



The University of Tehran Press

Intelligent Detection of Surface Defects in Solar Panels Using Deep Learning-Based Computer Vision

Ali Jelokhani 

School of Electrical and Computer Engineering, University of Tehran, Tehran, Iran. Email: Ali.jelokhani79@gmail.com

ARTICLE INFO

Article type:
Research Paper

Article History:

Received: 19 February 2026

Revised: 16 March 2026

Accepted: 22 May 2026

Published Online: 23 September 2026

Keywords:

Solar panel,
Solar panel defect detection,
Computer vision,
Deep learning.

ABSTRACT

With the increasing use of photovoltaic systems, monitoring the health of solar panels has become an important issue for sustainable operation, reducing maintenance costs, and improving energy generation efficiency. Defects such as breakage, dust, shadow, and surface variations can reduce panel performance and shorten its service life. Traditional inspection methods are generally time-consuming, dependent on human labor, and inefficient on a large scale. In this study, a computer vision and deep learning-based framework is proposed for detecting and localizing surface regions in solar panel images. The proposed model consists of a ResNet101 feature extractor, a Sobel operator-based edge extraction module, a spatial-channel attention module, an FPN/PAN multi-scale feature fusion structure, and a multi-scale detection head. The model output includes bounding boxes, region classes, and confidence scores; therefore, it can simultaneously identify multiple different regions in a single image. The dataset used in this study was constructed by combining images collected from public sources with manually collected data and includes four classes: breakage, clean region, dust, and shadow. The experimental results showed that the proposed model, achieving a Precision of 0.954, Recall of 0.939, F1-score of 0.946, mAP@0.5 of 0.966, and mAP@0.5:0.95 of 0.766, outperformed baseline models and recent photovoltaic defect detection methods. Furthermore, the ablation study results showed that combining edge information, spatial-channel attention, and multi-scale feature fusion plays an important role in improving detection and localization accuracy. Based on the results, the proposed method can be used as a practical solution for automatic monitoring, targeted cleaning, and predictive maintenance of solar panels.

Cite this article: Jelokhani, A. (2026). Intelligent Detection of Surface Defects in Solar Panels Using Deep Learning-Based Computer Vision. *Journal of Sustainable Energy Systems*, 5 (4), 665-689. DOI: <http://doi.org/10.22059/ses.2026.414893.1244>



© Authors retain the copyright and full publishing rights.
DOI: <http://doi.org/10.22059/ses.2026.414893.1244>

Publisher: University of Tehran Press.

Background and Objective

The rapid growth of photovoltaic systems has increased the need for reliable and automatic inspection of solar panels. Since solar panels are usually installed in outdoor environments, they are continuously exposed to dust, shadowing, illumination changes, surface contamination, breakage, and other physical or environmental effects. These conditions can reduce power generation, shorten module lifetime, and increase maintenance costs. Traditional inspection methods, such as manual visual inspection, electrical measurements, thermal imaging, and electroluminescence imaging, can be effective in specific situations; however, they often require expert operators, specialized equipment, controlled imaging conditions, or time-consuming analysis. These limitations make them less practical for large-scale and continuous monitoring of photovoltaic systems.

The objective of this study is to develop a computer vision and deep learning-based framework for detecting and localizing surface regions in solar panel images. Unlike image-level classification

methods that only determine the overall condition of an image, the proposed approach performs multi-class and multi-instance detection. Therefore, it can identify several regions in a single image and assign each region a bounding box, class label, and confidence score. This capability is important in real inspection scenarios, where a solar panel image may simultaneously contain breakage, dust, shadow, and clean regions.

Proposed Method

The proposed architecture is designed as a multi-scale object detection framework. The model receives an RGB image as input and produces a set of bounding boxes as output. Each detected box represents a surface region and includes its spatial coordinates, predicted class, and confidence score. The considered classes are breakage, clean region, dust, and shadow.

The model consists of several complementary components. First, input images are resized and normalized, and data augmentation operations such as rotation, brightness adjustment, contrast variation, scaling, shifting, and horizontal flipping are applied to improve generalization. Second, a Sobel-based edge extraction module is used to enhance structural information. Since defects such as cracks and breakage may appear as narrow and low-contrast patterns, edge information provides useful cues for detecting these regions. Third, ResNet101 is employed as the main feature extractor to learn deep hierarchical representations from the input image. Its residual connections help preserve gradient flow and support effective learning of complex visual patterns.

In addition, a spatial-channel attention module is integrated to emphasize informative feature channels and important spatial regions while reducing the influence of irrelevant background patterns, repeated solar cell grid lines, and reflections. To handle defects at different scales, an FPN/PAN-based multi-scale feature fusion structure is used. This component combines semantic deep features with spatially detailed shallow features, enabling the model to detect both small regions such as cracks and larger regions such as dust or shadow. Finally, a multi-scale detection head predicts bounding box coordinates, class labels, and confidence scores. During post-processing, low-confidence predictions are removed, and non-maximum suppression is applied to eliminate duplicate detections.

Dataset and Experimental Setup

The dataset used in this study was constructed by combining images collected from public sources and manually gathered samples. This strategy increased the diversity of the data in terms of illumination, viewpoint, panel appearance, shadow patterns, dust distribution, and surface conditions. Some images already contained annotations, while the remaining samples were manually annotated or corrected using bounding boxes. The dataset includes four classes: breakage, clean region, dust, and shadow. Although the clean region class is also annotated using bounding boxes, the main focus of the study is the detection and localization of surface regions relevant to inspection and maintenance.

The dataset was divided into training, validation, and test subsets using an 80:10:10 ratio. The proposed model was implemented in Python using the PyTorch framework and trained on a system equipped with an NVIDIA RTX 4060 GPU. The Adam optimizer was used with an initial learning rate of 1×10^{-4} , a batch size of 8, and 200 training epochs. The total loss function consisted of bounding box localization loss, classification loss, and objectness loss. For localization, CIoU loss was used because it considers overlap, center distance, and aspect-ratio consistency between predicted and ground-truth boxes.

Evaluation Metrics

The model was evaluated using standard object detection metrics, including Precision, Recall, F1-score, mAP@0.5, and mAP@0.5:0.95. Precision measures the proportion of correct detections among all predicted detections, while Recall measures the proportion of actual regions successfully detected by the model. F1-score provides a balanced measure between Precision and Recall. The mAP@0.5 metric evaluates detection performance at an IoU threshold of 0.5, while mAP@0.5:0.95 provides a stricter evaluation by averaging mAP values across IoU thresholds from 0.5 to 0.95.

Results and Discussion

The proposed model achieved strong performance on the test set, with a Precision of 0.954, Recall of

0.939, F1-score of 0.946, mAP@0.5 of 0.966, and mAP@0.5:0.95 of 0.766. These results show that the model provides a suitable balance between detection accuracy and recall. The high mAP@0.5 value indicates strong overall detection capability, while the mAP@0.5:0.95 value confirms acceptable localization accuracy under stricter IoU thresholds.

Class-wise evaluation showed stable performance across all classes. The clean region class achieved the highest performance due to its more uniform visual appearance. Breakage was more challenging because broken or cracked regions are often narrow, low-contrast, and visually similar to the grid lines of solar cells. Dust and shadow were also challenging because their appearance depends on illumination, viewpoint, and surface conditions. Nevertheless, the proposed model successfully distinguished these classes with high overall accuracy.

The proposed method was compared with several baseline and recent detection models, including Faster R-CNN, SSD, RetinaNet, DETR, YOLOv11n, YOLO-PV, PV-YOLOv12n, MRA-YOLOv8, YOLOv8-FSD, BiFDRRep-YOLOv8n, DCD-YOLOv8s, and an improved YOLO11 model. The comparison results showed that the proposed model achieved the best performance across all evaluation metrics. This improvement demonstrates the effectiveness of combining deep feature extraction, edge information, spatial-channel attention, and multi-scale feature fusion for detecting surface regions in solar panel images.

An ablation study was also conducted to evaluate the contribution of each component. Removing the FPN/PAN feature fusion structure caused the largest performance degradation, confirming the importance of multi-scale representation. Removing the Sobel-based edge module reduced the ability to detect narrow and boundary-based regions, especially breakage. Removing the spatial-channel attention module also decreased performance by weakening the model's focus on informative regions. These results confirm that each component contributes to the final detection accuracy.

Conclusion

This study presented a deep learning-based computer vision framework for intelligent detection and localization of surface regions in solar panel images. The proposed model combines ResNet101 feature extraction, Sobel-based edge enhancement, spatial-channel attention, FPN/PAN multi-scale feature fusion, and a multi-scale detection head. The model can detect multiple regions in a single image and assign each region a bounding box, class label, and confidence score.

The experimental results demonstrate that the proposed method outperforms baseline detectors and recent photovoltaic defect detection models. The proposed framework can be used as a practical tool for automatic monitoring, targeted cleaning, inspection prioritization, and predictive maintenance of solar panel systems. Future work can focus on expanding the dataset, incorporating drone-based or thermal images, developing lighter versions for edge devices, and evaluating the method in real-time field inspection scenarios.



انتشارات دانشگاه تهران

فصلنامه سیستم‌های انرژی پایدار

سایت نشریه: <https://ses.ut.ac.ir>

شاپا الکترونیک: ۸۶۹۳-۲۹۸۰

تشخیص هوشمند عیوب سطحی پنل‌های خورشیدی با رویکرد بینایی ماشین مبتنی بر یادگیری عمیق

علی جلوخانی

دانشکده کامپیوتر، دانشکده فنی دانشگاه تهران، رایانامه: Ali.jelokhani@gmail.com

چکیده

اطلاعات مقاله

نوع مقاله:

پژوهشی

تاریخ‌های مقاله:

تاریخ دریافت: ۱۴۰۴/۱۱/۳۰

تاریخ بازنگری: ۱۴۰۴/۱۲/۲۵

تاریخ پذیرش: ۱۴۰۵/۰۳/۰۱

تاریخ انتشار: ۱۴۰۵/۰۷/۰۱

کلیدواژه:

پنل خورشیدی،

تشخیص عیوب پنل‌های خورشیدی،

بینایی ماشین،

یادگیری عمیق .

با گسترش استفاده از سامانه‌های فتوولتائیک، پایش سلامت پنل‌های خورشیدی به یکی از موضوعات مهم در بهره‌برداری پایدار، کاهش هزینه‌های نگهداری و افزایش بازده تولید انرژی تبدیل شده است. عیوبی مانند شکستگی، گردوغبار، سایه و تغییرات سطحی می‌توانند باعث افت عملکرد پنل و کاهش عمر مفید آن شوند. روش‌های سنتی بازرسی معمولاً زمان‌بر، وابسته به نیروی انسانی و در مقیاس‌های بزرگ غیرکارآمد هستند. در این پژوهش، نوعی چارچوب مبتنی بر بینایی ماشین و یادگیری عمیق برای تشخیص و مکان‌یابی نواحی سطحی پنل‌های خورشیدی ارائه می‌شود. مدل پیشنهادی از استخراج گر ویژگی ResNet101، مازول استخراج لبه مبتنی بر عملگر سوبل، مازول توجه کانالی - فضایی، ساختار تلفیق ویژگی چندمقیاسی FPN/PAN و سر تشخیص چندمقیاسی تشکیل شده است. خروجی مدل شامل جعبه‌های محدودکننده، کلاس ناحیه و امتیاز اطمینان است و بنابراین، می‌تواند چندین ناحیه مختلف را به صورت هم‌زمان در یک تصویر شناسایی کند. مجموعه داده مورد استفاده از ترکیب تصاویر گردآوری شده از منابع عمومی و داده‌های جمع‌آوری شده دستی تشکیل شده و کلاس‌های شکستگی، ناحیه سالم، گردوغبار و سایه را شامل می‌شود. نتایج آزمایش‌ها نشان داد مدل پیشنهادی با دستیابی به دقت ۰/۹۵۴، بازخوانی ۰/۹۳۹، امتیاز F1 برابر با ۰/۹۴۶، مقدار mAP@0.5 برابر با ۰/۹۶۶ و مقدار mAP@0.5:0.95 برابر با ۰/۷۶۶، عملکرد بهتری نسبت به مدل‌های پایه و روش‌های جدید تشخیص عیوب فتوولتائیک دارد. همچنین، نتایج مطالعه حذف اجزا نشان داد ترکیب اطلاعات لبه‌ای، توجه کانالی - فضایی و تلفیق چندمقیاسی نقش مهمی در بهبود دقت تشخیص و مکان‌یابی دارد. بر اساس نتایج، روش پیشنهادی می‌تواند به عنوان یک راهکار کاربردی برای پایش خودکار، پاک‌سازی هدفمند، تعمیر و نگهداری پیش‌بینانه پنل‌های خورشیدی مورد استفاده قرار گیرد.

استناد: جلوخانی، علی (۱۴۰۵). تشخیص هوشمند عیوب سطحی پنل‌های خورشیدی با رویکرد بینایی ماشین مبتنی بر یادگیری عمیق. *فصلنامه سیستم‌های انرژی پایدار*، ۵ (۴) ۶۶۵-۶۸۹.

DOI: <http://doi.org/10.22059/ses.2026.414893.1244>

ناشر: مؤسسه انتشارات دانشگاه تهران.

© نویسندگان.

DOI: <http://doi.org/10.22059/ses.2026.414893.1244>



۱. مقدمه

گسترش سامانه‌های فتوولتائیک در سال‌های اخیر، پنل‌های خورشیدی را به یکی از اجزای کلیدی زیرساخت انرژی‌های تجدیدپذیر تبدیل کرده است. با وجود این، عملکرد بلندمدت این سامانه‌ها به شدت تحت تأثیر سلامت سطحی و ساختاری پنل‌ها قرار دارد. عیوبی مانند ترک‌های سطحی، شکستگی سلول، لکه، گردوغبار، آلودگی، خراش، داغ‌شدگی موضعی، تغییر رنگ و آسیب‌های مکانیکی می‌توانند باعث افت توان تولیدی، کاهش طول عمر ماژول و افزایش هزینه‌های نگهداری شوند [۱ و ۲]. از آنجا که نیروگاه‌های خورشیدی معمولاً در مقیاس وسیع و در محیط‌های باز نصب می‌شوند، پایش دستی پنل‌ها فرایندی زمان‌بر، پرهزینه و مستعد خطای انسانی است. به همین دلیل، توسعه روش‌های خودکار برای تشخیص و مکان‌یابی عیوب پنل‌های خورشیدی به یکی از موضوعات مهم در نگهداری هوشمند سامانه‌های فتوولتائیک تبدیل شده است [۱].

روش‌های سنتی بازرسی پنل‌های خورشیدی معمولاً بر مشاهده چشمی، اندازه‌گیری‌های الکتریکی، تصویربرداری حرارتی، الکترومومینسانس یا تحلیل‌های دستی متکی هستند. اگرچه این روش‌ها در برخی شرایط دقت مناسبی دارند، اما در کاربردهای عملی با محدودیت‌هایی مانند نیاز به تجهیزات تخصصی، وابستگی به کارشناس، دشواری پردازش حجم زیاد داده، حساسیت به شرایط تصویربرداری و ناتوانی در تشخیص سریع عیوب کوچک مواجه‌اند [۲ و ۳]. در مقابل، روش‌های مبتنی بر بینایی ماشین و یادگیری عمیق امکان استخراج خودکار ویژگی‌ها، تشخیص چندکلاسه عیوب، مکان‌یابی ناحیه آسیب‌دیده و پردازش تصاویر در مقیاس بالا را فراهم می‌کنند. به‌ویژه شبکه‌های کانولوشنی و مدل‌های تشخیص شیء، به دلیل توانایی در یادگیری الگوهای پیچیده تصویری، در سال‌های اخیر برای شناسایی عیوب فتوولتائیک در تصاویر RGB، حرارتی و الکترومومینسانس مورد استفاده قرار گرفته‌اند [۴].

با این حال، تشخیص عیوب سطحی پنل‌های خورشیدی همچنان مسئله‌ای چالش‌برانگیز است. بسیاری از عیوب، اندازه کوچک، مرز نامشخص، کنتراست پایین یا شباهت ظاهری زیادی با بافت طبیعی سلول و سطح پنل دارند. برای نمونه، ترک‌های ریز، خراش‌های کم‌رنگ یا لکه‌های پراکنده ممکن است در تصاویر معمولی به‌سختی از خطوط شبکه سلولی، بازتاب نور، سایه، گردوغبار یا تغییرات زاویه دید تفکیک شوند. از طرف دیگر، برخی عیوب به صورت ناحیه‌ای و گسترده ظاهر می‌شوند، در حالی که برخی دیگر بسیار کوچک و نقطه‌ای هستند. بنابراین، یک مدل کارآمد باید توانایی تشخیص عیوب کوچک و چندمقیاسه را داشته باشد و بتواند موقعیت مکانی آن‌ها را در قالب جعبه‌های محدودکننده مشخص کند [۵ و ۶].

در حوزه تشخیص عیوب فتوولتائیک، بخش قابل توجهی از مطالعات بر تصاویر الکترومومینسانس و حرارتی تمرکز داشته‌اند؛ زیرا این تصاویر می‌توانند برخی عیوب داخلی یا حرارتی را بهتر آشکار کنند. با این وجود، استفاده از چنین داده‌هایی همواره نیازمند تجهیزات تصویربرداری خاص، شرایط کنترل‌شده یا زیرساخت پرهزینه است. در بسیاری از کاربردهای عملی، به‌خصوص در بازرسی میدانی، تصاویر RGB یا تصاویر اخذشده با دوربین‌های متداول در دسترس‌تر و کم‌هزینه‌تر هستند، اما این تصاویر به دلیل تغییرات نور، بازتاب، سایه، زاویه دید، آلودگی سطح، تفاوت پس‌زمینه و کیفیت متغیر تصویر، دشواری‌های بیشتری برای مدل ایجاد می‌کنند [۷]. از این رو، طراحی سامانه‌ای که بتواند با داده‌های تصویری متداول، عیوب سطحی پنل‌های خورشیدی را با دقت قابل قبول تشخیص دهد و محل آن‌ها را مشخص کند، می‌تواند از نظر صنعتی و اقتصادی ارزشمند باشد.

از منظر معماری، مدل‌های تشخیص شیء مبتنی بر یادگیری عمیق به دلیل توانایی در شناسایی هم‌زمان نوع عیب و موقعیت مکانی آن، برای بازرسی خودکار پنل‌های خورشیدی گزینه‌ای مناسب محسوب می‌شوند. در میان این مدل‌ها، خانواده YOLO به دلیل سرعت زیاد، ساختار یک‌مرحله‌ای و قابلیت استفاده در کاربردهای نزدیک به بلادرنگ، جایگاه ویژه‌ای در مسائل تشخیص صنعتی دارد [۸]. در کنار آن، استفاده از شبکه‌های عمیق مانند ResNet برای استخراج ویژگی، ماژول‌های توجه، و راهکارهای تلفیق ویژگی چندمقیاسه می‌تواند به بهبود عملکرد مدل در برابر عیوب کوچک، کم‌کنتراست و دارای شکل نامنظم کمک کند [۹ و ۱۰]. مطالعات جدید در حوزه تشخیص عیوب فتوولتائیک نیز نشان داده‌اند بهبود معماری YOLO از طریق توجه، تلفیق ویژگی چندمقیاسه و بهینه‌سازی بخش گردن یا سر تشخیص، می‌تواند دقت مکان‌یابی و شناسایی عیوب پنل را افزایش دهد.

در پژوهش حاضر، یک چارچوب مبتنی بر بینایی ماشین و یادگیری عمیق برای تشخیص خودکار عیوب سطحی پنل‌های خورشیدی ارائه می‌شود. هدف این چارچوب، شناسایی عیوب موجود در تصویر و تعیین موقعیت آن‌ها با استفاده از جعبه‌های محدودکننده است. به خلاف روش‌هایی که تنها به طبقه‌بندی تصویر به دو حالت سالم و معیوب بسنده می‌کنند، رویکرد تشخیص شیء علاوه بر تعیین وجود عیب، محل تقریبی آن را نیز مشخص می‌سازد. این ویژگی برای کاربردهای عملی اهمیت زیادی دارد؛ زیرا دانستن محل عیب می‌تواند در تصمیم‌گیری برای تعمیر، پاک‌سازی، تعویض یا بررسی تخصصی پنل مورد استفاده قرار گیرد.

نوآوری‌ها و دستاوردهای اصلی این پژوهش به صورت زیر خلاصه می‌شوند:

۱. ارائه یک چارچوب هوشمند مبتنی بر بینایی ماشین و یادگیری عمیق برای تشخیص خودکار عیوب سطحی پنل‌های خورشیدی؛
 ۲. مکان‌یابی عیوب در قالب جعبه‌های محدودکننده به منظور فراهم‌سازی اطلاعات مکانی قابل استفاده در فرایند نگهداری و تعمیرات؛
 ۳. بهره‌گیری از رویکردهای یادگیری عمیق برای تشخیص عیوب کوچک، کم‌کنتراست و چندمقیاسه در تصاویر پنل خورشیدی؛
 ۴. ارزیابی مدل پیشنهادی با معیارهای رایج تشخیص شیء، از جمله Precision، Recall، F1-score و mAP؛
 ۵. فراهم کردن بستری برای توسعه سامانه‌های پایش خودکار و نگهداری پیش‌بینانه در نیروگاه‌ها و سامانه‌های خورشیدی.
- نوآوری اصلی این پژوهش در معرفی یک مازول منفرد جدید مانند ResNet101، سوبل یا توجه نیست؛ زیرا این اجزا به صورت مستقل در پژوهش‌های مختلف استفاده شده‌اند. سهم اصلی این پژوهش در طراحی و به‌کارگیری یک چارچوب ترکیبی و مسئله‌محور برای تشخیص چندکلاسه و چندنمونه‌ای نواحی سطحی پنل خورشیدی است. در این چارچوب، اطلاعات لبه‌ای استخراج‌شده با عملگر سوبل در دو مرحله، یعنی در سطح ورودی و در بخش تلفیق ویژگی چندمقیاسی، با ویژگی‌های عمیق ادغام می‌شود. همچنین، ترکیب استخراج‌گر عمیق ResNet101، توجه کانالی - فضایی، ساختار FPN/PAN و مازول هم‌پیچی اتساعی چندمقیاسی برای مسئله‌ای به کار گرفته شده است که در آن نواحی هدف از نظر اندازه، شکل، کنتراست و شباهت با الگوهای طبیعی سطح پنل تنوع زیادی دارند. بنابراین، تمایز اصلی روش پیشنهادی در نحوه تلفیق هدفمند این اجزا برای تشخیص و مکان‌یابی هم‌زمان نواحی شکستگی، گردوغبار، سایه و ناحیه سالم در تصاویر پنل خورشیدی است. بنابراین، پژوهش حاضر با تمرکز بر تشخیص و مکان‌یابی خودکار نواحی سطحی پنل‌های خورشیدی، تلاش می‌کند راهکاری کاربردی برای پایش هوشمند و نگهداری مؤثرتر سامانه‌های فتوولتائیک ارائه دهد.

۲. پیشینه پژوهش

۲-۱. پایش هوشمند و تشخیص عیب در سامانه‌های فتوولتائیک

پایش سلامت پنل‌های خورشیدی یکی از الزامات بهره‌برداری پایدار از سامانه‌های فتوولتائیک است. این پنل‌ها در محیط‌های باز در معرض عواملی مانند گردوغبار، رطوبت، تغییرات دمایی، سایه‌اندازی، ضربه، ترک، شکستگی شیشه، داغ‌شدگی موضعی و تخریب تدریجی مواد قرار دارند که می‌توانند موجب کاهش توان تولیدی، افزایش هزینه تعمیر و کاهش عمر مفید سامانه شوند. تانگ و همکاران [۱] نشان دادند گسترش نیروگاه‌های فتوولتائیک، نیاز به روش‌های هوشمند و خودکار برای پایش و نگهداری را جدی‌تر کرده است. همچنین، تکفان و سلامه [۲] تأکید کردند که روش‌های سنتی، مانند بازرسی دستی و اندازه‌گیری‌های الکتریکی، برای سامانه‌های بزرگ‌مقیاس کافی نیستند و استفاده از هوش مصنوعی می‌تواند در تشخیص و طبقه‌بندی خطاها نقش مهمی داشته باشد.

روش‌های رایج بازرسی پنل‌های خورشیدی شامل مشاهده چشمی، روش‌های الکتریکی، تصویربرداری الکترومینیسانس، تصویربرداری حرارتی، فوتولومینسانس، تصویربرداری RGB و روش‌های مبتنی بر پهپاد یا سامانه‌های رباتیک هستند. جلال و

همکاران [۳] نشان دادند هر یک از این روش‌ها برای آشکارسازی نوع خاصی از عیوب مناسب‌اند؛ برای نمونه، تصاویر الکترومینیسانس در تشخیص ترک‌ها و عیوب داخلی و تصاویر حرارتی در شناسایی ناهنجاری‌های حرارتی کاربرد بیشتری دارند. با وجود این، بسیاری از این روش‌ها به تجهیزات تخصصی، شرایط تصویربرداری کنترل‌شده یا نیروی متخصص نیاز دارند؛ از این‌رو پژوهش‌های جدید به سمت روش‌های یادگیری عمیق، تصاویر قابل دسترس‌تر و تشخیص شیء حرکت کرده‌اند.

۲-۲. یادگیری ماشین و یادگیری عمیق در تشخیص عیوب پنل خورشیدی

در پژوهش‌های اولیه، تشخیص خودکار عیوب پنل‌های خورشیدی بیشتر در قالب طبقه‌بندی تصویر انجام می‌شد. دیتش و همکاران [۱۱] با استفاده از تصاویر الکترومینیسانس، عملکرد روش‌های مبتنی بر ویژگی‌های دستی و شبکه‌های عصبی کانولوشنی را برای طبقه‌بندی سلول‌های معیوب مقایسه کردند و نشان دادند شبکه‌های عمیق توان بیشتری در استخراج ویژگی‌های مؤثر دارند. تانگ و همکاران [۱۲] نیز روشی مبتنی بر یادگیری عمیق برای شناسایی خودکار عیوب ماژول‌های فتوولتائیک با استفاده از تصاویر الکترومینیسانس ارائه کردند. این مطالعات از کارهای پایه‌ای در کاربرد یادگیری عمیق برای عیب‌یابی تصویری پنل‌های خورشیدی محسوب می‌شوند.

در کنار تصاویر الکترومینیسانس، تصاویر حرارتی نیز برای تشخیص داغ‌شدگی موضعی و خطاهای عملکردی استفاده شده‌اند. اکرم و همکاران [۱۳] از یادگیری عمیق برای تشخیص خودکار عیوب ماژول‌های فتوولتائیک در تصاویر مادون قرمز بهره گرفتند. با این حال، استفاده از تصاویر الکترومینیسانس و حرارتی در محیط واقعی ممکن است به دلیل هزینه تجهیزات و محدودیت شرایط تصویربرداری دشوار باشد. به همین دلیل، در سال‌های اخیر توجه بیشتری به تصاویر RGB و عیوب سطحی شده است. رحمان [۷] عیوب سطحی و آلودگی پنل‌های خورشیدی را در چند کلاس مانند پنل سالم، گردوغبار، عیب، آسیب فیزیکی و برف بررسی کرد. این رویکرد از نظر تمرکز بر داده‌های تصویری قابل دسترس‌تر به پژوهش حاضر نزدیک است، اما روش‌های طبقه‌بندی معمولاً محل دقیق عیب را مشخص نمی‌کنند و برای تعمیر، پاک‌سازی یا بررسی موضعی پنل اطلاعات مکانی کافی ارائه نمی‌دهند.

۲-۳. تشخیص شیء و اهمیت مدل‌های YOLO در عیب‌یابی پنل خورشیدی

تشخیص شیء علاوه بر تعیین نوع عیب، موقعیت آن را نیز در قالب جعبه محدودکننده مشخص می‌کند؛ بنابراین، برای کاربردهای عملی عیب‌یابی پنل خورشیدی اهمیت زیادی دارد. در میان مدل‌های تشخیص شیء، خانواده YOLO به دلیل سرعت بالا و ساختار یک‌مرحله‌ای، جایگاه ویژه‌ای در بازاری صنعتی دارد. ردمون و همکاران [۸] با معرفی YOLO، تشخیص شیء را به صورت یک مسئله رگرسیون مستقیم برای پیش‌بینی جعبه‌ها و کلاس‌ها مطرح کردند. در کنار YOLO، رن و همکاران [۱۴] با Faster R-CNN، لیو و همکاران [۱۵] با SSD، لین و همکاران [۱۶] با RetinaNet و کاربون و همکاران [۱۷] با DETR، مسیرهای مهم دیگری در توسعه مدل‌های تشخیص شیء ارائه کردند.

در حوزه فتوولتائیک نیز مدل‌های تشخیص شیء، به‌ویژه نسخه‌های مختلف YOLO، به طور گسترده استفاده شده‌اند. لانگ و لو [۴] یک آشکارساز عیب سلول فتوولتائیک را با ترکیب YOLOv8، مکانیزم توجه و ترنسفورمر ارائه کردند. پان و همکاران [۱۸] نیز مدلی مبتنی بر YOLO با تلفیق تطبیقی ویژگی‌ها برای کاهش تشخیص‌های ازدست‌رفته و خطاهای مثبت در عیوب پنل‌های فتوولتائیک معرفی کردند. جیا و همکاران [۱۹] با بهبود VarifocalNet، مسئله تعادل میان سرعت و دقت را بررسی کردند و کو و همکاران [۲۰] مدلی با توانایی استخراج دانش توپولوژیک از ساختار سلول فتوولتائیک ارائه دادند. این مطالعات نشان می‌دهند تشخیص عیوب پنل خورشیدی از مدل‌های عمومی تشخیص شیء به سمت معماری‌های تخصصی‌تر حرکت کرده است.

۲-۴. مدل‌های جدید YOLO برای تشخیص عیوب فتوولتائیک

در سال‌های ۲۰۲۴ و ۲۰۲۵، بخش مهمی از پژوهش‌ها بر بهبود نسخه‌های مختلف YOLO برای تشخیص عیوب پنل و سلول

خورشیدی متمرکز شده‌اند. کای و لو [۵] مدل YOLO-PV را بر پایه YOLO11n و تلفیق ویژگی چندمقیاسه ارائه کردند و محمد و همکاران [۶] نیز یک مدل بهینه‌شده مبتنی بر YOLO را برای تشخیص عیوب فتوولتائیک در تصاویر الکترولومینسانس معرفی کردند. این پژوهش‌ها نشان می‌دهند نسخه پایه YOLO معمولاً برای عیوب کوچک و کم‌کنتراست کافی نیست و باید با مازول‌هایی مانند توجه، گردن شبکه بهبودیافته، تلفیق ویژگی و سبک‌سازی معماری تقویت شود.

وانگ و همکاران [۲۱] مدل MRA-YOLOv8 را برای تقویت استخراج ویژگی در تشخیص عیوب سلول‌های فتوولتائیک پیشنهاد کردند. شوه و همکاران [۲۲] مدل SEC-YOLOv8 را با استفاده از ECA attention، مازول SPPELAN و CARAFE برای بهبود دقت و کارایی محاسباتی توسعه دادند. چن و همکاران [۲۳] نیز مدلی سبک‌وزن برای تشخیص عیوب سلول‌های فتوولتائیک معرفی کردند. همچنین، لو و همکاران [۲۴] تشخیص عیوب کوچک در مازول‌های فتوولتائیک را هدف قرار دادند، المه‌باشی و همکاران [۲۵] مدلی برای تشخیص عیوبی مانند گردوغبار، فضولات پرندگان و ترک‌ها ارائه کردند و جیانگ و لیو [۲۶] یک مدل بهبودیافته برای تشخیص و شناسایی عیوب پنل‌های خورشیدی معرفی کردند. روند این مطالعات نشان‌دهنده حرکت پژوهش‌ها به سمت مدل‌های دقیق‌تر، سبک‌تر و مقاوم‌تر برای شرایط واقعی است.

۲-۵. تصاویر RGB، حرارتی، چندحسی و شرایط میدانی

اگرچه تصاویر الکترولومینسانس برای تشخیص عیوب داخلی مفید هستند، استفاده از آن‌ها در بسیاری از کاربردهای میدانی دشوار است. در مقابل، تصاویر RGB، حرارتی و داده‌های پهنپادی از نظر عملیاتی در دسترس‌ترند، اما با چالش‌هایی مانند تغییر نور، بازتاب سطح پنل، سایه، زاویه دید، آلودگی محیطی و کیفیت متغیر تصویر همراه هستند. منگ و شو [۲۷] روشی سبک مبتنی بر مکانیزم توجه برای تشخیص هم‌زمان چند خطا در مازول‌های فتوولتائیک ارائه کردند و دیمیش [۲۸] نیز مدلی سبک و توجه‌محور برای تشخیص ناهنجاری‌های حرارتی در بازرسی پهنپادی سامانه‌های خورشیدی معرفی کرد. این مطالعات نشان می‌دهند مدل‌های عیب‌یابی باید علاوه بر دقت، از نظر سرعت، وزن محاسباتی و قابلیت استفاده در شرایط میدانی نیز مناسب باشند.

۲-۶. چالش داده، عدم توازن کلاس‌ها و قابلیت تعمیم

یکی از چالش‌های مهم در تشخیص عیوب پنل خورشیدی، کمبود داده‌های برچسب‌خورده، عدم توازن کلاس‌ها و تفاوت میان داده‌های آموزشی و شرایط واقعی است. برخی عیوب مانند گردوغبار فراوان‌تر هستند، در حالی که عیوبی مانند ترک یا شکستگی شدید نمونه‌های کمتری دارند. این مسئله می‌تواند مدل را به سمت کلاس‌های پرتکرار متمایل کند و باعث افت عملکرد در عیوب نادر شود. هی و همکاران [۲۹] یک چارچوب خودتکامل‌یاب برای تشخیص عیوب فتوولتائیک در محیط‌های صنعتی ارائه کردند که مسئله تغییر توزیع داده و ظهور الگوهای جدید عیب را در نظر می‌گیرد. گرژچا و همکاران [۳۰] نیز ضعف کنتراست در تصاویر الکترولومینسانس را با تبدیل آگاه از عیب به نمایش RGB و یادگیری عمیق کم‌هزینه بررسی کردند. طلعت و همکاران [۳۱] از روش‌های یادگیری عمیق برای تشخیص و طبقه‌بندی خطاهای سامانه‌های خورشیدی استفاده کردند و سقیور و همکاران [۳۲] ترکیب هوش مصنوعی و تحلیل زمان - فرکانس را برای عیب‌یابی سامانه‌های فتوولتائیک به کار گرفتند.

موهانتی و همکاران [۳۳] نشان دادند بسیاری از روش‌های بینایی ماشین برای تشخیص عیوب سلول‌های خورشیدی همچنان با چالش‌هایی مانند نوع تصویر، روش جمع‌آوری داده، تفسیرپذیری و قابلیت تجاری‌سازی روبه‌رو هستند. زمان و خانم [۳۴] یک معماری سبک شبکه‌های عصبی کانولوشنی برای تشخیص عیوب پنل‌های خورشیدی پیشنهاد کردند و ژانگ و همکاران [۳۵] از جست‌وجوی معماری عصبی و تقطیر دانش برای طراحی شبکه‌ای سبک در تشخیص عیوب سلول‌های خورشیدی بهره گرفتند. این مطالعات نشان می‌دهند پایداری، تعمیم‌پذیری و قابلیت استقرار عملی به اندازه افزایش دقت خام اهمیت دارند.

۲-۷. نقش استخراج ویژگی، توجه و معماری‌های عمیق

تشخیص عیوب سطحی پنل‌های خورشیدی تنها به انتخاب مدل تشخیص شیء محدود نمی‌شود، بلکه کیفیت استخراج و تلفیق ویژگی‌ها نیز نقش مهمی دارد. عیوب سطحی ممکن است کوچک، کم‌کنتراست، نامنظم یا از نظر رنگ و بافت شبیه سطح پنل

باشند؛ بنابراین، مدل باید بتواند ویژگی‌های لبه‌ای، بافتی و معنایی را به صورت هم‌زمان در نظر بگیرد. هی و همکاران [۹] با معرفی ResNet نشان دادند اتصال‌های باقی‌مانده آموزش شبکه‌های عمیق را پایدارتر می‌کنند و امکان استخراج ویژگی‌های پیچیده‌تر را فراهم می‌سازند. روی و همکاران [۱۰] نیز با معرفی توجه فضایی - کانالی نشان دادند توجه هم‌زمان به کانال‌ها و موقعیت‌های مکانی می‌تواند در شبکه‌های تمام‌کانولوشنی مفید باشد. این ایده‌ها، اگرچه به طور مستقیم برای پنل خورشیدی طراحی نشده‌اند، اما در معماری‌های جدید تشخیص عیب برای تمرکز بر نواحی مهم، حفظ اطلاعات چندمقیاسی و کاهش اثر پس‌زمینه‌های پیچیده کاربرد دارند.

۲-۸. جمع‌بندی شکاف پژوهشی

با وجود پیشرفت‌های اخیر، چند چالش همچنان باقی مانده است: تمرکز بسیاری از پژوهش‌ها بر تصاویر الکترومگناطیس یا شرایط کنترل‌شده، دشواری تشخیص عیوب کوچک و کم‌کنتراست در تصاویر میدانی، نبود اطلاعات مکانی در روش‌های فقط طبقه‌بندی، و سنگین بودن برخی مدل‌ها برای کاربردهای عملی. بر این اساس، پژوهش حاضر با تمرکز بر تشخیص و مکان‌یابی عیوب سطحی پنل‌های خورشیدی در قالب جعبه‌های محدودکننده، تلاش می‌کند رویکردی کاربردی برای شناسایی نواحی شکستگی، گردوغبار، سایه و ناحیه سالم ارائه دهد.

۳. روش پیشنهادی

در این پژوهش، یک چارچوب تشخیص شیء مبتنی بر یادگیری عمیق برای تشخیص و مکان‌یابی عیوب سطحی پنل‌های خورشیدی ارائه می‌شود. هدف مدل پیشنهادی، شناسایی نوع عیب و تعیین موقعیت آن در تصویر با استفاده از جعبه‌های محدودکننده است. به خلاف روش‌های طبقه‌بندی که تنها وضعیت کلی تصویر را به صورت سالم یا معیوب تعیین می‌کنند، روش پیشنهادی علاوه بر تشخیص عیب، محل وقوع آن را نیز مشخص می‌سازد. این ویژگی برای کاربردهای عملی مانند پاک‌سازی، تعمیر، تعویض، اولویت‌بندی بازرسی و نگهداری پیش‌بینانه اهمیت زیادی دارد.

عیوب سطحی پنل‌های خورشیدی از نظر اندازه، شکل، شدت و کنتراست تنوع زیادی دارند. برخی عیوب مانند ترک‌های ریز، خراش و شکستگی جزئی، کوچک و کم‌کنتراست هستند؛ در حالی که عیوبی مانند آلودگی، گردوغبار، شکستگی شیشه یا آسیب فیزیکی بخش بزرگ‌تری از پنل را درگیر می‌کنند. افزون بر این، خطوط شبکه سلول‌های خورشیدی، بازتاب نور، سایه، تغییر زاویه تصویربرداری و بافت تکرارشونده سطح پنل، فرایند تشخیص را دشوار می‌سازند. برای پاسخ به این چالش‌ها، معماری پیشنهادی از استخراج ویژگی عمیق، تقویت اطلاعات لبه‌ای، توجه کانالی - فضایی و تلفیق ویژگی‌های چندمقیاسی استفاده می‌کند. چارچوب کلی مدل پیشنهادی شامل شش بخش اصلی است:

پیش‌پردازش و افزایش داده، ماژول استخراج لبه، استخراج گر ویژگی ResNet101، ماژول توجه کانالی - فضایی، شبکه تلفیق ویژگی چندمقیاسی و سر تشخیص چندمقیاسی.

در معماری پیشنهادی، ResNet101 نقش استخراج گر ویژگی یا Encoder را به عهده دارد. از آنجا که مسئله این پژوهش از نوع تشخیص شیء است، مدل دارای کدگشا کلاسیک مشابه شبکه‌های بخش‌بندی نیست؛ بلکه ویژگی‌های استخراج‌شده پس از تقویت با اطلاعات لبه‌ای و ماژول توجه، وارد بخش گردن شبکه یعنی ساختار FPN/PAN می‌شوند. این بخش وظیفه تلفیق ویژگی‌های چندمقیاسی را به عهده دارد و سپس، سر تشخیص چندمقیاسی، خروجی‌های مربوط به سطوح P_3 ، P_4 و P_5 را تولید می‌کند. در هر سطح، مختصات جعبه محدودکننده، کلاس ناحیه و امتیاز اطمینان پیش‌بینی می‌شود.

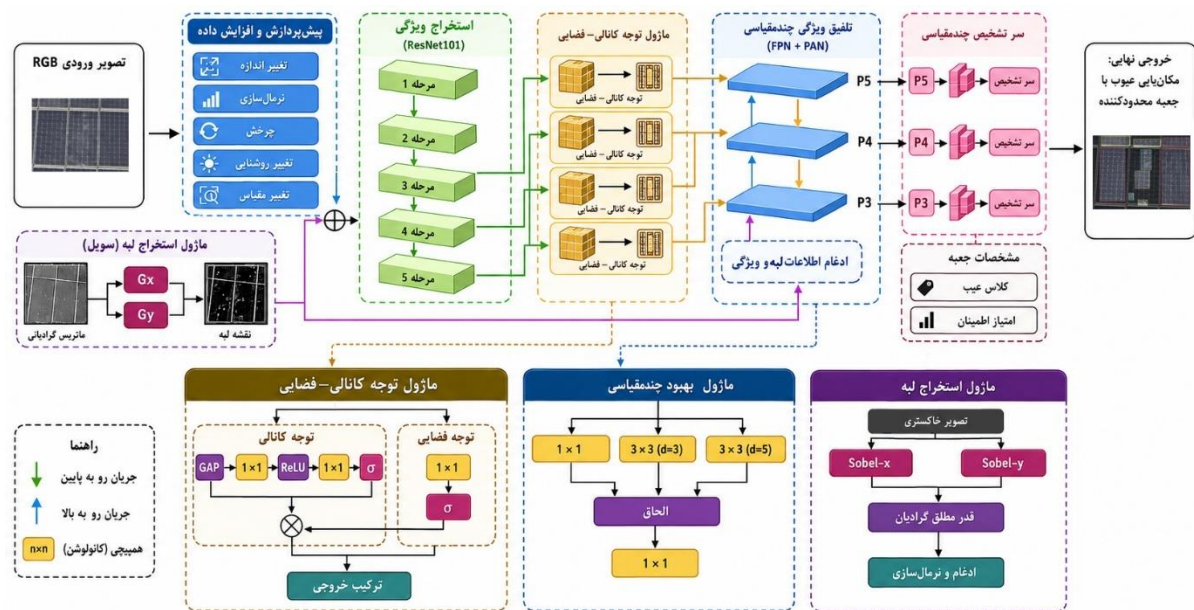
نمای کلی معماری مدل پیشنهادی و ارتباط میان اجزای آن در شکل ۱ نشان داده شده است. همان‌طور که در این شکل مشاهده می‌شود، تصویر ورودی پس از پیش‌پردازش و استخراج اطلاعات لبه‌ای وارد استخراج گر ویژگی ResNet101 می‌شود. ویژگی‌های حاصل پس از تقویت توسط ماژول توجه کانالی - فضایی و تلفیق در ساختار چندمقیاسی FPN/PAN، به سر تشخیص چندمقیاسی ارسال می‌شوند تا جعبه‌های محدودکننده، کلاس ناحیه و امتیاز اطمینان پیش‌بینی شوند.

۳-۱. ورودی، پیش‌پردازش و افزایش داده

ورودی مدل یک تصویر رنگی از سطح پنل خورشیدی است که به صورت رابطه ۱ تعریف می‌شود:

$$I \in R^{H \times W \times 3} \quad (1)$$

در این رابطه، H و W به ترتیب ارتفاع و عرض تصویر هستند و عدد ۳ نشان‌دهنده سه کانال رنگی RGB است. در مرحله پیش‌پردازش، ابتدا همه تصاویر به اندازه ثابت تغییر داده می‌شوند. سپس، مقادیر پیکسلی نرمال‌سازی می‌شوند تا اثر تفاوت روشنایی و شدت رنگ در تصاویر مختلف کاهش یابد.



شکل ۱. معماری مدل پیشنهادی برای تشخیص و مکان‌یابی چندکلاسه عیوب سطحی پنل‌های خورشیدی

برای افزایش توان تعمیم مدل و کاهش بیش‌برازش، روی داده‌های آموزشی که متشکل از چندین دیتا و همچنین داده‌های جمع‌آوری شده است، افزایش داده انجام می‌شود. این مرحله شامل تغییر مقیاس، چرخش محدود، تغییر روشنایی، تغییر کنتراست، جابه‌جایی و وارون‌سازی افقی است. تصویر نهایی پس از پیش‌پردازش و افزایش داده به صورت رابطه ۲ تعریف می‌شود:

$$I' = T(I) \quad (2)$$

در این رابطه، $T(\cdot)$ مجموعه عملیات پیش‌پردازش و افزایش داده را نشان می‌دهد و I' تصویر آماده‌شده برای ورود به شبکه است.

۳-۲. ماژول استخراج لبه با عملگر سوبل

در تصاویر پنل خورشیدی، بخشی از عیوب سطحی مانند ترک، خراش و شکستگی جزئی بیشتر در قالب تغییرات لبه‌ای و ساختاری ظاهر می‌شوند. به همین دلیل، در معماری پیشنهادی از یک ماژول استخراج لبه مبتنی بر عملگر سوبل استفاده می‌شود. این ماژول اطلاعات ساختاری تصویر را استخراج کرده و در مراحل بعدی با ویژگی‌های عمیق ترکیب می‌کند. ابتدا تصویر RGB به تصویر خاکستری تبدیل می‌شود (رابطه ۳):

$$I_{gray} = Gray(I') \quad (3)$$

سپس گرادیان‌های افقی و عمودی تصویر با استفاده از کرنل‌های سوبل محاسبه می‌شوند (رابطه ۴):

$$\begin{aligned} G_x &= I_{gray} * S_x \\ G_y &= I_{gray} * S_y \end{aligned} \quad (4)$$

در این روابط، S_x و S_y کرنل‌های سوبل در جهت افقی و عمودی هستند و علامت * عمل هم‌پیکری را نشان می‌دهد. نقشه نهایی لبه از ترکیب گرادیان‌های افقی و عمودی به دست می‌آید (رابطه ۵):

$$G = \sqrt{G_x^2 + G_y^2} \quad (5)$$

نقشه لبه G به مدل کمک می‌کند تا نواحی دارای تغییرات ساختاری را بهتر شناسایی کند. در معماری پیشنهادی، این نقشه به صورت یک راهنمای کمکی با ویژگی‌های عمیق ترکیب می‌شود و نقش آن تقویت تشخیص عیوب باریک، کم‌کنتراست و مرزی است.

۳-۳. استخراج ویژگی با ResNet101

در این پژوهش، از ResNet101 به عنوان استخراج‌گر اصلی ویژگی استفاده می‌شود. دلیل انتخاب ResNet101، توان بالای آن در یادگیری نمایش‌های عمیق و سلسله‌مراتبی از تصویر است. اتصال‌های باقی‌مانده در این شبکه باعث می‌شوند جریان گرادیان در لایه‌های عمیق حفظ شود و شبکه بتواند بدون افت شدید در آموزش، ویژگی‌های پیچیده‌تری را از تصویر استخراج کند.

انتخاب ResNet101 در این پژوهش با هدف دستیابی به نمایش‌های عمیق‌تر و پایدارتر از نواحی سطحی پنل خورشیدی انجام شده است. اگرچه معماری‌هایی مانند CSPDarknet یا EfficientNet از نظر محاسباتی سبک‌تر هستند و در بسیاری از آشکارسازهای جدید استفاده می‌شوند، تمرکز اصلی این پژوهش در مرحله حاضر بیش از آنکه بر استقرار روی سخت‌افزارهای کم‌مصرف باشد، بر افزایش دقت تشخیص عیوب کوچک، کم‌کنتراست و دارای مرز نامشخص بوده است. در تصاویر پنل خورشیدی، برخی نواحی مانند شکستگی‌های باریک یا آلودگی‌های پراکنده، از نظر ظاهری شباهت زیادی با خطوط طبیعی سلول، بازتاب نور یا سایه دارند. بنابراین، استفاده از استخراج‌گر عمیق‌تر باعث می‌شود مدل بتواند ویژگی‌های سطح پایین، بافتی و معنایی را به صورت غنی‌تر استخراج کند.

برای بررسی اثر این انتخاب، در مطالعه حذف اجزاء، نسخه‌ای از مدل با استخراج‌گر سبک‌تر نیز ارزیابی شد. نتایج نشان داد جایگزینی ResNet101 با استخراج‌گر سبک‌تر باعث کاهش Precision از ۰/۹۵۴ به ۰/۹۳۰، Recall از ۰/۹۳۹ به ۰/۹۰۵ و $mAP@0.5$ از ۰/۹۶۶ به ۰/۹۳۶ شد. بنابراین، در پیاده‌سازی فعلی، افزایش عمق ResNet101 باعث بهبود معنادار دقت تشخیص و مکان‌یابی شده است. با این حال، از آنجا که ResNet101 هزینه محاسباتی بالاتری دارد، توسعه نسخه‌های سبک‌تر برای استقرار روی تجهیزات لبه‌ای به عنوان مسیر آینده پژوهش در نظر گرفته شده است.

پس از ترکیب تصویر پیش‌پردازش‌شده با نقشه لبه، ورودی تقویت‌شده (I_e) وارد استخراج‌گر ویژگی ResNet101 می‌شود. خروجی مراحل مختلف ResNet101 به صورت مجموعه‌ای از نقشه‌های ویژگی تعریف می‌شود:

$$\{F_1, F_2, F_3, F_4, F_5\} = \text{ResNet101}(I_e) \quad (6)$$

در این رابطه، F_1 و F_2 شامل ویژگی‌های کم‌عمق‌تر مانند لبه، بافت و جزئیات مکانی هستند؛ در حالی که F_3, F_4, F_5 ویژگی‌های عمیق‌تر و معنایی‌تر را در بر می‌گیرند. برای تشخیص عیوب سطحی، هر دو نوع ویژگی اهمیت دارند، زیرا عیوب کوچک به جزئیات مکانی نیاز دارند و عیوب پیچیده‌تر با ویژگی‌های معنایی بهتر شناسایی می‌شوند. در معماری پیشنهادی، ویژگی‌های عمیق‌تر F_3, F_4, F_5 به شبکه تلفیق ویژگی چندمقیاسی ارسال می‌شوند.

۳-۴. ماژول توجه کانالی - فضایی

همه بخش‌های تصویر پنل خورشیدی به یک اندازه مهم نیستند. بخش زیادی از تصویر شامل سطح سالم پنل، خطوط تکراری سلول‌ها، بازتاب نور یا پس‌زمینه است. در مقابل، ناحیه عیب ممکن است کوچک، کم‌کنتراست و از نظر ظاهری شبیه سطح سالم باشد. به همین دلیل، در معماری پیشنهادی از ماژول توجه کانالی - فضایی استفاده می‌شود تا شبکه روی کانال‌ها و نواحی مکانی مهم‌تر تمرکز کند.

اگر F یک نقشه ویژگی ورودی باشد، توجه کانالی به صورت رابطه ۷ محاسبه می‌شود:

$$A_c = \sigma \left(\text{Conv}_{1 \times 1} \left(\text{ReLU} \left(\text{Conv}_{1 \times 1} \left(\text{GAP}(F) \right) \right) \right) \right) \quad (7)$$

در این رابطه، GAP میانگین‌گیری سراسری، $\text{Conv}_{1 \times 1}$ هم‌پیچی یک‌دریک، ReLU تابع فعال‌سازی و σ تابع سیگموئید است. توجه فضایی نیز به صورت رابطه ۸ محاسبه می‌شود:

$$A_s = \sigma \left(\text{Conv}_{1 \times 1} (F) \right) \quad (8)$$

در نهایت، خروجی تقویت‌شده ماژول توجه از ترکیب توجه کانالی و فضایی به دست می‌آید:

$$F_{att} = F \odot A_c + F \odot A_s \quad (9)$$

در این رابطه، $\odot$ ضرب عنصر به عنصر را نشان می‌دهد. این ماژول باعث می‌شود ویژگی‌های مربوط به عیب تقویت شوند و اثر نواحی غیرمرتبط مانند خطوط طبیعی سلول‌ها، بازتاب نور و پس‌زمینه کاهش یابد.

در معماری پیشنهادی، اطلاعات لبه‌ای استخراج‌شده توسط عملگر سوبل در دو مرحله با جریان اصلی ویژگی ترکیب می‌شود. در مرحله نخست، نقشه لبه پس از نرمال‌سازی و تطبیق ابعاد، با تصویر پیش‌پردازش‌شده در سطح ورودی ترکیب می‌شود و ورودی تقویت‌شده را تشکیل می‌دهد:

$$I_e = I' \oplus G \quad (10)$$

در این رابطه، $\oplus$ عمل الحاق کانالی را نشان می‌دهد. بنابراین، استخراج‌گر ویژگی ResNet101 علاوه بر اطلاعات رنگی تصویر، به اطلاعات ساختاری و مرزی اولیه نیز دسترسی دارد. برای سازگاری ورودی چهاربعدی با ResNet101، لایه هم‌پیچی نخست متناسب با تعداد کانال‌های ورودی اصلاح می‌شود و وزن کانال لبه با میانگین وزن‌های کانال‌های RGB مقداردهی اولیه می‌شود.

در مرحله دوم، نقشه لبه به عنوان یک شاخه کمکی در بخش تلفیق ویژگی چندمقیاسی نیز استفاده می‌شود. در این مرحله، G پس از تغییر اندازه و تطبیق کانالی با نقشه‌های ویژگی چندمقیاسی، در بخش FPN/PAN ادغام می‌شود تا اطلاعات مرزی در سطوح مختلف ویژگی حفظ شود. این طراحی باعث می‌شود مدل هم در لایه‌های اولیه به نشانه‌های لبه‌ای حساس باشد و هم در مرحله تلفیق چندمقیاسی بتواند از اطلاعات ساختاری برای تشخیص عیوب باریک، کم‌کنتراست و مرزی استفاده کند.

۳-۵. تلفیق ویژگی چندمقیاسی

عیوب سطحی پل خورشیدی در مقیاس‌های مختلف ظاهر می‌شوند. ترک‌ها و خراش‌های ریز معمولاً کوچک و باریک هستند، در حالی که شکستگی، آلودگی یا آسیب فیزیکی ممکن است ناحیه بزرگ‌تری از تصویر را در بر بگیرند. بنابراین، استفاده از تنها یک سطح ویژگی برای تشخیص همه عیوب کافی نیست.

در معماری پیشنهادی، از ساختار FPN + PAN برای تلفیق ویژگی‌های چندمقیاسی استفاده می‌شود. این ساختار ویژگی‌های عمیق، معنایی و کم‌وضوح را با ویژگی‌های کم‌عمق، مکانی و پرجزئیات ترکیب می‌کند. در این بخش، علاوه بر نقشه‌های ویژگی استخراج‌شده از ResNet101، نقشه لبه G نیز به صورت یک شاخه کمکی وارد فرایند تلفیق ویژگی می‌شود. هدف از این کار، حفظ اطلاعات مرزی و ساختاری در سطوح مختلف ویژگی است.

برای ادغام اطلاعات لبه‌ای، ابتدا نقشه لبه G برای هر سطح ویژگی تغییر اندازه داده شده و با یک هم‌پیچی 1×1 از نظر تعداد کانال با همان سطح تطبیق داده می‌شود:

$$G_k = \text{Resize} \left(\text{Conv}_{1 \times 1} (G), \text{size}(F_k) \right), k \in \{3, 4, 5\} \quad (11)$$

سپس، نقشه لبه تطبیق‌یافته با ویژگی‌های همان سطح الحاق شده و خروجی به ساختار FPN/PAN ارسال می‌شود:

$$\{P_3, P_4, P_5\} = \text{FPN} / \text{PAN} (F_3 \oplus G_3, F_4 \oplus G_4, F_5 \oplus G_5) \quad (12)$$

در این روابط، G_k نسخه تغییراندازه‌یافته و تطبیق‌یافته نقشه لبه برای سطح ویژگی F_k است و $\oplus$ عمل الحاق کانالی را نشان می‌دهد. خروجی این بخش سه نقشه ویژگی در مقیاس‌های مختلف است؛ P_3 برای تشخیص عیوب کوچک، P_4 برای

عیوب متوسط و P_5 برای عیوب بزرگ‌تر استفاده می‌شود. به این ترتیب، اطلاعات لبه‌ای علاوه بر مرحله ورودی، در بخش تلفیق ویژگی نیز حضور دارد و به حفظ مرز عیوب در مقیاس‌های مختلف کمک می‌کند. برای تقویت میدان دید شبکه و دریافت اطلاعات در مقیاس‌های مختلف، از ماژول بهبود چندمقیاسی استفاده می‌شود. این ماژول شامل سه شاخه هم‌پپی با اندازه و نرخ گشادگی متفاوت است:

$$\begin{aligned} M_1 &= \text{Conv}_{1 \times 1}(F) \\ M_3 &= \text{Conv}_{3 \times 3}^{d=3}(F) \\ M_5 &= \text{Conv}_{3 \times 3}^{d=5}(F) \end{aligned} \quad (13)$$

سپس خروجی این سه شاخه الحاق شده و با یک هم‌پپی یک‌دریک فشرده‌سازی می‌شود:

$$\begin{aligned} M_{cat} &= \text{Concat}(M_1, M_3, M_5) \\ F_{ms} &= \text{Conv}_{1 \times 1}(M_{cat}) \end{aligned} \quad (14)$$

در این روابط، d نرخ گشادگی هم‌پپی را نشان می‌دهد. این ماژول باعث می‌شود مدل بتواند هم عیوب کوچک و هم عیوب گسترده‌تر را با دقت بیشتری تشخیص دهد. در معماری پیشنهادی، این ماژول پس از ادغام اطلاعات لبه‌ای و پیش از ارسال ویژگی‌ها به سر تشخیص چندمقیاسی استفاده می‌شود.

۳-۶. سر تشخیص چندمقیاسی

پس از تلفیق ویژگی‌ها، سه نقشه ویژگی P_3, P_4, P_5 وارد سر تشخیص چندمقیاسی می‌شوند. هر سر تشخیص، برای مقیاس مربوط به خود، مختصات جعبه محدودکننده، کلاس عیب و امتیاز اطمینان را پیش‌بینی می‌کند. خروجی هر سطح به صورت رابطه ۱۵ تعریف می‌شود:

$$Y_k = \text{Head}(P_k), k \in \{3, 4, 5\} \quad (15)$$

خروجی نهایی مدل از ترکیب پیش‌بینی‌های سه مقیاس به دست می‌آید:

$$\hat{Y} = \{Y_3, Y_4, Y_5\} \quad (16)$$

هر پیش‌بینی شامل بردار زیر است (رابطه ۱۷):

$$\hat{y}_i = (\hat{x}_i, \hat{y}_i, \hat{w}_i, \hat{h}_i, \hat{c}_i, \hat{s}_i) \quad (17)$$

در این رابطه، $(\hat{x}_i, \hat{y}_i)$ مرکز جعبه، $\hat{w}_i, \hat{h}_i$ عرض و ارتفاع جعبه، $\hat{c}_i$ کلاس پیش‌بینی‌شده و $\hat{s}_i$ امتیاز اطمینان هستند.

امتیاز نهایی هر پیش‌بینی از ترکیب احتمال کلاس و اطمینان وجود عیب به دست می‌آید:

$$\hat{s}_i = \hat{p}_i \times \hat{o}_i \quad (18)$$

در این رابطه، $\hat{p}_i$ احتمال تعلق جعبه به کلاس عیب و $\hat{o}_i$ اطمینان مدل نسبت به وجود عیب در آن جعبه است.

در مرحله پس‌پردازش، ابتدا جعبه‌هایی که امتیاز اطمینان آن‌ها کمتر از آستانه تعیین‌شده است حذف می‌شوند. سپس برای حذف جعبه‌های تکراری، از روش **سرکوب غیر بیشینه** استفاده می‌شود. در این روش، اگر چند جعبه روی یک عیب واحد هم‌پوشانی زیادی داشته باشند، جعبه‌ای که بیشترین امتیاز اطمینان دارد حفظ می‌شود.

۳-۷. تابع زیان

از آنجا که مدل پیشنهادی یک مدل تشخیص شیء است، تابع زیان آن از سه بخش اصلی تشکیل می‌شود: زیان مکان‌یابی جعبه، زیان طبقه‌بندی و زیان اطمینان. تابع زیان کلی به صورت رابطه ۱۹ تعریف می‌شود:

$$L_{total} = \lambda_{box} L_{box} + \lambda_{cls} L_{cls} + \lambda_{obj} L_{obj} \quad (19)$$

در این رابطه، L_{box} زیان مکان‌یابی جعبه، L_{cls} زیان طبقه‌بندی، L_{obj} زیان اطمینان، و λ_{box} ، λ_{cls} ، λ_{obj} ضرایب وزن‌دهی هر بخش هستند.

برای زیان مکان‌یابی، از زیان مبتنی بر CioU استفاده می‌شود؛ زیرا این زیان علاوه بر هم‌پوشانی جعبه پیش‌بینی‌شده و جعبه واقعی، فاصله مرکز جعبه‌ها و نسبت ابعاد آن‌ها را نیز در نظر می‌گیرد:

$$L_{box} = L_{CioU}$$

$$L_{CioU} = 1 - IoU + \frac{\rho^2(b, \hat{b})}{c^2} + \alpha v \quad (20)$$

در این رابطه، IoU میزان هم‌پوشانی جعبه واقعی و پیش‌بینی‌شده، $\rho^2(b, \hat{b})$ فاصله مربعی مرکز دو جعبه، c قطر کوچک‌ترین جعبه دربرگیرنده هر دو جعبه و v اختلاف نسبت ابعاد جعبه‌ها است.

برای کاهش اثر عدم توازن کلاس‌ها، زیان طبقه‌بندی با استفاده از آنتروپی متقاطع وزن‌دار محاسبه می‌شود. در این روش، به کلاس‌هایی که نمونه‌های کمتری در مجموعه آموزش دارند وزن بیشتری اختصاص داده می‌شود تا مدل به سمت کلاس‌های پرتکرار، به‌ویژه ناحیه سالم، متمایل نشود:

$$L_{cls} = -\sum_{k=1}^C w_k y_k \log(\hat{p}_k) \quad (21)$$

در این رابطه، C تعداد کلاس‌ها، y_k برچسب واقعی کلاس k و $\hat{p}_k$ احتمال پیش‌بینی‌شده برای آن کلاس و w_k وزن متناظر با کلاس k است. وزن کلاس‌ها بر اساس فراوانی نسبی نمونه‌ها در مجموعه آموزش تعیین می‌شود؛ به گونه‌ای که کلاس‌های کم‌نمونه‌تر، مانند شکستگی، سهم بیشتری در زیان طبقه‌بندی داشته باشند.

زیان اطمینان نیز با آنتروپی متقاطع دودویی محاسبه می‌شود:

$$L_{obj} = -[o \log(\hat{o}) + (1-o) \log(1-\hat{o})] \quad (22)$$

در این رابطه، o برچسب واقعی وجود داشتن یا نداشتن عیب و $\hat{o}$ مقدار اطمینان پیش‌بینی‌شده توسط مدل است.

۳-۸. خروجی نهایی مدل

خروجی نهایی مدل پیشنهادی مجموعه‌ای از جعبه‌های محدودکننده است که محل عیوب سطحی پنل خورشیدی را مشخص می‌کنند. هر خروجی شامل مختصات جعبه، کلاس عیب و امتیاز اطمینان است:

$$Output = \left\{ \left(\hat{x}_i, \hat{y}_i, \hat{w}_i, \hat{h}_i, \hat{c}_i, \hat{s}_i \right) \right\}_{i=1}^N \quad (23)$$

در این رابطه، N تعداد نواحی تشخیص‌داده‌شده در تصویر است. این خروجی برای کاربردهای عملی بسیار مناسب است، زیرا علاوه بر تشخیص وجود عیوب، محل آن‌ها را نیز مشخص می‌کند و امکان استفاده از نتایج در سامانه‌های نگهداری، تعمیر، پاک‌سازی و پایش خودکار را فراهم می‌سازد.

مدل پیشنهادی محدود به یک پیش‌بینی واحد برای کل تصویر نیست، بلکه یک آشکارساز چندکلاسه و چندنمونه‌ای است. بنابراین، اگر در یک تصویر چند ناحیه با وضعیت‌های متفاوت مانند شکستگی، آلودگی، سایه یا ناحیه سالم وجود داشته باشد، مدل برای هر یک جعبه‌ای مستقل تولید کرده و کلاس متناظر آن را مشخص می‌کند. این ویژگی باعث می‌شود خروجی مدل برای کاربردهای عملی مانند پایش خودکار، پاک‌سازی، تعمیر، تعویض و اولویت‌بندی بازرسی پنل‌های خورشیدی قابل استفاده باشد.

۴. آزمایش‌ها و نتایج

۴-۱. مجموعه‌داده و آماده‌سازی تصاویر

برای آموزش و ارزیابی مدل پیشنهادی، از مجموعه‌داده عمومی Solar Panel Fault Detection منتشرشده در بستر Roboflow Universe استفاده شد. این مجموعه‌داده برای مسئله تشخیص شیء تهیه شده و برچسب‌های آن در قالب جعبه‌های محدودکننده و فرمت YOLO ارائه شده‌اند. نسخه اولیه مجموعه‌داده شامل ۵۱۲ تصویر از پنل‌های خورشیدی است و

چهار کلاس شکستگی، ناحیه سالم، گردوغبار و سایه را شامل می‌شود. از آنجا که در هر تصویر ممکن است چندین ناحیه مختلف وجود داشته باشد، تعداد نمونه‌های برچسب‌خورده بیشتر از تعداد تصاویر است.

برچسب‌گذاری داده‌ها بر اساس جعبه محدودکننده انجام شد. در برخی تصاویر، برچسب‌ها به صورت آماده در مجموعه داده وجود داشتند و در برخی موارد، جعبه‌های محدودکننده به صورت دستی بازبینی، اصلاح یا تکمیل شدند. برای کنترل کیفیت برچسب‌ها، نمونه‌های دارای جعبه نامعتبر، کلاس نامشخص یا مختصات خارج از محدوده تصویر اصلاح یا حذف شدند. کلاس‌های مورد استفاده در این پژوهش شامل شکستگی، ناحیه سالم، گردوغبار و سایه هستند.

وجود کلاس ناحیه سالم با هدف کاهش تشخیص‌های مثبت کاذب و آموزش مرز تصمیم دقیق‌تر میان نواحی طبیعی و نواحی غیرعادی در سطح پنل در نظر گرفته شد. در تصاویر پنل خورشیدی، الگوهایی مانند خطوط شبکه سلولی، بافت یکنواخت سطح پنل و بازتاب‌های ملایم نور ممکن است با عیوب کم‌کنتراست اشتباه گرفته شوند. بنابراین، برچسب‌گذاری نواحی سالم به مدل کمک می‌کند الگوهای طبیعی سطح پنل را نیز یاد بگیرد و همه تغییرات تصویری را به اشتباه به عنوان عیب تشخیص ندهد.

نمونه‌هایی از تصاویر مجموعه داده و نحوه برچسب‌گذاری نواحی شکستگی، ناحیه سالم، گردوغبار و سایه با استفاده از جعبه‌های محدودکننده در شکل ۲ نشان داده شده است. همان‌طور که در این شکل مشاهده می‌شود، ممکن است در هر تصویر چند ناحیه با کلاس‌های متفاوت به صورت هم‌زمان وجود داشته باشد؛ از این رو، مسئله پژوهش حاضر از نوع تشخیص چندکلاسه و چندنمونه‌ای است.



شکل ۲. نمونه‌هایی از مجموعه داده پنل‌های خورشیدی به همراه جعبه‌های محدودکننده مربوط به کلاس‌های شکستگی، ناحیه سالم، گردوغبار و سایه

مجموعه داده اولیه به سه بخش آموزش، اعتبارسنجی و آزمون با نسبت ۸۰، ۱۰ و ۱۰ درصد تقسیم شد. بر این اساس، ۴۱۰ تصویر برای آموزش، ۵۱ تصویر برای اعتبارسنجی و ۵۱ تصویر برای آزمون استفاده شد. برای جلوگیری از نشت داده، افزایش داده تنها روی بخش آموزش اعمال شد و تصاویر اعتبارسنجی و آزمون بدون افزایش داده ماندند. روش‌های افزایش داده شامل چرخش محدود، تغییر روشنایی، تغییر کنتراست، تغییر مقیاس، جابه‌جایی و وارون‌سازی افقی بودند. با اعمال این عملیات، تعداد مؤثر تصاویر آموزشی از ۴۱۰ به ۱۶۴۰ تصویر افزایش یافت و تعداد کل تصاویر مؤثر مورد استفاده در فرایند آموزش و ارزیابی به ۱۷۴۲ تصویر رسید.

تعداد کل جعبه‌های محدودکننده در داده اولیه برابر با ۳۸۶۴ نمونه بود. پس از افزایش داده بخش آموزش، تعداد مؤثر نمونه‌های برچسب‌خورده به ۱۳۱۳۷ جعبه محدودکننده رسید. توزیع نمونه‌ها در کلاس‌های مختلف در جدول ۱ ارائه شده است. تقسیم داده‌ها نیز به صورت آگاه از کلاس انجام شد تا هر چهار کلاس در بخش‌های آموزش، اعتبارسنجی و آزمون حضور داشته باشند.

جدول ۱. توزیع نمونه‌های برچسب‌خورده

کلاس	تعداد جعبه‌ها در داده اولیه	تعداد مؤثر جعبه‌ها پس از افزایش داده آموزش	نسبت تقریبی از کل
شکستگی	۸۱۲	۲۷۶۲	٪۲۱
ناحیه سالم	۱۲۴۶	۴۲۳۷	٪۳۲.۳
گردوغبار	۱۰۳۸	۳۵۲۸	٪۲۶.۹
سایه	۷۶۸	۲۶۱۰	٪۱۹.۹
مجموع	۳۸۶۴	۱۳۱۳۷	٪۱۰۰

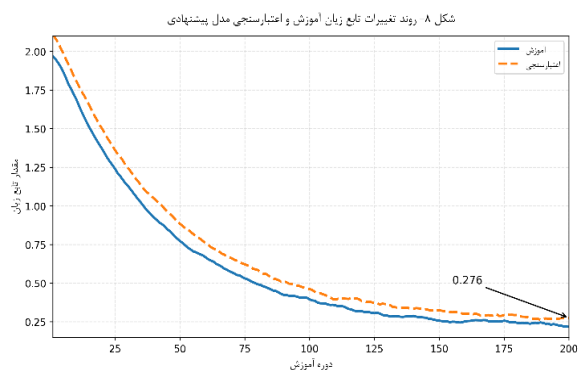
۴-۲. تنظیمات پیاده‌سازی و آموزش

پیاده‌سازی مدل پیشنهادی در محیط پایتون و با استفاده از چارچوب یادگیری عمیق PyTorch انجام شد. آموزش مدل روی سامانه‌ای مجهز به پردازنده گرافیکی NVIDIA RTX 4060 انجام گرفت. برای کنترل مصرف حافظه و افزایش سرعت پردازش، تصاویر به صورت دسته‌ای وارد شبکه شدند. اندازه دسته برابر با ۸ و تعداد دوره‌های آموزش برابر با ۲۰۰ در نظر گرفته شد. بهترین وزن‌های مدل بر اساس عملکرد روی داده اعتبارسنجی ذخیره شدند.

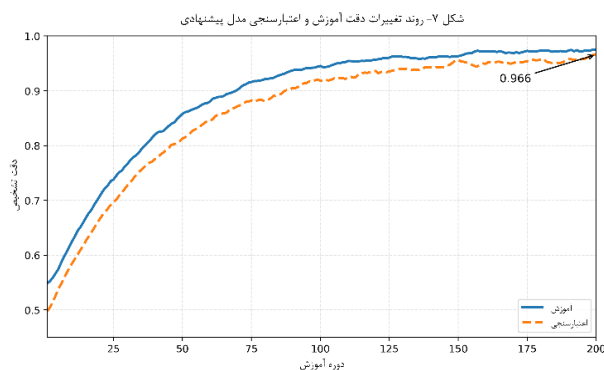
برای شفاف‌سازی پیکربندی مدل، معماری پیشنهادی به صورت یک آشکارساز چندمقیاسی طراحی شد که در آن ResNet101 به عنوان کدکننده و ساختار FPN/PAN به عنوان بخش تلفیق ویژگی یا گردن شبکه استفاده می‌شود. به خلاف مدل‌های بخش‌بندی، در این پژوهش کدگشا بازسازی‌کننده تصویر یا ماسک وجود ندارد؛ زیرا خروجی مدل به صورت جعبه محدودکننده، کلاس و امتیاز اطمینان تعریف شده است. پس از استخراج ویژگی توسط ResNet101، نقشه‌های ویژگی عمیق در سطوح مختلف با اطلاعات لبه‌ای سوبل و مازول توجه کانالی - فضایی تقویت می‌شوند. سپس ویژگی‌های حاصل وارد FPN/PAN می‌شوند تا اطلاعات معنایی عمیق و جزئیات مکانی در چند مقیاس ترکیب شوند. در نهایت، سر تشخیص چندمقیاسی روی سطوح P_3 تا P_5 اعمال می‌شود تا نواحی کوچک، متوسط و بزرگ به صورت هم‌زمان شناسایی شوند.

تنظیمات آموزشی نیز در همه آزمایش‌ها ثابت نگه داشته شد. تصاویر ورودی پس از تغییر اندازه و نرمال‌سازی وارد شبکه شدند. آموزش مدل با اندازه دسته ۸، تعداد ۲۰۰ دوره، بهینه‌ساز Adam و نرخ یادگیری اولیه 10^{-4} انجام شد. بهترین وزن‌های مدل بر اساس عملکرد روی داده اعتبارسنجی انتخاب شدند. تابع زیان شامل سه بخش مکان‌یابی جعبه، طبقه‌بندی و اطمینان بود؛ برای مکان‌یابی از CIOU و برای طبقه‌بندی از آنتروپی متقاطع وزن‌دار استفاده شد تا اثر عدم توازن کلاس‌ها کاهش یابد.

روند تغییرات تابع زیان و دقت در مجموعه‌های آموزش و اعتبارسنجی، به ترتیب در شکل‌های ۳ و ۴ ارائه شده است. مطابق شکل ۳، مقدار تابع زیان در هر دو مجموعه با افزایش دوره‌های آموزش به طور پیوسته کاهش یافته و در دوره‌های پایانی به مقادیر نزدیک و نسبتاً پایداری رسیده است. شکل ۴ نیز افزایش تدریجی دقت آموزش و اعتبارسنجی و نزدیک شدن مقادیر آن‌ها به یک سطح پایدار را نشان می‌دهد. نزدیکی روندهای آموزش و اعتبارسنجی بیانگر همگرایی مناسب مدل و نبود شکاف شدید میان عملکرد مدل روی این دو مجموعه است.



شکل ۴. روند تغییرات دقت آموزش و اعتبارسنجی مدل پیشنهادی طی دوره‌های آموزش



شکل ۵. روند تغییرات دقت آموزش و اعتبارسنجی مدل پیشنهادی طی دوره‌های آموزش

۳-۴. معیارهای ارزیابی

از آنجا که مسئله این پژوهش از نوع تشخیص شیء است، معیارهای ارزیابی بر اساس کیفیت تشخیص و مکان‌یابی جعبه‌های محدودکننده انتخاب شدند. در این پژوهش از معیارهای دقت، بازخوانی، امتیاز $F1$ ، $mAP@0.5$ و $mAP@0.5:0.95$ استفاده شد. این معیارها در پژوهش‌های جدید تشخیص عیوب فتوولتائیک نیز به طور گسترده استفاده شده‌اند. معیار دقت نشان می‌دهد چه نسبتی از تشخیص‌های مدل درست بوده‌اند:

$$Precision = \frac{TP}{TP + FP} \quad (24)$$

معیار بازخوانی نشان می‌دهد مدل چه نسبتی از نمونه‌های واقعی را شناسایی کرده است:

$$Recall = \frac{TP}{TP + FN} \quad (25)$$

امتیاز $F1$ میانگین هماهنگ دقت و بازخوانی است:

$$F1 = 2 \times \frac{Precision \times Recall}{Precision + Recall} \quad (26)$$

همچنین، معیار mAP برای ارزیابی کلی عملکرد مدل در تشخیص و مکان‌یابی استفاده شد:

$$AP_k = \int_0^1 Precision_k(Recall) dRecall \quad (27)$$

میانگین دقت متوسط برای همه کلاس‌ها به صورت رابطه ۲۸ محاسبه می‌شود:

$$mAP = \frac{1}{C} \sum_{k=1}^C AP_k \quad (28)$$

در این رابطه، C تعداد کلاس‌ها و AP_k دقت متوسط مربوط به کلاس k است. معیار $mAP@0.5$ عملکرد مدل را در آستانه هم‌پوشانی ۰/۵ بررسی می‌کند، در حالی که $mAP@0.5:0.95$ میانگین عملکرد مدل را در چند آستانه هم‌پوشانی سخت‌گیرانه‌تر نشان می‌دهد.

برای محاسبه هم‌پوشانی میان جعبه واقعی و جعبه پیش‌بینی‌شده از معیار IoU استفاده می‌شود:

$$IoU = \frac{Area(B \cap \hat{B})}{Area(B \cup \hat{B})} \quad (29)$$

در این رابطه، B جعبه واقعی و $\hat{B}$ جعبه پیش‌بینی‌شده است. همچنین، $B \cap \hat{B}$ ناحیه اشتراک دو جعبه و $B \cup \hat{B}$ ناحیه اجتماع آن‌ها را نشان می‌دهد.

معیار $mAP@0.5:0.95$ میانگین mAP را در آستانه‌های مختلف IoU از ۰/۵ تا ۰/۹۵ با گام ۰/۰۵ محاسبه می‌کند:

$$mAP@0.5:0.95 = \frac{1}{10} \sum_{t \in \{0.50, 0.55, \dots, 0.95\}} mAP@t \quad (30)$$

در این پژوهش، $mAP@0.5$ برای ارزیابی توان کلی مدل در تشخیص نواحی هدف و $mAP@0.5:0.95$ برای بررسی دقت مکانی جعبه‌های محدودکننده استفاده شد. هرچه مقدار $mAP@0.5:0.95$ بالاتر باشد، مدل در مکان‌یابی دقیق نواحی سطحی پنل خورشیدی عملکرد بهتری دارد.

۴-۴. نتایج کمی مدل پیشنهادی

نتایج کمی مدل پیشنهادی روی مجموعه آزمون در جدول ۲ ارائه شده است. همان‌طور که مشاهده می‌شود، مدل پیشنهادی توانسته است در تشخیص و مکان‌یابی نواحی سطحی پنل خورشیدی عملکرد مناسبی به دست آورد. مقدار بالای $mAP@0.5$ نشان می‌دهد مدل در تشخیص کلی نواحی هدف عملکرد دقیقی دارد. همچنین، مقدار مناسب $mAP@0.5:0.95$ نشان می‌دهد جعبه‌های پیش‌بینی شده از نظر مکانی با برچسب‌های واقعی انطباق قابل قبولی دارند.

جدول ۲. نتایج ارزیابی مدل پیشنهادی روی مجموعه آزمون

مقدار	معیار ارزیابی
۰.۹۵۴	Precision
۰.۹۳۹	Recall
۰.۹۴۶	F1-score
۰.۹۶۶	mAP@0.5
۰.۷۶۶	mAP@0.5:0.95

نتایج نشان می‌دهد مدل پیشنهادی علاوه بر تشخیص صحیح نواحی مختلف، توانسته است تعادل مناسبی میان دقت و بازخوانی برقرار کند. این موضوع برای کاربردهای عملی اهمیت زیادی دارد؛ زیرا در پایش پنل‌های خورشیدی، هم کاهش تشخیص‌های اشتباه و هم جلوگیری از دست رفتن نواحی مهم ضروری است.

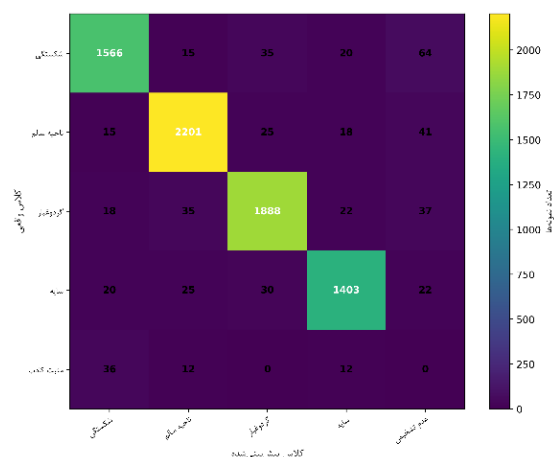
۴-۵. عملکرد مدل به تفکیک کلاس‌ها

برای بررسی دقیق‌تر عملکرد مدل، نتایج به تفکیک کلاس‌ها در جدول ۳ ارائه شده است. این تحلیل نشان می‌دهد مدل در تشخیص هر یک از کلاس‌های شکستگی، ناحیه سالم، گردوغبار و سایه چه عملکردی داشته است.

جدول ۳. عملکرد مدل پیشنهادی به تفکیک کلاس‌ها

کلاس	Precision	Recall	F1-score	AP@0.5
شکستگی	۰.۹۴۶	۰.۹۲۱	۰.۹۳۳	۰.۹۵۷
ناحیه سالم	۰.۹۶۲	۰.۹۵۷	۰.۹۵۹	۰.۹۷۴
گردوغبار	۰.۹۵۶	۰.۹۴۴	۰.۹۵۰	۰.۹۶۹
سایه	۰.۹۵۱	۰.۹۳۵	۰.۹۴۳	۰.۹۶۳
میانگین	۰.۹۵۴	۰.۹۳۹	۰.۹۴۶	۰.۹۶۶

شکل ۵ ماتریس آشفتگی مدل پیشنهادی را در سطح جعبه‌های محدودکننده نشان می‌دهد. در این ماتریس، چهار کلاس اصلی شامل شکستگی، ناحیه سالم، گردوغبار و سایه در نظر گرفته شده‌اند. علاوه بر آن، ستون «عدم تشخیص» برای نمایش نمونه‌هایی که در برچسب واقعی وجود داشته‌اند، اما توسط مدل شناسایی نشده‌اند، و ردیف «مثبت کاذب» برای نمایش پیش‌بینی‌هایی که متناظر واقعی نداشته‌اند، اضافه شده است. همان‌طور که مشاهده می‌شود، بیشترین خطاها مربوط به شکستگی‌های باریک و کم‌کتراست و همچنین، شباهت ظاهری میان گردوغبار و سایه است.



شکل ۵. ماتریس آشفتگی مدل پیشنهادی برای کلاس‌های شکستگی، ناحیه سالم، گردوغبار و سایه

بیشترین عملکرد مربوط به کلاس ناحیه سالم است، زیرا این کلاس معمولاً الگوی ظاهری یکنواخت‌تری دارد. در مقابل، کلاس شکستگی چالش بیشتری ایجاد می‌کند، زیرا شکستگی‌ها معمولاً باریک، کم‌کنتراست و از نظر ظاهری شبیه خطوط شبکه سلول‌های خورشیدی هستند. با وجود این، استفاده از اطلاعات لبه‌ای و تلفیق ویژگی‌های چندمقیاسی باعث شده است مدل در تشخیص شکستگی نیز عملکرد قابل قبولی داشته باشد. کلاس‌های گردوغبار و سایه نیز به دلیل وابستگی به شرایط نور، زاویه دید و تغییرات سطح پنل، چالش برانگیز هستند؛ با این حال نتایج نشان می‌دهد مدل توانسته است میان این نواحی و سطح سالم پنل تمایز مناسبی ایجاد کند.

۴-۶. مقایسه با مدل‌های پایه و روش‌های جدید

برای ارزیابی دقیق‌تر عملکرد مدل پیشنهادی، نتایج آن با چند آشکارساز پایه و تعدادی از مدل‌های جدید تشخیص عیوب فتوولتائیک مقایسه شد. مدل‌های مقایسه‌ای شامل YOLO-، YOLOv11n، DETR، RetinaNet، SSD، Faster R-CNN، PV، PV-YOLOv12n، MRA-YOLOv8، YOLOv8-FSD، BiFDRRep-YOLOv8n، DCD-YOLOv8s و مدل بهبودیافته YOLO11 هستند. انتخاب این مدل‌ها به این دلیل انجام شد که هم آشکارسازهای کلاسیک و عمومی و هم روش‌های جدید و نزدیک به حوزه فتوولتائیک در مقایسه حضور داشته باشند.

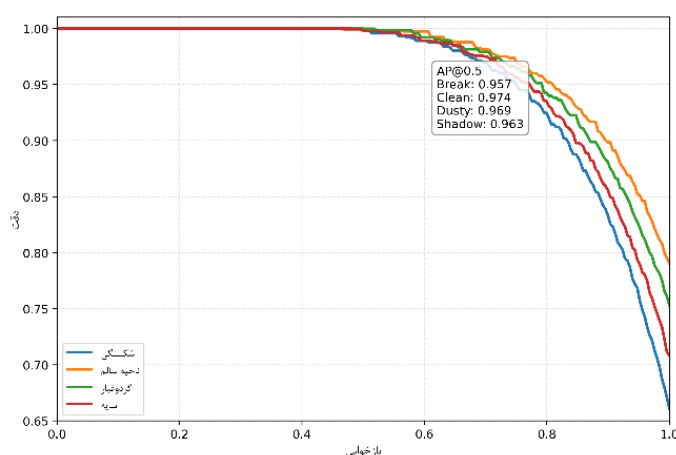
برای رعایت انصاف در مقایسه، همه مدل‌های پایه و مدل‌های جدید مورد مقایسه روی همان مجموعه داده این پژوهش و با تقسیم‌بندی یکسان آموزش، اعتبارسنجی و آزمون ارزیابی شدند. در تمام آزمایش‌ها، اندازه ورودی، روش‌های پیش‌پردازش، راهبرد افزایش داده، تعداد دوره‌های آموزش، اندازه دسته، معیار انتخاب بهترین وزن‌ها و معیارهای ارزیابی ثابت نگه داشته شد. به منظور کاهش اثر تصادفی بودن مقاردهای اولیه و نوسان‌های فرایند آموزش، هر مدل چند بار با تنظیمات یکسان اجرا شد و نتایج نهایی بر اساس بهترین عملکرد پایدار روی داده اعتبارسنجی و سپس ارزیابی روی مجموعه آزمون گزارش شد.

در مواردی که برخی مدل‌ها دارای تنظیمات معماری اختصاصی بودند، ساختار اصلی و تنظیمات پیشنهادی همان مدل حفظ شد، اما شرایط داده، تقسیم‌بندی، افزایش داده و معیارهای ارزیابی برای همه مدل‌ها یکسان در نظر گرفته شد. بنابراین، اعداد گزارش شده در جدول ۴ مربوط به اجرای همه مدل‌ها روی مجموعه داده مورد استفاده در این پژوهش است و فقط از مقالات اصلی نقل نشده‌اند. از این رو، این نتایج لزوماً با مقادیر گزارش شده در منابع اصلی یکسان نیستند، زیرا نوع داده، تعداد کلاس‌ها، شرایط تصویربرداری، کیفیت برچسب‌ها و تقسیم‌بندی آموزش/آزمون در هر پژوهش متفاوت است.

منحنی دقت - بازخوانی مدل پیشنهادی در شکل ۶ نشان می‌دهد مدل در بیشتر کلاس‌ها توانسته است تعادل مناسبی میان دقت و بازخوانی برقرار کند. سطح بالای منحنی‌ها بیانگر آن است که مدل علاوه بر کاهش تشخیص‌های اشتباه، توانسته است بخش قابل توجهی از نواحی واقعی را نیز شناسایی کند.

جدول ۴. مقایسه عملکرد مدل پیشنهادی با آشکارسازهای پایه و مدل‌های جدید

مدل	Precision	Recall	F1-score	mAP@0.5	mAP@0.5:0.95
Faster R-CNN[14]	۰.۸۷۶	۰.۸۴۲	۰.۸۵۹	۰.۸۸۴	۰.۶۰۳
SSD[15]	۰.۸۶۱	۰.۸۲۴	۰.۸۴۲	۰.۸۶۹	۰.۵۸۵
RetinaNet[16]	۰.۸۸۹	۰.۸۵۴	۰.۸۷۱	۰.۸۹۷	۰.۶۳۹
DETR[17]	۰.۹۰۲	۰.۸۷۲	۰.۸۸۷	۰.۹۱۲	۰.۶۶۱
YOLOv11n[26]	۰.۹۳۲	۰.۹۱۳	۰.۹۲۲	۰.۹۴۳	۰.۷۲۶
YOLO-PV[5]	۰.۹۳۸	۰.۹۱۹	۰.۹۲۸	۰.۹۵۱	۰.۷۳۹
PV-YOLOv12n[6]	۰.۹۴۱	۰.۹۲۳	۰.۹۳۲	۰.۹۵۳	۰.۷۴۴
MRA-YOLOv8[21]	۰.۹۳۶	۰.۹۱۶	۰.۹۲۶	۰.۹۴۸	۰.۷۳۵
YOLOv8-FSD[23]	۰.۹۲۹	۰.۹۰۷	۰.۹۱۸	۰.۹۴۱	۰.۷۲۱
BiFDRRep-YOLOv8n[24]	۰.۹۴۴	۰.۹۲۸	۰.۹۳۶	۰.۹۵۶	۰.۷۵۱
DCD-YOLOv8s[25]	۰.۹۴۷	۰.۹۳۰	۰.۹۳۸	۰.۹۵۷	۰.۷۵۳
Improved YOLO11[26]	۰.۹۴۹	۰.۹۳۲	۰.۹۴۰	۰.۹۵۸	۰.۷۵۵
مدل پیشنهادی	۰.۹۵۴	۰.۹۳۹	۰.۹۴۶	۰.۹۶۶	۰.۷۶۶



شکل ۶. منحنی دقت - بازخوانی مدل پیشنهادی برای کلاس‌های مختلف

همان‌طور که در جدول ۴ مشاهده می‌شود، مدل پیشنهادی در همه معیارها بهترین عملکرد را به دست آورده است. مدل‌های عمومی مانند Faster R-CNN، SSD، RetinaNet و DETR به دلیل ناتوانی نسبی در تشخیص نواحی کوچک و کم‌کنتراست عملکرد پایین‌تری داشتند، در حالی که مدل‌های جدید مبتنی بر YOLO نتایج نزدیک‌تری به مدل پیشنهادی ارائه کردند. برتری مدل پیشنهادی را می‌توان به ترکیب اطلاعات لبه‌ای، توجه کانالی - فضایی و تلفیق چندمقیاسی نسبت داد که باعث بهبود تشخیص نواحی باریک، نامنظم و دارای کنتراست پایین شده است.

۴-۷. مطالعه حذف اجزای مدل

برای بررسی نقش هر بخش از معماری پیشنهادی، مطالعه حذف اجزا انجام شد. در این آزمایش، هر بار یکی از اجزای اصلی مدل حذف شد و مدل با همان شرایط آموزشی و همان تقسیم‌بندی داده‌ها دوباره ارزیابی شد. هدف از این آزمایش، مشخص کردن سهم هر بخش در عملکرد نهایی مدل است.

اجزای بررسی‌شده شامل ماژول استخراج لبه سوبل، ماژول توجه کانالی - فضایی، ماژول بهبود چندمقیاسی، ساختار تلفیق ویژگی FPN/PAN و استخراج گر ویژگی ResNet101 هستند.

جدول ۵. نتایج مطالعه حذف اجزای مدل پیشنهادی

مدل	Precision	Recall	F1-score	mAP@0.5	mAP@0.5:0.95
بدون ماژول استخراج لبه	۰.۹۳۸	۰.۹۱۴	۰.۹۲۶	۰.۹۴۶	۰.۷۳۹
بدون توجه کانالی - فضایی	۰.۹۴۲	۰.۹۲۰	۰.۹۳۱	۰.۹۵۱	۰.۷۳۷
بدون ماژول کانولوشن گشادشده	۰.۹۴۶	۰.۹۲۵	۰.۹۳۵	۰.۹۵۵	۰.۷۴۴
بدون ماژول بهبود چندمقیاسی	۰.۹۴۰	۰.۹۱۷	۰.۹۲۸	۰.۹۴۸	۰.۷۳۳
بدون تلفیق FPN/PAN	۰.۹۳۴	۰.۹۰۹	۰.۹۲۱	۰.۹۴۱	۰.۷۱۸
بدون ResNet101 و با استخراج گر سبک	۰.۹۳۰	۰.۹۰۵	۰.۹۱۷	۰.۹۳۶	۰.۷۱۱
مدل کامل پیشنهادی	۰.۹۵۴	۰.۹۳۹	۰.۹۴۶	۰.۹۶۶	۰.۷۶۶

برای پاسخ دقیق‌تر به نقش ماژول کانولوشن گشادشده، یک آزمایش اضافی نیز انجام شد که در آن ساختار FPN/PAN حفظ شد، اما شاخه‌های کانولوشنی با نرخ گشادگی ($d=1,3,5$) حذف شدند. نتایج نشان داد حذف این ماژول مقدار [mAP@0.5](#) را از ۰/۹۶۶ به ۰/۹۵۵ و مقدار $mAP@0.5:0.95$ را از ۰/۷۶۶ به ۰/۷۴۴ کاهش می‌دهد. این افت عملکرد نشان می‌دهد کانولوشن‌های گشادشده تنها تکرار عملکرد FPN/PAN نیستند، بلکه با افزایش میدان دید مؤثر شبکه، به تشخیص نواحی دارای شکل نامنظم، شکستگی‌های کشیده و آلودگی‌های پراکنده کمک می‌کنند. با این حال، افت ناشی از حذف کامل FPN/PAN بیشتر است؛ بنابراین، FPN/PAN نقش اصلی را در تلفیق مقیاس‌ها دارد و ماژول کانولوشن گشادشده به عنوان یک تقویت‌کننده مکمل برای بهبود زمینه چندمقیاسی عمل می‌کند.

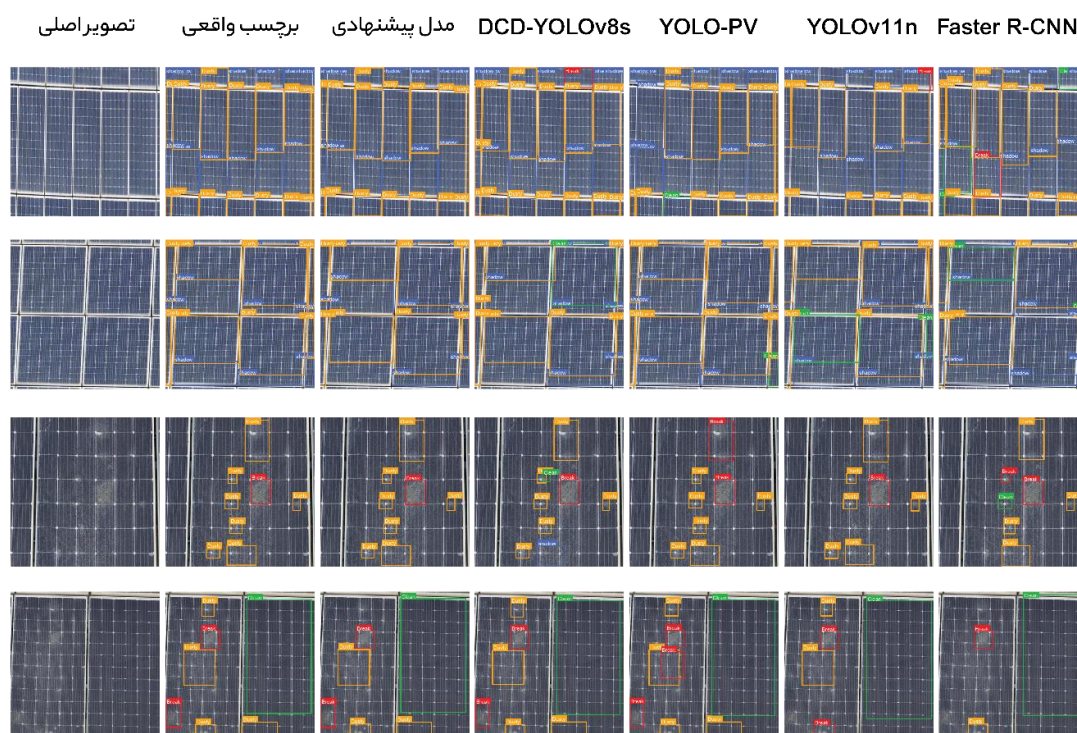
نتایج جدول ۵ نشان می‌دهد مدل کامل پیشنهادی بهترین عملکرد را در همه معیارها به دست آورده است. حذف ساختار FPN/PAN بیشترین افت را در $mAP@0.5$ و $mAP@0.5:0.95$ ایجاد کرده است. این موضوع نشان می‌دهد تلفیق ویژگی‌های چندمقیاسی برای تشخیص نواحی با اندازه‌های متفاوت نقش اساسی دارد. حذف ماژول استخراج لبه نیز باعث کاهش عملکرد در تشخیص نواحی باریک و مرزی، به‌ویژه شکستگی‌ها، شده است.

حذف توجه کانالی - فضایی باعث افت دقت و بازخوانی می‌شود، زیرا مدل توان کمتری در تمرکز بر نواحی مهم تصویر و کاهش اثر پس‌زمینه، بازتاب و خطوط تکراری پنل خواهد داشت. همچنین، جایگزینی ResNet101 با استخراج گر سبک‌تر موجب کاهش کیفیت ویژگی‌های عمیق و افت عملکرد کلی شده است. بنابراین، ترکیب هم‌زمان استخراج ویژگی عمیق، اطلاعات لبه‌ای، توجه کانالی - فضایی و تلفیق چندمقیاسی بیشترین اثر را در افزایش دقت تشخیص و مکان‌یابی نواحی سطحی پنل خورشیدی ایجاد کرده است.

۴-۸. تحلیل کیفی نتایج

برای بررسی بصری عملکرد مدل، چند نمونه از خروجی‌های تشخیص در شکل ۷ ارائه شده است. نتایج کیفی نشان می‌دهند مدل پیشنهادی در بسیاری از تصاویر توانسته است نواحی مربوط به شکستگی، گردوغبار، سایه و ناحیه سالم را با جعبه‌های مناسب مشخص کند. همچنین، در تصاویری که چند ناحیه مختلف به صورت هم‌زمان وجود دارند، مدل چند جعبه مستقل تولید کرده و برای هر یک کلاس متناظر را پیش‌بینی کرده است.

در برخی موارد، خطاهای مدل بیشتر در نواحی رخ می‌دهد که مرز میان کلاس‌ها واضح نیست. برای نمونه، سایه‌های شدید ممکن است با گردوغبار یا تغییر روشنایی سطح پنل اشتباه گرفته شوند. همچنین، در مواردی که شکستگی بسیار باریک یا کم‌کنتراست است، احتمال از دست رفتن آن افزایش می‌یابد. با این حال، استفاده از اطلاعات لبه‌ای و ویژگی‌های چندمقیاسی باعث شده است مدل در بیشتر نمونه‌ها موقعیت ناحیه مورد نظر را با دقت قابل قبول تشخیص دهد.



شکل ۷. نمونه‌هایی از تشخیص چندکلاسه و چندجبه‌ای مدل پیشنهادی روی تصاویر پنل خورشیدی

۴-۹. ارزیابی پایداری مدل در شرایط تصویری مختلف

برای بررسی پایداری مدل در برابر تغییرات رایج در تصاویر میدانی، عملکرد آن در شرایط تصویری مختلف ارزیابی شد. در این آزمایش، نمونه‌هایی با تغییر روشنایی، تغییر کنتراست، نویز، تاری و سایه شدید بررسی شدند. هدف از این ارزیابی، بررسی توان مدل در حفظ عملکرد هنگام مواجهه با شرایطی است که در بازرسی واقعی پنل‌های خورشیدی رخ می‌دهند.

جدول ۶. ارزیابی پایداری مدل پیشنهادی در شرایط تصویری مختلف

شرایط تصویر	Precision	Recall	F1-score	mAP@0.5
تصویر اصلی	۰.۹۵۴	۰.۹۳۹	۰.۹۴۶	۰.۹۶۶
تغییر روشنایی	۰.۹۴۸	۰.۹۳۱	۰.۹۳۹	۰.۹۵۷
تغییر کنتراست	۰.۹۴۶	۰.۹۲۹	۰.۹۳۷	۰.۹۵۴
نویز	۰.۹۳۸	۰.۹۱۶	۰.۹۲۷	۰.۹۴۴
تاری	۰.۹۲۳	۰.۸۹۹	۰.۹۱۱	۰.۹۲۸
سایه شدید	۰.۹۳۴	۰.۹۱۰	۰.۹۲۲	۰.۹۳۸

نتایج جدول ۶ نشان می‌دهد مدل پیشنهادی در برابر تغییرات روشنایی و کنتراست افت محدودی دارد. این موضوع نشان می‌دهد مدل تنها به شدت رنگ یا روشنایی وابسته نیست و توانسته است ویژگی‌های ساختاری و مکانی نواحی را نیز یاد بگیرد. بیشترین افت عملکرد مربوط به حالت تاری است، زیرا در این وضعیت بخشی از اطلاعات لبه‌ای و جزئیات ظریف از بین می‌رود. همچنین سایه شدید باعث کاهش نسبی دقت می‌شود، زیرا مرز میان سایه، گردوغبار و سطح سالم در برخی نمونه‌ها مبهم می‌شود. با این حال، افت عملکرد در همه شرایط در محدوده قابل قبول باقی مانده و این موضوع نشان‌دهنده پایداری مدل پیشنهادی در شرایط تصویری متنوع است.

۵. بحث و نتیجه‌گیری

در این پژوهش، یک چارچوب مبتنی بر بینایی ماشین و یادگیری عمیق برای تشخیص و مکان‌یابی نواحی سطحی پنل‌های خورشیدی ارائه شد. هدف اصلی مدل پیشنهادی، شناسایی هم‌زمان چند ناحیه در یک تصویر و تعیین کلاس، موقعیت و امتیاز اطمینان هر ناحیه در قالب جعبه محدودکننده بود. به خلاف روش‌های طبقه‌بندی که تنها وضعیت کلی تصویر را مشخص می‌کنند، روش پیشنهادی اطلاعات مکانی قابل استفاده برای پایش، پاک‌سازی، تعمیر و نگهداری پنل‌های خورشیدی را فراهم می‌کند.

معماری پیشنهادی از استخراج گر ویژگی ResNet101، ماژول استخراج لبه مبتنی بر عملگر سوبل، ماژول توجه کانالی - فضایی، ساختار تلفیق ویژگی چندمقیاسی FPN/PAN و سر تشخیص چندمقیاسی تشکیل شده است. استفاده از ResNet101 باعث استخراج ویژگی‌های عمیق و پایدار از تصویر شد. ماژول سوبل اطلاعات لبه‌ای و ساختاری را برای تشخیص نواحی باریک و کم‌کنتراست تقویت کرد. همچنین، ماژول توجه کانالی - فضایی تمرکز مدل را بر نواحی مهم‌تر افزایش داد و ساختار FPN/PAN امکان تشخیص نواحی در مقیاس‌های مختلف را فراهم ساخت.

نتایج آزمایش‌ها نشان داد مدل پیشنهادی با دستیابی به دقت ۰/۹۵۴، بازخوانی ۰/۹۳۹، امتیاز F1 برابر با ۰/۹۴۶، مقدار [mAP@0.5](#) برابر با ۰/۹۶۶ و مقدار $mAP@0.5:0.95$ برابر با ۰/۷۶۶، عملکرد مناسبی در تشخیص و مکان‌یابی نواحی سطحی پنل خورشیدی دارد. مقایسه با مدل‌های پایه و روش‌های جدید نشان داد مدل پیشنهادی در همه معیارهای ارزیابی عملکرد بالاتری دارد. همچنین، نتایج تفکیکی کلاس‌ها نشان داد مدل توانسته است کلاس‌های شکستگی، ناحیه سالم، گردوغبار و سایه را با دقت قابل قبول تشخیص دهد.

مطالعه حذف اجزای مدل نشان داد هر یک از بخش‌های معماری پیشنهادی در بهبود عملکرد نهایی نقش دارند. حذف ساختار تلفیق ویژگی چندمقیاسی بیشترین افت عملکرد را ایجاد کرد و نشان داد تشخیص عیوب با اندازه‌های متفاوت نیازمند ترکیب ویژگی‌ها در چند سطح است. همچنین، حذف ماژول استخراج لبه و توجه کانالی - فضایی باعث کاهش توان مدل در تشخیص نواحی مرزی، کم‌کنتراست و کوچک شد. این نتایج نشان می‌دهد ترکیب ویژگی‌های عمیق، اطلاعات لبه‌ای، توجه و تلفیق چندمقیاسی برای تشخیص دقیق نواحی سطحی پنل‌های خورشیدی مؤثر است.

بر اساس نتایج به‌دست‌آمده، روش پیشنهادی می‌تواند به عنوان یک ابزار کاربردی در سامانه‌های پایش خودکار پنل‌های خورشیدی مورد استفاده قرار گیرد. خروجی مدل، شامل کلاس و موقعیت هر ناحیه، می‌تواند در فرایندهایی مانند پاک‌سازی هدفمند، تعمیر، تعویض، اولویت‌بندی بازرسی و نگهداری پیش‌بینانه به کار گرفته شود. در پژوهش‌های آینده، می‌توان با افزایش حجم داده‌ها، استفاده از تصاویر پهپادی، افزودن داده‌های حرارتی، بررسی مدل‌های سبک‌تر و پیاده‌سازی روی سخت‌افزارهای کم‌مصرف، قابلیت استقرار عملی مدل را بهبود داد.

منابع

- [1] W. Tang, Q. Yang, Z. Dai, and W. Yan, "Module defect detection and diagnosis for intelligent maintenance of solar photovoltaic plants: Techniques, systems and perspectives," *Energy*, vol. 297, p. 131222, 2024/06/15/ 2024, doi: <https://doi.org/10.1016/j.energy.2024.131222>.
- [2] A. Thakfan and Y. Bin Salamah, "Artificial-Intelligence-Based Detection of Defects and Faults in Photovoltaic Systems: A Survey," *Energies*, vol. 17, no. 19, p. 4807, 2024. [Online]. Available: <https://www.mdpi.com/1996-1073/17/19/4807>.
- [3] M. Jalal, I. U. Khalil, and A. u. Haq, "Deep learning approaches for visual faults diagnosis of photovoltaic systems: State-of-the-Art review," *Results in Engineering*, vol. 23, p. 102622, 2024/09/01/ 2024, doi: <https://doi.org/10.1016/j.rineng.2024.102622>.
- [4] D. Lang and Z. Lv, "A PV cell defect detector combined with transformer and attention mechanism," *Scientific Reports*, vol. 14, no. 1, p. 20671, 2024/09/05 2024, doi: 10.1038/s41598-024-72019-5.
- [5] W. Cai and H. Lv, "YOLO-PV: An Enhanced YOLO11n Model with Multi-Scale Feature Fusion for Photovoltaic Panel Defect Detection," *Energies*, vol. 18, no. 20, p. 5489, 2025. [Online]. Available: <https://www.mdpi.com/1996-1073/18/20/5489>.
- [6] A. Mohamed *et al.*, "Optimized YOLO based model for photovoltaic defect detection in electroluminescence images," *Scientific Reports*, vol. 15, no. 1, p. 32955, 2025/09/26 2025, doi: 10.1038/s41598-025-13956-7.
- [7] A. Rahman, "Solar Panel Surface Defect and Dust Detection: Deep Learning Approach," *Journal of Imaging*, vol. 11, no. 9, p. 287, 2025. [Online]. Available: <https://www.mdpi.com/2313-433X/11/9/287>.
- [8] J. Redmon, S. Divvala, R. Girshick, and A. Farhadi, "You Only Look Once: Unified, Real-Time Object Detection," in *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 27-30 June 2016 2016, pp. 779-788, doi: 10.1109/CVPR.2016.91 .
- [9] K. He, X. Zhang, S. Ren, and J. Sun, "Deep Residual Learning for Image Recognition," in *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR ۲۰۱۶)*, (June 2016 2016, pp. 770-778, doi: 10.1109/CVPR.2016.90 .
- [10] A. Guha Roy, N. Navab, and C. Wachinger, "Concurrent Spatial and Channel Squeeze & Excitation in Fully Convolutional Networks," 03/07 2018, doi: 10.48550/arXiv.1803.02579.
- [11] S. Deitsch *et al.*, "Automatic classification of defective photovoltaic module cells in electroluminescence images," *Solar Energy*, vol. 185, pp. 455-468, 2019/06/01/ 2019, doi: <https://doi.org/10.1016/j.solener.2019.02.067>.
- [12] W. Tang, Q. Yang, K. Xiong, and W. Yan, "Deep learning based automatic defect identification of photovoltaic module using electroluminescence images," *Solar Energy*, vol. 201, pp. 453-460, 2020/05/01/ 2020, doi: <https://doi.org/10.1016/j.solener.2020.03.049>.
- [13] M. W. Akram, G. Li, Y. Jin ,X. Chen, C. Zhu, and A. Ahmad, "Automatic detection of photovoltaic module defects in infrared images with isolated and develop-model transfer deep learning," *Solar Energy*, vol. 198, pp. 175-186, 2020/03/01/ 2020, doi: <https://doi.org/10.1016/j.solener.2020.01.055>.
- [14] S. Ren, K. He, R. Girshick, and J. Sun, "Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 39, no. 6, pp. 1137-1149, 2017, doi: 10.1109/TPAMI.2016.2577031.
- [15] W. Liu *et al.*, "SSD: Single Shot MultiBox Detector," in *Computer Vision – ECCV 2016*, Cham, B. Leibe, J. Matas, N. Sebe, and M. Welling, Eds., 2016// 2016: Springer International Publishing, pp. 21-37 .
- [16] T. Y. Lin, P. Goyal, R. Girshick, K. He, and P. Dollár, "Focal Loss for Dense Object Detection," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 42, no. 2, pp. 318-327, 2020, doi: 10.1109/TPAMI.2018.2858826.
- [17] N. Carion, F. Massa, G. Synnaeve, N. Usunier ,A. Kirillov, and S. Zagoruyko, "End-to-End Object Detection with Transformers," in *Computer Vision – ECCV 2020*, Cham, A. Vedaldi, H. Bischof, T. Brox, and J.-M. Frahm, Eds., 2020// 2020: Springer International Publishing, pp. 213-229 .

- [18] W. Pan, X. Sun, Y. Wang, Y. Cao, Y. Lang, and Y. Qian, "Enhanced photovoltaic panel defect detection via adaptive complementary fusion in YOLO-ACF," *Scientific Reports*, vol. 14, no. 1, p. 26425, 2024/11/02 2024, doi: 10.1038/s41598-024-75772-9.
- [19] Y. Jia, G. Chen, and L. Zhao, "Defect detection of photovoltaic modules based on improved VarifocalNet," *Scientific Reports*, vol. 14, no. 1, p. 15170, 2024/07/02 2024, doi: 10.1038/s41598-024-66234-3.
- [20] Z. Qu *et al.*, "A photovoltaic cell defect detection model capable of topological knowledge extraction," *Scientific Reports*, vol. 14, no. 1, p. 21904, 2024/09/19 2024, doi: 10.1038/s41598-024-72717-0.
- [21] N. Wang, S. Huang, X. Liu, Z. Wang, Y. Liu, and Z. Gao, "MRA-YOLOv8: A Network Enhancing Feature Extraction Ability for Photovoltaic Cell Defects," *Sensors*, vol. 25, no. 5, p. 1542, 2025. [Online]. Available: <https://www.mdpi.com/1424-8220/25/5/1542>.
- [22] H. Xue, L. Liu, Q. Wu, J. He, and Y. Fan, "Defect Detection Algorithm for Photovoltaic Cells Based on SEC-YOLOv8," *Processes*, vol. 13, no. 8, p. 2425, 2025. [Online]. Available: <https://www.mdpi.com/2227-9717/13/8/2425>.
- [23] C. Chen, Z. Chen, H. Li, Y. Wang, G. Lei, and L. Wu, "Research on Defect Detection in Lightweight Photovoltaic Cells Using YOLOv8-FSD," *Sensors*, vol. 25, no. 3, p. 843, 2025. [Online]. Available: <https://www.mdpi.com/1424-8220/25/3/843>.
- [24] Y. Lu, C. Du, X. Li, S. Liang, Q. Zhang, and Z. Zhao, "A High-Precision Defect Detection Approach Based on BiFDRRep-YOLOv8n for Small Target Defects in Photovoltaic Modules," *Energies*, vol. 18, no. 9, p. 2299, 2025. [Online]. Available: <https://www.mdpi.com/1996-1073/18/9/2299>.
- [25] M. Al-Mahbashi *et al.*, "An effective approach to improving photovoltaic defect detection using the new DCD-YOLOv8s model," *Scientific Reports*, vol. 15, no. 1, p. 38308, 2025/11/03 2025, doi: 10.1038/s41598-025-22307-5.
- [26] W. Jiang and W. Liu, "Study on defect detection and identification of solar photovoltaic (PV) panels based on improved YOLO11," *Scientific Reports*, 2026/04/24 2026, doi: 10.1038/s41598-026-48824-5.
- [27] S. Meng and T. Xu, "Inspection Method Enabled by Lightweight Self-Attention for Multi-Fault Detection in Photovoltaic Modules," *Electronics*, 2025.
- [28] M. Dhimish, "HOTSPOT-YOLO: A Lightweight Deep Learning Attention-Driven Model for Detecting Thermal Anomalies in Drone-Based Solar Photovoltaic Inspections," *ArXiv*, vol. abs/2508.18912, 2025.
- [29] H. He *et al.*, "A Self-Evolving Defect Detection Framework for Industrial Photovoltaic Systems," *arXiv preprint arXiv:2603.14869*, 2026.
- [30] D. Grzechca, F. Ez-Zahiri, Ł. Chruszczyk, and F. Bian, "Defect-Aware RGB Representation and Resolution-Efficient Deep Learning for Photovoltaic Failure Detection in Electroluminescence Images," *Applied Sciences*, vol. 16, no. 4, p. 2148, 2026. [Online]. Available: <https://www.mdpi.com/2076-3417/16/4/2148>.
- [31] F. M. Talaat, M. Salem, and W. M. Shaban, "AI-driven fault detection and classification in photovoltaic systems using deep learning techniques," *Scientific Reports*, vol. 16, no. 1, p. 8727, 2026/03/10 2026, doi: 10.1038/s41598-026-40246-7.
- [32] A. Seghiour *et al.*, "Fault detection and diagnosis in photovoltaic systems using artificial intelligence and time–frequency analysis," *Scientific Reports*, vol. 16, no. 1, p. 10056, 2026/02/20 2026, doi: 10.1038/s41598-026-39386-7.
- [33] S. R. Mohanty, M. U. Maruf, V. Singh, and Z. Ahmad, "Machine learning approaches for automatic defect detection in photovoltaic systems," *Solar Energy*, vol. 298, p. 113672, 2025/09/15/ 2025, doi: <https://doi.org/10.1016/j.solener.2025.113672>.
- [34] E. Zaman and R. Khanam, *PV-faultNet: Optimized CNN Architecture to detect defects resulting efficient PV production*. 2024.
- [35] J. Zhang, X. Chen, H. Wei, and K. Zhang, "A lightweight network for photovoltaic cell defect detection in electroluminescence images based on neural architecture search and knowledge distillation," *Applied Energy*, vol. 355, p. 122184, 2024/02/01/ 2024, doi: <https://doi.org/10.1016/j.apenergy.2023.122184>.