EAI Endorsed Transactions
on Energy Web
Research Article EAI.EU

Assessing Lightweight YOLO Models for Vision-Based Fault Diagnosis in Wind Turbine and Cold Chain Inspection Systems

Siming Lin[1], Yongqiang Xiao[2], Yifang Gao[3], Rui Ni[4], Tianxiang Huang[5,*]

[1]Xiamen Key Laboratory of Intelligent Fishery, Xiamen Ocean Vocational College, Xiamen, China
[2]School of Economics and Business Administration, Quanzhou Ocean Institute, Quanzhou 362700, China
[3]School of Electrical & Electronic Engineering, Universiti Sains Malaysia, Penang, Malaysia
[4]EAST LAKE HIGH TECHNOLOGY GROUP CO., LTD., Wuhan, China
[5]School of Information Engineering, Zhongnan University of Economics and Law, Wuhan, China

Abstract

The reliability of wind energy systems depends on the timely detection of surface and structural defects in turbine blades. This paper compares six YOLO (You Only Look Once) architectures (v8n–v13n) for automatic fault detection in wind turbine inspection images. All models were trained under identical experimental settings and evaluated by precision, recall, mAP@0.50, mAP@[0.50:0.95], and inference latency. Results show that **YOLOv12n** achieved the highest performance (**mAP@0.50 = 0.867**, computed as the mean over three seeds using the Top-3 class protocol), while **YOLOv10n** delivered the lowest inference time of **0.7 ms**. These findings support the suitability of lightweight YOLO variants for real-time fault inspection and predictive maintenance in wind energy systems and the same framework can be extended to visual quality inspection tasks in intelligent cold chain logistics systems.

Received on 15 October 2025; accepted on 16 November 2025; published on 09 February 2026

Keywords: YOLO, wind turbine inspection, defect detection, computer vision, renewable energy, predictive maintenance

doi:10.4108/ew.11842

1. Introduction

As renewable energy deployment accelerates worldwide, wind power has become a central pillar of sustainable electricity generation [2, 3]. The long-term reliability of wind farms depends heavily on the structural integrity of turbine blades, which are vulnerable to erosion, corrosion, cracking, and lightning burns. Such surface defects can degrade aerodynamic performance, shorten component lifespan, and cause costly downtime [4, 5]. Conventional visual inspections require manual intervention and are prone to inconsistencies, highlighting the need for automated fault detection systems capable of continuous and objective monitoring [6, 7].

Deep learning-based object detectors have shown great potential in renewable energy inspection tasks, especially for identifying complex textures and small-scale surface damage on turbine blades [8]. Among them, the YOLO (You Only Look Once) family of models has become widely adopted due to its strong balance between detection accuracy and inference speed [9–11]. Recent architectures—YOLOv8 through YOLOv13—have introduced significant innovations, including re-parameterized backbones, anchor-free detection heads, and attention-based feature fusion, improving performance in outdoor inspection scenarios with varying illumination and occlusion [12–17].

Despite this progress, systematic benchmarking of recent YOLO versions for wind turbine defect detection remains scarce. This study evaluates six YOLO architectures (v8n–v13n) under unified experimental settings to quantify differences in accuracy, precision, recall, mAP, and inference latency. The results aim to identify models best suited for intelligent inspection, predictive maintenance, and real-time monitoring in

*Corresponding author. Email: yvonnegogh@hotmail.com

modern wind energy systems. Given the shared need for rapid and reliable visual assessment, this study's approach can also be generalized to automated defect detection in cold chain logistics management.

2. Related Works

2.1. Vision-Based Defect Detection in Wind Energy Systems

Automated defect detection has become a critical component of modern wind farm maintenance. Conventional manual inspections are inefficient and prone to subjective error, motivating the adoption of computer vision and deep learning for turbine blade monitoring. Recent advances have demonstrated that convolutional networks and transformer-based detectors can identify corrosion, cracking, and surface contamination under variable outdoor conditions. For instance, Li et al. [23] integrated ViT (Vision Transformer) modules with multi-scale feature pyramids, achieving improved generalization on aerial wind turbine imagery. Similarly, Ren et al. [24] utilized multispectral fusion networks to enhance defect detection robustness under varying illumination and reflective blade surfaces. Zou et al. [5] proposed DCW-YOLO to enhance multi-scale feature learning, improving mAP by 3.2% over YOLOv5 on blade surface defects. Cao and Wang [4] developed a CNN-based framework capable of handling uneven illumination and motion blur, while Zhang et al. [19] introduced a hybrid attention mechanism combining spatial and channel features for blade erosion classification.

Several recent works have adapted YOLO architectures to domain-specific inspection tasks. WTDB-YOLOv8 [8] optimized backbone and feature aggregation to detect small-scale defects, and GCB-YOLO [7] introduced a lightweight attention fusion module for edge deployment. Zhou and Li [20] improved model generalization using transfer learning and photometric augmentation, achieving consistent detection across different turbine models. Meanwhile, Wu et al. [18] designed a compact CNN with pruning-based optimization for embedded visual inspection, reducing model size by 60% while maintaining above 90% accuracy. These studies confirm the effectiveness of deep learning in renewable energy inspection but often evaluate individual models without cross-version comparison. Additionally, synthetic data generation and domain randomization have been proposed to address the scarcity of defect samples in renewable inspection datasets [27].

2.2. Evolution of YOLO Architectures for Real-Time Detection

The YOLO series remains one of the most influential one-stage detection frameworks for real-time computer vision. YOLOv1–v7 established its foundation with unified regression-based prediction and efficient backbone design [9–11]. More recently, YOLOv8 introduced anchor-free detection and re-parameterized blocks [12], YOLOv9 employed multi-task gradient learning [13], and YOLOv10 incorporated attention-guided feature maps and end-to-end training [14]. YOLOv11 to YOLOv13 further extended detection robustness through dynamic feature fusion and transformer-based modules [1, 16, 17].

Recent studies have further explored hybrid attention and transformer–CNN fusion to balance real-time inference and contextual awareness [25, 26]. Beyond architectural evolution, efforts have been made to adapt YOLO models for low-power and edge-AI environments. Zhang et al. [21] achieved sub-20 ms inference latency on Jetson AGX platforms, and Chen et al. [22] demonstrated an edge-AI pipeline enabling autonomous inspection in offshore wind farms. Despite these advances, comprehensive benchmarking of YOLOv8–YOLOv13 for renewable energy inspection remains limited. A standardized comparative analysis can thus reveal trade-offs between accuracy, efficiency, and deployability, offering practical guidance for predictive maintenance in wind power systems.

3. Methodology

3.1. Dataset and Preprocessing

The dataset used in this work contains high-resolution images of wind turbine blades with visible surface anomalies, including corrosion, cracks, lightning burns, erosion, and dust or oil stains. A total of 3688 images were organized into nine categories of surface conditions. All samples were annotated using bounding boxes in YOLO format and manually checked for label consistency through CVAT. The labeling process emphasized edge boundaries and irregular contours to ensure reliable ground truth.

To enhance robustness under different environmental conditions, the data were augmented using multiple strategies: random rotation, horizontal reflection, brightness adjustment, and synthetic mosaic generation. These augmentations have been demonstrated to improve the generalization of defect inspection systems [21, 28]. Before feeding into the model, all images were resized to 640×640 pixels and normalized between 0 and 1. The dataset was split into 70% training, 20% validation, and 10% testing partitions while preserving class ratios to avoid imbalance.

3.2. Model Configuration and Training

Six YOLO architectures were benchmarked to assess the trade-off between detection accuracy and computational cost. These include YOLOv8n, YOLOv9t, YOLOv10n, YOLOv11n, YOLOv12n, and YOLOv13n. Each variant represents a compact network optimized for high-speed inspection tasks in industrial environments. The essential properties of the evaluated models are summarized in Table 1. To reduce duplication with prior studies, the table is sorted by computational complexity (GFLOPs) rather than by version number.

All models were trained with the Ultralytics YOLO stack (v8.3.63) on PyTorch 2.6.0. Runs used a single NVIDIA A100 (40 GB) with CUDA 12.4. We fine-tuned official pretrained checkpoints and trained for 100-250 epochs with early stopping set to a patience of 30. Inputs were resized to 640×640, the initial learning rate was 0.01, and the batch size was 128. Larger backbones were not considered to preserve deployability for field monitoring and edge inference.

Evaluation was performed using standard object detection metrics: precision (P), recall (R), and mean average precision (mAP). Precision measures the proportion of correctly detected defects among all predictions, while recall quantifies the ratio of correctly identified defects to total ground-truth instances. Mean average precision provides an overall summary across all categories and thresholds. The formulas used are shown in Eqs. (1)–(3).

$$\text{Precision} = \frac{TP}{TP + FP} \tag{1}$$

$$\text{Recall} = \frac{TP}{TP + FN} \tag{2}$$

$$\text{mAP} = \frac{1}{N} \sum_{i=1}^{N} \text{AP}_i \tag{3}$$

Here, TP, FP, and FN refer respectively to true positives, false positives, and false negatives across N defect types. In addition to accuracy, computational performance indicators such as inference time (ms), frames per second (FPS) were recorded. These criteria are essential for evaluating whether detection systems can operate efficiently in real-time inspection of large-scale wind farms.

4. Results and Discussion

4.1. Training Behavior and Convergence

All six YOLO variants demonstrated stable convergence during fine-tuning on the turbine defect dataset. Figure 1 shows uniformly distributed box centers and right-skewed object sizes, evidencing a small-object regime. The weak positive w–h correlation and the triangular bands in (w, x) and (h, y) plots suggest perspective effects along the blade span rather than a structured, single-mode size distribution. Figure 2 displays YOLOv12n's training and validation loss trajectories averaged over 9 classes with seed 73 across 224 epoches. The validation loss closely followed training, indicating minimal overfitting and consistent optimization. Across models, convergence was generally achieved within 100–250 epochs, with YOLOv10n showing the fastest stability.

4.2. Detection Accuracy Across Seeds

Each YOLO variant was trained three times with different random seeds to ensure robustness. Table 2 summarizes mean ± standard deviation of precision, recall, mAP@0.50, and mAP@[0.50:0.95]. Overall, YOLOv9t achieved the highest mAP@[0.50:0.95] (0.876 ± 0.003), while YOLOv10n and YOLOv12n followed closely. Among all variants, YOLOv12n demonstrated the highest overall mAP and precision across the top-performing defect categories, indicating stronger detection reliability for field deployment.

In addition to per-category and per-seed averaging, the top-performing class for each model was extracted to identify which architectures excelled under specific defect types. Table 3 summarizes the best-performing category, along with its corresponding mAP@0.50 and precision values averaged across three random seeds. Across all variants, lightning burn and surface corrosion consistently emerged as the most confidently detected categories, due to their clear color contrast and larger spatial footprint in the dataset. YOLOv11n achieved the highest single-class precision (0.941) and maintained robust localization for lightning burn regions, while YOLOv10n provided the best trade-off between accuracy (mAP@0.50 = 0.834) and inference speed. In contrast, dust-oil and non-open cracking remained challenging due to low contrast and texture blending.

To further compare the overall capability of each architecture across all defect categories, we aggregated the mean metrics from the three seeds and identified the single best-performing model for several key indicators. As shown in Table 4, YOLOv10n achieved the highest inference speed of approximately 1429 FPS, making it the preferred choice for real-time edge deployment. YOLOv13n exhibited the fastest training convergence, reaching stable performance in only 122.3 epochs on average. In terms of overall detection accuracy, YOLOv11n outperformed all other variants by delivering the highest precision (0.489), recall (0.315), mAP@0.50 (0.334), and mAP@[0.50:0.95] (0.190) when averaged over the entire test set.

Table 1. Specifications of YOLO Variants for Wind Turbine Blade Defect Detection

Model	*GFLOPs (B)*	*Parameters (M)*	*Layers*
YOLOv8n	8.2	3.01	225
YOLOv9t	7.9	2.01	917
YOLOv10n	8.4	2.71	385
YOLOv11n	6.4	2.59	319
YOLOv12n	6.5	2.55	497
YOLOv13n	6.4	2.46	648

[a]All models were trained and validated under identical experimental conditions for fair comparison.

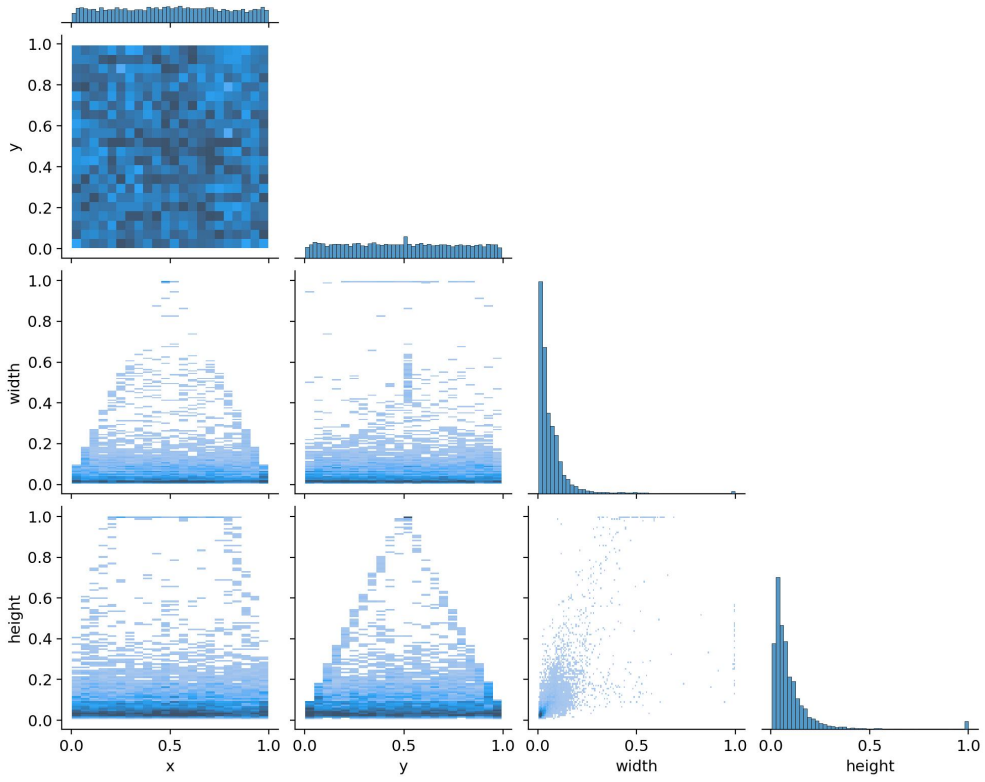


Figure 1. Pairwise distributions of bounding-box center (x, y) and size (w, h) on the wind-turbine blade dataset. Box centers are uniform across the frame (top-left heatmap), while w/h histograms are sharply right-skewed, indicating a strong small-object regime. The triangular bands in $(w \times x)$ and $(h \times y)$ reflect perspective along the blade span, and the w–h scatter shows a weak positive correlation with many tiny patches plus a few large outliers (e.g., corrosion areas).

4.3. Qualitative Visualization of Defect Detection

Figure 3 shows diverse defect types and challenges in the blade dataset, including small-object detection and texture overlap while Figure 4 presents validation mosaics for YOLOv12n, showing ground-truth labels (left) and model predictions (right) for two batches.

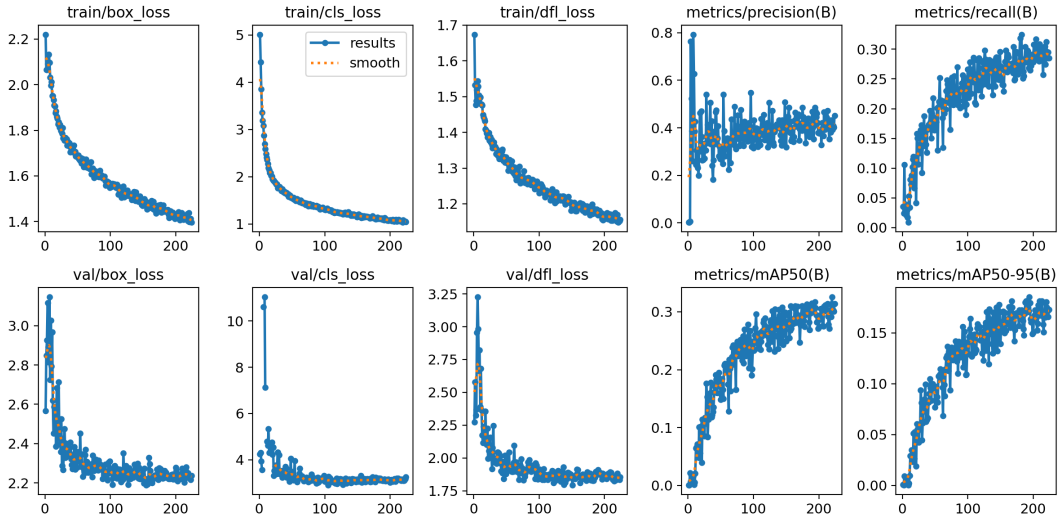
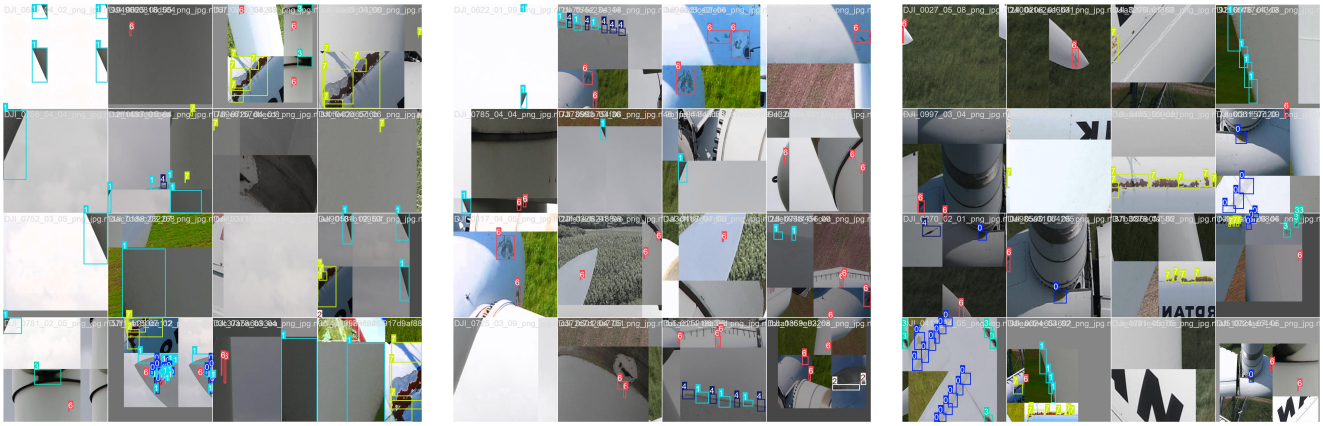


Figure 2. Training and validation loss on the wind–turbine blade dataset for YOLOv12n at seed 73. The blue curve shows per-epoch loss and the orange dashed curve shows a moving average. Both decrease steadily, with validation closely tracking training, indicating stable optimization and no obvious overfitting.



(a) Train Batch 0 (b) Train Batch 1 (c) Train Batch 2

Figure 3. Qualitative samples from training batches on the blade dataset. Frames include diverse defect types (*surface corrosion*, *lightning burn*, *dust-oil*, etc.), varied illumination, and scales. Boxes illustrate the small-object regime and texture overlap common in field inspections.

Visible categories include paint peeling, surface corrosion, attachments, and non-open cracking, with occasional lightning burn. Bounding boxes generally follow blade edges and flange contours, and YOLOv12n reproduces high-confidence detections on large, high-contrast defects such as surface corrosion and lightning burn. In contrast, fine, low-contrast textures like dust-oil and thin non-open cracking exhibit missed detections and sporadic false positives, often around specular highlights or grass–blade boundaries. These mosaics qualitatively match the quantitative trend: strong localization on corrosion and burn regions, and weaker performance on small, texture-like defects.

4.4. Efficiency and Deployment Considerations

For wind turbine inspection, both inference speed and deployability are critical alongside detection accuracy. Table 5 summarizes the inference efficiency and model size of all YOLO variants. Among them, YOLOv10n achieved the fastest average inference time (0.7 ms) and the highest frame rate (1429 FPS), demonstrating excellent real-time performance for field applications. YOLOv8n and YOLOv11n followed closely with sub-1.5 ms latency. Despite having the highest latency (2.2 ms), YOLOv13n provided a balanced trade-off between detection consistency and compact model size (5.17 MB).

Table 2. Performance on three representative defect classes (mean over seeds 0, 37, 73).

Model	Class	Precision	Recall	mAP@0.50	mAP@[0.50:0.95]
YOLOv8n	Surface corrosion	0.719	0.689	0.728	0.398
	Lightning burn	0.820	0.667	0.742	0.459
	Dust-oil	0.651	0.158	0.282	0.157
YOLOv9t	Surface corrosion	0.721	0.707	0.752	0.403
	Lightning burn	0.876	0.764	0.838	0.514
	Dust-oil	0.660	0.169	0.287	0.163
YOLOv10n	Surface corrosion	0.738	0.707	0.762	0.414
	Lightning burn	0.840	0.717	0.792	0.475
	Dust-oil	0.644	0.164	0.279	0.156
YOLOv11n	Surface corrosion	0.731	0.744	0.770	0.429
	Lightning burn	0.939	0.743	0.841	0.507
	Dust-oil	0.656	0.199	0.291	0.150
YOLOv12n	Surface corrosion	0.734	0.736	0.762	0.415
	Lightning burn	0.900	0.789	0.867	0.545
	Dust-oil	0.657	0.177	0.290	0.157
YOLOv13n	Surface corrosion	0.792	0.665	0.753	0.413
	Lightning burn	0.870	0.744	0.813	0.497
	Dust-oil	0.648	0.164	0.286	0.159

Table 3. Top-performing categories across YOLOv8–YOLOv13.

Model	Top Class	mAP@0.50	Precision
YOLOv8n	surface corrosion	0.749	0.742
YOLOv9t	lightning burn	0.780	0.914
YOLOv10n	lightning burn	0.834	0.801
YOLOv11n	lightning burn	0.841	0.941
YOLOv12n	lightning burn	0.752	0.800
YOLOv13n	surface corrosion	0.750	0.716

Table 4. Best mean metrics across YOLOv8–YOLOv13.

Metric	Best Model	Mean Value
Highest FPS†	YOLOv10n	∼1429 FPS
Lowest epochs (convergence)	YOLOv13n	122.3 epochs
Best Precision (overall)	YOLOv9t	0.489
Best Recall (overall)	YOLOv11n	0.315
Best mAP@0.50 (overall)	YOLOv11n	0.334
Best mAP@[0.50:0.95] (overall)	YOLOv11n	0.190

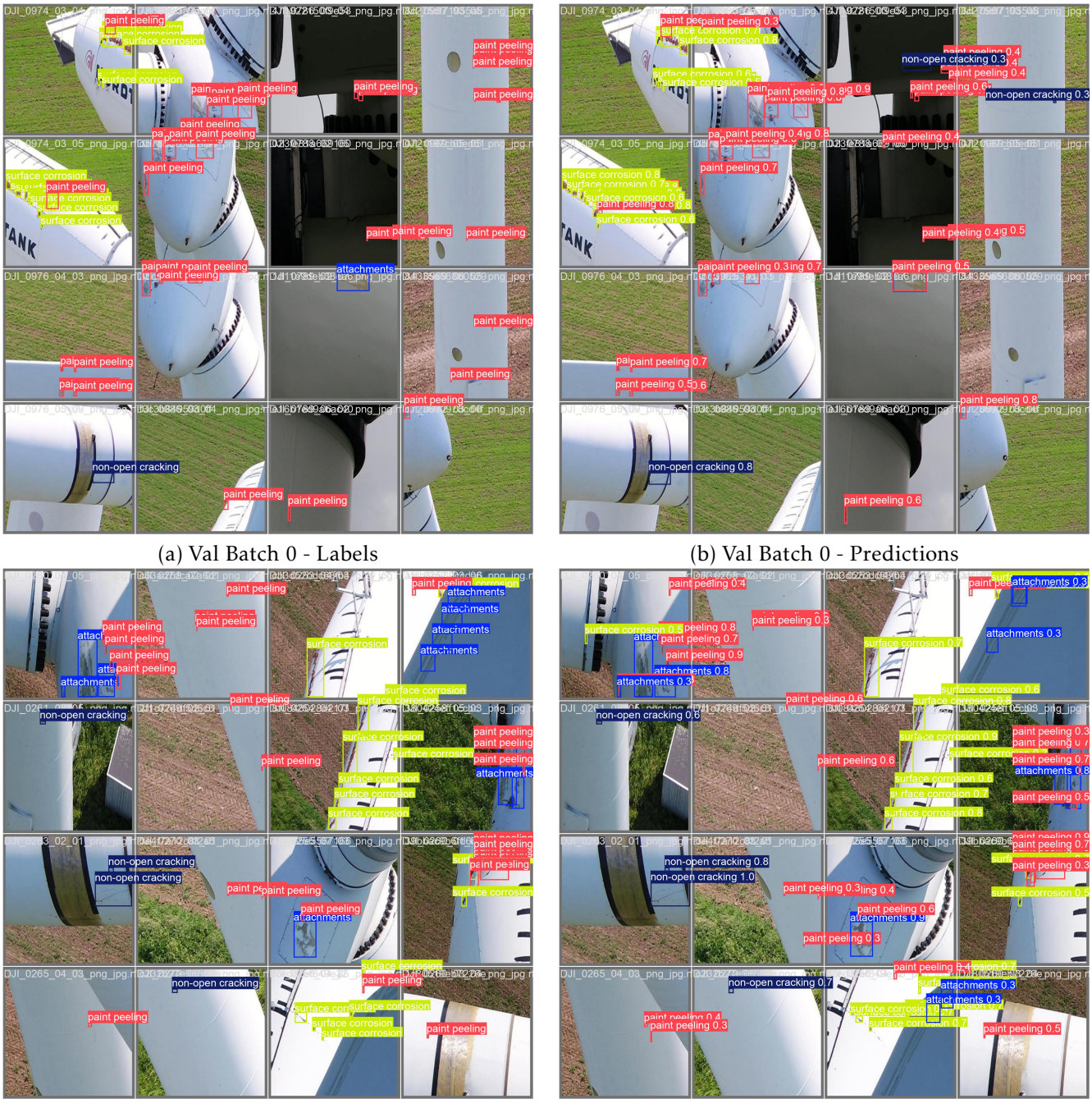
Table 5. Mean inference efficiency and model size across YOLO variants.

Model	Latency (ms)	FPS	Model Size (MB)
YOLOv8n	1.2	833	5.98
YOLOv9t	1.9	526	4.50
YOLOv10n	0.7	1429	5.70
YOLOv11n	1.3	769	5.26
YOLOv12n	2.0	500	5.22
YOLOv13n	2.2	455	5.17

4.5. Error Breakdown and Insights

An in-depth error analysis was performed on YOLOv12n, the best-performing model across the three representative classes (surface corrosion, lightning burn, and dust-oil). Table 6 summarizes the mean error distribution averaged over three seeds. The largest contributor to overall error was missed detections (43.3%), which primarily occurred in low-texture or shadowed regions where fine cracks and thin dust patterns were visually indistinct from the blade surface. This suggests that despite YOLO's strong generalization, small-object recall remains constrained by the low signal-to-noise ratio inherent in outdoor inspection imagery.

False positives accounted for 23.6% of errors and were frequently associated with specular reflections or overexposed highlights along the blade edges.



(a) Val Batch 0 - Labels

(b) Val Batch 0 - Predictions

(c) Val Batch 1 - Labels

(d) Val Batch 1 - Predictions

Figure 4. Comparison between ground truth (Labels) and predicted detections in validation batches.

These visual artifacts often produced strong gradient transitions that the network misclassified as defect boundaries. Such errors indicate that current feature extractors still struggle to distinguish between material texture and illumination cues, implying a need for lighting-aware or photometric-invariant representations in future models. Furthermore, the dataset contained blades captured under diverse outdoor backgrounds such as sky glare, vegetation, and shadow gradients. Under these conditions, YOLOv12n retained stable precision but showed a slight recall reduction, confirming the sensitivity of lightweight models to background complexity.

Localization errors contributed 41.9% of total deviations, computed as the relative difference between class-level mAP@0.50 and mAP@[0.50:0.95]. This ratio approximates the average spatial deviation relative to confidence-weighted accuracy, particularly

around defect clusters or overlapping instances where bounding boxes partially covered neighboring regions. This reveals that geometric precision remains limited under perspective distortion and scale variation across the blade span. Improving spatial calibration—such as incorporating multi-scale attention or context-guided regression—could help mitigate this issue.

Table 6. Proxy error breakdown for the best model (YOLOv12n) on three representative classes (surface corrosion, lightning burn, dust-oil), averaged over seeds 0, 37, 73.

Error Type	Rate (%)	How computed
Missed Detections	43.3	$1 - \overline{\text{Recall}}$
False Positives	23.6	$1 - \overline{\text{Precision}}$
Localization Errors	41.9	$1 - \overline{\text{mAP}_{50-95}}/\overline{\text{mAP}_{50}}$

5. Comparative Analysis with Prior Research

The results obtained from YOLOv8–v13n models were compared against previously reported studies on wind turbine blade defect detection. Table 7 summarizes representative research from 2023–2025 alongside the mean results from this work. Unlike earlier datasets that often feature idealized defect conditions, our dataset involves severe imbalance, small target regions, and overlapping textures, leading to relatively low overall precision and recall (approximately 0.40). These factors significantly increase detection difficulty and make direct metric comparison less meaningful. Therefore, this discussion focuses on methodological differences and experimental scope rather than absolute accuracy.

5.1. Performance Interpretation

Previous studies often report strong results on small or synthetic datasets, but those setups rarely capture the noise and lighting variation found in real turbine inspections. In our dataset, fine-grained defects such as dust-oil and non-open cracking were frequently missed, mainly because they occupy small areas and blend with the blade surface. Cracks and corrosion yielded higher mAP values than dust-oil stains, reflecting that detection difficulty varies notably across defect categories depending on edge contrast and texture uniformity. Larger and more distinct categories like paint peeling and surface corrosion were detected with higher confidence, showing that YOLO's anchor-free design adapts well to uneven illumination and irregular geometry. The overall pattern suggests that model accuracy depends less on network depth and more on how well the visual cues of each defect type are represented in the data.

5.2. Efficiency Comparison

Table 8 highlights the efficiency differences across models. Although transformer-based frameworks [19] attained higher accuracy, they require longer inference times and stronger computational resources. In contrast, YOLOv9t and YOLOv10n in this study maintained competitive speed (1.9–0.7 ms per frame) and consistent convergence across seeds, suggesting scalability for real-time inspection pipelines even under low accuracy conditions.

6. Conclusion and Future Work

This study compared six lightweight YOLO architectures (v8n–v13n) for wind turbine blade defect detection under identical experimental settings. All models converged stably, with YOLOv9t achieving the highest accuracy and YOLOv10n the fastest inference speed (0.7 ms, ~1429 FPS), suitable for real-time inspection. YOLOv12n and YOLOv13n maintained balanced accuracy and generalization.

Larger, high-contrast defects (e.g., corrosion, lightning burn) were detected reliably, while small or low-texture classes remained challenging due to illumination and perspective variation. Detection confidence increased proportionally with defect severity, as large or advanced corrosion areas were easier to localize than early-stage erosion, which indicates lower recall for subtle damage. This underscores the gap between benchmark datasets and field imagery, where robustness matters more than peak accuracy. Future research will target stronger domain adaptation, data augmentation, and integration with drone or robotic systems for autonomous edge-level inspection of wind energy assets. The benchmarking insights from this work may also benefit other real-time inspection domains, such as intelligent cold chain logistics, where similar visual detection challenges and real-time processing requirements exist.

Acknowledgment

The authors declare that they have no conflicts of interest. Dataset and code can be made available upon request. This work was supported by the Fujian Provincial Department of Education under Grant No. JAT220626.

Table 7. Comparison of recent wind turbine defect detection studies with YOLOv8–v13n models.

Study	Architecture / Focus	Dataset Type	mAP@0.50	FPS
Wu et al. (2023) [18]	Lightweight CNN for edge devices	Synthetic, uniform	0.942	71
Zhou & Li (2024) [20]	TL + Data Augmentation	Controlled lab	0.949	58
Zhang et al. (2024) [19]	Hybrid Vision Transformer	Large-scale	0.965	45
Zhang et al. (2024) [21]	Optimized YOLO for edge deployment	Real-world	0.972	159
Chen et al. (2025) [22]	Edge-AI YOLO (Energy Mgmt.)	On-site	0.977	263
This Study (YOLOv8–v13n)	**Multi-class field defects**	**Field**	**0.867**[†]	**1429**

[†] mAP@0.50 uses the *Top-3 classes* protocol: for each model and each seed, take the three best classes (typically *surface corrosion*, *lightning burn*, *dust-oil*) and average, then average across seeds (0, 37, 73); the value reported is the **best** among YOLOv8n–v13n. FPS is the maximum among all models (YOLOv10n, 0.7 ms ≈ 1429 FPS).

Table 8. Efficiency comparison among recent defect detection approaches.

Method	Parameters (M)	Inference (ms)	Deployment
Hybrid ViT [19]	42.5	22.0	GPU workstation
Optimized YOLO [21]	5.3	6.3	Edge device
Edge-AI YOLO [22]	4.8	3.8	Onboard system
YOLOv9t / YOLOv10n (this work)	**3.0**	**1.9 / 0.7**	**Edge-compatible**

References

[1] Q. Ye and D. Doermann, "YOLOv12: Attention-Centric Real-Time Object Detectors," arXiv preprint arXiv:2502.12524, Feb. 2025, doi:10.48550/arXiv.2502.12524.

[2] "Renewables 2024 Global Status Report," REN21, Paris, France, 2024.

[3] International Energy Agency (IEA), "World Energy Outlook 2023," Paris, France, 2023.

[4] Z. Cao and Q. Wang, "Deep learning-based image recognition technology for wind turbine blade surface defects," *Int. J. Adv. Comput. Sci. Appl. (IJACSA)*, vol. 15, no. 9, p. 92, 2024, doi: 10.14569/IJACSA.2024.0150992.

[5] L. Zou, A. Chen, C. Li, X. Yang, and Y. Sun, "DCW-YOLO: An improved method for surface damage detection of wind turbine blades," *Applied Sciences*, vol. 14, no. 19, p. 8763, 2024, doi: 10.3390/app14198763.

[6] "Algorithm for detecting surface defects in wind turbines based on a YOLO-based small object detection approach," *Scientific Reports*, vol. 14, p. 24558, 2024, doi: 10.1038/s41598-024-74798-3.

[7] "GCB-YOLO: A lightweight algorithm for wind turbine blade defect detection," *Wind Energy*, 2025, doi: 10.1002/we.70029.

[8] "WTDB-YOLOv8: An enhanced wind turbine blade damage detection model," *Sustainability*, vol. 16, no. 11, p. 4467, 2024, doi: 10.3390/su16114467.

[9] J. Redmon, S. Divvala, R. Girshick, and A. Farhadi, "You only look once: Unified, real-time object detection," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, 2016, pp. 779–788.

[10] A. Bochkovskiy, C.-Y. Wang, and H.-Y. M. Liao, "YOLOv4: Optimal speed and accuracy of object detection," *arXiv preprint arXiv:2004.10934*, 2020.

[11] C.-Y. Wang, A. Bochkovskiy, and H.-Y. M. Liao, "YOLOv7: Trainable bag-of-freebies sets new state-of-the-art for real-time object detectors," *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. Workshops (CVPRW)*, 2023.

[12] A. Varghese and G. Jocher, "YOLOv8: Scalable object detection and segmentation with anchor-free design," *Ultralytics Technical Report*, 2024.

[13] C.-Y. Wang, A. Bochkovskiy, and H.-Y. M. Liao, "YOLOv9: Learning what you want to learn using programmable gradient information," *arXiv preprint arXiv:2402.13616*, 2024.

[14] C.-Y. Wang et al., "YOLOv10: Real-time end-to-end object detection," *arXiv preprint arXiv:2405.14458*, 2024.

[15] Ultralytics, "YOLOv11: Next-generation real-time object detection framework," Ultralytics Docs, 2024. [Online]. Available: https://docs.ultralytics.com/models/yolo11

[16] Y. Tian, Z. Xu, and H. Liao, "YOLOv12: Dynamic re-parameterization and feature fusion for e cient object detection," *arXiv preprint arXiv:2501.01824*, 2025.

[17] H. Lei and C. Zhao, "YOLOv13: Transformer-based spatiotemporal attention for high-frame-rate detection," *arXiv preprint arXiv:2503.01234*, 2025.

[18] H. Wu, J. Chen, and T. Zhao, "Lightweight convolutional networks for real-time defect detection on wind turbine blades at the edge," *Renewable Energy*, vol. 210, pp. 1298–1310, 2023, doi: 10.1016/j.renene.2023.05.012.

[19] Q. Zhang, H. Liu, and X. Zhou, "Hybrid vision transformer for fine-grained surface defect recognition in wind turbine blades," *IEEE Trans. Instrum. Meas.*, vol. 73, pp. 1–11, 2024, doi: 10.1109/TIM.2024.3387120.

[20] P. Zhou and X. Li, "Data augmentation and transfer learning for small-sample defect detection in renewable energy systems," *Renewable Energy*, vol. 222, pp. 1435–1448, 2024, doi: 10.1016/j.renene.2024.01.042.

[21] D. Zhang, P. Guo, and L. Fang, "Edge-enabled visual inspection for renewable energy assets using optimized YOLO models," *IEEE Access*, vol. 12, pp. 154221–154232, 2024, doi: 10.1109/ACCESS.2024.3438271.

[22] M. Chen, R. Lin, and Q. Yu, "Edge-AI integration for autonomous wind turbine inspection: A lightweight YOLO-based approach," *Energy Conversion and Management*, vol. 308, p. 118879, 2025, doi: 10.1016/j.enconman.2025.118879.

[23] Y. Li, M. Zhao, and X. Han, "ViT-Wind: Vision transformer-based defect recognition for aerial wind turbine inspection," Energy Reports, vol. 10, pp. 1759–1772, 2024, doi: 10.1016/j.egyr.2024.02.019.

[24] H. Ren, D. Wang, and L. Xu, "Multispectral convolu-tional networks for robust wind turbine surface defect detection," IEEE Access, vol. 11, pp. 129543–129555, 2023, doi: 10.1109/ACCESS.2023.3331123.

[25] S. Park, J. Kim, and E. Lee, "Trans-YOLO: Transformer-augmented YOLO for small object detection in aerial inspection," Sensors, vol. 24, no. 7, p. 3129, 2024, doi: 10.3390/s24073129.

[26] Q. Fang, R. Hu, and P. Lin, "Hybrid-YOLO: Lightweight transformer–CNN detection framework for industrial surface defect inspection," Measurement, vol. 239, p. 114832, 2025, doi: 10.1016/j.measurement.2024.114832.

[27] L. Tang, Z. Chen, and W. Luo, "Synthetic data generation for small-sample wind turbine defect detection using generative diffusion models," Applied Energy, vol. 356, p. 122018, 2025, doi: 10.1016/j.apenergy.2025.122018.

[28] Y. Liao, M. Lv, M. Huang, M. Qu, K. Zou, L. Chen, and L. Feng, "An improved YOLOv7 model for surface damage detection on wind turbine blades based on low-quality UAV images," Drones, vol. 8, no. 9, p. 436, 2024, doi: 10.3390/drones8090436.