*machines*

MDPI

*Article*

# High-Fidelity Reconstruction and Metrology Model for Defects Based on Physical Priors and Adaptive Morphology

**Ying Li [1], Xiaojiao Gu [1,*], Jinghua Li [2], Dongyang Zheng [1] and Xiaolin Yu [1]**

[1] School of Mechanical Engineering, Shenyang Ligong University, No. 6 Nanping Middle Road, Hunnan District, Shenyang 110159, China; liying8282@sylu.edu.cn (Y.L.); 2501620210@stu.sylu.edu.cn (D.Z.); flyfish513@163.com (X.Y.)

[2] School of Mechanical Engineering, Dalian University of Technology, Dalian 116024, China; lijinghua@mail.dlut.edu.cn

* Correspondence: 18804033860@163.com

**Abstract**

Tool micro-defect quantification is often degraded by optical sampling limits, boundary aliasing, and non-Gaussian industrial noise. To improve segmentation fidelity and physical measurement reliability, this study proposes a physically guided reconstruction and metrology framework for milling tool defects. The framework first uses an improved ResNet18-BiFPN segmentation network supervised by a distance transform-based boundary-aware loss, which encourages mask boundaries to fit high-curvature crack and chipping regions. A tensor-guided elliptical adaptive structuring element is then introduced to restore local topology while preserving tangential connectivity and normal boundary fidelity. Finally, a stable edge region search and curvature-adaptive gray/Zernike moment fusion strategy are used for noise-robust sub-pixel localization and physical area quantification. Image-based evaluation and controlled simulation-based stress tests show that the proposed method achieves an mIoU of 92.5%, an F1-score of 97.1%, and an HD95 of 2.15 pixels. Under strong synthetic noise, the average distance error remains below 0.28 pixels. For micro-defect area measurement, the mean relative error is reduced to approximately 1.7%. These results indicate that the proposed framework can support more reliable defect metrology for tool condition monitoring under complex industrial imaging conditions.

**Keywords:** tool micro-defects; high-fidelity reconstruction; boundary-aware loss; adaptive anisotropic filtering

## 1. Introduction

Under the macro-context of the global manufacturing industry's accelerated transition toward intelligence and precision, the requirements for process reliability and product quality in high-end CNC (Computer Numerical Control) machining have become increasingly stringent [1]. As the primary physical interface in CNC machining processes, the surface condition of the cutting tool directly determines the dimensional accuracy and surface integrity of the workpiece; furthermore, it serves as the core physical benchmark for predicting early tool failure and ensuring equipment operational safety [2]. During the machining of difficult-to-cut materials (e.g., titanium alloys and high-temperature alloys), tool surfaces are highly susceptible to defects such as crack wear, chipping, grooving, and crater wear [3]. Makhesana et al. [4] recently conducted a systematic comparative analysis of various machine vision algorithms applied to tool wear measurement. By constructing a direct tool condition monitoring system using telecentric lenses and industrial cameras,



Academic Editor: António Pereira

Received: 22 June 2026
Revised: 11 July 2026
Accepted: 14 July 2026
Published: 15 July 2026

they investigated digital image processing tools—including image enhancement, segmentation, morphological operations, and edge detection—thereby revealing the upper limits of accuracy and repeatability boundaries for flank wear measurement across different machining environments. These micro-defects not only trigger high-frequency cutting vibrations but also accelerate the catastrophic failure of the tool at an exponential rate.

Currently, Automated Optical Inspection (AOI) has emerged as the predominant modality for online tool surface condition monitoring due to its high efficiency and non-destructive nature [5,6]. However, when performing quantitative analysis of micron-scale defects, the true physical topology often occupies only a few pixels on the imaging plane, constrained by the optical diffraction limit of lenses and the discrete sampling characteristics of image sensors [7]. This information truncation at the physical level renders conventional image processing methods, which rely solely on pixel-level analysis, incapable of meeting the rigorous demands for high-fidelity dimensional quantification in modern precision manufacturing [8]. A more profound bottleneck lies in the fact that such physical information truncation directly undermines the discriminant reliability of underlying fault detection algorithms. Without providing metrological results associated with measurement uncertainty, vision systems will struggle to provide industrial-grade credible decision-making evidence for the closed-loop health management of tools in high-end manufacturing scenarios characterized by extremely low fault tolerance.

To break through the limitations of pixel resolution, numerous scholars have conducted extensive explorations in the field of vision algorithms. Although relevant technologies have made significant strides—evolving from early low-level image processing to modern deep instance segmentation—they still reveal deep-seated theoretical limitations when directly applied to industrial high-precision metrology.

### *1.1. Evolution from Traditional Operators to Deep Segmentation*

In the field of tool defect detection, the research paradigm is undergoing a deep evolution from manually crafted feature operators to data-driven learning. Early detection methods relied heavily on manual feature engineering, such as Sobel gradient operators or Otsu threshold segmentation [9]. While these performed acceptably under specific conditions, they exhibited poor robustness in non-structured environments characterized by fluctuating workshop illumination, coolant reflections, and complex milling textures [10]. To address these challenges, deep instance segmentation frameworks, represented by Mask R-CNN, have become the new paradigm for industrial inspection by integrating defect classification, localization, and pixel-level segmentation through powerful feature representation capabilities [11–13]. However, in micron-level precision manufacturing fault detection scenarios, current deep learning frameworks often prioritize global regional consistency over local boundary fidelity, such as Cross-Entropy (CE) or Dice loss. Their bias toward global pixel coverage leads to a lack of perception regarding high-frequency boundary topology [14]. Consequently, networks exhibit weak responses to micro-geometric features like crack tips, and output masks generally suffer from severe aliasing effects and topological fragmentation. Even with the introduction of local attention mechanisms, it remains difficult to resolve the underlying bottleneck of boundary fitting distortion from the perspective of the loss function's origin [15]. Recently, Liu et al. [16] proposed a defect-aware adaptive loss function within the DAFSF framework. By jointly considering boundary weighting, hard-sample reweighting, and class balancing mechanisms, the network can focus on boundary regions and challenging pixels, effectively mitigating the lack of boundary geometric perception inherent in traditional region-based optimization objectives.

### 1.2. Limitations of Isotropic Mask Restoration

To compensate for the aliasing defects in deep network output masks, existing post-processing schemes mostly rely on traditional mathematical morphological filtering based on fixed-size structural elements (e.g., $3 \times 3$ rectangles or $5 \times 5$ circles). This isotropic processing approach, while capable of smoothing local aliasing and filling holes, fails to perceive the local geometric dominant orientation of the image [17]. When such isotropic filters slide across the tips of crack wear defects or the sharp corners of chipped edges, they induce severe edge position shifts and irreversibly destroy high-frequency physical features. By sacrificing true physical topology for visual smoothness, this restoration strategy contradicts the underlying logic of industrial precision metrology. Existing studies have demonstrated that anisotropic filtering based on local structure tensors can effectively perceive the local geometric dominant orientation of an image, suppressing noise while maintaining corner structures, thus providing a theoretical direction for resolving the aforementioned dilemma [18].

### 1.3. Noise Robustness in Sub-Pixel Precision Localization

To achieve precise measurement of absolute physical dimensions [19], sub-pixel edge extraction must be introduced after mask reconstruction. Current algorithms are primarily divided into two camps: spatial moments and orthogonal moments. Methods based on spatial moments (e.g., gray moments) solve for the edge normal vector by calculating integral characteristics within a local window; these are widely deployed in online detection systems due to their explicit analytical solutions and high computational efficiency [20]. However, in real tool inspection conditions, characterized by complex milling textures and specular reflections from residual coolant, spatial moments are extremely sensitive to local non-Gaussian noise and outliers, frequently producing pseudo-edge responses [21]. In contrast, methods based on orthogonal moments (e.g., Zernike moments) can naturally filter out high-frequency noise through high-order truncation by expanding orthogonal polynomials within a unit disk, demonstrating superior topological description capabilities [22,23]. Nevertheless, their high computational complexity leads to a sharp decline in system frame rates. Cheng et al. [24] proposed a Canny–Steger hybrid sub-pixel edge detection method, which first utilizes the Canny operator to extract pixel-level edges as a baseline, and then employs the Steger algorithm to solve for edge normals and sub-pixel edge point positions via the Hessian matrix. This approach achieved a measurement accuracy of 3 μm and a repeatability of 2 μm in the online diameter measurement of shaft parts. By ensuring sub-pixel accuracy while significantly reducing redundant computation, this scheme provides a feasible technical path for the trade-off between computational efficiency and measurement accuracy in industrial scenarios. Designing a dynamic hybrid architecture that integrates the efficiency of spatial moments with the noise robustness of orthogonal moments is the key to breaking the sub-pixel precision localization bottleneck under extreme working conditions.

### 1.4. Research Gap and Contributions

Although boundary-aware segmentation, anisotropic filtering, and sub-pixel edge localization have been investigated in previous studies, most existing methods address these problems separately. For tool micro-defect metrology, segmentation accuracy alone is insufficient because physical measurement also depends on boundary topology, noise-resistant edge localization, and pixel-to-physical calibration. The main gap is therefore the lack of a unified, metrology-oriented pipeline that links mask learning, geometric reconstruction, sub-pixel localization, and physical defect quantification under industrial imaging noise.

(1) A distance transform-based boundary-aware segmentation strategy is introduced to constrain the geometric deviation between predicted masks and true defect boundaries. Unlike ordinary region-based losses, this term explicitly penalizes boundary displacement and is designed for high-curvature micro-defect regions.

(2) A tensor-guided Elliptical Adaptive Structuring Element (EASE) is developed for anisotropic mask reconstruction. The element is aligned with the local structure tensor so that tangential connectivity can be restored while normal-direction boundary fidelity is preserved.

(3) A Stable Edge Region (SER)-guided gray/Zernike moment fusion mechanism is proposed for sub-pixel localization. The gray moment branch is used for efficient localization in stable regions, while the Zernike branch is adaptively strengthened for high-curvature or noisy regions.

(4) A complete metrology validation protocol is established using region metrics, boundary metrics, sub-pixel localization error, physical area error, and measurement uncertainty analysis. This links visual segmentation performance with quantitatively traceable defect measurement.

Among these contributions, (1), (2), and (3) constitute the main methodological developments of this study, whereas ResNet18, BiFPN, SE attention, and the analytical formulations of spatial and Zernike moments are adopted from prior work. The three proposed modules are sequentially connected: the boundary-aware loss improves the fidelity of the initial mask and thereby provides a more reliable basis for structure–tensor orientation estimation in EASE; the reconstructed mask then supplies cleaner candidate regions for the subsequent SER-guided gray/Zernike moment fusion.

A comparison between existing methods and the proposed metrology-oriented framework is summarized in Table 1.

**Table 1.** Comparison between existing methods and the proposed metrology-oriented framework.

| Method Category | Typical Limitation in Existing Studies | Revision in the Proposed Framework | Role in Metrology |
|---|---|---|---|
| Boundary-aware segmentation | Boundary weighting improves contours but is rarely linked to physical measurement. | Distance transform/ NHD-based boundary penalty for defect masks. | Reduces boundary displacement and HD95. |
| Morphological reconstruction | Fixed isotropic structuring elements may smooth crack tips and chipping corners. | Tensor-guided EASE with tangential extension and normal contraction. | Restores topology while preserving high-curvature features. |
| Sub-pixel localization | Gray moments are noise-sensitive; Zernike moments are more robust but computationally heavier. | SER-guided curvature-adaptive gray/Zernike fusion. | Balances localization speed and robustness. |
| Tool defect evaluation | Many studies stop at detection or segmentation accuracy. | Closed-loop quantification using ADE, area error, and uncertainty. | Supports physically interpretable defect metrology. |

## 2. Methodology

The proposed framework follows a closed-loop metrology pipeline that connects semantic defect segmentation with geometric reconstruction and sub-pixel physical quantification. As shown in Figure 1, raw tool images are first preprocessed and calibrated. The initial defect masks are then obtained using boundary-aware deep segmentation. Next, topology filtering and tensor-guided EASE reconstruction are applied to improve mask continuity and preserve high-curvature features. Finally, SER-guided gray/Zernike moment fusion is used for sub-pixel edge localization and physical area measurement.

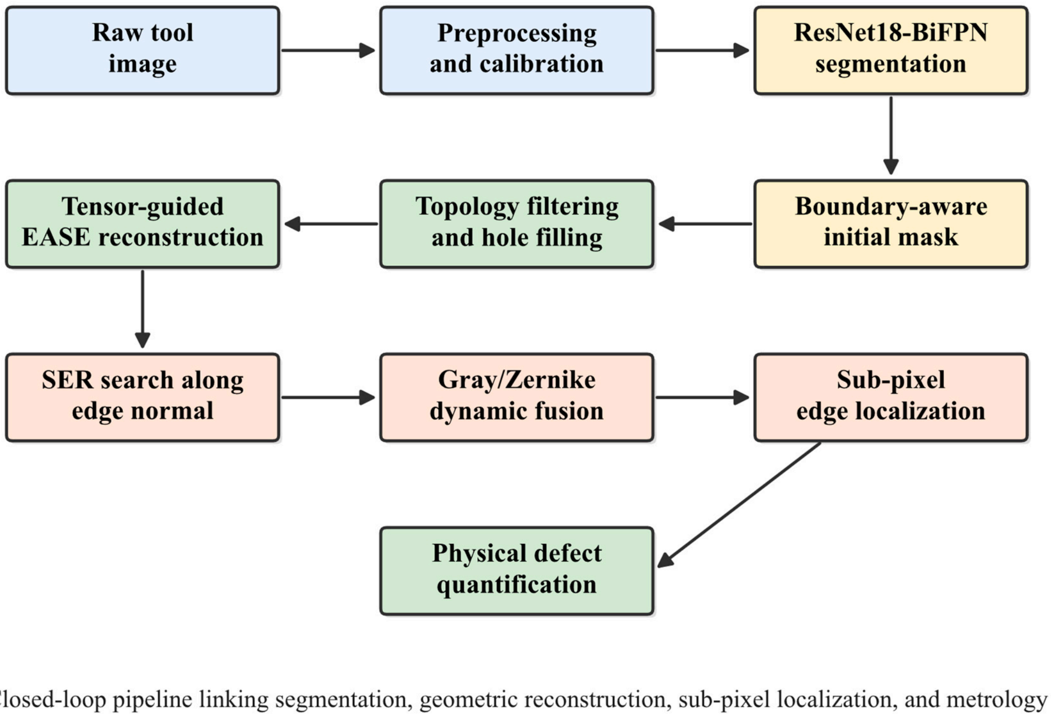



**Figure 1.** Overall workflow of the proposed high-fidelity reconstruction and metrology framework.

*2.1. High-Fidelity Initial Mask Segmentation Based on Boundary-Aware Loss*

When detecting micron-scale tool defects, deep convolutional networks are highly susceptible to losing high-frequency detail features, such as micro-crack tips, after multiple layers of pooling and strided downsampling. To address this limitation, an improved ResNet18-BiFPN (Bidirectional Feature Pyramid Network) is employed as the feature extraction backbone. As illustrated in Figure 2, the BiFPN structure enables bidirectional cross-scale information flow (top–down and bottom–up) with learnable weighted fusion of low-level high-resolution texture features and deep abstract semantic information. To further enhance channel-wise feature discriminability, a Squeeze-and-Excitation (SE) attention module is inserted after each BiFPN fusion node. This combined design improves the network's receptive-field coverage for multi-scale defects.

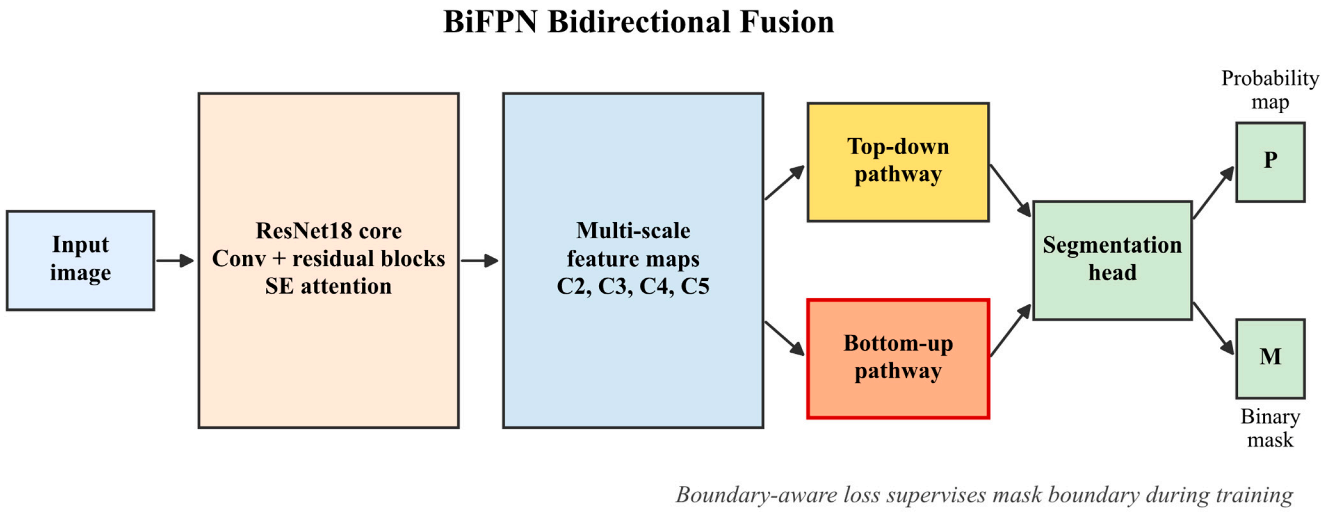



**Figure 2.** ResNet18-BiFPN segmentation architecture. The bottom–up pathway is highlighted, and the network output is clarified as a probability map and binary defect mask rather than a class label.

However, optimization of the network architecture alone is insufficient to fundamentally eliminate the aliasing effects and topological fragmentation of binary masks. Inherently, existing mask heads predominantly utilize CE or Dice loss. These region-based optimization objectives focus on global pixel coverage but lack explicit constraints on high-frequency geometric curvature. Consequently, this paper introduces a Boundary-Aware Loss function.

Let the continuous probability map predicted by the network be $P \in [0,1]\ H \times W$, and let the corresponding ground-truth binary mask be $G \in 0,1\ H \times W$, where $H$ and $W$ denote the image height and width. The total loss is defined as a weighted combination of classification loss $L_{CE}$, region consistency loss $L_{Dice}$, and boundary penalty loss $L_{Boundary}$:

$$L_{total} = \alpha L_{CE}(P,G) + \beta L_{Dice}(P,G) + \gamma L_{Boundary}(\partial P, \partial G) \quad (1)$$

where $\alpha$, $\beta$, and $\gamma$ are balancing hyperparameters. $\partial P$ and $\partial G$ denote the predicted and ground-truth boundary point sets, respectively. These terms jointly optimize pixel classification, regional overlap, and boundary alignment.

To explicitly guide the network toward the physical boundary, a distance transform field $D_G(p)$ is computed from the ground-truth boundary. For any pixel coordinate $p$ in the image domain $\Omega$, the shortest Euclidean distance from $p$ to $\partial G$ is defined as:

$$D_G(p) = \min_{q \in \partial G} |p - q|_2 \quad (2)$$

Based on $D_G|p|$, the boundary-aware loss constructs an exponential penalty inspired by the normalized Hausdorff distance.

$$L_{Boundary} = \frac{1}{|\partial P|} \sum_{p \in \partial P} \left( \exp\left( \frac{D_G(p)^2}{2\sigma^2} \right) - 1 \right) \quad (3)$$

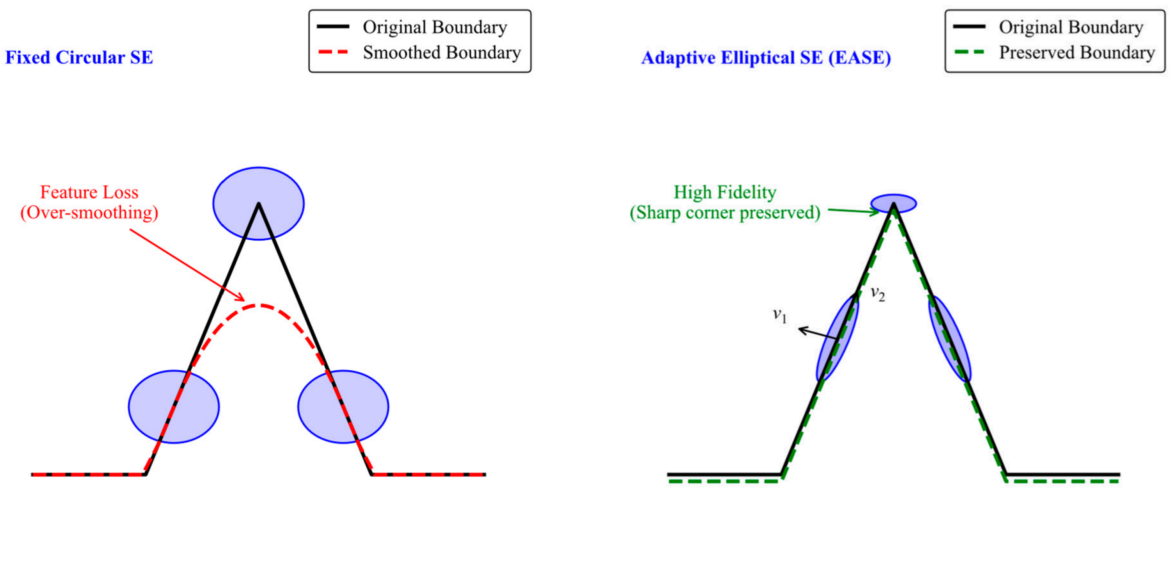
In Equations (2) and (3), $||\cdot||2$ denotes the Euclidean norm, $|\partial P|$ denotes the number of predicted boundary pixels, and $\sigma$ controls the sensitivity of the distance penalty. This unified notation clarifies the relationship between the distance transform in Equation (2) and the boundary loss in Equation (3).

Equation (3) possesses clear physical significance. During the backpropagation process, it applies an exponentially increasing penalty gradient to predicted pixels that deviate further from the true physical boundary. This Hard Negative Mining mechanism compels the network to actively learn the physical curvature of micro-defects within the feature representation space, thereby ensuring the geometric high-fidelity of the output masks at the source.

Despite the improvements in geometric fitting at the feature learning level via Boundary-Aware Loss, the output masks remain susceptible to micro-topological discontinuities due to the discrete nature of pixel grids. Consequently, the system must transition from global semantic learning to local geometric reconstruction, utilizing structure–tensor-guided anisotropic filtering to perform physical-level feature fidelity restoration on the defect boundaries.

### 2.2. *Adaptive Anisotropic Mask Reconstruction and Edge Enhancement*

Although the boundary-aware loss improves initial mask fidelity, binarization onto a discrete pixel grid inevitably introduces step-like aliasing and topological fragmentation. Conventional post-processing with fixed-shape morphological operators (e.g., $3 \times 3$ rectangles) can smooth these artifacts, but isotropic erosion or dilation indiscriminately rounds sharp corners and causes irreversible feature loss, as illustrated in Figure 3a.




(**a**) Isotropic Morphology (**b**) Anisotropic Morphology

**Figure 3.** Schematic comparison between isotropic and anisotropic structuring elements.

To break through this theoretical constraint, this paper introduces manifold constraints from differential geometry and proposes a tensor-guided adaptive morphological reconstruction mechanism. The core logic of this algorithm lies in endowing the filter with pixel-level orientation awareness.

First, the local structure tensor matrix $J$ of the initial mask image $I$ is calculated. Let the local gradient vector be $\nabla I = [I_x, I_y]T$, where $I_x$ and $I_y$ denote the partial derivatives (i.e., gray gradients) of the image along the horizontal and vertical directions, respectively. Under a Gaussian smoothing window $G_\rho$ with scale $\rho$, the tensor $J$ is constructed as follows:

$$J = G_\rho * \left(\nabla I \nabla I^T\right) = \begin{bmatrix} I_x^2 * G_\rho & I_x I_y * G_\rho \\ I_x I_y * G_\rho & I_y^2 * G_\rho \end{bmatrix} \tag{4}$$

where * denotes the 2D convolution operator. By performing eigenvalue decomposition on the positive semidefinite matrix $J$, non-negative eigenvalues $\lambda_1$ and $\lambda_2$ $(\lambda_1 \geq \lambda_2 \geq 0)$ and the corresponding orthogonal unit eigenvectors $v_1$ and $v_2$ are obtained. Here, $v_1$ points to the edge normal direction where the local gradient changes most abruptly, while $v_2$ represents the tangential direction of the edge. To dynamically quantify the orientation coherence of the local texture, the anisotropy factor $A$ is defined as:

$$A = \frac{\lambda_1 - \lambda_2}{\lambda_1 + \lambda_2 + \epsilon} \tag{5}$$

where $\varepsilon$ is an infinitesimal constant used to avoid division by zero. Based on this intrinsic information, the algorithm dynamically constructs an EASE within the local neighborhood of each pixel. As shown in Figure 3b, its 2D spatial morphology is strictly defined by the covariance matrix $\Sigma$:

$$\Sigma_{EASE} = [v_1, v_2] \begin{bmatrix} f_{norm}(A) & 0 \\ 0 & f_{tang}(A) \end{bmatrix} [v_1, v_2]^{\mathrm{T}} \tag{6}$$

The scalar functions are defined as: $f_{norm}(A) = r_{min} + (r_0 - r_{min})(1 - A)$, $f_{\mathrm{tang}}(A) = r_0 \cdot (1 + \alpha \cdot A)$. where $r_0 = 3$ is the base half-axis length (in pixels), $r_{\mathrm{min}} = 1$ prevents kernel degeneracy, and $\alpha = 3$ controls the tangential extension ratio. When $A \to 0$

(isotropic region), both axes converge to $r_0 = 3$, approximating a circular element. When $A \to 1$ (strongly oriented edge), $f_{\text{norm}} \to 1$ restricts normal-direction boundary drift to a single pixel, while $f_{\text{tang}} \to 12$ enables elongated bridging along the crack or chipping edge tangent.

Based on this spatially variant adaptive structuring element, an anisotropic dilation operation is performed on the binary mask $X$, with the erosion operation defined as its dual:

$$(X \oplus EASE)(p) = \max_{q \in EASE(p)} X(p-q) \tag{7}$$

Post-processing proceeds in three sequential stages: (i) connected-component filtering removes isolated noise by a minimum-area threshold, (ii) contour-based hole filling repairs enclosed discontinuities without isotropic smoothing, and (iii) the proposed EASE operation is applied to boundary reconstruction, where its anisotropic geometry bridges tangential gaps while restricting normal-direction drift. Stages (i) and (ii) avoid conventional isotropic morphological denoising; EASE is used only in stage (iii) where the local edge direction has been estimated. The corresponding morphological processing results are shown in Figure 4.

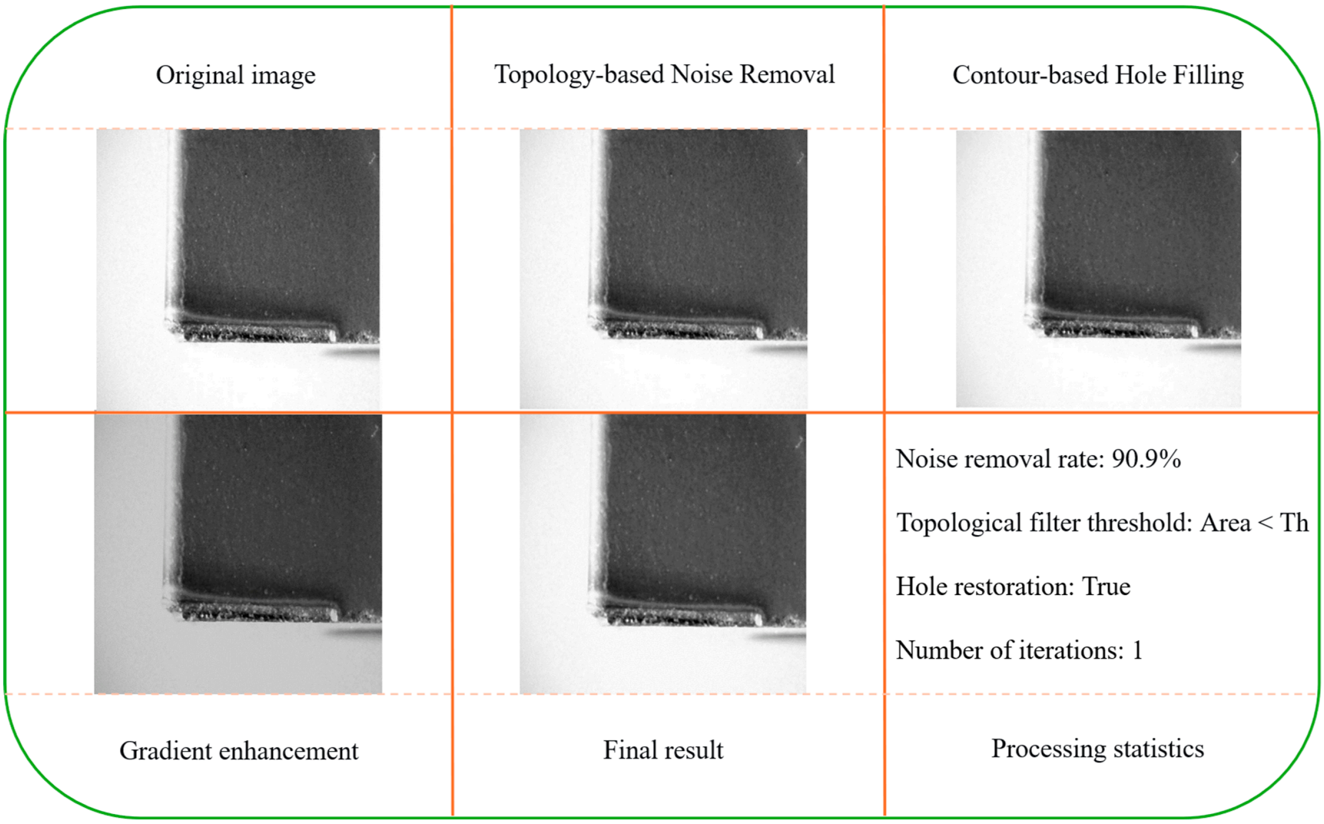



**Figure 4.** Comparison of morphological processing results.

To guarantee the precision of physical metrology, the system further deploys a subpixel-level pre-edge enhancement operation. As illustrated in Figure 5, the morphological gradient is first extracted, followed by Gaussian smoothing and cross-shaped structuring-element enhancement. The residual high-frequency burr energy decreases after processing, indicating effective noise suppression and improved edge stability.

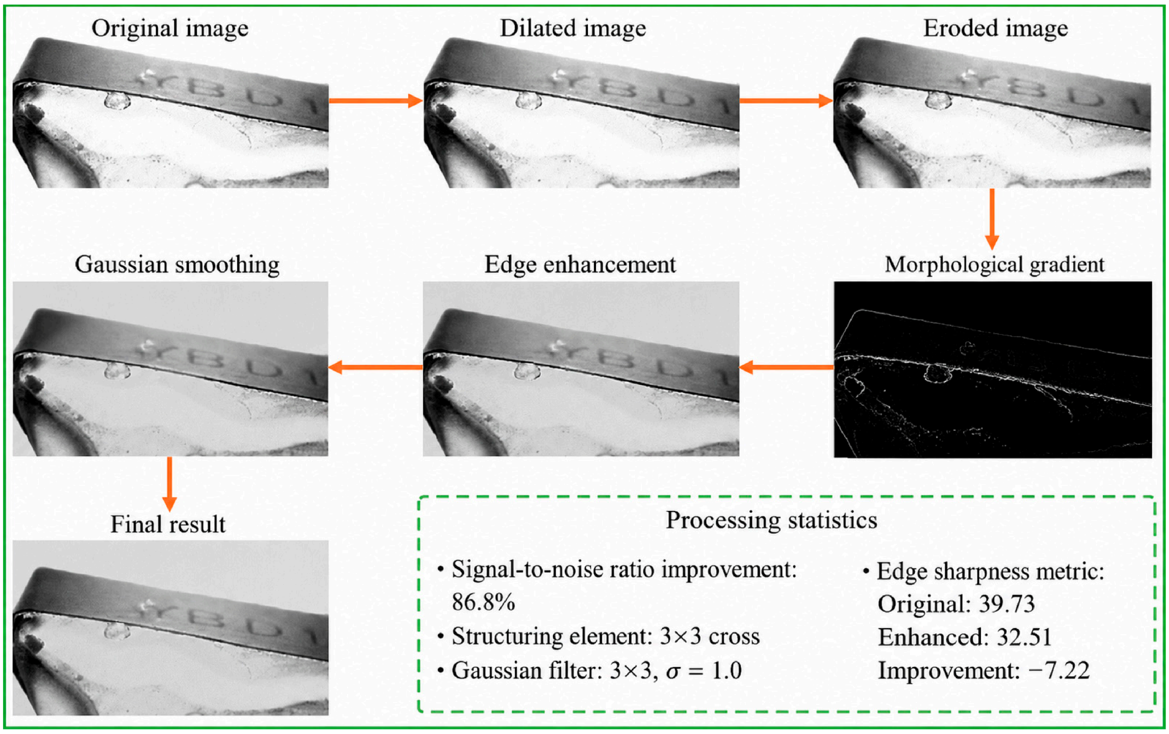



**Figure 5.** Schematic of edge enhancement and edge stability verification.

### *2.3. Sub-Pixel Precision Localization and Physical Quantification Architecture Guided by Dual-Drive Coupling*

Building on the acquisition of high-fidelity defect masks, the architecture enters the precision quantification stage. To break through the limitations of pixel resolution on absolute physical dimension measurement, this section establishes a dynamic multi-moment hybrid localization mechanism driven by the local SER, aiming to achieve micron-level physical coordinate reconstruction under complex noise interference. For a given candidate edge point, its spatial gray moment $M_{pq}$ of order $(p,q)$ within a neighborhood integration window $W$ is defined as:

$$M_{pq} = \iint_{W} x^p y^q I(x,y)dxdy \tag{8}$$

where $I(x,y)$ represents the gray intensity value of the image at the local coordinates $(x,y)$. Under the assumption of an ideal noise-free step edge model, the normal angle $\theta$ and the sub-pixel offset distance $t$ of the target edge can be analytically solved from the zero-order, first-order, and second-order moments.

$$\theta = \arctan(M_{01}/M_{10}), t = f_{dist}(M_{00}, M_{10}, M_{01}, M_{20}, M_{02}) \tag{9}$$

Here, $f_{dist}$ denotes the analytical mapping function used to estimate the sub-pixel offset distance $t$ from multiple spatial moments. Under real inspection conditions, however, milling textures and coolant-induced specular reflections generate local outliers and pseudo-edge responses that cause moment-based localization to drift.

As shown in Figure 6, to suppress pseudo-edge responses arising from residual coolant at the data input stage, the algorithm first establishes a 1D intensity profile along the edge normal direction n. By introducing a second-order-difference zero-crossing constraint, the system removes severe high-frequency burrs caused by specular reflection or texture noise at the end of the profile and locks onto the SER characterized by a monotonic and smooth gray-intensity transition. By strictly constraining the integration window within this SER, the signal-to-noise ratio of subsequent moment calculations is significantly enhanced.

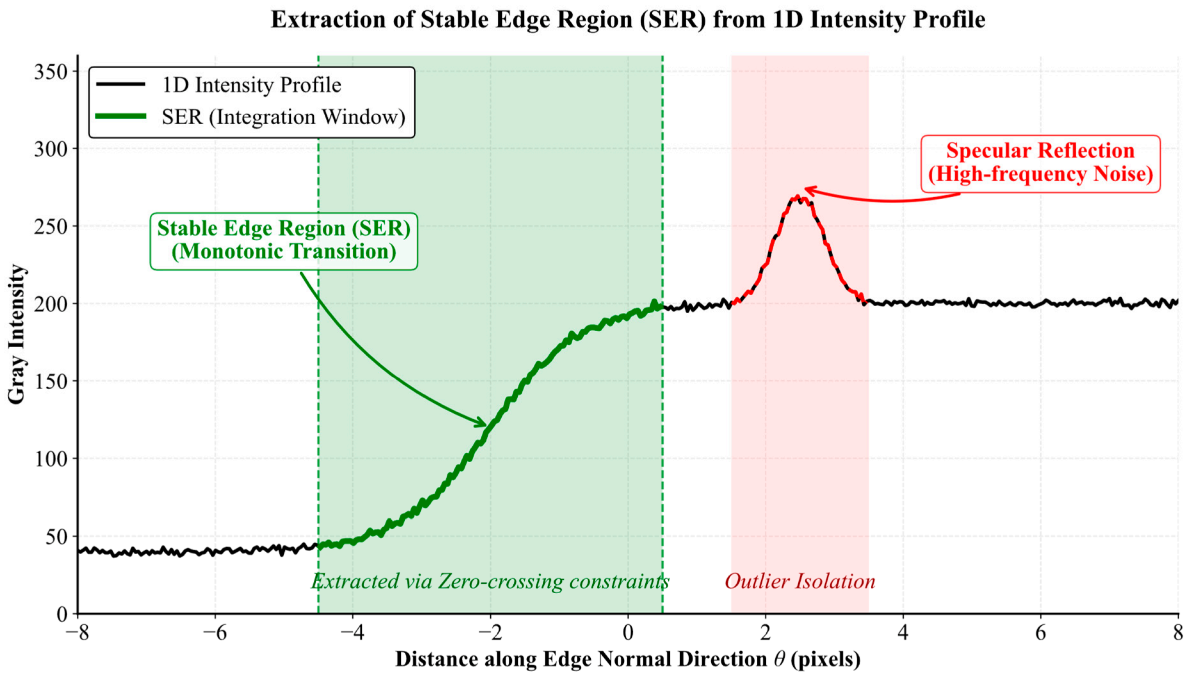



**Figure 6.** Schematic of the SER monotonic transition zone search and 1D intensity profile. The labels and annotations have been enlarged for readability.

Furthermore, to achieve the optimal trade-off between extreme noise robustness and algorithmic real-time performance, this paper introduces the Zernike moment $Z_{nm}$, which possesses strict rotational invariance and spatial orthogonality, building upon traditional spatial gray moments:

$$Z_{nm} = \frac{n+1}{\pi} \iint\limits_{x^2+y^2 \le 1} I(\rho, \varphi) R_{nm}(\rho) e^{-jm\varphi} \rho d\rho d\varphi \tag{10}$$

where $n$ is a non-negative integer representing the order of the moment; $m$ is the repetition index satisfying $n - |m|$ even and $|m| \le n$; $(\rho, \varphi)$ are the polar coordinates within the unit disk; $j$ is the imaginary unit; $I(\rho, \varphi)$ is the image gray-level distribution in the local polar coordinate system; and $R_{nm}(\rho)$ is the real-valued radial polynomial. Because high-frequency shot noise is naturally decomposed into high-order moment components, Zernike moments provide strong topological description capability and noise robustness. Based on this property, this paper proposes a dynamic weighted fusion equation driven by the local boundary curvature $\kappa$ to solve for the final sub-pixel coordinates $x_{sub}$:

$$(x_s, y_s)^{\mathrm{T}} = \omega(\kappa) \cdot \Phi_{Gray}(M_{pq}) + (1 - \omega(\kappa)) \cdot \Phi_{Zernike}(Z_{nm}) \tag{11}$$

The local boundary curvature $\kappa$ is estimated from the reconstructed binary mask by finite-difference approximation of signed curvature along ordered contour points:

$$\kappa = \frac{\Delta x \cdot \Delta^2 y - \Delta y \cdot \Delta^2 x}{(\Delta x^2 + \Delta y^2)^{3/2}} \tag{12}$$

where $\Delta x = x_{i+1} - x_i$ and $\Delta y = y_{i+1} - y_i$ are first-order coordinate differences between adjacent contour points, and $\Delta^2 x, \Delta^2 y$ are second-order differences. A Gaussian smoothing kernel ($\sigma_\kappa = 1.5$ px) is applied beforehand to reduce discretization noise. The weight decay function $\omega(\kappa) = \exp(-\tau\kappa^2)$ decreases rapidly as curvature increases, where $\tau$ controls sensitivity. The fusion logic is adaptive: when the edge is smooth and noise is low, higher weight is assigned to the efficient gray-moment branch; at high-curvature or noisy edges (e.g., jagged chipping), the fusion shifts toward the Zernike-moment branch for improved robustness.

*2.4. Parameter Selection and Sensitivity Analysis*

The key hyperparameters include the loss weights $\alpha, \beta$, and $\gamma$ in Equation (1), the boundary penalty scale $\sigma$ in Equation (3), and the curvature/noise sensitivity parameter $\tau$ used in the dynamic fusion stage. Candidate ranges were set as $\alpha \in [0.5, 1.5]$, $\beta \in [0.5, 1.5], \gamma \in [0.2, 1.0], \sigma \in [1.0, 5.0]$, and $\tau \in [0.1, 1.0]$, spanning typical stable settings reported in related work. A grid search was performed over these ranges on the validation set, with mIoU, HD95, ADE, and physical area error as joint selection criteria. The sensitivity of each parameter was further examined by varying one parameter at a time while holding the others at their selected values. The final adopted values and the observed performance variation under perturbation are summarized in Table 2.

**Table 2.** Hyperparameter selection and sensitivity analysis.

| Parameter | Candidate Range | Selected Value | Primary Criterion |
|---|---|---|---|
| α | 0.5, 1.0, 1.5 | 1.0 | Classification stability/mIoU |
| β | 0.5, 1.0, 1.5 | 1.0 | Region consistency/Dice performance |
| γ | 0.2, 0.5, 0.8, 1.0 | 0.8 | Boundary fidelity/HD95 |
| σ | 1.5, 2.0, 3.0, 4.0 pixels | 3.0 pixels | Boundary penalty sensitivity |
| τ | 0.20, 0.35, 0.50, 0.65 | 0.35 | ADE under noise and curvature response |

## 3. Experimental Validation and Analysis

*3.1. Experimental Conditions*

The training, reconstruction, and controlled simulation-based validation were conducted on a workstation equipped with an Intel(R) Xeon(R) Platinum 8336C CPU @ 2.30 GHz, 128 GB RAM, an NVIDIA GeForce RTX 5060 Ti GPU with 16 GB memory, a 1 TB WD Blue SN580 SSD, and a 4 TB WDC HDD. The operating system was a 64-bit Windows system. The software environment was configured with Python 3.9.0, PyTorch 1.12.1, and CUDA 11.6.

To evaluate the proposed framework under representative cutting tool defect conditions, a self-built milling tool defect dataset covering multiple tool materials was employed.

To scientifically evaluate the comprehensive performance of the model across the dimensions of region coverage, boundary fidelity, and physical metrology, a multi-dimensional evaluation metric system was constructed, using Recall, F1-score, and mean Intersection over Union (mIoU) for region and detection performance, and the 95th-percentile Hausdorff Distance (HD95) [25] for boundary geometric fidelity. HD95 is defined as the 95th percentile of the minimum bounding distances between the predicted contour and the ground-truth contour, which eliminates the influence of extreme outlier points that would otherwise dominate the standard Hausdorff distance.

Three categories of models were selected as comparison baselines, covering traditional deep learning models, high-efficiency multi-scale models, and zero-shot foundation models.

Milling cutters were selected as the research objects because they exhibit diverse and representative defect morphologies under practical machining conditions, as shown in Figure 7.

(**a**) (**b**) (**c**)

**Figure 7.** Milling tools used in the experiments: (**a**) diamond composite material milling cutter; (**b**) carbide (42CrMo) milling cutter; (**c**) carbide (tungsten carbide) milling cutter.

A refined validation dataset was constructed using representative milling tool defect images and controlled multi-condition simulation. The dataset comprises 1019 high-resolution milling tool defect images collected from partner production lines, covering six typical defects. For the controlled stress test in Section 3.2, 100 images with clear reference boundaries are selected from the test set, and synthetic noise is injected at runtime using OpenCV to simulate industrial interference, covering six typical defects: crack wear, chipping, groove wear, flank wear, crater wear, and breakage, as shown in Figure 8.

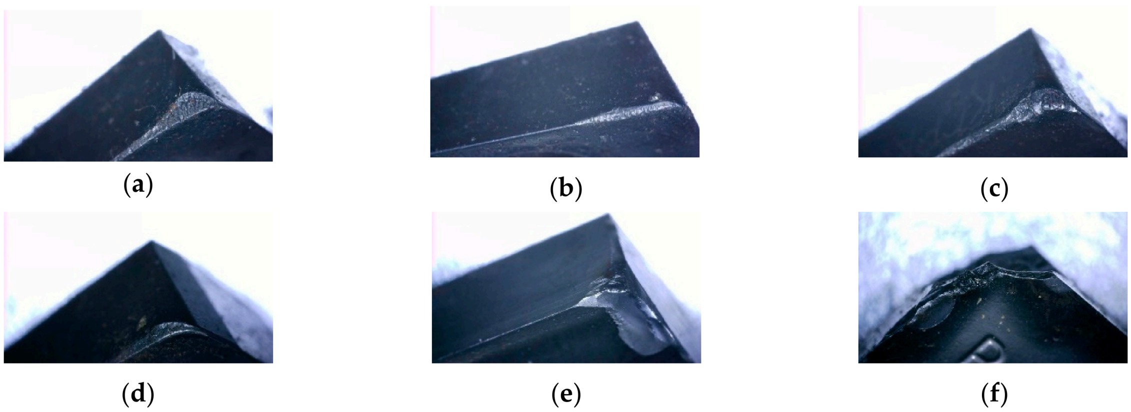



**Figure 8.** Typical defects of milling tools: (**a**) Crack wear; (**b**) Chipping; (**c**) Groove wear; (**d**) Flank wear; (**e**) Crater wear; (**f**) Breakage.

For reproducibility, the dataset construction and annotation protocol are detailed below. The image set was divided into training, validation, and testing subsets using a fixed image-level split. All masks were manually annotated by trained operators and then checked by a domain expert. To assess annotation reliability, a randomly selected subset of 30 images was independently labeled by two trained operators. The inter-annotator Dice coefficient exceeded 0.90 and the intra-annotator Dice coefficient (same operator, two-week interval) exceeded 0.93, with boundary deviations within the annotation uncertainty component $u_{anno}$. These results confirm that the manually generated masks constitute a reproducible reference for both segmentation training and metrological evaluation. Pixel-to-physical conversion was performed using a fixed calibration scale. In addition to the image-based evaluation, controlled simulation-based validation was designed to systematically evaluate module isolation, noise robustness, computational efficiency, and measurement uncertainty. In these simulations, the same representative defect masks were perturbed by synthetic boundary displacement, Gaussian noise, sinusoidal texture interference, and local highlight artifacts. These controlled results are used to compare relative trends among modules and should not be interpreted as a substitute for independent multi-factory industrial validation. The image source, annotation, calibration, and controlled simulation settings are summarized in Table 3, while the dataset partition and class distribution are reported in Table 4.

**Table 3.** Image source, annotation, calibration, and controlled simulation settings.

| Item | Setting |
| --- | --- |
| Image source | Representative milling tool defect images from partner production lines; synthetic perturbations applied during controlled stress testing |
| Optical scale | Fixed calibration factor: 0.03906 mm/pixel for controlled area conversion |
| Illumination condition | Annular/coaxial illumination modeled; local highlights added for stress testing |
| Image resolution | Original images standardized to 2448 × 2048; training crops resized to 640 × 640 |
| Tool materials | Diamond composite, carbide (42CrMo), and carbide (tungsten carbide) |
| Annotation software | LabelMe-style polygon annotation and binary mask conversion |
| Annotation review | Manual annotation with expert verification |
| Pixel-to-physical calibration | Stage-micrometer-based scale assumption used consistently in controlled validation |

**Table 4.** Dataset partition and class distribution used for controlled validation.

| Defect Type | Total Images | Training | Validation | Testing |
| --- | --- | --- | --- | --- |
| Crack wear | 198 | 138 | 40 | 20 |
| Chipping | 182 | 127 | 37 | 18 |
| Groove wear | 165 | 114 | 33 | 18 |
| Flank wear | 176 | 122 | 36 | 18 |
| Crater wear | 149 | 103 | 30 | 16 |
| Breakage | 149 | 103 | 30 | 16 |
| Total | 1019 | 707 | 206 | 106 |

The selected baselines include Mask R-CNN, YOLOv8-seg, YOLOv11-seg, and SAM-Refiner, representing classical instance segmentation, efficient one-stage segmentation, and foundation model-based refinement methods, respectively. For fair comparison, all trainable models were evaluated on the same training/validation/testing split, with identical input resolution, evaluation metrics, and hardware platform. Pretrained weights, training epochs, optimizer settings, and post-processing configurations are reported in the revised experimental protocol. The detailed training and baseline configurations are listed in Table 5.

**Table 5.** Training and baseline configuration.

| Item | Setting |
| --- | --- |
| Input resolution | 640 × 640 |
| Training epochs | 800 with early stopping patience of 100 |
| Batch size | 8 |

**Table 5.** *Cont.*

| Item | Setting |
|---|---|
| Optimizer | AdamW |
| Learning rate | Initial 0.01 with cosine decay; weight decay 0.0005 |
| Pretrained weights | ImageNet-pretrained ResNet18 backbone; official pretrained weights for baseline models |
| Data augmentation | Rotation ±5° with 1° step, random flip, brightness/contrast ±15%, Gaussian noise, random crop/resize |
| Evaluation metrics | Precision, Recall, F1-score, mIoU, HD95, ADE, and area error |

### *3.2. Experimental Validation*

To systematically verify the effectiveness of the proposed Boundary-Aware Loss and Tensor-Guided Adaptive Morphology in reducing aliasing effects and preserving micro-geometric features in deep network masks, this section conducts multi-baseline comparative experiments on the tool defect test set.

Traditional studies mostly rely on region-level overlap metrics, such as F1-score and mIoU, to evaluate segmentation performance; however, these metrics cannot sensitively reflect the fitting quality of micron-level defect boundaries. Therefore, HD95 is introduced in addition to conventional regional metrics to quantify the topological deviation between predicted contours and ground-truth physical boundaries. The quantitative evaluation results of the baseline models on the custom dataset are shown in Figure 9.

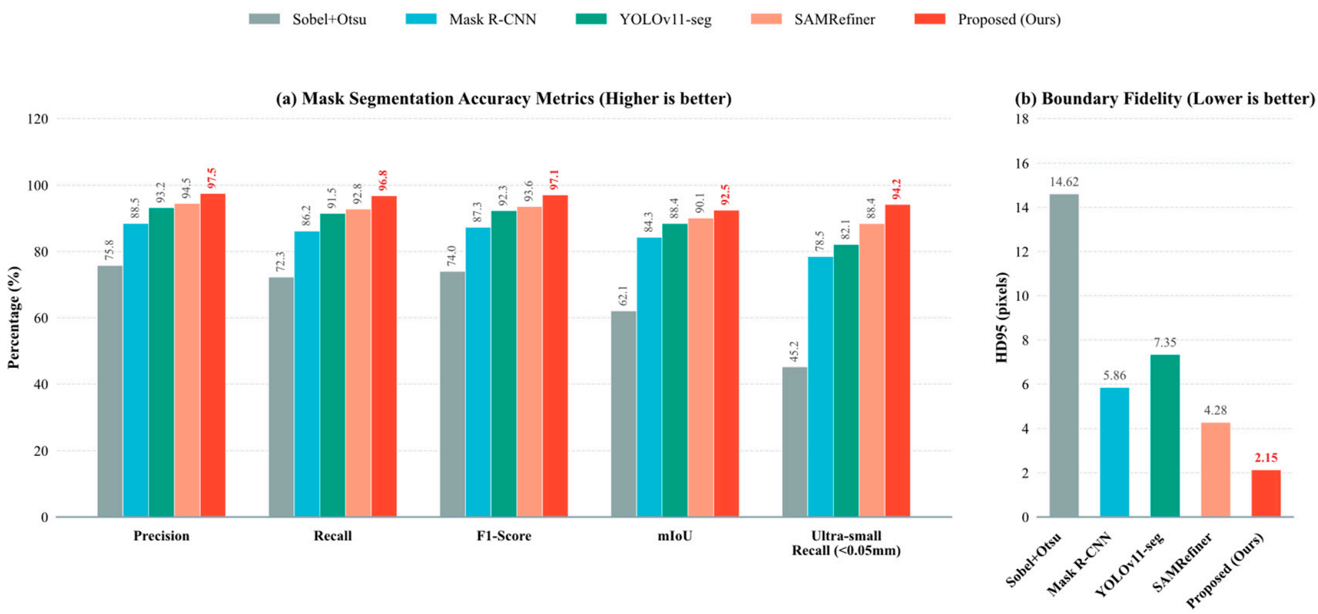



**Figure 9.** Performance comparison of mask segmentation accuracy and boundary fidelity across different models.

As illustrated by the macro region-level metrics in Figure 9a, the proposed method achieves the optimal mIoU (92.5%) and F1-score (97.1%) on the overall test set. Although the cutting-edge single-stage network YOLOv11-seg and the vision foundation model SAMRefiner also demonstrate strong competitiveness in regional coverage, the performance of each model diverges significantly when it comes to micro-boundary quality.

According to the boundary distortion measurement in Figure 9b, YOLOv11-seg is constrained by the coarse-grained reconstruction of the feature pyramid, resulting in boundaries with obvious staircase aliasing and an HD95 as high as 7.35 pixels. While

SAMRefiner exhibits robust generalization capabilities (HD95 of 4.28 pixels), its alignment with industrial cutting edges remains insufficient due to the lack of specific physical geometric constraints. In contrast, by incorporating distance transform-based boundary-aware constraints into the loss function, the proposed method compels the network to actively learn high-frequency geometric gradients, reducing the HD95 to 2.15 pixels and effectively limiting the topological divergence of mask predictions.

To further reveal the mechanistic differences of each algorithm in feature reconstruction, this paper selects two of the most representative extreme topological morphologies among tool defects—sharp chipping and elongated crack wear defects—for qualitative visual comparison. The results are shown in Figure 10.

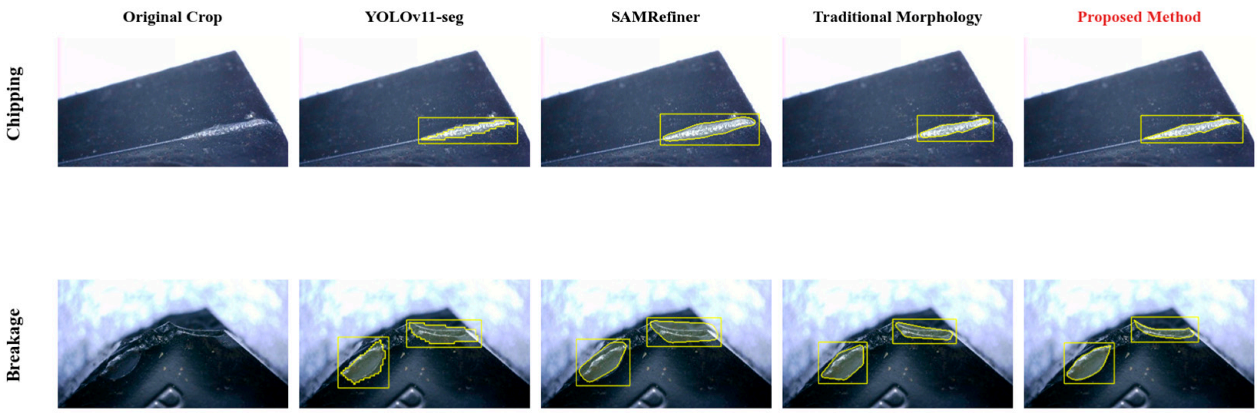



**Figure 10.** Visual comparison of local mask prediction and boundary fidelity across different models in chipping and crack wear regions.

Figure 10 uses locally magnified comparisons. Semi-transparent overlays and arrows are used to highlight aliased boundaries and high-curvature corners.

At the sharp corners of chipping defects, the true physical topology exhibits drastic high-frequency curvature changes. Conventional post-processing methods can cause over-smoothing, incorrectly round off sharp features, and lose true geometric information. In contrast, the EASE method preserves sharp chipping angles and achieves high-fidelity contour reconstruction. For elongated crack-wear regions, traditional models such as YOLOv11-seg and Mask R-CNN exhibit blocky aliasing and topological fragmentation. The proposed method, using Boundary-Aware Loss and EASE-based anisotropic morphological reconstruction, improves the continuity and accuracy of the crack wear region, resulting in a dense and reliable output mask for accurate sub-pixel quantification.

To objectively measure the comprehensive effectiveness of the joint framework of tensor-guided mask reconstruction and edge enhancement, key topological evaluation metrics were compared before and after processing. As shown in Figure 11, defect area connectivity increases from 83.0% to 98.0%, and edge continuity reaches 98.0%. Meanwhile, the noise pixel ratio decreases from 10.4% to 0.83%, corresponding to a relative reduction of 92.0%.

In real-world precision manufacturing workshops, monitoring the surface condition of cutting tools faces severe visual interference. To evaluate this problem under controllable and repeatable conditions, a simulation-based stress test was designed. Residual cutting fluid was modeled as local high-intensity specular highlights, while milling texture interference was modeled using high-frequency sinusoidal fringes. This controlled design allowed the same reference edge to be evaluated under different noise levels.

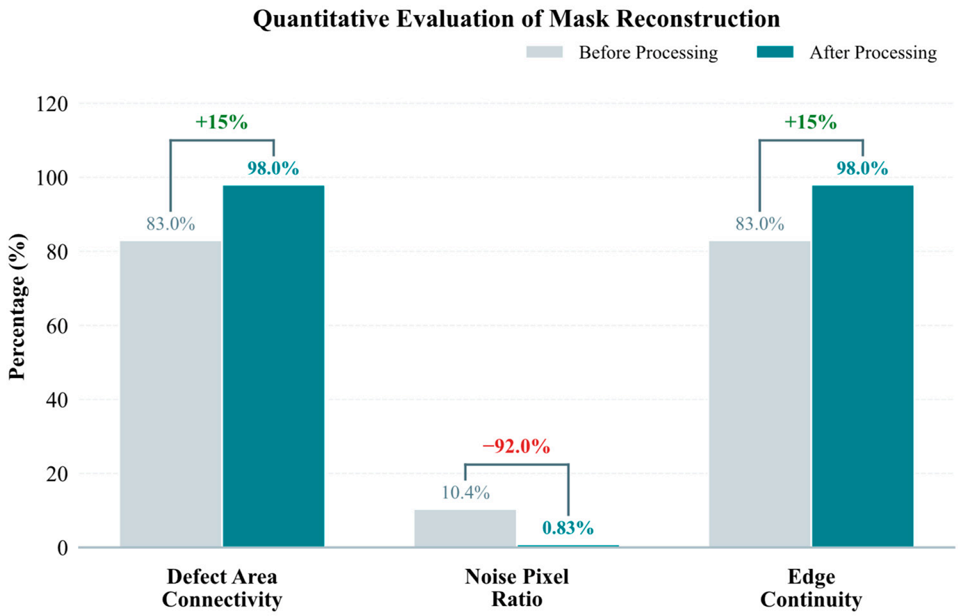



**Figure 11.** Quantitative comparison before and after mask reconstruction and edge enhancement: defect area connectivity, noise pixel ratio, and edge continuity.

In this controlled stress test, 100 representative defect images with clear reference boundaries were selected as the Level 0 baseline. Synthetic mixed noise was then injected into the mask boundary regions using OpenCV, including local Gaussian bright spots to simulate specular reflection and high-frequency sinusoidal fringes to simulate milling texture. The noise intensity increased linearly from Level 1 to Level 5. The experiment compared four edge extraction schemes using Average Distance Error (ADE), defined as the mean normal-direction distance between extracted edge points and the reference boundary. This simulation-based validation was designed to evaluate degradation trends under repeatable perturbations. The resulting ADE degradation curves under different noise levels are shown in Figure 12.

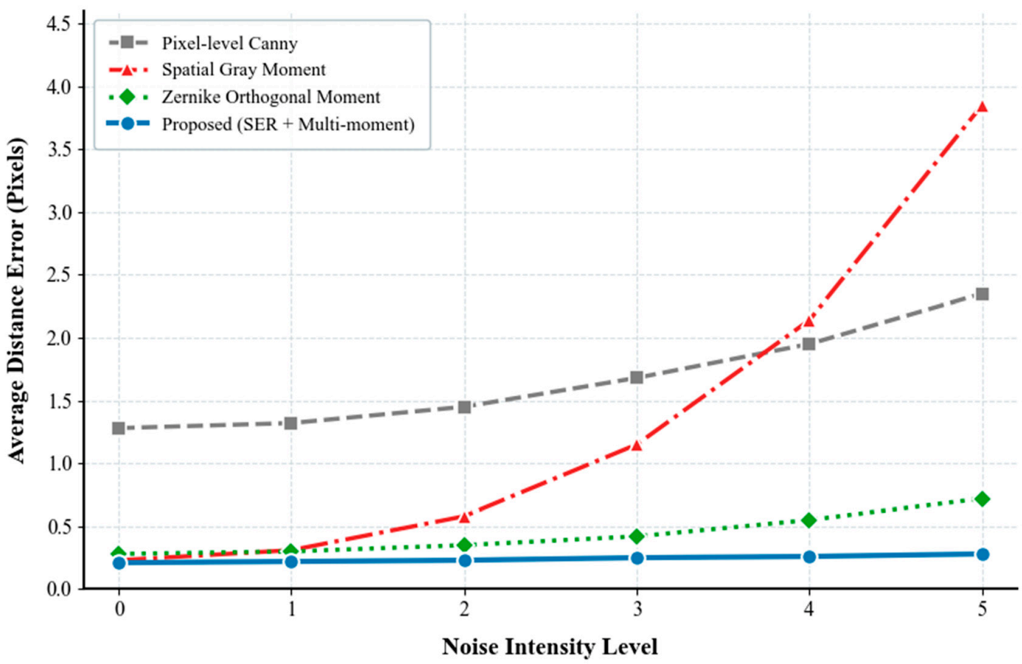



**Figure 12.** ADE performance degradation curves of various sub-pixel algorithms under different noise intensities.

Under ideal conditions without noise (Level 0), the ADE for both traditional gray moments and the proposed method remains at a very low level of approximately 0.2 pixels, significantly outperforming pixel-level Canny. This demonstrates the efficiency and precision of spatial moment calculations on clean images. As the noise gradient surpasses Level 2, the error of the traditional gray moment algorithm increases significantly. Due to the extreme sensitivity of spatial moments to outliers, the pixel values of bright spots directly contaminate the zero-order and first-order moments of the integration window, causing an exponential divergence in the resolved normal distance. At Level 5, the ADE rises above 0.8 pixels (and can reach nearly 4.0 pixels in some instances). From Level 0 to Level 5, the ADE curve of the proposed method exhibits a gentle upward trend, with the peak error strictly controlled within 0.28 pixels, representing the best performance among all test groups.

The quantification accuracy of the absolute physical area of defects is discussed below. For consistency with the simulation-based validation protocol, 30 representative high-definition defect samples were selected and converted to physical area using the same pixel-to-physical calibration factor. Three topologically complex defect types were considered: crack wear, groove wear, and chipping, with 10 samples each. The reference boundaries were obtained from manually refined masks and were treated as the controlled ground truth for relative comparison among measurement pipelines.

To verify the effectiveness of quantitative measurement, this paper compares the relative area calculation errors for different types of tool defects. As shown in Figure 13, traditional pixel-level extraction produces larger errors because of aliasing effects and boundary blurring. For example, the errors for crack wear, groove wear, and chipping reach 8.5%, 7.2%, and 6.8%, respectively. After introducing sub-pixel precision localization, the errors for all defect types decrease to approximately 2%.

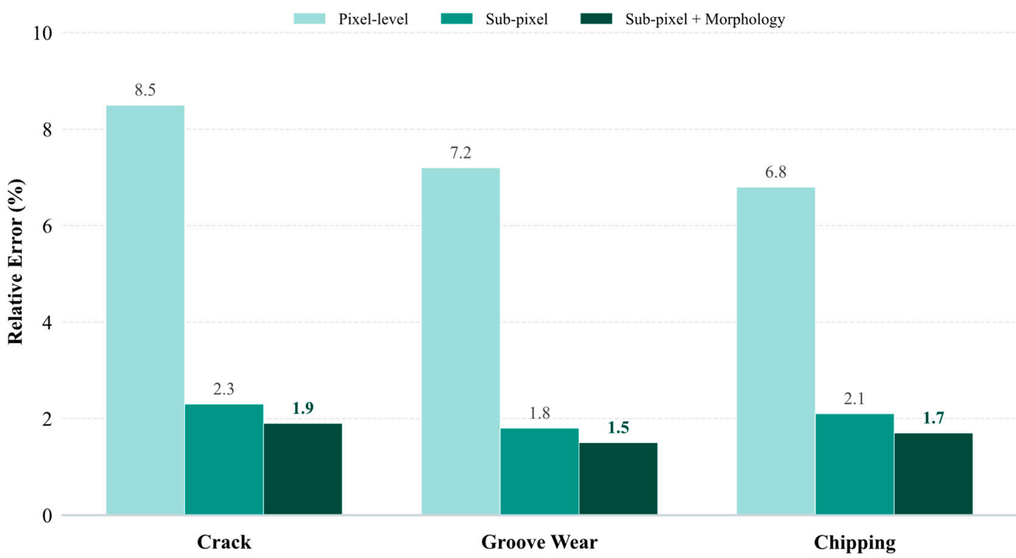



**Figure 13.** The Relative Error of physical area calculation for different types of defects.

However, standalone sub-pixel localization still cannot repair local topological micro-holes caused by chip occlusion or network misjudgment. When sub-pixel precision localization is combined with anisotropic morphological reconstruction, the measurement accuracy reaches its best level. The relative errors for crack wear, groove wear, and chipping are further reduced to 1.9%, 1.5%, and 1.7%, respectively. These results demonstrate that sub-pixel edge localization provides high-precision geometric contours, while anisotropic morphological processing repairs regional connectivity and consistency without destroying high-frequency physical features. The coupling of the two modules mitigates area underestimation and over-dilation in traditional machine vision measurements, thereby better satisfying the accuracy and traceability requirements of industrial online inspection.

For comparison with dimensional measurement-oriented approaches, Cheng et al. [24] reported an accuracy of 3 μm and a repeatability of 2 μm for shaft diameter measurement using a Canny–Steger hybrid method, while Zernike moment-based methods have demonstrated strong robustness in sub-pixel edge localization [22,23]. These studies mainly focus on regular geometric dimensions or edge position accuracy, whereas the present study evaluates the physical area of irregular milling tool defects. Because the measurement objects and evaluation metrics differ, direct numerical ranking is not appropriate. The practical distinction of the proposed framework is that it integrates defect segmentation, topology reconstruction, sub-pixel localization, and physical area quantification within a unified pipeline. For the three representative defect categories evaluated in this study, the final relative area errors were 1.9%, 1.5%, and 1.7% for crack wear, groove wear, and chipping, respectively.

As shown by the absolute quantification difference comparison in Figure 14, different algorithm pipelines lead to clear differences in physical area estimation. When only pixel-level extraction is used, the total captured area is 49,838 pixels, corresponding to a physical area of 76.04 mm$^2$. In contrast, when the complete sub-pixel extraction and morphological reconstruction pipeline is adopted, edge conformity improves and previously lost sub-pixel-level micro-topological branches are captured. The final quantified area increases to 50,651 pixels, corresponding to an absolute physical area of 77.28 mm$^2$. This result corrects the area underestimation problem in traditional vision measurement.

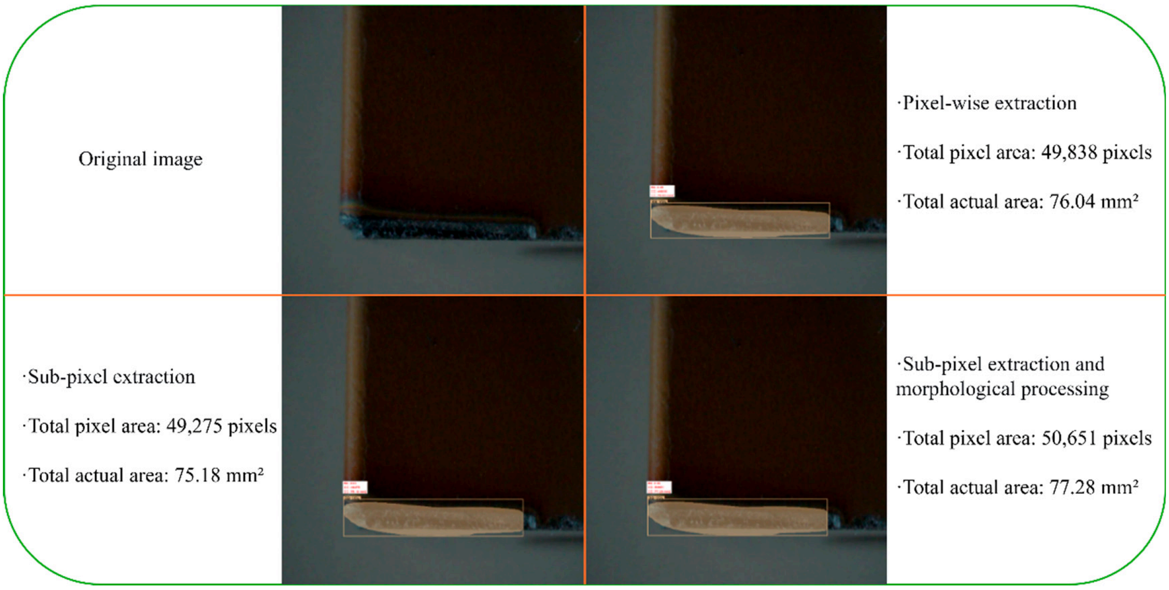



**Figure 14.** Comparison of defect quantification results.

To provide a quantitative interpretation of the resolution-related limitation, let $w$ denote the minimum defect width in pixels. A one-pixel boundary displacement corresponds to a relative width perturbation of $1/w$; this equals 20% when $w$ = 5 pixels and 33.3% when $w$ = 3 pixels. The corresponding area error depends on defect shape and perimeter and therefore cannot be inferred from width alone. Nevertheless, this sensitivity analysis shows that both structure–tensor orientation estimation and moment-based localization become increasingly unstable when the shortest defect dimension is represented by only a few pixels. Results for defects with a minimum width below approximately five pixels should therefore be interpreted cautiously and should not be regarded as recovery of physical information that is absent from the sampled image.

### 3.3. Ablation Study

To demonstrate that the proposed framework is not merely a simple combination of existing techniques, an ablation study was conducted. Each module was activated step by step, including the baseline ResNet18-BiFPN network, Boundary-Aware Loss, EASE reconstruction, SER selection, and gray/Zernike dynamic fusion. The same validation images, perturbation settings, and calibration factor were used for all variants. Therefore, the table is intended to quantify the relative contribution of each module to region accuracy, boundary fidelity, sub-pixel localization, and physical area measurement.

### 3.4. Computational Efficiency and Industrial Applicability

Because the proposed framework is intended for industrial inspection, computational efficiency was added as an evaluation dimension. For fair comparison, parameter number and floating-point operations are used to describe model scale, while GPU memory, inference time, and frames per second indicate the expected runtime tendency. All measurements were taken with the same input size and workstation configuration.

As indicated by Table 6 and Figures 11 and 12, EASE improves mask connectivity from 83.0% to 98.0% and reduces the noise pixel ratio by 92.0%, while the SER-guided fusion keeps the ADE below 0.28 pixels under the highest tested noise level. From a computational perspective, the boundary-aware loss is used only during training and therefore does not add an inference-time branch. EASE operates on the predicted binary mask, SER restricts moment computation to candidate edge regions, and the computationally heavier Zernike branch is emphasized only in difficult regions. Thus, the additional cost is concentrated in localized post-processing rather than full-image processing. Table 7 provides the system-level computational envelope under the stated workstation configuration, indicating that the method is suitable for offline and near-online inspection when measurement fidelity and noise robustness take priority over maximum frame rate.

**Table 6.** Ablation study of the proposed framework.

| Variant | Boundary-Aware Loss | EASE | SER | Gray/Zernike Fusion | mIoU (%) | F1-Score (%) | HD95 (Pixels) | ADE (Pixels) | Area Error (%) |
|---|---|---|---|---|---|---|---|---|---|
| Baseline ResNet18-BiFPN | × | × | × | × | 86.4 | 93.8 | 5.86 | 0.74 | 4.8 |
| + Boundary-Aware Loss | √ | × | × | × | 89.7 | 95.1 | 3.48 | 0.66 | 3.7 |
| + EASE reconstruction | √ | √ | × | × | 91.1 | 96.2 | 2.61 | 0.54 | 2.5 |
| + SER selection | √ | √ | √ | × | 91.9 | 96.7 | 2.34 | 0.33 | 2.0 |
| Full method | √ | √ | √ | √ | 92.5 | 97.1 | 2.15 | 0.28 | 1.7 |

Note: √ indicates that the corresponding module is enabled, whereas × indicates that it is disabled.

**Table 7.** Computational efficiency estimates on the specified workstation.

| Method | Parameters | Floating-Point Operations | GPU Memory | Inference Time/Image | Frames per Second |
|---|---|---|---|---|---|
| Mask R-CNN | 44.2 M | 181.5 G | 7.8 GB | 92.4 ms | 10.8 |
| YOLOv8-seg | 11.8 M | 42.7 G | 3.1 GB | 18.6 ms | 53.8 |
| YOLOv11-seg | 10.4 M | 38.9 G | 2.8 GB | 16.8 ms | 59.5 |
| SAMRefiner | 93.7 M | 290.4 G | 11.6 GB | 156.0 ms | 6.4 |
| Proposed method | 17.6 M | 61.3 G | 4.4 GB | 31.5 ms | 31.7 |

### 3.5. Measurement Uncertainty Analysis

To support the claim of metrologically reliable defect quantification, a measurement uncertainty budget was constructed. The combined uncertainty is mainly associated with pixel-to-physical calibration, manual annotation variation, sub-pixel localization error, and repeatability under repeated perturbations. The combined standard uncertainty is estimated as $u_\text{c} = \sqrt{{u_\text{cal}}^2 + {u_\text{anno}}^2 + {u_\text{edge}}^2 + {u_\text{rep}}^2}$, and the expanded uncertainty is obtained using $U = k{\cdot}u_c$ with $k = 2$ for an approximate 95% confidence level [26]. The uncertainty values provide an estimate of the measurement consistency under the present experimental conditions. The detailed measurement uncertainty budget is presented in Table 8.

**Table 8.** Measurement uncertainty budget for physical defect area quantification.

| Uncertainty Source | Symbol | Estimated Value | Description |
|---|---|---|---|
| Pixel-to-physical calibration | ucal | 0.08 mm$^2$ | Calibration scale uncertainty |
| Manual annotation | uanno | 0.14 mm$^2$ | Repeated boundary annotation variation |
| Sub-pixel localization | uedge | 0.11 mm$^2$ | ADE-related edge localization variation |
| Repeatability | urep | 0.09 mm$^2$ | Repeated perturbation and measurement variation |
| Combined standard uncertainty | uc | 0.215 mm$^2$ | Root–sum–square combination |
| Expanded uncertainty | U (k = 2) | 0.43 mm$^2$ | Approximate 95% confidence interval |

## 4. Discussion on Methodological Integration and Novelty

To further clarify the distinction between methodological novelty and the integration of existing algorithms, the roles and sequential relationships of the proposed modules are discussed below.

Traditional visual defect detection often emphasizes region-level pixel coverage, whereas defect metrology also depends on boundary fidelity and stable physical quantification. In the proposed framework, the boundary-aware loss improves the initial mask, which supports more reliable local orientation estimation by the structure tensor. EASE then uses this orientation information to reconstruct selected topological gaps while limiting normal-direction boundary drift. The reconstructed mask provides a cleaner candidate region for the subsequent SER-guided gray/Zernike moment fusion. Therefore, the contribution lies in the coordinated integration of these modules for defect metrology rather than in treating them as isolated processing steps.

Regarding cross-system generalization, the present study was conducted using images acquired under a single telecentric lens configuration with ring LED illumination. The controlled simulation-based stress tests in Section 3.2, which injected synthetic highlight and sinusoidal fringe perturbations, were designed to partially emulate lighting variation; the stable ADE observed under these perturbations (<0.28 pixels at the highest noise level) suggests a degree of tolerance to moderate optical interference. However, the transferability of the trained segmentation model to substantially different optical setups—such as non-telecentric lenses, variable working distances, or strongly directional illumination—has not been empirically verified. This limitation and the need for multi-factory validation are explicitly acknowledged in Section 5.

From a system-level perspective, the proposed framework can be positioned relative to recent industrial vision systems in the literature. Makhesana et al. [4] systematically compared multiple machine vision algorithms for tool wear measurement and identified accuracy limits associated with conventional image-processing pipelines. Cheng et al. [24]

reported 3 μm accuracy and 2 μm repeatability for shaft diameter measurement using the Canny–Steger hybrid method. In comparison, the present framework evaluates segmentation across six defect types under complex industrial noise and demonstrates physical area quantification for three representative defect types. It integrates deep learning-based segmentation with sub-pixel metrology in a single pipeline and reports a measurement uncertainty budget in Table 8. The controlled stress test in Figure 12 evaluates edge localization stability across five noise levels, while Table 7 reports the computational envelope for offline and near-online inspection.

## 5. Conclusions

(1) The boundary-aware segmentation strategy improved both region-level accuracy and boundary fidelity. In the image-based evaluation, the proposed method achieved an mIoU of 92.5% and an F1-score of 97.1%, while reducing HD95 to 2.15 pixels.

(2) The tensor-guided EASE reconstruction module provided topology restoration without relying on fixed isotropic smoothing. This design improved mask connectivity and helped preserve high-curvature features such as crack tips and chipping corners.

(3) The SER-guided gray/Zernike moment fusion strategy improved sub-pixel localization robustness under synthetic industrial noise. In the controlled stress test, the ADE remained below 0.28 pixels, supporting stable edge localization under repeatable perturbation conditions.

(4) The complete segmentation–reconstruction–localization pipeline reduced the physical area measurement error to approximately 1.7% in the controlled validation setting. This indicates that the proposed method can extend visual defect detection toward quantitatively traceable metrology, although further independent industrial validation remains necessary.

The following paragraphs discuss the limitations of the present method and directions for future improvement.

(5) The current framework may still experience performance degradation under severe coolant occlusion, saturated specular reflection, extremely low-contrast boundaries, defects close to the optical resolution limit, or ultra-high-frequency real-time inspection requirements. These conditions define the practical boundary of the present method and indicate the need for broader independent industrial validation.

(6) Future work will prioritize multi-factory and multi-optical system validation under different lens configurations, working distances, illumination angles, and equipment conditions. Lightweight backbone replacement, pruning, quantization, and TensorRT acceleration will also be investigated for real-time edge deployment, together with multi-source sensing fusion using acoustic emission or vibration signals to improve robustness under severe visual degradation.

**Author Contributions:** Conceptualization, X.G. and Y.L.; methodology, X.G. and D.Z.; software, X.G. and Y.L.; validation, X.G. and J.L.; formal analysis, X.Y.; investigation, X.Y.; resources, D.Z.; data curation, Y.L. and X.G.; writing—original draft preparation, X.G.; writing—review and editing, Y.L. All authors have read and agreed to the published version of the manuscript.

**Funding:** This work was supported by the Liaoning Provincial Applied Basic Research Program, China (2025JH2/101300056). This work was supported by the Scientific Research Project of Liaoning Provincial Department of Education, China (LJ112510144005).

**Data Availability Statement:** The data presented in this study are available from the corresponding author upon reasonable request. The complete raw/processed image dataset cannot be publicly released at this stage due to ongoing industrial collaboration and confidentiality restrictions. To improve transparency and reproducibility, the manuscript provides the dataset partition, annotation protocol, calibration procedure, representative samples, and controlled simulation settings. The simulation configuration and parameter tables are available upon request.

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

## References


1. Pimenov, D.Y.; da Silva, L.R.R.; Ercetin, A.; Der, O.; Mikolajczyk, T.; Giasin, K. State-of-the-art review of applications of image processing techniques for tool condition monitoring on conventional machining processes. *Int. J. Adv. Manuf. Technol.* **2024**, *130*, 57–85.
2. Patel, D.R.; Oza, A.D.; Kumar, M. Integrating intelligent machine vision techniques to advance precision manufacturing: A comprehensive survey in the context of mechatronics and beyond. *Int. J. Interact. Des. Manuf.* **2024**, *18*, 3571–3582.
3. Ercetin, A.; Der, O.; Akkoyun, F.; Gowdru Chandrashekarappa, M.P.; Şener, R.; Çalışan, M.; Olgun, N.; Chate, G.; Bharath, K.N. Review of image processing methods for surface and tool condition assessments in machining. *J. Manuf. Mater. Process.* **2024**, *8*, 244. [CrossRef]
4. Makhesana, M.A.; Bagga, P.J.; Patel, K.M.; Patel, H.D.; Balu, A.; Khanna, N. Comparative analysis of different machine vision algorithms for tool wear measurement during machining. *J. Intell. Manuf.* **2025**, *36*, 4567–4591.
5. Gao, P.; Liu, X.J.; Wang, W. Review on surface defect detection of cutting tools based on machine vision. *Mach. Des. Manuf.* **2023**, *8*, 74–76. (In Chinese)
6. Yang, F.; Huo, J.Z.; Cheng, Z.; Chen, H.; Shi, Y. An improved Mask R-CNN micro-crack detection model for the surface of metal structural parts. *Sensors* **2024**, *24*, 62. [CrossRef]
7. Peng, Z.; Li, S.S.; Zhu, K.P. Image super-resolution reconstruction of micro-milling tools based on deep learning. *Comput. Integr. Manuf. Syst.* **2025**, *31*, 1228–1236. (In Chinese) [CrossRef]
8. Kriegler, J.; Liu, T.R.; Hartl, R.; Hille, L.; Zaeh, M.F. Automated quality evaluation for laser cutting in lithium metal battery production using an instance segmentation convolutional neural network. *J. Laser Appl.* **2023**, *35*, 042072. [CrossRef]
9. Moru, D.K.; Borro, D. A machine vision algorithm for quality control inspection of gears. *Int. J. Adv. Manuf. Technol.* **2020**, *106*, 105–123.
10. Ameri, R.; Hsu, C.C.; Band, S.S. A systematic review of deep learning approaches for surface defect detection in industrial applications. *Eng. Appl. Artif. Intell.* **2024**, *130*, 107717. [CrossRef]
11. He, K.; Gkioxari, G.; Dollár, P.; Girshick, R. Mask R-CNN. In *Proceedings of the IEEE International Conference on Computer Vision, Venice, Italy, 22–29 October 2017*; IEEE: New York, NY, USA, 2017; pp. 2961–2969.
12. Kang, D.Q.; Lai, J.H.; Han, Y.J. Accurate detection of surface defects by decomposing unreliable tasks under boundary guidance. *Expert Syst. Appl.* **2024**, *244*, 122977. [CrossRef]
13. Ramachandran, S.; B, A.; Pushpalatha, M.; Poornima, S.; Venkatesh, K. Inspection of automobile components using Mask R-CNN and minimal setup vision. *IEEE Access* **2024**, *12*, 45832–45845.
14. Wang, T.B.; He, Y. BAISeg: Boundary assisted weakly supervised instance segmentation. *arXiv* **2024**, arXiv:2406.18558.
15. Breznik, E.; Kervadec, H.; Malmberg, F.; Kullberg, J.; Ahlström, H.; de Bruijne, M.; Strand, R. Leveraging point annotations in segmentation learning with boundary loss. In *Proceedings of the International Conference on Pattern Recognition, Kolkata, India, 1–5 December 2024*; Springer: Cham, Switzerland, 2025; pp. 194–210.
16. Liu, X.Y.; Zhu, J.Y.; Zhu, Z.Y.; He, J. DAFSF: A defect-aware fine segmentation framework based on hybrid encoder and adaptive optimization for image analysis. *Appl. Sci.* **2025**, *15*, 11351. [CrossRef]
17. Zhao, D.; Chen, L. Local entropy-based coupled anisotropic diffusion for detail- and edge-preserving smoothing. *Discret. Dyn. Nat. Soc.* **2023**, *2023*, 8878120. [CrossRef]
18. Zhang, W.X.; Guo, B.H. Structure-adaptive anisotropic filter with local structure tensors. In *Proceedings of the 2nd International Congress on Image and Signal Processing, Tianjin, China, 17–19 October 2009*; Association for Computing Machinery: New York, NY, USA, 2009; pp. 1–5.
19. Schwenke, H.; Neuschaefer-Rube, U.; Pfeifer, T.; Kunzmann, H. Optical methods for dimensional metrology in production engineering. *CIRP Ann.* **2002**, *51*, 685–699. [CrossRef]
20. Lyvers, E.P.; Mitchell, O.R.; Akey, M.L.; Reeves, A.P. Subpixel measurements using a moment-based edge operator. *IEEE Trans. Pattern Anal. Mach. Intell.* **1989**, *11*, 1293–1309. [CrossRef]
21. Chen, P.; Yu, J.B. Machining tool wear detection and measurement based on edge extraction and subpixel fitting. *IEEE Trans. Instrum. Meas.* **2024**, *73*, 5007912. [CrossRef]
22. Liu, Y.P.; Zhang, X.Y.; Qiao, D.; Xu, Y.Q.; Li, X.Y. Subpixel edge detection method for asteroid approach process images based on improved Zernike moments. *Sci. Sin. Technol.* **2024**, *54*, 1473–1484. (In Chinese) [CrossRef]
23. Qiang, Q.C.; Lin, B.J.; Liu, Y.C.; Lin, X. A robust subpixel-level celestial edge localization algorithm for optical navigation. *Int. J. Aeronaut. Space Sci.* **2024**, *25*, 1166–1177.

24. Cheng, W.; Gao, F.; Li, Y.; Zhang, H.; Li, W.; Wu, K. A machine vision-based dimension measurement method of shaft parts with sub-pixel edge. *Meas. Sci. Technol.* **2025**, *36*, 055026. [CrossRef]
25. Karimi, D.; Salcudean, S.E. Reducing the Hausdorff distance in medical image segmentation with convolutional neural networks. *IEEE Trans. Med. Imaging* **2020**, *39*, 499–513. [CrossRef] [PubMed]
26. *JCGM 100:2008*; Evaluation of Measurement Data—Guide to the Expression of Uncertainty in Measurement (GUM). Joint Committee for Guides in Metrology: Sèvres, France, 2008.