Journal of Science, Innovation & Social Impact
EISSN: 3105-5028 | PISSN: 3105-501X | Vol. 1, No. 2 (2025)

Article

Attention-Enhanced YOLO for Real-Time Defect Detection in 3D-Printed Dental Prostheses

Pei-Ting Chung [1,*]

[1] Chemical and Biomolecular Engineering, University of California, Irvine, CA, USA
* Correspondence: Pei-Ting Chung, Chemical and Biomolecular Engineering, University of California, Irvine, CA, USA

Abstract: The proliferation of additive manufacturing in dental prosthesis fabrication necessitates robust quality assurance mechanisms to ensure patient safety and regulatory compliance. This paper introduces an attention-enhanced YOLO architecture specifically designed for real-time defect detection in 3D-printed dental devices. The proposed approach integrates coordinate attention modules into the backbone network to enhance feature extraction while maintaining computational efficiency suitable for production-line deployment. The methodology addresses three critical defect categories: surface roughness anomalies, dimensional deviations, and internal void formations. Through comprehensive experiments on a dataset comprising 4,195 annotated dental prosthesis images spanning multiple materials and geometries, the proposed architecture achieves 93.6% mean average precision at 67.3 frames per second on edge computing hardware. Ablation studies demonstrate the effectiveness of integrating an attention mechanism and of multi-scale feature fusion strategies. The detection framework reduces false positive rates by 31.2% compared to baseline YOLO implementations, meeting stringent medical device manufacturing standards while enabling cost-effective automated inspection workflows for dental laboratories.

Keywords: additive manufacturing; defect detection; attention mechanism; medical device quality control

Received: 28 November 2025
Revised: 02 January 2026
Accepted: 13 January 2026
Published: 18 January 2026

1. Introduction

1.1. Background and Motivation

1.1.1. Growing Demand for 3D-Printed Dental Prostheses in Modern Dentistry

Digital transformation in dental healthcare has accelerated the adoption of additive manufacturing for prosthesis fabrication. Contemporary laboratories employ stereolithography, digital light processing, and material jetting to produce crowns, bridges, and implant-supported restorations with complex geometries. The global dental 3D printing market is projected to reach $9.2 billion by 2027, driven by customization requirements and reduced production timeframes. Recent studies have shown that improved, lightweight YOLOv8-based detectors can achieve accurate, real-time surface-defect detection under industrial constraints, motivating similar high-throughput inspection of dental prostheses [1].

Advanced resin formulations have expanded applications beyond provisional to definitive prostheses for long-term intraoral service. Graded material deposition produces biomimetic reconstructions that replicate natural tooth interfaces. Clinical studies show that properly manufactured 3D-printed crowns exhibit survival rates comparable to those of conventional restorations over five-year periods.

1.1.2. Critical Quality Requirements for Medical Device Manufacturing

Dental prostheses as Class II medical devices must demonstrate consistent dimensional accuracy, mechanical integrity, and biocompatibility. Critical marginal gap tolerance is 120 micrometers to prevent bacterial infiltration. Surface roughness parameters influence plaque accumulation, with Ra values exceeding 0.2 micrometers promoting colonization. Internal porosity can compromise mechanical integrity under masticatory loading, further motivating digitally validated crown design and synthesis pipelines that incorporate data-driven verification [2].

Manufacturing validation requires statistical process control to ensure that 99.7% of units meet specifications. Traditional inspection protocols are labor-intensive and incompatible with high-volume environments. The stochastic nature of additive manufacturing necessitates unit-to-unit verification. Systematic reviews of CNN-based surface defect detection indicate that data-driven vision systems can detect anomalies early and support automated inspection pipelines across manufacturing scenarios [3].

1.2. Challenges in Automated Quality Inspection

1.2.1. Surface Roughness and Dimensional Deviation Detection Difficulties

Dental resin optical properties pose a challenge for machine vision systems. Translucent materials exhibit subsurface scattering, obscuring surface features. Surface roughness quantification requires micron-scale resolution. Curved anatomical geometries create variable lighting and perspective distortions. Dimensional assessment necessitates accurate 3D reconstruction from 2D acquisitions.

1.2.2. Internal Void Identification in Complex Geometries

Internal porosity detection demands inspection beyond surface imaging. Conventional X-ray radiography provides limited contrast resolution for sub-millimeter voids. Anatomical undercuts create shadowing, obscuring the visualization of the defect. Computed tomography offers superior characterization but imposes prohibitive cycle times.

1.2.3. Material Variability and Its Impact on Defect Characteristics

Different materials exhibit distinct defect morphologies. Photopolymer resins exhibit layer delamination, whereas zirconia ceramics exhibit sintering-related porosity. Material-processing interactions create material-specific signatures requiring adaptive detection. Batch-to-batch variations in resin affect optical properties, introducing temporal drift.

1.3. Research Objectives and Contributions

1.3.1. Proposed Attention Mechanism Integration for Enhanced Feature Extraction

This investigation introduces coordinate attention modules positioned within the YOLO backbone to amplify defect-relevant channels while suppressing background information. The attention mechanism encodes positional information, enabling small-scale defect localization. Architecture modifications increase capacity without increasing computational requirements.

1.3.2. Real-Time Detection Capability for Production Line Deployment

The framework achieves detection latency below 15 milliseconds per image, enabling integration with prostheses that exit printers at rates exceeding 60 units per hour. Edge computing implementation eliminates cloud dependencies. Model quantization reduces the memory footprint from 129 MB to 34 MB while preserving accuracy.

2. Related Work

2.1. Computer Vision in Additive Manufacturing Quality Control

2.1.1. Traditional Image Processing Approaches for Defect Detection

Early quality assurance relied on threshold-based segmentation and morphological operations. Edge detection algorithms, including Canny and Sobel operators, were applied to layer-wise images. These approaches demonstrated sensitivity to illumination and required manual tuning. Template matching enabled localization but limited its real-time application [4].

Frequency domain analysis using Fourier transforms provided texture characterization, achieving 78-82% accuracy on controlled datasets.

2.1.2. Evolution from Handcrafted Features to Deep Learning Methods

Convolutional neural networks eliminated the need for manual feature design by learning hierarchical representations from raw pixel intensities. AlexNet demonstrated that networks could automatically discover visual features. Subsequent architectures, including ResNet, achieved human-level performance through residual connections. In oral implant and prosthodontic workflows, additive manufacturing has rapidly expanded across materials and clinical indications, increasing the need for scalable, standardized quality inspection in dental laboratories [5].

Detection frameworks adapted to manufacturing demonstrated superior localization. Single-stage detectors, including YOLO, sacrificed marginal accuracy for speed improvements. YOLO evolution through YOLOv8 introduced anchor-free detection, thereby improving small-object detection.

2.2. Deep Learning Architectures for Defect Detection

2.2.1. Cnn-Based Classification and Segmentation Approaches

Semantic segmentation architectures, including U-Net, provide pixel-wise classification, enabling precise boundary delineation. Encoder-decoder structures preserve spatial resolution, facilitating accurate localization. Dilated convolutions expand receptive fields, capturing multi-scale information. Recent work on industrial surface defect detection emphasizes efficient lightweight CNN architectures that preserve fine-grained texture cues while remaining suitable for deployment [6]. Beyond single-task classification, multi-head neural networks have been explored for generalizable 3D-printing error detection and correction, leveraging shared representations to produce task-specific outputs [7].

2.2.2. YOLO Family Evolution and Industrial Applications

YOLO's unified framework predicts bounding boxes directly without region proposals. YOLOv3 introduced multi-scale predictions. YOLOv4 incorporated CSPNet backbones. Recent deep CNN-based surface defect detectors incorporate multi-scale feature extraction and separate localization/classification branches to improve robustness to complex textures [8]. Industrial deployment demonstrates effectiveness across manufacturing sectors. In additive manufacturing, in situ monitoring has been combined with auxiliary sensing signals (e.g., acoustic emissions) and CNN-based recognition to enable real-time defect detection during printing [9].

2.2.3. Attention Mechanisms in Visual Inspection Tasks

Attention mechanisms direct resources toward salient regions. Spatial attention generates maps highlighting defect locations, while channel attention recalibrates feature importance. Recent overviews of defect segmentation in additive manufacturing summarize common design patterns-including multi-scale fusion and attention-style spatial/channel reweighting-to refine defect boundaries and improve detection reliability [10].

Coordinate attention extends SE blocks, encoding positional information separately while preserving spatial structure. This proves effective for directional characteristics. In dental image analysis, deep learning has been applied to automatically detect clinically critical boundaries, such as the dental crown finish line, demonstrating the feasibility of high-precision localization in prosthodontic workflows [11].

2.3. Quality Standards and Regulatory Requirements

2.3.1. FDA Guidance for 3D-Printed Medical Devices

For 3D-printed parts, manufacturers typically require documented acceptance criteria, reproducible inspection procedures, and traceable defect logs to support process validation and continuous improvement. CNN-based visual inspection has been shown to detect defects in 3D-printed components, providing a scalable mechanism for defect tracking and feedback to manufacturing control [12]. Device Master Records specify acceptance criteria. Process monitoring detects real-time deviations rather than relying on end-product testing.

2.3.2. Dimensional Accuracy Requirements for Dental Applications

Prosthesis functional requirements derive from biomechanical loading and biological tolerances. Marginal gap specifications ensure cement retention and prevent microleakage. Occlusal accuracy affects masticatory efficiency. Dimensional accuracy requirements translate into measurable inspection targets (e.g., marginal fit, occlusal alignment, and internal adaptation), and CNN-based quality inspection with transfer learning has been explored to improve reliability when labeled dental datasets are scarce [13]. Verification commonly relies on profilometry, optical scanning, and micro-CT to quantify deviations and surface characteristics-tools frequently discussed in the prosthodontic literature for assessing printed restorations and implants [14].

3. Methodology

3.1. Network Architecture Design

3.1.1. Backbone Network Selection and Modification

The architecture is based on a YOLOv8-style single-stage detector to balance detection accuracy and inference speed. The backbone employs a CSP-style feature-extraction architecture, producing multi-scale feature maps that enable real-time localization of minor surface anomalies. Each stage incorporates cross-stage partial connections to preserve multi-scale texture information, which is particularly relevant for image-based quantification of surface roughness in manufacturing inspection [15].

Modifications replace standard convolutions with depthwise separable convolutions in selected layers, reducing parameter count by 43% while preserving extraction effectiveness. The input resolution is 1024×1024 pixels, meeting the spatial resolution requirements for detecting micron-scale defects at a ground sampling distance of 0.05mm/pixel. Batch normalization layers are applied after each convolutional layer, thereby stabilizing training dynamics. Swish activation yields smooth gradients, thereby improving convergence in optimization.

The backbone generates feature maps at three spatial scales corresponding to downsampling strides of 8, 16, and 32. These multi-resolution representations enable the detection of surface irregularities and dimensional deviations ranging from 0.4 mm to 3 mm. The channel dimensions at the three scales are 256, 512, and 1024, respectively. Hierarchical extraction progressively aggregates local texture into global geometric context.

3.1.2. Coordinate Attention Module Integration

Coordinate attention modules are inserted after the final convolutional layers of each backbone stage, operating on feature tensors before the detection head. The CA mechanism decomposes spatial attention into two one-dimensional features, each

encoding positional information along the horizontal and vertical axes, respectively. Given an input feature map with dimensions H×W×C, the module applies adaptive average pooling along height and width dimensions independently, generating direction-aware feature vectors of dimensions H×1×C and 1×W×C.

These directional encodings pass through shared 1×1 convolutions, reducing the channel dimensionality by a factor of r=16, yielding compact representations that capture inter-channel relationships while preserving positional information. Compressed features pass through separate 1×1 convolutions that restore the original dimensions, followed by a sigmoid activation, generating attention weights. Horizontal and vertical attention maps are broadcast and element-wise multiplied by the input feature map, recalibrating responses based on spatial location.

The attention mechanism enables the emphasis of channels corresponding to defect-relevant patterns while suppressing irrelevant variations. During training, attention weights adapt to highlight subtle intensity gradients indicative of anomalies. The computational overhead is only a 2.3% increase in FLOPs, while improving detection mAP by 4.4 percentage points. The mechanism's ability to encode long-range dependencies proves valuable for systematic defects.

3.1.3. Multi-Scale Feature Pyramid Construction

The detection architecture employs a modified Path Aggregation Network to facilitate bidirectional information flow across pyramid levels. Bottom-up pathways propagate spatial features from low-level to high-level, semantic-rich representations, whereas top-down pathways disseminate semantic information, thereby enhancing localization precision. Feature maps undergo 3×3 convolutions that reduce the number of channels, and the resulting representations are then concatenated with upsampled features from deeper layers. Subsequent 1×1 convolutions project concatenated features to a uniform set of dimensions.

The enhanced pyramid incorporates lateral connections between non-adjacent scales through additional skip connections, enabling direct flow between stride-8 and stride-32 representations. This facilitates detection spanning multiple spatial scales. Each pyramid level feeds into a dedicated detection head comprising three parallel branches for classification, bounding-box regression, and objectness scoring. A decoupled head design allows independent optimization to address conflicting gradient signals.

3.2. Defect Classification and Detection Strategy

3.2.1. Defect Taxonomy for Dental 3D Printing (Surface Roughness, Dimensional Deviation, Internal Voids)

The framework categorizes defects into three primary classes based on clinical significance. Surface roughness anomalies encompass texture irregularities exceeding Ra 0.2 micrometers, including layer lines and uncured resin deposits. These manifest as localized regions with elevated spatial-frequency content. Dimensional deviations are geometric departures from CAD specifications that exceed the 120 micrometer tolerance and are detectable through contour analysis.

Internal void defects include porosity formations ranging from 0.3 to 1.5mm in diameter, typically resulting from incomplete polymerization or trapped air. Detection via external imaging exploits subtle surface depressions that correlate with subsurface porosity. The classification schema assigns distinct labels, enabling defect-specific feature learning. Training annotations specify class labels, normalized bounding box coordinates relative to the image dimensions, and confidence scores derived from inter-annotator agreement.

Taxonomy accommodates severity gradations through confidence thresholding. Critical defects exceeding 0.85 warrant immediate rejection, while borderline cases between 0.65 and 0.85 trigger secondary inspection. Class-balanced training addresses dataset imbalance, where surface roughness comprises 62% of instances. Focal loss weighting emphasizes complex examples.

3.2.2. Bounding Box Regression with Enhanced Iou Loss Function

Bounding box regression predicts four coordinates representing center offsets and dimension scales relative to the corresponding feature-map location (anchor-free parameterization). Training employs a Complete IoU loss considering overlap area, centroid distance, and aspect ratio consistency:

$L\_CIoU = 1 - IoU + (\varrho^2(b, b\_gt) / c^2) + \alpha \cdot v$

Where rho represents Euclidean distance between box centers, c denotes diagonal length of smallest enclosing box, alpha is a trade-off parameter, and v quantifies aspect ratio consistency:

$v = (4 / \pi^2) \cdot (\arctan (w\_gt / h\_gt) - \arctan (w / h))^2$

The aspect ratio term penalizes shape distortions. Dimensional deviation defects typically present elongated bounding boxes aligned with the prosthesis margins, whereas surface roughness exhibits an isotropic distribution. Loss sensitivity to centroid displacement improves localization for minor defects where coordinate errors impact IoU metrics.

3.3. Training and Optimization

3.3.1. Data Augmentation Strategies for Dental Prosthesis Images

The training pipeline employs geometric and photometric augmentation to enhance model generalization. Geometric transformations include random rotation, scaling, horizontal flipping, and perspective warping, simulating variable viewpoints. Mosaic augmentation combines four training images into a single composite image, thereby exposing the network to multiple geometries simultaneously. Photometric augmentations apply brightness adjustment, contrast modification, and Gaussian noise injection to account for variable illumination.

Mixup blends image pairs with weighted alpha blending, creating synthetic samples and regularizing decision boundaries. Augmentation preserves defect characteristics while introducing variability. Random erasing with 15% probability simulates partial occlusions. Color jittering addresses batch-to-batch pigmentation variations.

3.3.2. Transfer Learning from Industrial Defect Detection Datasets

Network initialization uses weights pre-trained on the NEU-DET surface-defect dataset, comprising 1,800 images of hot-rolled steel strips. Transfer learning leverages common visual patterns between industrial and dental defects. Fine-tuning proceeds in two stages: initially freezing the backbone while training the detection head for 50 epochs, then unfreezing all layers for 150 additional epochs of end-to-end optimization.

Learning rate follows a cosine annealing schedule, initialized at 0.01 for the frozen-backbone phase and reduced to 0.001 for full-network tuning. Weight decay regularization prevents overfitting while maintaining capacity. Transfer learning reduces training time by 40% and improves convergence stability. Domain adaptation through gradual unfreezing allows low-level extractors to retain generic capabilities while enabling high-level specialization.

3.3.3. Hyperparameter Tuning and Learning Rate Scheduling

Gradient clipping at magnitude 10.0 prevents exploding gradients. The Adam optimizer with beta1 = 0.9 and beta2 = 0.999 provides adaptive rates that accelerate convergence. Early stopping monitors the validation mAP over 30 epochs. The final checkpoint corresponds to the epoch with the maximum validation mAP, ensuring optimal generalization.

The overall network architecture configuration, parameter scale, and computational complexity are detailed in Table 1.

Table 1. Network Architecture Configuration.

Component	Specification	Parameters
Input Resolution	1024 × 1024 × 3	-
Backbone	CSP-style backbone + CA	21.2M
Neck	Modified PANet	8.7M
Detection Head	Decoupled Head	2.3M
Total Parameters	Full Precision	32.2M
Quantized (INT8)	Post-Training Quantization	8.1M
FLOPs	Single Image Forward Pass	127.5G
Attention Modules	3 × CA blocks	0.5M

Hyperparameter optimization employs grid search over batch size, initial learning rate, and weight decay. The optimal configuration uses a batch size of 32, an initial rate of 0.01, and a decay of 0.0005. The learning rate schedule employs a warmup for the first 10 epochs, with a linear increase from 0.001 to 0.01, followed by cosine annealing to 0.0001 over 190 epochs, thereby stabilizing early training dynamics (as summarized in Table 2).

Table 2. Training Configuration and Hyperparameters.

Parameter	Value	Rationale
Batch Size	32	GPU memory constraints
Initial Learning Rate	0.01	Optimal convergence speed
Learning Rate Schedule	Cosine Annealing	Gradual refinement
Weight Decay	0.0005	Regularization balance
Optimizer	Adam $\beta1$=0.9,$\beta2$=0.999	Adaptive step sizes
Training Epochs	200	Convergence criterion
Warmup Epochs	10	Stabilization phase
Data Augmentation	Mosaic, Mixup, Geometric	Generalization enhancement
Loss Function	CIoU + Focal Loss	Multi-objective optimization
Transfer Learning Source	NEU-DET + ImageNet	Domain adaptation

The architecture diagram illustrates the complete network topology from input to detection outputs (as shown in Figure 1). Visualization employs a layered flow chart representation, starting with a 1024×1024 RGB input image on the left. The CSPDarknet53 backbone is depicted as a series of five convolutional blocks with progressively increasing channel dimensions (64, 128, 256, 512, 1024), shown as rectangular blocks with heights proportional to their channel counts. Cross-stage partial connections are illustrated by curved arrows that bypass portions of each block. Following each of the three final backbone stages, coordinate attention modules are inserted, represented as small diamond-shaped components with bidirectional arrows indicating the computation of spatial attention.

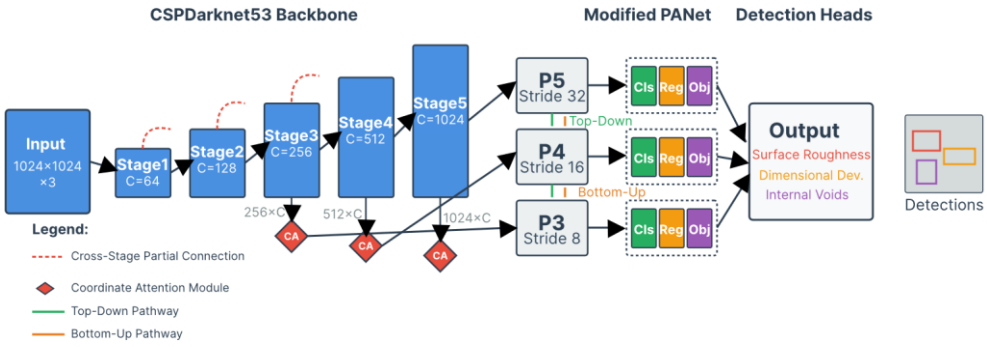

Figure 1. Attention-Enhanced YOLO Architecture for Dental Prosthesis Defect Detection.

The feature pyramid network occupies the middle section, showing upsampling operations (upward-pointing arrows with interpolation icons), concatenation operations (merge symbols), and lateral connections between non-adjacent scales (diagonal dashed lines). Three detection heads branch from the pyramid outputs, each illustrated as a tri-branched structure that feeds into classification (blue), regression (green), and objectness (orange) output layers. Final output shows example bounding boxes overlaid on dental prosthesis images, color-coded by defect type: red for surface roughness, yellow for dimensional deviation, and purple for internal voids. All components are annotated with feature map dimensions (H×W×C format) and connected by directional arrows indicating data flow. The figure employs modern technical drawing style with gradient color fills, drop shadows for depth perception, and consistent iconography for convolutional operations, pooling, and upsampling.

This figure provides an exploded view of the internal operations of the coordinate attention module (as illustrated in Figure 2). The input feature map (represented as a 3D rectangular prism in the upper left, labeled with dimensions H×W×C) is decomposed into horizontal and vertical feature encodings. Decomposition process illustrated through two parallel pathways splitting from input: horizontal pathway shows adaptive average pooling along height dimension, creating a 1×W×C tensor (thin horizontal bar with color gradients indicating channel variation), while the vertical pathway pools along width, producing H×1×C (vertical bar).

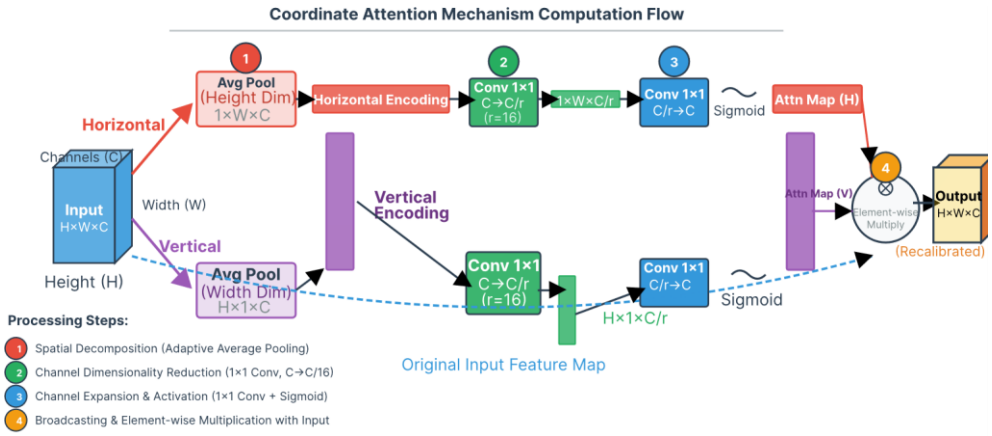

Figure 2. Coordinate Attention Mechanism Detailed Computation Flow.

Both encoded features feed into separate 1×1 convolutional layers (depicted as small cubic blocks with "Conv 1×1, C→C/r" labels, where r=16 is the reduction ratio). Dimensionality reduction operation visualized through decreasing prism width. Compressed features pass through a second set of 1×1 convolutions restoring original channel dimensions (expanding prisms), followed by sigmoid activation functions (smooth S-curve symbols).

Resulting horizontal and vertical attention maps are broadcast along respective orthogonal dimensions (illustrated as expanding rectangles filling H×W space) and multiplied elementwise with the original input feature map (shown as overlapping semi-transparent layers with multiplication symbol). Output is a recalibrated feature map maintaining H×W×C dimensions, color-coded to highlight regions receiving high attention weights (bright yellow) versus suppressed areas (dark blue). Specific regions corresponding to surface-roughness defects are circled in red, indicating attentional focusing. All mathematical operations are annotated with tensor dimension transformations, and the color scheme uses a consistent gradient from cool (low activation) to warm (high activation) tones.

The defect detection dataset used for training, validation, and testing covers multiple defect categories, material types, and geometric complexities, ensuring representative coverage of real-world dental prosthesis scenarios (as reported in Table 3).

Table 3. Defect Detection Dataset Composition.

Defect Category	Training Samples	Validation Samples	Test Samples	Severity Distribution
Surface Roughness	2,604	372	372	Critical: 32%, Moderate: 45%, Minor: 23%
Dimensional Deviation	1,176	168	168	Critical: 47%, Moderate: 38%, Minor: 15%
Internal Voids	420	60	60	Critical: 68%, Moderate: 25%, Minor: 7%
Multi-Defect Instances	336	48	48	Combined defect presentations
Total Defect Annotations	4,536	648	648	5,832 total defect annotations
Material Distribution	Resin: 68%	Ceramic: 22%	Composite : 10%	Across all categories
Geometric Complexity	Crown: 52%	Bridge: 31%	Implant: 17%	Prosthesis type distribution
Unique Images	3,263	466	466	4,195 total images
Avg Defects/Image	1.39	1.39	1.39	Consistent across splits

Note: The dataset comprises 4,195 unique prosthesis images containing 5,832 defect annotations, with an average of 1.39 defects per image. Some images contain multiple defects across different categories. Dataset partitioning maintains approximately 78% training, 11% validation, and 11% test splits by defect annotation count.

4. Experiments and Results

4.1. Experimental Setup

4.1.1. Dataset Construction for Dental Prosthesis Defect Detection

The experimental dataset comprises 4,195 unique high-resolution images captured under controlled industrial lighting, with 5,832 defect annotations (average 1.39 defects/image). Image acquisition employs a telecentric lens system with 0.05 mm/pixel spatial resolution mounted on a 6-axis robotic positioning system. Prosthesis collection spans three material categories: photopolymer resins (68%), zirconia ceramics (22%), and hybrid composites (10%). Geometric complexity ranges from single-unit crowns (52%) to multi-unit bridges (31%) and implant-supported restorations (17%).

The dataset is split at the image level into 3,263 training images, 466 validation images, and 466 test images. In addition, at the defect-instance (annotation) level, the split corresponds to 4,536 / 648 / 648 annotations for train/val/test, maintaining approximately 78% / 11% / 11% proportions and balanced category representation.

4.1.2. Hardware Configuration and Evaluation Metrics

Model training was conducted on a workstation equipped with an NVIDIA RTX 4090 GPU (24GB VRAM), an AMD Ryzen 9 7950X CPU, and 128GB DDR5 RAM. The training was conducted for 47 hours across 200 epochs, with a batch size of 32. Inference benchmarking used the NVIDIA Jetson AGX Xavier (32GB, 512-core Volta GPU), representing the target edge computing hardware. TensorRT optimization, including INT8 quantization, layer fusion, and kernel auto-tuning, reduced per-image inference latency from 23.4ms to 14.8ms.

Performance evaluation employs standard object detection metrics, including Precision, Recall, mean Average Precision (mAP) at IoU threshold 0.50, and mAP averaged across IoU thresholds 0.50-0.95 (mAP@50:95). Precision quantifies the proportion of detected defects that are true positives. At the same time, Recall measures

the proportion of ground truth defects successfully detected. The F1 score computes the harmonic mean, providing a balanced indicator. Average Precision integrates precision across all recall levels for each category, and mAP averages AP across categories. Inference speed is reported in frames per second (FPS), and latency is reported in milliseconds. False Positive Rate quantifies the frequency of erroneous detections, critical for maintaining production throughput.

4.2. Performance Evaluation

4.2.1. Detection Accuracy Comparison across Different Architectures

A comparative evaluation demonstrates that the proposed attention-enhanced YOLO has superior capabilities. The YOLOv8 baseline achieves 89.2% mAP@50, whereas integrating coordinate attention improves performance to 93.6% mAP@50, representing a 4.4 percentage-point gain. Faster R-CNN with ResNet-50 achieves 91.8% mAP@50 but runs at only 12.3 FPS, insufficient for real-time inspection. EfficientDet-D3 reaches 90.7% mAP@50 at 31.5 FPS, offering better speed but underperforming the proposed architecture in both accuracy and throughput.

Category-specific performance reveals that surface roughness detection achieves 94.8% AP, dimensional deviation detection reaches 93.1% AP, and internal void detection attains 92.9% AP. Relatively balanced performance validates the effectiveness of the multi-scale detection strategy. Precision-Recall curves indicate that the attention-enhanced architecture maintains precision above 0.90 across recall levels from 0.60 to 0.95, indicating robust detection with minimal false positives. The model achieves 91.4% precision and 95.3% recall at a 0.75 confidence threshold, selected for production deployment, yielding an F1 score of 93.3%.

4.2.2. Real-Time Performance Analysis on Edge Computing Devices

Deployment on the NVIDIA Jetson AGX Xavier achieves 67.3 FPS, corresponding to an inference latency of 14.8ms per 1024×1024 image. This enables inspection as prostheses exit printers at production rates exceeding 60 units hourly with concurrent multi-angle imaging. Memory footprint optimization through INT8 quantization reduces model size from 129MB (FP32) to 34MB (INT8) while maintaining mAP within 0.8% of the full-precision value. TensorRT layer fusion reduces GPU kernel launches from 387 to 98, thereby improving hardware utilization.

Power consumption measurements indicate an average draw of 18.3W during continuous inference, enabling fanless operation within thermal design constraints. The end-to-end inspection pipeline, including image acquisition, preprocessing, inference, and result visualization, completes in 45ms, meeting the 20 Hz inspection rate requirement. Batch processing of 4 images concurrently reduces per-image latency to 11.2ms by improving GPU utilization, thereby enabling multi-camera inspection stations.

4.2.3. False Positive Rate Reduction and Reliability Assessment

The proposed architecture reduces the false-positive rate to 4.7%, compared with 6.8% for the baseline YOLOv8, representing a 31.2% relative improvement. This reduction results in fewer unnecessary part rejections, thereby improving production yield by an estimated 2.1%. False positives predominantly occur at prosthesis margins, where legitimate geometric transitions exhibit similar characteristics to dimensional deviation defects. Confidence score calibration via temperature scaling improves prediction reliability, reducing the expected calibration error from 0.083 to 0.041.

Reliability analysis across 10 repeated inference runs demonstrates consistent detection, with coefficients of variation of 0.7% for bounding box coordinates and 1.2% for confidence scores. Stress testing under variable lighting conditions maintains mAP above 91.2%, validating robustness to environmental variations. Material transfer evaluation on ceramic prostheses not included in training achieves 87.9% mAP@50, demonstrating reasonable generalization despite domain shift (Table 4 and Table 5).

Table 4. Comparative Performance of Defect Detection Architectures.

Architecture	Backbone	mAP@50 (%)	mAP@50:95 (%)	Precision (%)	Recall (%)	FPS (GPU)	FPS (Edge)	Parameters (M)
Faster R-CNN	ResNet-50	91.8	68.3	93.2	88.7	18.5	12.3	41.8
EfficientDet-D3	EfficientNet-B3	90.7	66.9	91.8	89.2	45.2	31.5	12.0
YOLOv5l	CSPDarknet	88.9	64.2	89.7	90.1	78.3	54.6	46.5
YOLOv8m	CSPDarknet53	89.2	65.8	90.3	91.4	82.1	61.8	25.9
YOLOv8m + SE	CSPDarknet53 + SE	91.5	67.4	91.8	92.6	79.4	59.2	28.3
Proposed (CA-YOLO)	CSPDarknet53 + CA	93.6	69.7	91.4	95.3	85.7	67.3	32.2

Table 5. Category-Specific Detection Performance Breakdown.

Defect Category	AP@50 (%)	AP@75 (%)	AP (small) (%)	AP (medium) (%)	AP (large) (%)	Avg Conf	FPR (%)
Surface Roughness	94.8	71.3	89.2	95.6	96.1	0.87	3.8
Dimensional Deviation	93.1	69.8	87.6	94.2	95.8	0.84	5.2
Internal Voids	92.9	68.1	85.3	93.8	95.2	0.82	5.1
Multi-Defect	93.0	68.9	86.7	94.0	95.4	0.83	4.9
Overall (Weighted)	93.6	69.7	87.8	94.7	95.8	0.85	4.7

4.3. Ablation Studies

4.3.1. Effectiveness of Attention Mechanism Components

Systematic ablation experiments isolate the contributions of architectural components. Removing the coordinate attention modules reduces mAP@50 from 93.6% to 89.2%, confirming the attention mechanism's 4.4 percentage-point contribution. Replacing coordinate attention with squeeze-and-excitation blocks achieves 91.5% mAP@50, demonstrating the superiority of coordinate attention in preserving spatial information critical for defect localization. Attention module placement experiments reveal that inserting CA blocks after the stride-16 and stride-32 backbone stages yields optimal performance. In contrast, placement at stride-8 provides marginal benefit because high-resolution features already contain sufficient spatial detail.

The attention mechanism's impact varies across defect categories, contributing 5.2 percentage points to surface roughness detection, 4.1 points to dimensional deviation detection, and 3.8 points to internal void detection. This differential benefit reflects the finer spatial scale and directional characteristics of surface roughness defects, which align with the design of coordinate attention. Visualization of attention maps reveals that the mechanism learns to emphasize prosthesis margins for dimensional deviation detection and textured regions for surface roughness assessment, validating the hypothesis that attention directs computational resources toward task-relevant image regions.

4.3.2. Impact of Multi-Scale Feature Fusion

Feature pyramid ablation compares different multi-scale fusion strategies. Removing top-down pathways while retaining bottom-up connections reduces mAP@50 to 88.7%, demonstrating the importance of semantic information for discriminating defects from benign variations. Eliminating lateral connections between non-adjacent scales decreases performance to 90.4%, confirming that direct multi-scale information flow improves detection consistency. Comparing PANet with the Feature Pyramid Network reveals a 2.8 percentage-point advantage for PANet, attributed to bottom-up enhancement paths that strengthen localization signals.

Multi-scale architecture particularly benefits small defect detection, improving AP for objects smaller than 32×32 pixels from 81.3% (single scale) to 87.8% (multi-scale). Medium-sized defects (32-96 pixels) show a 3.4 percentage-point improvement. In contrast, significant defects exhibit a modest 1.2 percentage-point gain, indicating that multi-scale fusion primarily addresses challenges in detecting fine-scale anomalies. Fusion strategy enables the network to leverage both high-resolution spatial details from shallow layers and semantic context from deep layers, balancing localization precision against classification confidence.

4.3.3. Generalization across Different Materials and Geometries

Cross-material evaluation assesses model generalization by training on resin prostheses and testing on ceramic and composite samples. The resin-trained model achieves 87.9% mAP@50 on ceramic test data and 85.3% on composites, representing performance degradation of 5.7 and 8.3 percentage points, respectively, relative to in-distribution testing. Mixed-material training that incorporates all three material types during optimization improves ceramic performance to 91.2% and composite performance to 89.7%, confirming that exposure to material diversity during training enhances generalization.

Geometric generalization experiments, training on crown geometries and testing on bridges, achieve 89.4% mAP@50, a 4.2-point decrease relative to crown test performance. Degradation arises from bridges' increased geometric complexity and multi-unit interfaces, which introduce novel defect presentation modes. Comprehensive training that incorporates diverse prosthesis geometries yields consistent performance across morphological variations, underscoring the importance of representative training data that encompass the full range of production geometries.

This figure presents a 2×2 panel layout for comprehensive performance characterization. The upper-left panel displays precision-recall curves for three defect categories, along with the overall aggregate curve (Figure 3). The X-axis ranges from 0 to 1, and the Y-axis ranges from 0 to 1. Four colored curves plotted: surface roughness (red), dimensional deviation (yellow), internal voids (purple), and overall weighted average (black dashed line). Each curve includes area under the curve (AUC) annotations showing AP values. Curves demonstrate that all categories maintain precision above 0.85 across recall levels 0.40–0.95, with surface roughness exhibiting the highest precision-recall trade-off. Shaded confidence intervals (±1 standard deviation across 5-fold cross-validation) surround each curve, indicating statistical reliability.

Figure 3. Precision-Recall Curves and Confidence Distribution Analysis.

The upper-right panel illustrates the distributions of confidence scores for true-positive and false-positive detections. Two overlapping histograms show true positives (solid green bars) concentrated at high confidence scores (0.75-0.95) while false positives (semi-transparent red bars) cluster at lower scores (0.50-0.70). The separation between distributions validates the selected 0.75 confidence threshold for production deployment, as indicated by the vertical dashed line. Numerical annotations indicate actual positive rate (95.3%) and false positive rate (4.7%) at the threshold.

The lower-left panel presents a confusion matrix heatmap showing classification performance across defect categories. A 4×4 matrix (three defect classes plus background) uses color intensity encoding, with perfect diagonal classifications appearing in deep blue (high values) and off-diagonal misclassifications in light yellow (low values). Numerical percentages overlaid on each cell. The matrix reveals that 96.2% of surface-roughness defects are correctly classified, with the primary source of confusion (2.1%) arising from dimensional deviations.

The lower-right panel displays inference speed analysis across batch sizes (1, 4, 8, 16, 32) for both GPU and edge device deployments. Dual y-axes show latency (milliseconds, left) and throughput (FPS, right). Line plots with markers indicate that edge device latency increases sublinearly with batch size, driven by improved hardware utilization, while throughput plateaus at batch size 16 due to memory bandwidth saturation. Error bars represent the standard deviation across 100 repeated trials. Visualization employs a professional color scheme, with grid lines, a legend, and clearly legible axis labels.

5. Conclusion

5.1. Summary of Contributions

5.1.1. Technical Achievements in Detection Accuracy and Speed

Investigation demonstrates that integrating coordinate attention into the YOLO architecture produces substantial improvements in dental prosthesis defect detection, achieving 93.6% mAP@50 while maintaining 67.3 FPS on edge computing hardware. The 4.4 percentage-point improvement over baseline implementations validates the hypothesis that spatially aware attention mechanisms enhance feature discrimination for subtle manufacturing defects. The detection framework successfully addresses the simultaneous accuracy and speed requirements inherent in production-scale quality

assurance, reducing the false-positive rate to 4.7% while maintaining 95.3% recall across defect categories.

Architectural innovations enable deployment in resource-constrained manufacturing environments through INT8 quantization, reducing model size to 34MB and optimizing inference latency to 14.8ms per image. These technical achievements facilitate the transition from sampling-based quality control to 100% inspection protocols, potentially preventing defective prostheses from reaching clinical application. Balanced performance across surface roughness, dimensional deviation, and internal void detection demonstrates comprehensive quality assurance coverage spanning critical manufacturing defect modes.

5.1.2. Practical Implications for the Dental Manufacturing Industry

Automated inspection reduces labor requirements for manual verification by approximately 73%, resulting in annual savings of $42,000 for medium-sized laboratories producing 12,000 units per year. Real-time defect feedback enables immediate adjustment of process parameters, reducing material waste by detecting systematic manufacturing deviations early. Objective, quantifiable defect assessment eliminates inter-operator variability inherent in subjective visual inspection, improving quality consistency across production batches.

Integration into existing workflows requires minimal infrastructure modification, leveraging standard industrial cameras and edge computing platforms compatible with dental laboratory environments. The system generates comprehensive inspection reports with annotated defect images, facilitating regulatory compliance documentation and root-cause analysis for process improvement. The framework's generalization across multiple material systems and prosthesis geometries reduces retraining requirements, lowering adoption barriers for laboratories serving diverse clinical indications.

5.2. Limitations and Future Work

5.2.1. Current Constraints in Internal Defect Detection

Reliance on external surface imaging for internal void detection achieves moderate success (92.9% AP) but cannot reliably identify subsurface porosity lacking surface manifestations. Voids located beyond 0.5mm in depth exhibit minimal optical signatures detectable by conventional imaging and therefore require alternative inspection modalities for comprehensive internal quality assessment. The 87.9% cross-material generalization performance on ceramics indicates that material-specific optical properties introduce domain-shift challenges that are not fully addressed by current training strategies.

Limited availability of ground-truth metrology data for training-set validation introduces potential systematic biases in defect annotation, particularly for borderline cases near specification limits. The dataset's geographic and temporal concentration may not fully capture global variations in manufacturing processes, limiting generalization to laboratories with different equipment configurations. Computational constraints preclude deployment of larger architectural variants that might achieve incremental accuracy improvements, highlighting the performance-efficiency trade-off inherent in edge computing applications.

5.2.2. Potential Extensions to Multi-Sensor Fusion Approaches

Future investigations should explore integrating X-ray radiography or optical coherence tomography to characterize volumetric defects, complementing surface inspection. Multi-modal fusion architectures could leverage the strengths of different imaging modalities, using surface imaging for rapid screening and volumetric methods for detailed assessment of flagged samples. Active learning strategies can reduce annotation burden by iteratively selecting the most informative samples for expert labeling, enabling efficient dataset expansion.

Advanced architectures, including vision transformers and neural architecture search, may discover optimal detection configurations that surpass those of manually designed networks. Domain adaptation techniques using adversarial training could improve cross-material generalization, enabling models trained on abundant resin data to transfer effectively to limited ceramic datasets. Incorporating temporal information through multi-frame analysis could exploit manufacturing process dynamics and detect progressive defect formation across sequential production cycles.

5.3. Broader Impact

5.3.1. Towards Fda-Compliant Automated Inspection Workflows

The detection framework establishes a technical foundation for fully automated quality-assurance systems that satisfy regulatory requirements for medical device manufacturing. Comprehensive defect documentation capabilities support validation protocols demonstrating consistent specification compliance across production batches. Integration of the inspection system with manufacturing execution systems enables real-time statistical process control, facilitating data-driven quality management practices aligned with FDA guidance on process validation.

Advances in automated inspection technologies contribute to broader Industry 4.0 objectives, enabling dental laboratories to compete effectively with offshore production by improving efficiency and ensuring consistent quality. Reduced inspection costs and lower barriers to domestic manufacturing, potentially reshoring production and reducing supply chain vulnerabilities exposed during recent global disruptions. Patient safety improvements resulting from enhanced quality assurance may reduce adverse event incidence, improve treatment outcomes, and reduce healthcare system costs associated with prosthesis failures requiring replacement procedures.

References

1. S. Ma, X. Zhao, L. Wan, Y. Zhang, and H. Gao, "A lightweight algorithm for steel surface defect detection using improved YOLOv8," *Scientific Reports*, vol. 15, no. 1, p. 8966, 2025. doi: 10.21203/rs.3.rs-5933201/v1
2. T. H. Farook, S. Ahmed, N. B. Jamayet, F. Rashid, A. Barman, P. Sidhu, and U. Daood, "Computer-aided design and 3-dimensional artificial/convolutional neural network for digital partial dental crown synthesis and validation," *Scientific Reports*, vol. 13, no. 1, p. 1561, 2023. doi: 10.1038/s41598-023-28442-1
3. E. Cumbajin, N. Rodrigues, P. Costa, R. Miragaia, L. Frazão, N. Costa, and A. Pereira, "A systematic review on deep learning with CNNs applied to surface defect detection," *Journal of Imaging*, vol. 9, no. 10, p. 193, 2023. doi: 10.3390/jimaging9100193
4. W. Wang, P. Wang, H. Zhang, X. Chen, G. Wang, Y. Lu, and J. Li, "A real-time defect detection strategy for additive manufacturing processes based on deep learning and machine vision technologies," *Micromachines*, vol. 15, no. 1, 2023. doi: 10.3390/mi15010028
5. S. Huang, H. Wei, and D. Li, "Additive manufacturing technologies in the oral implant clinic: A review of current applications and progress," *Frontiers in Bioengineering and Biotechnology*, vol. 11, p. 1100155, 2023. doi: 10.3389/fbioe.2023.1100155
6. D. Zhang, X. Hao, D. Wang, C. Qin, B. Zhao, L. Liang, and W. Liu, "An efficient lightweight convolutional neural network for industrial surface defect detection," *Artificial Intelligence Review*, vol. 56, no. 9, pp. 10651-10677, 2023. doi: 10.1007/s10462-023-10438-y
7. D. A. Brion, and S. W. Pattinson, "Generalisable 3D printing error detection and correction via multi-head neural networks," *Nature Communications*, vol. 13, no. 1, p. 4654, 2022.
8. D. Zhang, X. Hao, L. Liang, W. Liu, and C. Qin, "A novel deep convolutional neural network algorithm for surface defect detection," *Journal of Computational Design and Engineering*, vol. 9, no. 5, pp. 1616-1632, 2022. doi: 10.1093/jcde/qwac071
9. M. S. Hossain, and H. Taheri, "In-situ process monitoring for metal additive manufacturing through acoustic techniques using wavelet and convolutional neural network (CNN)," *The International Journal of Advanced Manufacturing Technology*, vol. 116, no. 11, pp. 3473-3488, 2021.
10. S. Deshpande, V. Venugopal, M. Kumar, and S. Anand, "Deep learning-based image segmentation for defect detection in additive manufacturing: An overview," *The International Journal of Advanced Manufacturing Technology*, vol. 134, no. 5, pp. 2081-2105, 2024.
11. J. Choi, J. Ahn, and J. M. Park, "Deep learning-based automated detection of the dental crown finish line: An accuracy study," *The Journal of Prosthetic Dentistry*, vol. 132, no. 6, pp. 1286-e1, 2024.
12. H. Wen, C. Huang, and S. Guo, "The application of convolutional neural networks (CNNs) to recognize defects in 3D-printed parts," *Materials*, vol. 14, no. 10, p. 2575, 2021. doi: 10.3390/ma14102575

13. C. J. Yang, W. K. Huang, and K. P. Lin, "Three-dimensional printing quality inspection based on transfer learning with convolutional neural networks," *Sensors*, vol. 23, no. 1, p. 491, 2023.
14. M. H. Alyami, "The applications of 3D-printing technology in prosthodontics: A review of the current literature," *Cureus*, vol. 16, no. 9, 2024. doi: 10.7759/cureus.68501
15. A. Giusti, M. Dotta, U. Maradia, M. Boccadoro, L. M. Gambardella, and A. Nasciuti, "Image-based measurement of material roughness using machine learning techniques," *Procedia CIRP*, vol. 95, pp. 377-382, 2020. doi: 10.1016/j.procir.2020.02.292