEE Trans Evol Computat*. 2024;28(3):570–81. doi:10.1109/tevc.2023.3314766.
11. Ma L, Kang H, Yu G, Li Q, He Q. Single-domain generalized predictor for neural architecture search system. *IEEE Trans Comput*. 2024;73(5):1400–13. doi:10.1109/tc.2024.3365949.
12. Chu X, Zhou T, Zhang B, Li J. Fair DARTS: eliminating unfair advantages in differentiable architecture search. In: *Computer vision—ECCV 2020*. Cham, Switzerland: Springer International Publishing; 2020. p. 465–80. doi:10.1007/978-3-030-58555-6_28.

13. Xiong Q, Gattozzi AL, Feng X, Penney CE, Zhang C, Ji S, et al. Development of a fault detection and localization algorithm for photovoltaic systems. *IEEE J Photovoltaics*. 2023;13(6):958–67. doi:10.1109/jphotov.2023.3306073.
14. AL-Jubori HN, AL-Darraji I, Jerbi H. Defect detection using thermography camera techniques: a review. *Alkej*. 2024;20(4):70–88. doi:10.22153/kej.2024.03.002.
15. Alabsi M, Liao Y, Nabulsi AA. Bearing fault diagnosis using deep learning techniques coupled with handcrafted feature extraction: a comparative study. *J Vib Control*. 2021;27(3–4):404–14. doi:10.1177/1077546320929141.
16. Abid A, Khan MT, Iqbal J. A review on fault detection and diagnosis techniques: basics and beyond. *Artif Intell Rev*. 2021;54(5):3639–64. doi:10.1007/s10462-020-09934-2.
17. Arnaudo E, Blanco G, Monti A, Bianco G, Monaco C, Pasquali P, et al. A comparative evaluation of deep learning techniques for photovoltaic panel detection from aerial images. *IEEE Access*. 2023;11:47579–94. doi:10.1109/access.2023.3275435.
18. Yousif H, Al-Milaji Z. Fault detection from PV images using hybrid deep learning model. *Sol Energy*. 2024;267:112207. doi:10.1016/j.solener.2023.112207.
19. Montagnon T, Giffard-Roisin S, Dalla Mura M, Marchandon M, Pathier E, Hollingsworth J. Sub-pixel displacement estimation with deep learning: application to optical satellite images containing sharp displacements. *J Geophys Res Mach Learn Comput*. 2024;1(4):e2024JH000174. doi:10.1029/2024JH000174.
20. Chen T, Zhang X, Hamann B, Wang D, Zhang H. A multi-level feature integration network for image inpainting. *Multimed Tools Appl*. 2022;81(27):38781–802. doi:10.1007/s11042-022-13028-2.
21. Yu D, Liu X, Ning J, Wang S, Zhu C, Zhao W. Deep reinforcement learning-based AI task offloading in resource-constrained IIoT computing environments. *IEEE Internet Things J*. 2025;12(24):54256–73. doi:10.1109/JIOT.2025.3620126.
22. Akubathini P, Chouksey S, Satheesh HS. Evaluation of machine learning approaches for resource constrained IIoT devices. In: *Proceedings of the 2021 13th International Conference on Information Technology and Electrical Engineering (ICITEE)*; 2021 Oct 14–15; Chiang Mai, Thailand. p. 74–9. doi:10.1109/icitee53064.2021.9611880.
23. Özüpak Y. Real-time detection of photovoltaic module faults using a hybrid machine learning model. *Sol Energy*. 2025;302:114014. doi:10.1016/j.solener.2025.114014.
24. Atkinson PM. Sub-pixel target mapping from soft-classified, remotely sensed imagery. *Photogramm Eng Remote Sensing*. 2005;71(7):839–46. doi:10.14358/pers.71.7.839.
25. Xuan Y, Zhang S, Chen L, Zhang H. Improved interpolation with sub-pixel relocation method for strong barrel distortion. *Signal Process*. 2023;203:108795. doi:10.1016/j.sigpro.2022.108795.
26. Ma J, Jiang X, Fan A, Jiang J, Yan J. Image matching from handcrafted to deep features: a survey. *Int J Comput Vis*. 2021;129(1):23–79. doi:10.1007/s11263-020-01359-2.
27. Stuhlsatz A, Lippel J, Zielke T. Feature extraction with deep neural networks by a generalized discriminant analysis. *IEEE Trans Neural Netw Learning Syst*. 2012;23(4):596–608. doi:10.1109/tnnls.2012.2183645.
28. De K, Pedersen M. Impact of colour on robustness of deep neural networks. In: *Proceedings of the 2021 IEEE/CVF International Conference on Computer Vision Workshops (ICCVW)*; 2021 Oct 11–17; Montreal, BC, Canada. p. 21–30. doi:10.1109/iccvw54120.2021.00009.
29. Al-Jarrah OY, Shatnawi AS, Shurman MM, Ramadan OA, Muhaidat S. Exploring deep learning-based visual localization techniques for UAVs in GPS-denied environments. *IEEE Access*. 2024;12:113049–71. doi:10.1109/ACCESS.2024.3440064.
30. Liu HI, Galindo M, Xie H, Wong LK, Shuai HH, Li YH, et al. Lightweight deep learning for resource-constrained environments: a survey. *ACM Comput Surv*. 2024;56(10):1–42. doi:10.1145/3657282.
31. Mustafa Abro GE, Ali A, Ali Memon S, Din Memon T, Khan F. Strategies and challenges for unmanned aerial vehicle-based continuous inspection and predictive maintenance of solar modules. *IEEE Access*. 2024;12:176615–29. doi:10.1109/ACCESS.2024.3505754.
32. Aghaei M, Kolahi M, Nedaei A, Venkatesh NS, Esmailifar SM, Moradi Sizkouhi AM, et al. Autonomous intelligent monitoring of photovoltaic systems: an in-depth multidisciplinary review. *Progress Photovoltaics*. 2025;33(3):381–409. doi:10.1002/pip.3859.

33. Mishra R, Gupta H. Transforming large-size to lightweight deep neural networks for IoT applications. *ACM Comput Surv.* 2023;55(11):1–35. doi:10.1145/3570955.
34. Uzel H, Özüpak Y, Alpsalaz F, Aslan E. Optimized ANN-RF hybrid model with optuna for fault detection and classification in power transmission systems. *Sci Rep.* 2026;16:1495. doi:10.1038/s41598-025-31008-y.
35. Ren P, Xiao Y, Chang X, Huang PY, Li Z, Chen X, et al. A comprehensive survey of neural architecture search: challenges and solutions. *ACM Comput Surv.* 2022;54(4):1–34. doi:10.1145/3447582.
36. Ma L, Li N, Zhu P, Tang K, Khan A, Wang F, et al. A novel fuzzy neural network architecture search framework for defect recognition with uncertainties. *IEEE Trans Fuzzy Syst.* 2024;32(5):3274–85. doi:10.1109/tfuzz.2024.3373792.
37. Huang B, Abtahi A, Aminifar A. Energy-aware integrated neural architecture search and partitioning for distributed Internet of Things (IoT). *IEEE Trans Circuits Syst Artif Intell.* 2024;1(2):257–71. doi:10.1109/TCASAI.2024.3493036.
38. Zheng C, Wu W, Chen C, Yang T, Zhu S, Shen J, et al. Deep learning-based human pose estimation: a survey. *ACM Comput Surv.* 2024;56(1):1–37. doi:10.1145/3603618.
39. Wang Y, Shen J, Hu TK, Xu P, Nguyen T, Baraniuk R, et al. Dual dynamic inference: enabling more efficient, adaptive, and controllable deep inference. *IEEE J Sel Top Signal Process.* 2020;14(4):623–33. doi:10.1109/jstsp.2020.2979669.
40. Millendorf M, Obropta E, Vadhavkar N. Infrared solar module dataset for anomaly detection. In: *Proceedings of the Eighth International Conference on Learning Representations*; 2020 Apr 26–May 1; Online. 101545 p.
41. Pan B. Digital image correlation for surface deformation measurement: historical developments, recent advances and future goals. *Meas Sci Technol.* 2018;29(8):082001. doi:10.1088/1361-6501/aac55b.
42. Sarlin PE, DeTone D, Malisiewicz T, Rabinovich A. SuperGlue: learning feature matching with graph neural networks. In: *Proceedings of the 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*; 2020 Jun 13–19; Seattle, WA, USA. p. 4937–46. doi:10.1109/cvpr42600.2020.00499.
43. Cai H, Gan C, Wang T, Zhang Z, Han S. Once-for-all: train one network and specialize it for efficient deployment. arXiv:1908.09791. 2019. Available from: <https://arxiv.org/abs/1908.09791>.
44. Zhang Y, Zhang X, Tu D. Solar photovoltaic module defect detection based on deep learning. *Meas Sci Technol.* 2024;35(12):125404. doi:10.1088/1361-6501/ad7d28.
45. Ramadan EA, Moawad NM, Abouzalm BA, Sakr AA, Abouzaid WF, El-Banby GM. An innovative transformer neural network for fault detection and classification for photovoltaic modules. *Energy Convers Manag.* 2024;314:118718. doi:10.1016/j.enconman.2024.118718.
46. Guo J, Chong CF, Abreu PH, Mao C, Li J, Lam CT, et al. Reparameterization convolutional neural networks for handling imbalanced datasets in solar panel fault classification. *Eng Appl Artif Intell.* 2025;150:110541. doi:10.1016/j.engappai.2025.110541.