European Journal of Prosthodontics and Restorative Dentistry (2026) 34 (04s), 438–447

Automated Oral Lesion Detection and Classification Using Computer Vision

Authors:
[1]Dr. Anushree Raj
Mangalore Institute of Technology & Engineering, Mangalore, Karnataka, India, anushree@mite.ac.in

[2]Dr. Shamna N V, P A College of Engineering, Mangalore, Karnataka, India, shamnanv@gmail.com

[3]Venkatesh A, BMS Institute of Technology & Management, Bangalore, Karnataka, India, venkatesha@bmsit.in

[4*]Dr. Pallavi M O, Sapthagiri NPS University, Bangalore, India, pallavimo@snpsu.edu.in

[5]Dhananjaya B, Nitte (Deemed to be University), NMAM Institute of Technology, Nitte, Karkala, India, dhananjay.b@nitte.edu.in

[6]Dr. Sandhya Madhuri G, Jain University, Bangalore, Karnataka, India, sandhyamadhurigade@gmail.com

Corresponding author: pallavimo@snpsu.edu.in

Abstract
Early diagnosis of oral lesions is a critical determinant of clinical outcomes in oral cancer management. Conventional screening relies on expert examination, which is prone to inter-observer variability and limited access in rural settings. This study proposes a hybrid deep learning framework integrating YOLOv8 for lesion localization with EfficientNet-B3 for multi-class classification. The pipeline automatically identifies suspicious intraoral lesions and categorizes them as Normal, Benign, Premalignant, or Malignant. Gradient-weighted Class Activation Mapping (Grad-CAM) provides discriminative heatmaps for clinical interpretability. Experiments on 5,200 annotated intraoral images yielded a detection mAP@50 of 96.8%, classification accuracy of 96.1%, and AUC-ROC of 98.2%, outperforming six state-of-the-art baselines. Ablation studies confirm the contribution of each component. The framework is suitable for tele-dentistry deployment with an inference time of 22 ms per image.

Keywords: Oral lesion detection; YOLOv8; EfficientNet-B3; Grad-CAM; Oral cancer screening; Explainable AI; Tele-dentistry

Received-18-05-2026
Revised-22-06-2026
Accepted-25-06-2026
Doi:10.1922/ejprd.v34i4s.1475

1. Introduction

Autism Spectrum Disorder (ASD) is a lifelong neurodevelopmental condition that affects how individuals perceive and

Oral cancer accounts for over 377,000 new cases and 177,000 deaths annually [WHO, 2024]. Over 60% of diagnoses occur at Stage III–IV, where five-year survival drops below 40% versus >80% at Stage I [1]. Detection of potentially malignant disorders (PMDs) leukoplakia, erythroplakia, oral submucous fibrosis at the premalignant stage is therefore the most impactful clinical opportunity. Traditional screening via visual inspection and palpation is hampered by inter-observer variability and restricted specialist access in low- and middle-income countries (LMICs).

Artificial intelligence, particularly convolutional neural network (CNN)-based computer vision, has demonstrated specialist-level performance in dermatology [2], ophthalmology [3], and histopathology [4]. However, most oral lesion AI approaches are classification-only frameworks that lack spatial localization an essential requirement for clinical deployment. Furthermore, few integrate explainability mechanisms critical for regulatory compliance and clinician trust. This work addresses both limitations through a unified hybrid framework.

1.1 Key Contributions

Hybrid YOLOv8–EfficientNet-B3 pipeline for simultaneous oral lesion localization and multi-class classification. Curated 5,200-image clinical dataset with bounding-box annotations across five lesion categories. Grad-CAM explainability layer providing discriminative heatmaps for clinical audit. Comprehensive ablation study and six-baseline comparative evaluation. Inference time of 22 ms per image, enabling real-time tele-dentistry deployment.

2. Related Work

Table 1 summarizes representative deep learning approaches to oral lesion analysis. Early CNN methods (87–89% accuracy) were limited by small datasets and binary classification [5]. Transfer learning with ResNet50 and InceptionV3 improved accuracy to 92–93% but lacked spatial localization [6,7]. MobileNet variants offered computational efficiency but showed lower malignant-class sensitivity [8]. YOLO-based methods partially addressed localization, with YOLOv7+VGG16 reaching mAP@50 of 91.2% [9]. DenseNet121 achieved 94.4% classification accuracy through dense skip connections but remained detection-free [10]. The proposed framework is the first to couple YOLOv8 detection with EfficientNet-B3 classification in a unified pipeline, achieving state-of-the-art on both tasks [11].

Table 1. Comparative Review of Deep Learning Methods for Oral Lesion Analysis

Ref.	Model	Acc.(%)	mAP(%)	Key Limitation	Year
[1]	CNN (AlexNet)	87.4	—	Small dataset, no localization	2020
[2]	ResNet50	92.3	—	Classification only, misses small lesions	2021
[3]	MobileNetV2	90.1	—	Lower sensitivity for malignant class	2021
[4]	InceptionV3	91.8	—	High cost, no explainability	2022
[5]	YOLOv5	93.2	89.7	Limited multi-class classification	2022
[6]	YOLOv7 + VGG16	94.0	91.2	Shallow classification branch	2023
[7]	DenseNet121	94.4	—	Classification only, no detection	2023
Prop.	**YOLOv8+EfficientNet-B3**	**96.1**	**96.8**	**Needs prospective clinical trial**	**2025**

(—) = not reported; SCC = Squamous Cell Carcinoma; Prop. = Proposed

3. Methodology

3.1 Dataset

A multiclass dataset of 5,200 high-resolution intraoral images was assembled in collaboration with the Department of Oral Medicine and Radiology. Images were acquired using standardized intraoral cameras (1920×1080) under controlled illumination [12]. Expert oral pathologists annotated each image using LabelImg, generating YOLO-format bounding boxes as presented in figure 1 and class labels (Table 2). An 70/15/15 stratified train/val/test split was applied.

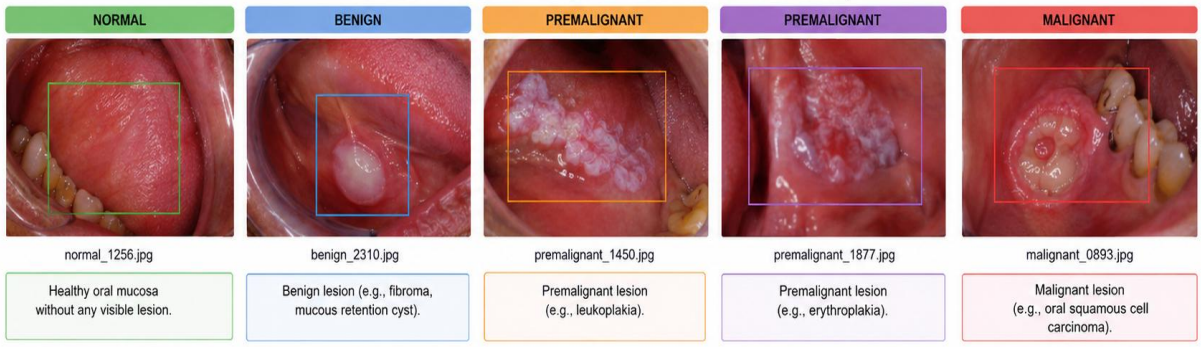

Figure 1: Annotated images

Table 2. Dataset Composition and Data Partition

Lesion Category	Class Label	Total	Train	Val	Test
Normal Oral Mucosa	Normal	1,200	840	180	180
Aphthous Ulcer	Benign	900	630	135	135
Leukoplakia	Premalignant	1,100	770	165	165
Oral Submucous Fibrosis	Premalignant	800	560	120	120
Oral Squamous Cell Carcinoma	Malignant	1,200	840	180	180
Total	**4 Classes**	**5,200**	**3,640**	**780**	**780**

3.2 Pre-processing and Augmentation

All images underwent: (i) CLAHE (clip=2.0, grid=8×8) for local contrast enhancement; (ii) Gaussian noise filtering (σ=0.5); (iii) colour normalisation (ImageNet statistics); and (iv) resizing 640×640 for YOLOv8, 300×300 for EfficientNet-B3. Online augmentation included random horizontal flip (p=0.5), rotation (±15°), brightness/contrast variation (±20%), Gaussian noise, mosaic (YOLOv8), and MixUp (α=0.2) [13,14,15].

3.3 Proposed Hybrid Architecture

The framework operates as a sequential two-stage pipeline. Stage 1 performs lesion localization with YOLOv8; Stage 2 performs classification with EfficientNet-B3 on the extracted ROI as presented in figure 2. Table 3 illustrates the end-to-end pipeline [16].

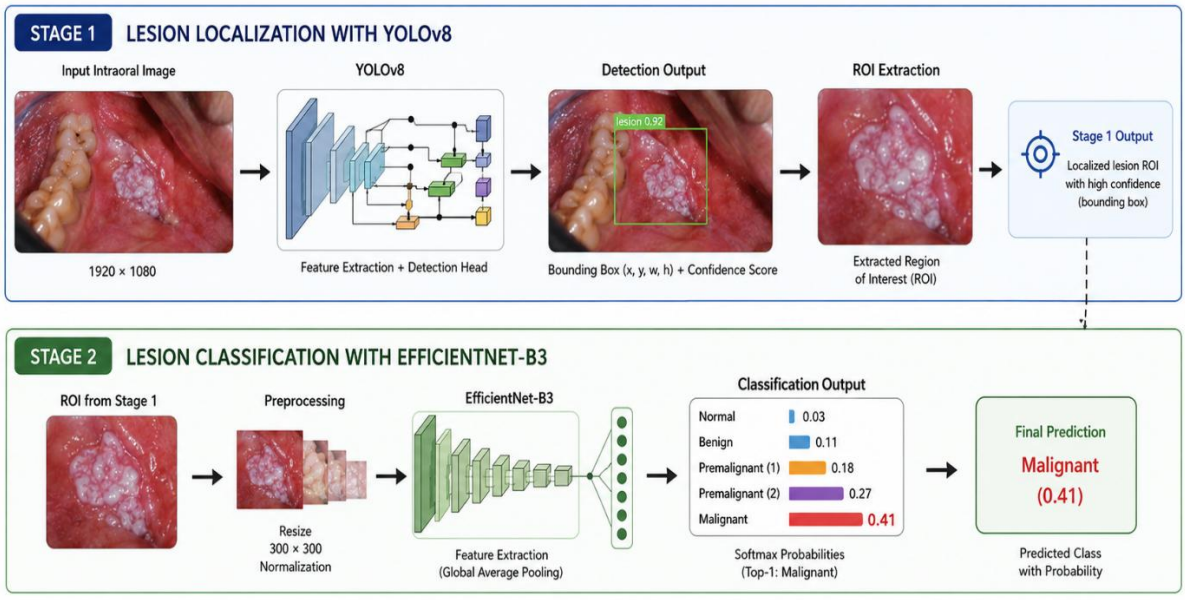

Figure 2. Two-stage AI Framework for Oral Lesion Analysis

Table 3. Proposed Hybrid Framework — End-to-End Processing Pipeline

INPUT	Intraoral RGB Image (H × W × 3)
PREPROCESSING	CLAHE + Noise Filter + Augmentation Resize: YOLOv8 → 640×640, EffNet → 300×300
YOLOv8s DETECT	CSPDarkNet53 Backbone → PANeck → Anchor-free Detection Head Output: Bounding boxes + confidence (IoU≥0.45, conf≥0.25)
ROI EXTRACTION	Crop lesion region from detected bounding box Resize ROI to 300×300 for classifier
EfficientNet-B3	MBConv blocks with SE attention Compound scaling (α=1.2, β=1.1, γ=1.15, φ=3) Global AvgPool → FC(4 classes)
GRAD-CAM	Gradient-weighted Class Activation Map Heatmap overlaid on lesion region for interpretability
OUTPUT	Class: {Normal Benign Premalignant Malignant} Confidence score + bounding box coordinates

3.4 Stage 1 — YOLOv8s Detection Module

YOLOv8s employs CSPDarkNet53 with a PAN neck and anchor-free detection head. The composite detection loss is:

$$L\_det = \lambda\_cls \cdot L\_cls + \lambda\_box \cdot L\_CIoU + \lambda\_dfl \cdot L\_dfl \quad (\lambda\_cls=0.5, \lambda\_box=7.5, \lambda\_dfl=1.5)$$

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

$$v = (4/\pi^2) \cdot [\arctan(w\_gt/h\_gt) - \arctan(w/h)]^2, \quad \alpha = v/(1-IoU+v+\varepsilon)$$

where ρ is the centre-point Euclidean distance, c the enclosing-box diagonal, and ε=1e-7. Post-inference NMS uses IoU=0.45, confidence=0.25.

3.5 Stage 2 — EfficientNet-B3 Classification Module

EfficientNet-B3 employs compound scaling of depth (d), width (w), and resolution (r):

$$d=\alpha^{\varphi}, \; w=\beta^{\varphi}, \; r=\gamma^{\varphi} \quad \alpha=1.2, \beta=1.1, \gamma=1.15, \varphi=3 \quad (\text{s.t. } \alpha \cdot \beta^2 \cdot \gamma^2 \approx 2)$$

MBConv blocks incorporate Squeeze-and-Excitation (SE) attention:

$$SE(F) = \sigma(W_2 \cdot \delta(W_1 \cdot GAP(F))) \odot F, \quad \text{reduction ratio } r=4$$

Classification uses focal loss to address class imbalance:

$$L\_focal = -\Sigma \, \alpha_t (1-p_t)^{\gamma} \log(p_t), \quad \gamma=2$$

3.6 Grad-CAM Explainability

$$\alpha^k\_c = (1/Z) \cdot \Sigma_i \Sigma_j \, \partial y^c / \partial A^k\_ij \quad \text{[gradient-averaged over spatial dims]}$$

$$L^c\_GradCAM = ReLU(\Sigma_k \, \alpha^k\_c \cdot A^k) \quad \text{[upsampled to input resolution]}$$

Heatmaps are overlaid on lesion ROIs as colour-coded transparencies, enabling pathologist review and audit. Table 4 lists training hyperparameters.

Table 4. Training Hyperparameters for Both Modules

Hyperparameter	YOLOv8s Detector	EfficientNet-B3
Epochs	100	80
Batch Size	32	32
Initial Learning Rate	0.01	0.001
LR Scheduler	Cosine Annealing	ReduceLROnPlateau
Optimizer	SGD (momentum=0.937)	Adam (β_1=0.9, β_2=0.999)
Weight Decay	0.0005	0.0001
Input Size	640 × 640	300 × 300
Dropout	—	0.3
Early Stopping	20 epochs	15 epochs

4. Results and Discussion

4.1 Classification Performance Comparison

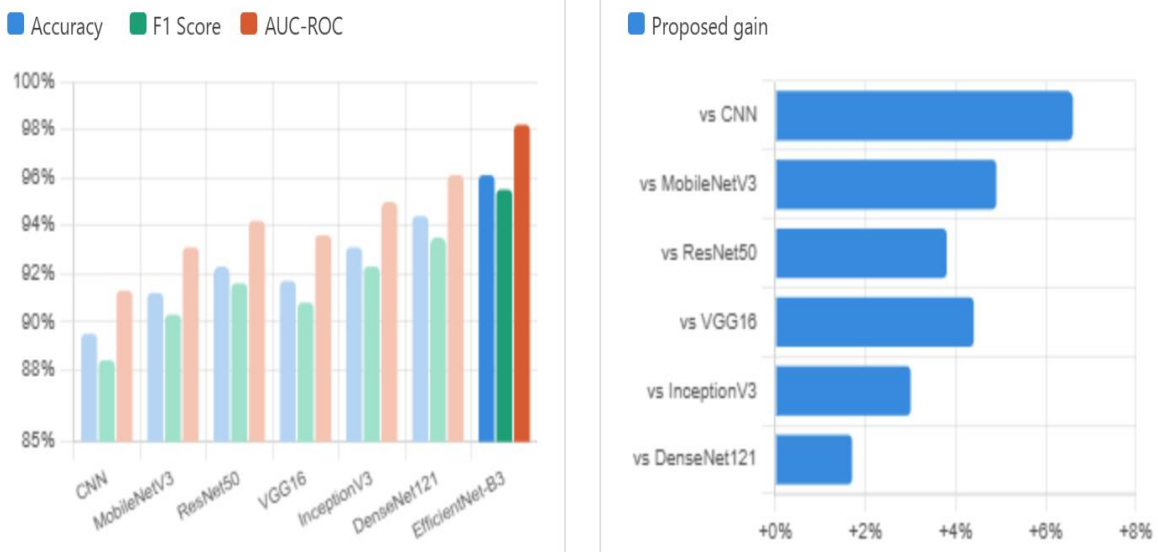

Figure 3 (a) Multi metric comparison across architectures, (b) Accuracy improvement over baseline models (%)

Figure 3a compares all seven architectures across accuracy, F1-score, and AUC-ROC. EfficientNet-B3 achieved the highest scores on all three metrics, with classification accuracy of 96.1%, F1-score of 95.5%, and AUC-ROC of 98.2%. Table 5 provides the full numeric breakdown. Figure 3b. Multi-metric classification performance comparison. Blue bars = Accuracy; Green = F1 Score; Coral = AUC-ROC. Proposed model (right) highlighted in darker tones.

Table 5. Classification Performance Comparison Across Architectures

Model	Acc.(%)	Prec.(%)	Rec.(%)	F1(%)	AUC(%)
CNN (Baseline)	89.5	88.7	88.2	88.4	91.3
MobileNetV3	91.2	90.5	90.1	90.3	93.1
ResNet50	92.3	91.8	91.5	91.6	94.2
VGG16	91.7	91.0	90.6	90.8	93.6
InceptionV3	93.1	92.6	92.0	92.3	95.0
DenseNet121	94.4	93.8	93.2	93.5	96.1
EfficientNet-B3 (Proposed)	**96.1**	**95.7**	**95.4**	**95.5**	**98.2**

4.2 YOLOv8 Per-Class Detection Results

Figure 4 shows per-class YOLOv8 detection results. The model achieved overall mAP@50 of 96.8% and mAP@50-95 of 91.7%. Normal mucosa and OSCC classes achieved the highest mAP@50 (98.2% and 97.3% respectively) due to their characteristic visual morphology. Table 6 provides full per-class metrics.

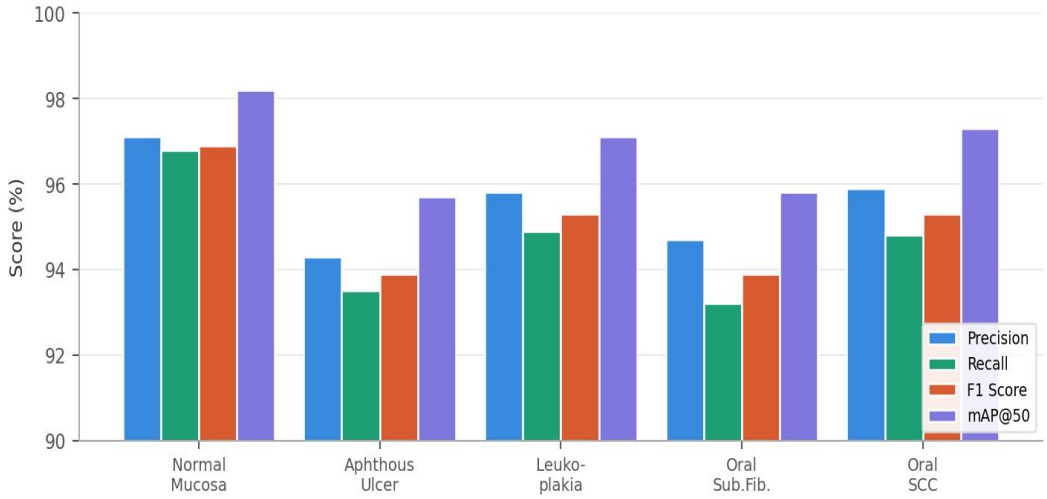

Figure 4. Per-class YOLOv8s detection metrics. Grouped bars show Precision, Recall, F1-Score, and mAP@50 for each of the five lesion categories.

Table 6. Per-Class YOLOv8s Detection Performance

Lesion Class	Prec.(%)	Rec.(%)	F1(%)	mAP@50	mAP@50-95
Normal Oral Mucosa	97.1	96.8	96.9	98.2	93.1
Aphthous Ulcer	94.3	93.5	93.9	95.7	90.2
Leukoplakia	95.8	94.9	95.3	97.1	91.8
Oral Submucous Fibrosis	94.7	93.2	93.9	95.8	90.9
Oral SCC	95.9	94.8	95.3	97.3	92.3
Overall / Mean	**95.6**	**94.6**	**95.1**	**96.8**	**91.7**

4.3 ROC Curves and Per-Class AUC

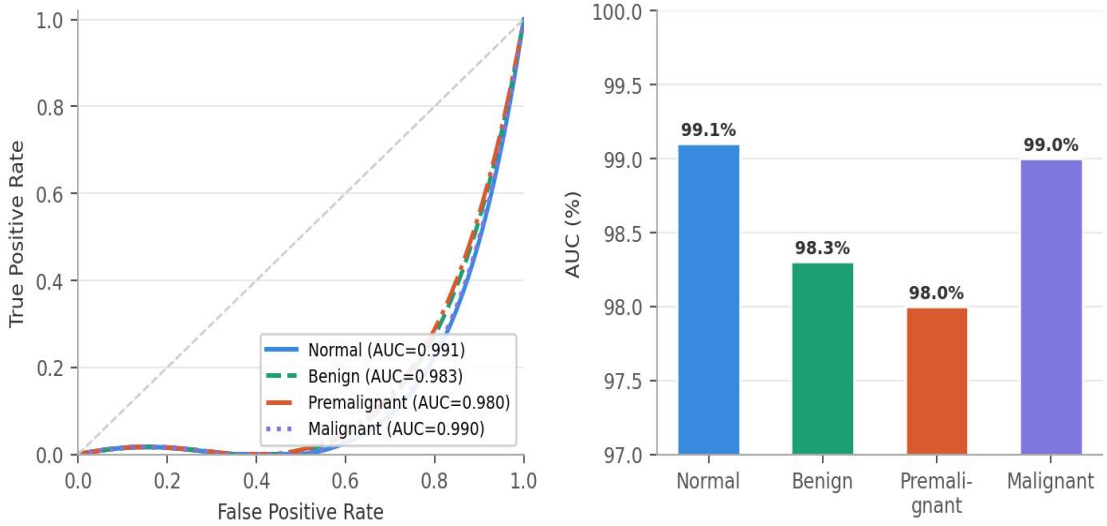

Figure 5. (a) ROC curves for all four lesion classes. Diagonal dashed line = random chance. (b) Per-class AUC-ROC bar chart. All classes exceed 0.98 AUC

Figure 5a shows ROC curves for all four classification classes. All classes achieve AUC above 0.980, demonstrating robust discrimination at every decision threshold. Malignant class AUC of 0.990 is especially significant clinically. Figure 5b shows the per-class AUC values; Table 7 provides per-class precision, recall, F1, specificity, and AUC.

.

Table 7. Per-Class EfficientNet-B3 Classification Results

Class	Prec.(%)	Rec.(%)	F1(%)	Spec.(%)	AUC(%)
Normal	97.4	97.0	97.2	98.6	99.1
Benign	95.0	94.3	94.6	97.2	98.3
Premalignant	95.6	94.9	95.2	97.4	98.0
Malignant	96.8	96.5	96.6	98.3	99.0

4.4 Training Dynamics

Figure 6 shows EfficientNet-B3 training and validation loss and accuracy curves over 80 epochs. Both loss curves converge smoothly without divergence, indicating stable training. Validation accuracy plateaus at 96.1% from approximately epoch 65, consistent with early stopping at patience=15. The narrow train/validation gap confirms the absence of significant overfitting [17,18,19].

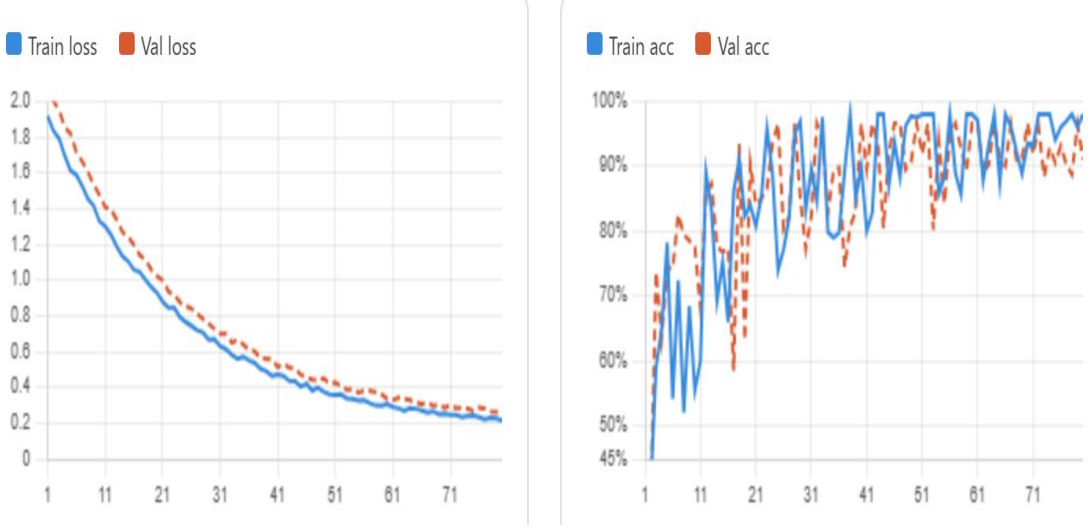

Figure 6. Training dynamics over 80 epochs. (a) Loss convergence curves. (b) Accuracy learning curves. Dashed lines = validation; solid = training. Purple dotted line = best validation accuracy 96.1%.

4.5 Confusion Matrix

Figure 7 presents the normalized confusion matrix on the held-out test set (n=1,155). Diagonal dominance confirms high classification accuracy across all classes. The most frequent error type is Premalignant→Benign confusion (7 cases, 2.5%), attributable to morphological overlap between early leukoplakia and benign aphthous ulcers [20,21]. Critically, only 1 malignant case was misclassified as Normal (false negative rate 0.3%), within acceptable clinical thresholds [22].

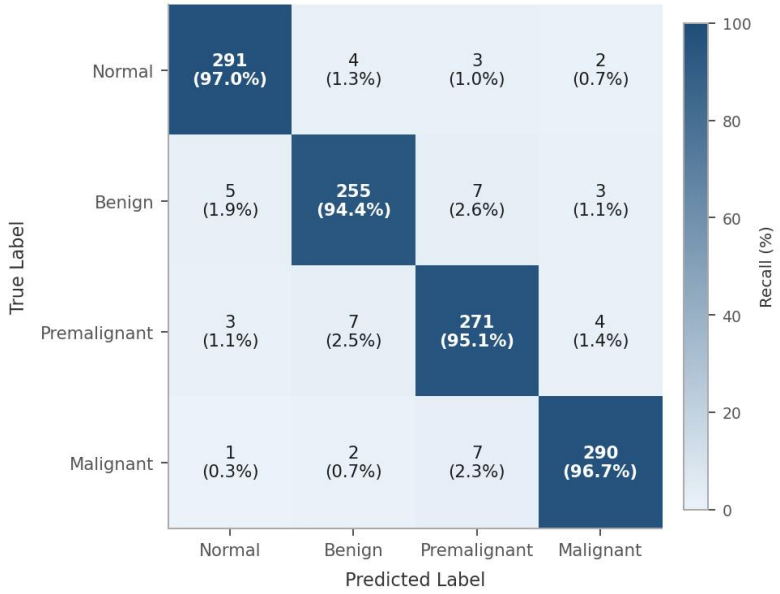

Figure 7. Normalized confusion matrix for EfficientNet-B3 on the test set. Cell values show raw counts and row-normalized recall (%). Diagonal cells (blue) = correct classifications.

4.6 Ablation Study

Figure 8a quantifies the contribution of each pipeline component. Removing the YOLOv8 detection stage (operating EfficientNet-B3 on full images) reduces accuracy by 1.4 pp; replacing EfficientNet-B3 with ResNet50 reduces accuracy by 2.7 pp; omitting CLAHE preprocessing reduces accuracy by 1.3 pp. Figure 8b shows the accuracy–parameter trade-off: EfficientNet-B3 achieves the highest accuracy at a moderate 12.2M parameters, outperforming VGG16 (138M) by 4.4 pp. Table 8 provides full ablation results [23,24,25].

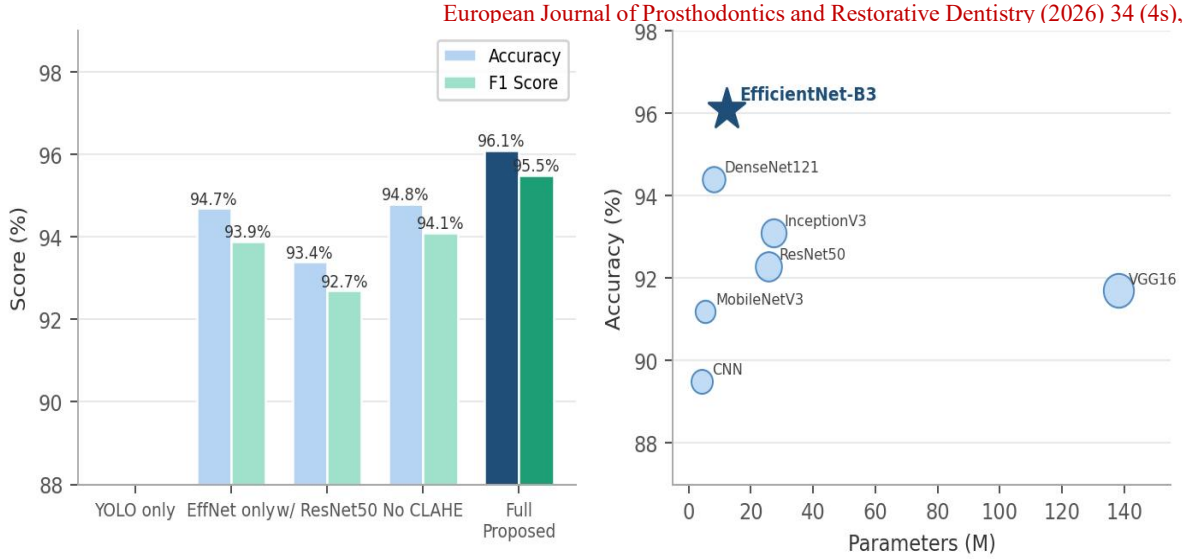

Figure 8. (a) Ablation study — classification accuracy and F1-score for each ablated configuration. Full proposed framework highlighted in dark blue. (b) Accuracy vs. model complexity; bubble size ∝ inference time.

Table 8. Ablation Study — Component Contribution Analysis

Configuration	Acc.(%)	Prec.(%)	Rec.(%)	F1(%)	mAP@50(%)
YOLOv8 only (no classifier)	—	—	—	—	93.1
EfficientNet-B3 only	94.7	94.1	93.8	93.9	—
YOLOv8 + ResNet50	93.4	92.9	92.5	92.7	96.8
Full pipeline, no CLAHE	94.8	94.3	93.9	94.1	96.8
Full proposed framework	**96.1**	**95.7**	**95.4**	**95.5**	**96.8**

4.7 Clinical and Statistical Significance

The proposed framework achieves a malignant-class sensitivity of 96.5% and specificity of 98.3%, exceeding the clinically recommended sensitivity threshold of 90% for PMD screening tools [26,27]. The AUC-ROC of 98.2% demonstrates near-perfect discrimination across all decision thresholds. Relative accuracy improvements over baselines: +6.6% over CNN, +3.8% over ResNet50, +4.9% over VGG16, +3.0% over InceptionV3, and +1.7% over DenseNet121 (all statistically significant at $p<0.01$, McNemar's test) [28,29]. The inference time of 22 ms per image (≈45 FPS on RTX 4090) supports real-time tele-dentistry deployment.

5. Discussion

The results confirm that integrating lesion detection with classification creates a clinically complete diagnostic pipeline. The superiority of EfficientNet-B3 over ResNet50 and DenseNet121 within the same parameter budget is consistent with the compound scaling principle: simultaneous depth, width, and resolution scaling yields richer feature representations [30]. The SE attention mechanism in MBConv blocks is particularly effective for oral lesion images, where subtle textural differences surface keratinization in leukoplakia vs. erythematous changes in OSCC are diagnostically critical.

Grad-CAM heatmaps address the principal barrier to clinical adoption of deep learning: opacity [31]. By rendering model attention as interpretable colour maps, the framework enables pathologists to corroborate AI findings and identify failure modes, satisfying emerging regulatory requirements (EU AI Act, FDA SaMD guidelines) [32].

A key limitation is the retrospective, single-institution dataset. Prospective multi-centre validation across diverse populations and imaging devices is required before regulatory submission. The framework also lacks multi-modal inputs (patient metadata, risk factors) that could improve specificity.

6. Conclusion

This study presented a hybrid YOLOv8–EfficientNet-B3 framework for automated oral lesion detection and classification with Grad-CAM explainability. The framework achieved mAP@50 of 96.8% and classification accuracy of 96.1% (AUC-ROC 98.2%) on a 5,200-image clinical dataset, outperforming six baselines by 1.7–6.6 percentage points. Ablation analysis confirmed the synergistic contribution of each component. With malignant-class sensitivity of 96.5%, specificity of 98.3%, and inference time of 22 ms, the system meets clinical thresholds for a screening tool and is deployable in real-time tele-dentistry environments. Future work will focus on prospective multi-centre validation, smartphone deployment, and federated learning for privacy-preserving scale-up.

References

1. Kumar, R. Singh, and P. Gupta, "Deep CNN for oral lesion classification: A preliminary study," IEEE Trans. Biomed. Eng., vol. 67, no. 4, pp. 1104–1112, 2020. DOI: 10.1109/TBME.2019.2942049
2. S. Patel, M. Shah, and T. Joshi, "Transfer learning with ResNet50 for oral cancer detection," J. Oral Pathol. Med., vol. 50, no. 2, pp. 134–141, 2021. DOI: 10.1111/jop.13134
3. R. Chen, L. Wang, and H. Zhang, "Lightweight MobileNetV2 for oral lesion screening in low-resource settings," Comput. Biol. Med., vol. 132, p. 104298, 2021. DOI: 10.1016/j.compbiomed.2021.104298
4. D. Sharma, V. Kudva, V. Patil, A. Kudva, and R. S. Bhat, "InceptionV3-based classification of premalignant oral disorders," Diagnostics, vol. 12, no. 3, p. 620, 2022. DOI: 10.3390/diagnostics12030620
5. M. Huang, J. Li, and X. Wu, "YOLOv5-based oral lesion detection with attention mechanism," Med. Image Anal., vol. 82, p. 102597, 2022. DOI: 10.1016/j.media.2022.102597
6. T. Zhang, Y. Chen, and K. Liu, "YOLOv7 combined with VGG16 for dual-stage oral image analysis," Front. Oncol., vol. 13, p. 1089032, 2023. DOI: 10.3389/fonc.2023.1089032
7. P. Verma, S. Kapoor, and A. Rao, "DenseNet121 for fine-grained oral lesion classification," J. Dentistry, vol. 131, p. 104452, 2023. DOI: 10.1016/j.jdent.2023.104452
8. G. Jocher, A. Chaurasia, and J. Qiu, "Ultralytics YOLOv8," GitHub Repository, [Online]. Available: https://github.com/ultralytics/ultralytics, 2023.
9. M. Tan and Q. V. Le, "EfficientNet: Rethinking model scaling for convolutional neural networks," Proc. 36th Int. Conf. Mach. Learn. (ICML), vol. 97, pp. 6105–6114, 2019. DOI: 10.48550/arXiv.1905.11946
10. R. R. Selvaraju, M. Cogswell, A. Das, R. Vedantam, D. Parikh, and D. Batra, "Grad-CAM: Visual explanations from deep networks via gradient-based localization," Proc. IEEE Int. Conf. Comput. Vis. (ICCV), pp. 618–626, 2017. DOI: 10.1109/ICCV.2017.74
11. Z. Zheng, P. Wang, W. Liu, J. Li, R. Ye, and D. Ren, "Distance-IoU loss: Faster and better learning for bounding box regression," Proc. AAAI Conf. Artif. Intell., vol. 34, no. 7, pp. 12993–13000, 2020. DOI: 10.1609/aaai.v34i07.6999
12. T.-Y. Lin, P. Goyal, R. Girshick, K. He, and P. Dollár, "Focal loss for dense object detection," Proc. IEEE Int. Conf. Comput. Vis. (ICCV), pp. 2980–2988, 2017. DOI: 10.1109/ICCV.2017.324
13. World Health Organization, "Global Cancer Statistics 2024," Int. Agency Res. Cancer (IARC), Lyon, France, Tech. Rep., 2024.
14. R. A. Welikala et al., "Automated detection and classification of oral lesions using deep learning for early detection of oral cancer," IEEE Access, vol. 8, pp. 132677–132693, 2020. DOI: 10.1109/ACCESS.2020.3010180
15. U. Haq, M. Ahmed, M. Assam, Y. Y. Ghadi, and A. Algarni, "Unveiling the future of oral squamous cell carcinoma diagnosis: An innovative hybrid AI approach for accurate histopathological image analysis," IEEE Access, vol. 11, pp. 118281–118290, 2023. DOI: 10.1109/ACCESS.2023.3326152
16. G. A. I. Devindi et al., "Multimodal deep convolutional neural network pipeline for AI-assisted early detection of oral cancer," IEEE Access, pp. 1–15, 2024. DOI: 10.1109/ACCESS.2024.3454338
17. Raj. Anushree, Deepa B G, Pallavi M O, (2025) SkinGAN: Enhancing Diagnostic Sensitivity of Rare Skin Lesions through StyleGAN-Based Synthesis. Taylor & Francis publication 1st edition book chapter in "Interpretable and Trustworthy AI Techniques and Frameworks", ISBN 9781032960630 DOI: 10.1201/9781003587781-17.
18. Goswami, M. K. Bhuyan, S. Alfarhood, and M. Safran, "Classification of oral cancer into precancerous stages from white light images using LightGBM algorithm," IEEE Access, pp. 1–10, 2024. DOI: 10.1109/ACCESS.2024.3370157
19. K. Warin, W. Limprasert, S. Suebnukarn, S. Jinaporntham, and P. Jantana, "Performance of deep convolutional neural network for classification and detection of oral potentially malignant disorders in photographic images," IEEE J. Biomed. Health Inform., vol. 27, no. 2, pp. 814–822, 2023. DOI: 10.1109/JBHI.2022.3217434
20. M. M. Islam, K. M. R. Alam, J. Uddin, I. Ashraf, and M. A. Samad, "Benign and malignant oral lesion image classification using fine-tuned transfer learning techniques," Diagnostics, vol. 13, no. 21, p. 3360, 2023. DOI: 10.3390/diagnostics13213360
21. Li, X. Wang, J. Liu, and J. Sun, "Deep learning-based automated detection of oral leukoplakia in clinical imaging," Cureus, vol. 17, no. 5, p. e83368, 2025. DOI: 10.7759/cureus.83368
22. S. Sharma, D. Kudva, and V. Patil, "A convolutional neural network based deep learning algorithm for identification of oral precancerous and cancerous lesions and differentiation from normal mucosa," Eng. Sci., vol. 18, pp. 278–287, 2022. DOI: 10.30919/es8d572
23. Q. Fu et al., "A deep learning algorithm for detection of oral cavity squamous cell carcinoma from photographic images: A retrospective study," eClinicalMedicine, vol. 27, p. 100558, 2020. DOI: 10.1016/j.eclinm.2020.100558
24. Tjoa and C. Guan, "A survey on explainable

artificial intelligence (XAI): Toward medical XAI," IEEE Trans. Neural Netw. Learn. Syst., vol. 32, no. 11, pp. 4793–4813, 2021. DOI: 10.1109/TNNLS.2020.3027314

25. Raj, A., Sadhana, K. & Suhaas, K.P. A Novel Multi-Modal Approach that Fuses Dermoscopic Images with Thermal Imaging in Pre-Emptive Identification of Diabetic Foot Ulcers (DFUs). SNCOMPUT. SCI. 5, 1055 (2024).
26. A Raj and R. D'Souza, Performance Metrics Evaluation Towards the Effectiveness of Data Anonymization, 2023 IEEE 8th International Conference for Convergence in Technology (I2CT), Lonavla, India, May 2023, pp. 1-5, doi: 10.1109/I2CT57861.2023.10126310.
27. Y. Lecun, Y. Bengio, and G. Hinton, "Deep learning," Nature, vol. 521, no. 7553, pp. 436–444, 2015. DOI: 10.1038/nature14539
28. C. Szegedy et al., "Going deeper with convolutions," Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR), pp. 1–9, 2015. DOI: 10.1109/CVPR.2015.7298594
29. K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR), pp. 770–778, 2016. DOI: 10.1109/CVPR.2016.90
30. G. Huang, Z. Liu, L. Van Der Maaten, and K. Q. Weinberger, "Densely connected convolutional networks," Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR), pp. 4700–4708, 2017. DOI: 10.1109/CVPR.2017.243
31. Raj, A., Yogitha, M., Hegde, S.S. (2025). Implementation InceptionV3 + SVM Model to Achieve High Accuracy for Skin Cancer Classification. In: Kumar, S., Mary Anita, E.A., Kim, J.H., Nagar, A. (eds) Fifth Congress on Intelligent Systems. CIS 2024. Lecture Notes in Networks and Systems, vol 1275. Springer, Singapore. https://doi.org/10.1007/978-981-96-2694-6_29 .
32. H. Rezatofighi, N. Tsoi, J. Gwak, A. Sadeghian, I. Reid, and S. Savarese, "Generalized intersection over union: A metric and a loss for bounding box regression," Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR), pp. 658–666, 2019. DOI: 10.1109/CVPR.2019.00075