Available Online: 1 April 2026 DOI: 10.54254/2977-3903/2026.32530

Research on a wind turbine blade surface defect detection method based on an improved YOLO11 model

Guangshuo Wang

School of Information Science, North China University of Technology, Beijing, China

415322955@qq.com

Abstract. Wind turbine blades are continuously exposed to environmental factors such as wind-blown sand, humidity, heat, and ultraviolet radiation. As service time increases, surface defects—including cracks, stains, and material peeling—gradually emerge. If not detected at an early stage, these defects may worsen over time and compromise the safety and operational stability of the turbine. In current operation and maintenance practices, blade inspection still relies heavily on manual methods, which not only suffer from limited efficiency but also involve risks associated with high-altitude work. Moreover, inspection results are often influenced by the experience of individual inspectors. In response to these practical challenges, this study focuses on the identification of surface defects on wind turbine blades and introduces targeted improvements within a deep learning–based detection framework. A single-stage object detection approach is adopted for model construction, with two key enhancements. First, a DySample dynamic upsampling module is incorporated into the feature fusion process to improve information transfer across features of different scales. Second, the FocalUIoU loss function is introduced in the bounding box regression stage to enhance the model's localization capability for small-scale defects and to improve optimization during training. Experimental results show that the improved model achieves a Precision of 0.874, a Recall of 0.793, and mAP@50 and mAP@50–95 of 0.856 and 0.544, respectively, on the test set. Compared with the baseline YOLO11 model, Recall is increased by approximately 15.6%, while mAP@50 and mAP@50–95 are improved by 6.5% and 7.9%, respectively. The model contains only 2.60 million parameters, indicating minimal change in overall scale while retaining the characteristics of a lightweight architecture. Under the premise of acceptable detection efficiency, the proposed method improves the recognition performance for multiple categories of blade defects and provides technical support for wind power equipment inspection and subsequent maintenance.

Keywords: wind turbine blade, defect detection, YOLO11, DySample, FocalUIoU

1. Introduction

Against the backdrop of low-carbon transition and the development of clean energy, wind power generation has maintained a rapid growth trajectory in recent years. Driven in particular by China's "dual carbon" goals, the installed capacity of wind power has continued to expand, and the deployment of wind farms has gradually extended to offshore areas as well as regions characterized by more complex environmental conditions, such

as extreme cold, high humidity, and high salinity. Wind turbine blades are critical components for converting wind energy into electrical power, and their service condition is directly related to both power generation efficiency and operational safety. Existing studies indicate that turbine blades are predominantly manufactured from composite materials. During long-term service, these materials are subjected to repeated cyclic loading, while also being continuously affected by environmental variations and fatigue accumulation, making them more susceptible to performance degradation and localized damage [1]. Research by Mishnaevsky et al. further suggests that under the combined effects of wind–sand erosion, ultraviolet radiation, and cyclic temperature and humidity changes, microcracks tend to form within the blade structure, which subsequently evolve into observable structural defects [2]. After examining the damage patterns of large-scale wind turbine blades, Wang et al. pointed out that issues such as cracking, delamination, erosion, and coating peeling are commonly observed in operational equipment [3]. In addition, studies have shown that once the leading edge of a blade undergoes erosion, its aerodynamic performance declines accordingly, and subsequent maintenance costs increase; in severe cases, structural failure may even occur [4]. From the perspective of operation and maintenance economics, Jin et al. argued that blade maintenance accounts for a significant proportion of the total lifecycle cost of wind turbines, highlighting the clear engineering value of improving the efficiency and reliability of blade defect detection [5].

In earlier stages, wind turbine blade inspection was primarily carried out through manual methods, including tower climbing, rope access operations, and ground-based telescope observation. These approaches present several notable limitations, such as low efficiency, high labor intensity, and considerable safety risks associated with high-altitude work. Furthermore, inspection outcomes largely depend on the experience of individual operators, making it difficult to establish standardized evaluation criteria. To enhance detection capability, some studies have explored the application of non-destructive testing methods—such as ultrasonic testing, acoustic emission, infrared thermography, and vibration signal analysis—to blade inspection tasks [6]. However, these methods typically require specialized equipment and involve relatively complex procedures, resulting in limited efficiency when applied to large-scale wind farm scenarios.

In recent years, Unmanned Aerial Vehicles (UAVs) have been increasingly adopted for wind power equipment inspection. Equipped with high-resolution imaging devices, UAVs are capable of capturing surface images of blades under non-contact conditions, enabling long-distance inspection of high-altitude structures. This approach not only reduces the risks associated with manual inspection but also improves operational efficiency [7]. On this basis, researchers have begun integrating UAV imagery with computer vision techniques, attempting to achieve automatic identification and localization of blade defects through image analysis, thereby providing new pathways for intelligent operation and maintenance of wind power equipment [8]. With the advancement of deep learning methods, visual detection models based on convolutional neural networks have demonstrated stronger feature extraction capabilities in complex environments, gradually becoming a key technical approach in wind turbine blade defect detection research [9].

From the perspective of algorithmic evolution, early studies on blade defect identification mainly relied on traditional image processing techniques or classical machine learning models, such as edge detection, texture analysis, and handcrafted features combined with classifiers. These methods are often sensitive to variations in illumination, background complexity, and shooting angles, making it difficult to maintain stable performance in real-world scenarios. With the development of deep learning, object detection models have been progressively introduced into industrial defect detection tasks. Among them, the You Only Look Once (YOLO) series adopts a single-stage, end-to-end detection framework, achieving relatively high detection accuracy while maintaining fast inference speed, and has therefore been widely applied in real-time detection tasks [10]. Building on this foundation, researchers have continuously refined and optimized the YOLO

network architecture, enabling it to perform effectively in tasks such as UAV-based aerial image analysis, small-object detection, and detection in complex background environments [11].

However, surface defects on wind turbine blades exhibit considerable complexity. On the one hand, different types of defects vary significantly in scale, with a high proportion of small targets; some defects also present elongated shapes with inconspicuous textures. On the other hand, UAV-acquired images are often affected by background clutter, illumination variations, and viewpoint shifts. As a result, baseline object detection models are prone to issues such as missed detections of small targets, inaccurate localization, and a relatively high rate of false positives in this task. In addition, if fixed interpolation methods are still used during the multi-scale feature reconstruction stage, the fusion of high-level semantic information and low-level detail information tends to be insufficient, further constraining model performance. Common IoU-based loss functions primarily emphasize the overlap between predicted and ground-truth bounding boxes, offering limited capability in distinguishing hard samples. Consequently, there remains room for improvement in bounding box regression accuracy under complex conditions.

2. Improved YOLO11 network architecture

Considering the specific characteristics of wind turbine blade surface defect detection, this study introduces targeted modifications to the original YOLO11 architecture and constructs a lightweight improved model on this basis. YOLO11 belongs to the relatively lightweight category within the YOLO series, featuring a compact structure and fast inference speed, which provides a solid foundation for industrial vision applications. However, when directly applied to tasks such as wind turbine blade inspection—where defect scales vary significantly and texture features are often inconspicuous—the original model still exhibits limitations in multi-scale feature representation and bounding box localization.

To address the detection requirements in complex defect scenarios, this study focuses on two key aspects of improvement. First, the DySample module is introduced during the feature upsampling stage, allowing sampling positions to dynamically adapt to feature distribution. This enhances the recovery of high-level semantic information and strengthens the model's multi-scale representation capability, particularly benefiting the detection of small-scale defects. Second, the FocalUIoU loss function is employed in the bounding box regression stage, integrating dynamic bounding box scaling with hard-sample weighting to improve the regression process and enhance localization accuracy in complex environments.

Based on these improvements, an enhanced YOLO11 model tailored for wind turbine blade defect detection is developed. The overall architecture retains the original lightweight design, with modifications limited to the incorporation of the DySample module at the upsampling positions within the feature pyramid and the adoption of the FocalUIoU loss function during training for bounding box regression optimization. As a result, the model size remains largely unchanged, while detection performance is further improved. The overall network structure is illustrated in Figure 1.

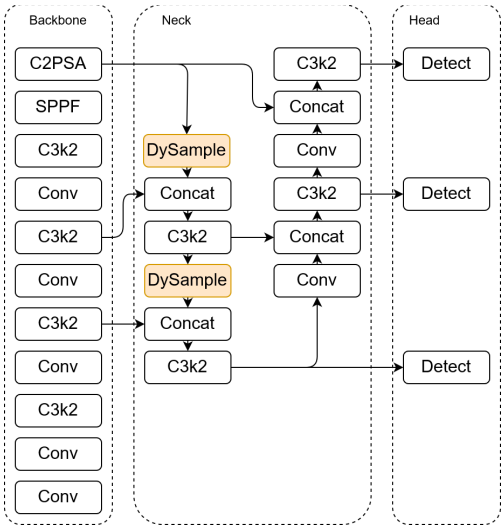

Figure 1. Overall network architecture

2.1. Dynamic upsampling: DySample

In object detection networks, upsampling is primarily used to restore the spatial resolution of deep feature maps, enabling their fusion with shallow features to support multi-scale object recognition. Conventional methods typically rely on fixed strategies such as nearest-neighbor interpolation, bilinear interpolation, or transposed convolution. However, these approaches lack adaptability to variations in feature distribution. Particularly in complex scenarios and small-object detection tasks, fixed upsampling often fails to adequately preserve high-level semantic information, thereby limiting fine-grained target recognition.

To improve information recovery during the upsampling stage, DySample proposes a dynamic sampling–based upsampling method. This approach learns spatial offsets of sampling points from the input feature map, allowing sampling locations to adaptively adjust according to feature distribution, thereby enabling a more flexible feature reconstruction process [12]. Compared with traditional methods based on fixed interpolation rules, DySample predicts sampling offsets through the network and dynamically refines the original sampling positions, allowing the model to more accurately preserve critical region information during feature recovery. The basic structure of the DySample module is shown in Figure 2. Essentially, it introduces learnable offsets into the upsampling process to improve reconstruction accuracy.

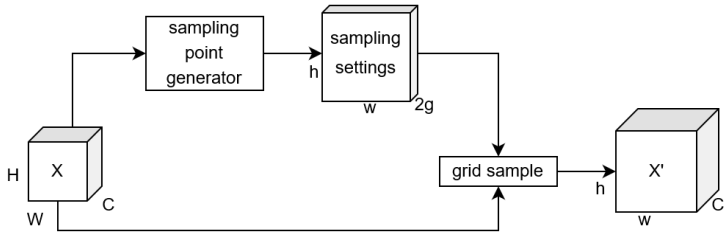

Figure 2. Schematic diagram of the DySample module

Let the input feature map be denoted as F, with spatial resolution $H \times W$. DySample first uses a convolutional layer to predict the sampling offset Δp, which describes the spatial displacement of sampling points relative to their original grid positions:

$$\Delta p = Conv(F) \tag{1}$$

After obtaining the offsets, dynamic sampling is performed on the input features based on the predicted sampling positions, which can be expressed as:

$$F'(p) = \sum_{i=1}^{k} w_i \cdot F(p + \Delta p_i) \tag{2}$$

where p denotes a sampling location in the feature map, Δp_i represents the dynamic offset of the corresponding sampling point, w_i is the interpolation weight, and k is the number of sampling points within the neighborhood. As illustrated in Figure 2, DySample adaptively adjusts sampling positions according to the spatial distribution of the input feature map, enabling more accurate recovery of critical feature information during upsampling. Compared with traditional fixed interpolation methods, this mechanism enhances feature reconstruction capability while maintaining relatively low computational overhead, allowing the network to focus more effectively on discriminative regions during feature recovery.

Within object detection models, DySample is placed at the upsampling positions of the Feature Pyramid Network (FPN), replacing conventional operations such as nearest-neighbor or bilinear interpolation. During multi-scale feature fusion, deep features tend to encode stronger semantic information, while shallow features preserve more spatial details. With the introduction of DySample, high-level semantic features gain greater flexibility in restoring spatial structures, resulting in more effective fusion across different scales. Since sampling positions can dynamically adapt to feature distribution, the model's ability to leverage complementary information across scales is also enhanced, leading to more comprehensive feature representation.

For the task of wind turbine blade surface defect detection, targets such as cracks, corrosion, and localized peeling are typically small in size; some also exhibit elongated shapes with weak texture cues. In addition, UAV-acquired images are often affected by complex backgrounds, illumination variations, and viewpoint changes. Under such conditions, fixed upsampling methods are more likely to lose fine-grained details during feature reconstruction, increasing the risk of missed detections for small-scale defects. By incorporating the DySample module into the detection network, key region information can be more accurately recovered during the upsampling stage, improving the fusion quality between high-level semantic features and low-level spatial details. This, in turn, enhances the model's capability to detect small-scale defect targets and provides more reliable feature representations for subsequent defect localization and recognition.

2.2. FocalUIoU bounding box regression loss function

In object detection training, the choice of bounding box regression loss function directly affects localization performance. Intersection over Union (IoU) is commonly used to measure the overlap between predicted and ground-truth boxes and serves as a fundamental basis for regression loss design. However, when there is no overlap between the predicted box and the ground-truth box, IoU degenerates to zero, resulting in significantly weakened gradient signals and reduced training stability. To address this limitation, a series of improved variants—such as GIoU, DIoU, and CIoU—have been proposed. These methods typically incorporate additional geometric constraints into the IoU formulation, including the minimum enclosing box, center point distance, and aspect ratio consistency, thereby providing more informative optimization signals and improving localization performance.

Although these approaches enhance regression accuracy to some extent, they primarily focus on extending geometric metrics. In practice, the distribution of predicted box quality during training is often highly imbalanced, with low-IoU samples significantly outnumbering high-IoU ones. As a result, the training process may be dominated by low-quality samples, which weakens the model's ability to further refine high-quality predictions and ultimately limits localization accuracy.

To address this issue, this study introduces the Unified IoU (UIoU) regression method [13]. Unlike traditional IoU-based losses that focus on geometric improvements, UIoU dynamically scales the sizes of predicted and ground-truth boxes to adjust the contribution of samples with different quality levels during training. This mechanism guides the model to gradually shift its attention from low-quality to high-quality samples. Let the predicted box and ground-truth box be defined as:

$$B_p = (x_p, y_p, w_p, h_p), B_g = (x_g, y_g, w_g, h_g) \tag{3}$$

where (x, y) denote the center coordinates, and w and h represent the width and height of the bounding boxes, respectively. Before computing IoU, UIoU applies a scale transformation to both boxes with respect to their centers:

$$B'_p = (x_p, y_p, rw_p, rh_p), B'_g = (x_g, y_g, rw_g, rh_g) \tag{4}$$

where r is a scaling factor (Ratio) that controls the degree of resizing. To enable the model to focus on different quality samples at different training stages, the scaling factor is dynamically adjusted over training epochs. In this study, a linear decay strategy is adopted:

$$r = -0.005 \times epoch + 2 \tag{5}$$

where $epoch$ denotes the current training iteration. This strategy allows different optimization priorities at different stages of training. In the early stage, when $r > 1$, both predicted and ground-truth boxes are enlarged, increasing the likelihood of overlap and reducing the influence of low-quality predictions on the loss function, thereby facilitating rapid learning of coarse object localization. As training progresses, r gradually decreases; when $r < 1$, the boxes are shrunk, making IoU variations more sensitive for high-quality predictions. This enhances the contribution of high-quality samples and improves localization precision.

After applying dynamic scaling, IoU and its variants are computed based on the scaled boxes. Taking CIoU as an example, it can be expressed as:

$$CIoU = IoU - (\frac{\rho^2(b_p, b_g)}{c^2} + \alpha v) \tag{6}$$

where IoU represents the overlap ratio between the predicted and ground-truth boxes, $\rho(b_p, b_g)$ denotes the Euclidean distance between their centers, c is the diagonal length of the minimum enclosing box, v measures aspect ratio consistency, and α is a balancing factor. By computing CIoU on the scaled boxes, a UIoU-based regression term incorporating dynamic scaling can be obtained. Compared with conventional IoU-based losses, this approach adjusts the weighting of samples during training, enabling the model to focus more on high-quality predictions in later stages and thereby improving localization accuracy.

To further emphasize hard samples during training, this study incorporates the Focal modulation mechanism into the UIoU framework, resulting in the FocalUIoU bounding box regression loss. Originally proposed for addressing class imbalance in classification tasks, the Focal mechanism reduces the contribution of easy samples while increasing the influence of hard ones. Following this idea, a modulation factor is constructed based on IoU:

$$M_{focal} = 1 - (IoU)^{\gamma} \tag{7}$$

where γ is a tunable parameter controlling the emphasis on hard samples. When IoU is small, the modulation factor becomes larger, increasing the contribution of difficult samples; conversely, when IoU is large, the influence of easy samples is reduced, preventing the model from overfitting to them.

Based on this formulation, the FocalUIoU loss function can be expressed as:

$$L_{FocalUIoU} = M_{focal} \cdot (1 - UIoU) \tag{8}$$

that is,

$$L_{FocalUIoU} = (1 - IoU^{\gamma}) \cdot (1 - UIoU) \tag{9}$$

In essence, this loss function integrates the dynamic scaling strategy of UIoU with the Focal modulation mechanism within a unified framework. With this design, the model maintains relatively fast convergence in the early stages of training, while in later stages it increasingly focuses on high-quality predictions, thereby further improving bounding box localization accuracy.

In the context of wind turbine blade surface defect detection, targets are typically small in size, exhibit significant shape variability, and are subject to strong background interference. These characteristics increase the likelihood of inaccurate localization or blurred boundaries. By introducing the FocalUIoU loss function, the training process assigns greater emphasis to difficult defect regions, improving both the accuracy and stability of predicted bounding boxes and providing more reliable candidate regions for subsequent defect recognition.

2.3. Summary

The YOLO series demonstrates clear advantages in real-time detection; however, in complex industrial defect scenarios, its performance is still constrained by limited small-object detection capability and suboptimal bounding box localization accuracy. The DySample module improves the upsampling process through dynamic sampling, enabling more effective recovery of spatial information from high-level semantic features and enhancing the representation of small-scale defect targets. Meanwhile, the FocalUIoU loss function refines the bounding box regression process through a hard-sample modulation mechanism, allowing the model to better optimize high-quality predictions during training and thereby improve localization performance. Based on this approach, this study integrates both the DySample module and the FocalUIoU loss function into the YOLO11 framework, ultimately constructing an improved model tailored for wind turbine blade surface defect detection.

3. Experimental results and analysis

To evaluate the effectiveness of the improved model in wind turbine blade surface defect detection, comparative experiments were conducted on a blade defect dataset. The proposed model was comprehensively compared with the baseline model and other methods. The evaluation focuses on several aspects, including detection accuracy, model complexity, and inference speed.

3.1. Experimental environment and parameter settings

All experiments were conducted on the Ubuntu 22.04 operating system. The hardware configuration includes an NVIDIA A100 GPU with 40 GB of memory and an Intel® Xeon® Gold 6248R CPU at 3.00 GHz. The CUDA version is 12.4, and the software environment is based on PyTorch 2.6.0.

During training, the input image resolution was uniformly set to 640 × 640, with a batch size of 64 and a total of 150 training epochs. The Stochastic Gradient Descent (SGD) optimizer was employed, which adapts the learning process across different training stages through gradient-based updates. The initial learning rate was set to 0.01, with a momentum of 0.937 and a weight decay coefficient of 0.0005.

To ensure fair and comparable results, all models were trained under identical conditions, including consistent data augmentation strategies and parameter settings. The augmentation techniques included random flipping, scale transformation, and Mosaic augmentation.

For evaluation metrics, this study adopts commonly used indicators in object detection tasks: Precision (P), Recall (R), mAP@50, and mAP@50–95. Precision measures the proportion of correctly predicted positive samples, Recall reflects the proportion of ground-truth targets successfully detected, and mAP provides a comprehensive assessment of detection performance across different IoU thresholds.

3.2. Dataset description

The dataset used in this study consists of wind turbine blade images captured by Unmanned Aerial Vehicles (UAVs). Based on these images, manual annotation was performed to construct a blade surface defect detection dataset. The dataset contains a total of 4,467 images and 10,688 annotated defect instances, covering seven typical defect categories: burning, crack, deformity, dirt, oil, peeling, and rusty.

These categories exhibit significant differences in morphological structure and texture characteristics. For example, crack defects typically appear as elongated shapes, whereas dirt-related defects often present irregular boundaries. Such variability increases the complexity of automatic detection.

For dataset partitioning, all images were randomly divided into training, validation, and test sets at a ratio of 8:1:1. The training set contains 3,573 images with 8,580 defect instances and is used for model parameter learning. The validation set includes 446 images with 1,049 defect instances and is used for hyperparameter tuning and model selection during training. The test set consists of 448 images with 1,059 defect instances and is used for final performance evaluation.

The dataset exhibits a notable class imbalance. The dirt and peeling categories account for the largest proportions, at 30.65% and 27.72%, respectively, indicating that these defects occur more frequently in real-world operating environments. In contrast, the burning and oil categories are relatively underrepresented, each accounting for only 4.68% and 4.69%. The remaining categories—crack, deformity, and rusty—fall within a moderate range. This imbalance introduces challenges during training and places higher demands on the model's generalization capability.

Overall, the dataset encompasses a variety of typical blade surface defect types and reflects, to a certain extent, the distribution characteristics of defects in real operating conditions. It provides a solid data foundation for training and evaluating object detection models in this domain.

3.3. Ablation study

To further analyze the impact of each proposed component on detection performance, ablation experiments were conducted based on the baseline YOLO11 model. By incrementally introducing different improvement strategies, the effects of the DySample upsampling module and the FocalUIoU loss function on wind turbine blade defect detection were systematically evaluated. All experiments were carried out under identical training environments and parameter settings to ensure fairness and comparability. The results are presented in Table 1.

According to the experimental results, the introduction of different modules into the YOLO11 baseline leads to performance improvements to varying degrees. When the original regression loss is replaced solely with UIoU, the model's mAP@50 increases from 0.804 to 0.835, while Recall improves from 0.686 to 0.742. This indicates that UIoU can effectively enhance the bounding box regression process and improve the model's localization capability for defect targets.

After incorporating the DySample dynamic upsampling module, the model shows a noticeable improvement in detecting small-scale defect targets. As shown in Table 1, mAP@50–95 increases from 0.504 to 0.531, and Recall rises to 0.768. These results suggest that DySample contributes to preserving fine-grained information during the feature reconstruction stage, thereby providing performance gains in complex defect scenarios.

When DySample and UIoU are applied simultaneously, the model performance is further improved. This demonstrates that enhancing both feature representation and bounding box regression jointly leads to better overall detection performance. Building on this, the introduction of the FocalUIoU loss function further improves the training process by placing greater emphasis on samples with larger localization errors, thereby strengthening the model's ability to learn from complex and difficult cases.

From the experimental results, the best performance is achieved when both DySample and FocalUIoU are integrated into YOLO11. Under this configuration, Precision and Recall reach 0.874 and 0.793, respectively, while mAP@50 attains 0.856 and mAP@50–95 reaches 0.544. Compared with the baseline YOLO11 model, mAP@50 and mAP@50–95 are improved by approximately 6.5% and 7.9%, respectively. Notably, these gains are achieved without a significant increase in model size, indicating that the combined modifications effectively enhance detection performance while maintaining a lightweight architecture.

Overall, the results show that DySample primarily improves the model's multi-scale feature representation capability, whereas FocalUIoU mainly optimizes the bounding box regression process, thereby enhancing localization accuracy. When combined, these two improvements enable the model to achieve superior performance in complex wind turbine blade defect detection scenarios.

Table 1. Results of ablation experiments

Model	P	R	mAP50	mAP50-95	Params (M)	Inference (ms)
YOLO11	0.872	0.686	0.804	0.504	2.584	0.5
YOLO11 + UIoU	0.863	0.742	0.835	0.515	2.584	0.5
YOLO11 + DySample	0.853	0.768	0.830	0.531	2.596	0.5
YOLO11 + DySample + UIoU	0.843	0.799	0.858	0.537	2.596	0.5
YOLO11 + FocalUIoU	0.874	0.772	0.849	0.532	2.584	0.5
YOLO11 + DySample + FocalUIoU	0.874	0.793	0.856	0.544	2.596	0.5

3.4. Comparison with mainstream algorithms

To further validate the effectiveness of the proposed model for wind turbine blade surface defect detection, comparative experiments were conducted against several mainstream object detection models, including YOLOv5n, YOLOv8n, and the baseline YOLO11. All models were evaluated under the same experimental conditions, using identical dataset splits, input resolutions, and training parameters, in order to minimize the influence of irrelevant factors. The results are presented in Table 2.

Table 2. Comparison results of mainstream algorithms

Model	P	R	mAP50	mAP50-95	Params (M)	Inference (ms)
YOLOv5n	0.835	0.741	0.816	0.502	2.504	0.4
YOLOv8n	0.844	0.769	0.823	0.516	3.007	0.4
YOLO11 (Baseline)	0.872	0.686	0.804	0.504	2.584	0.5
YOLO11 + DySample + FocalUIoU	0.874	0.793	0.856	0.544	2.596	0.5

As shown in Table 2, different detection models exhibit varying performance on this task. YOLOv5n achieves a Precision of 0.835 and a Recall of 0.741, with mAP@50 and mAP@50–95 of 0.816 and 0.502, respectively, indicating relatively moderate overall performance. YOLOv8n demonstrates improved detection accuracy compared to YOLOv5n, with Precision reaching 0.844 and Recall 0.769, while mAP@50 and

mAP@50–95 increase to 0.823 and 0.516. However, this improvement is accompanied by an increase in model size to 3.007 million parameters.

The baseline YOLO11 model achieves a Precision of 0.872, Recall of 0.686, mAP@50 of 0.804, and mAP@50–95 of 0.504 on this dataset. These results suggest that there remains room for improvement in handling complex wind turbine blade defect scenarios. After incorporating the DySample dynamic upsampling module and the FocalUIoU bounding box regression loss function, the model demonstrates a notable performance improvement. Precision increases to 0.874, Recall to 0.793, and mAP@50 and mAP@50–95 reach 0.856 and 0.544, respectively. Compared with the baseline YOLO11 model, the improved model achieves superior performance in blade defect detection tasks.

In terms of inference speed, all models operate at the millisecond level per image. YOLOv5n and YOLOv8n achieve approximately 0.4 ms per image, while YOLO11 and the proposed improved model require about 0.5 ms. Although the improved model shows a slight increase in inference time, it remains within an acceptable range for real-time detection and is suitable for practical applications such as UAV-based inspection.

Overall, the proposed model effectively improves detection performance for wind turbine blade surface defects while maintaining a lightweight architecture. It achieves a favorable balance among detection accuracy, model complexity, and inference speed.

4. Conclusion

To address the challenges posed by the complexity and difficulty of detecting surface defects on wind turbine blades, this study proposes a defect detection method based on YOLO11, integrating the DySample module and the FocalUIoU loss function. DySample enhances feature reconstruction during the upsampling stage, while FocalUIoU optimizes the bounding box regression process. Together, these components improve the model's detection capability in complex defect scenarios.

Experiments were conducted on a UAV-based wind turbine blade defect dataset, and the proposed method was compared with YOLOv5n, YOLOv8n, and the original YOLO11 model. The improved model demonstrates superior detection accuracy, achieving an mAP@50 of 0.856 and an mAP@50–95 of 0.544 on the test set, representing improvements of 6.5% and 7.9%, respectively, over the baseline YOLO11. The model contains only 2.596 million parameters, achieving notable performance gains while maintaining low complexity.

In summary, the proposed method enhances both the accuracy and stability of wind turbine blade surface defect detection and shows practical value for intelligent inspection of wind power equipment.

Future research may focus on further optimizing the network architecture to achieve greater model lightweighting while maintaining strong performance across various defect types. In addition, deploying the model on mobile or edge devices could enable fast and accurate real-time detection of blade surface defects in practical applications.

References

[1] Brøndsted, P., Lilholt, H., & Lystrup, A. (2005). Composite materials for wind power turbine blades. *Annual Review of Materials Research, 35*(1), 505–538.

[2] Mishnaevsky, L., Jr., Branner, K., Petersen, H. N., Beauson, J., McGugan, M., & Sørensen, B. F. (2017). Materials for wind turbine blades: An overview. *Materials, 10*(11), Article 1285.

[3] Wang, W., Xue, Y., He, C., & Zhao, Y. (2022). Review of the typical damage and damage-detection methods of large wind turbine blades. *Energies, 15*(15), Article 5672.

[4] Herring, R., Dyer, K., Martin, F., & Ward, C. (2019). The increasing importance of leading edge erosion and a review of existing protection solutions. *Renewable and Sustainable Energy Reviews, 115*, Article 109382.

[5] Jin, X. H., Sun, Y., Shan, J. H., & Li, C. (2017). Fault diagnosis and prognosis for wind turbines: An overview. *Chinese Journal of Scientific Instrument, 38*(5), 1041–1053.

[6] Murray, R. E., Beach, R., Barnes, D., Snowberg, D., & Mandell, D. (2021). Structural validation of a thermoplastic composite wind turbine blade with comparison to a thermoset composite blade. *Renewable Energy, 164*, 1100–1107.

[7] Zhang, S., He, Y., Gu, Y., & Yang, H. (2025). UAV based defect detection and fault diagnosis for static and rotating wind turbine blade: A review. *Nondestructive Testing and Evaluation, 40*(4), 1691–1729.

[8] Memari, M., Shakya, P., Shekaramiz, M., & Sebghati, M. (2024). Review on the advancements in wind turbine blade inspection: Integrating drone and deep learning technologies for enhanced defect detection. *IEEE Access, 12*, 33236–33282.

[9] Masita, K., Hasan, A. N., Shongwe, T., & Ntuli, N. (2025). Deep learning in defect detection of wind turbine blades: A review. *IEEE Access*. Advance online publication.

[10] Redmon, J., & Farhadi, A. (2018). *YOLOv3: An incremental improvement*. arXiv preprint. https: //arxiv.org/abs/1804.02767

[11] Bochkovskiy, A., Wang, C. Y., & Liao, H. Y. M. (2020). *YOLOv4: Optimal speed and accuracy of object detection*. arXiv preprint. https: //arxiv.org/abs/2004.10934

[12] Liu, W., Lu, H., Fu, H., & Cao, Z. (2023). Learning to upsample by learning to sample. In *Proceedings of the IEEE/CVF International Conference on Computer Vision* (ICCV 2023) (pp. 6027–6037). IEEE.

[13] Luo, X., Cai, Z., Shao, B., & Zhang, Y. (2024). *Unified-IoU: For high-quality object detection*. arXiv preprint. https: //arxiv.org/abs/2408.06636