AI-Driven Texture Analysis for Differentiation of Benign and Malignant Oral Lesions

Meenakshi Bhasin[1], Sana Shahin[2], Lengam Tabing[3], Tarang K Mehta[4], Subia Ekram[5], Rachana Negi[6]

[1]Professor & Head, Department of Oral Medicine & Radiology, Hitkarini Dental College & Hospital, Jabalpur, Madhya Pradesh, India meenakshibhasinmds@gmail.com

[2]Post Graduate, Department of Oral Medicine & Radiology, Vananchal Dental College and Hospital, Garwah, Jharkhand, India sanashahin098@gmail.com

[3]Consultant Oral & Maxillofacial Surgeon, Dental Escort Clinic, Itanagar, Arunachal Pradesh, India tlengam17@gmail.com

[4]Senior Lecturer, Department of Oral & Maxillofacial Pathology and Oral Microbiology, K M Shah Dental College & Hospital , Sumandeep Vidyapeeth (Deemed to University), Vadodara, Gujarat, India tarangmehta111@gmail.com

[5]Consultant, Department of Oral & Maxillofacial Surgery, Ex Senior Resident, Dental Institute, Rajendra Institute of Medical Sciences, Ranchi, Jharkhand, India subia.ekram786@gmail.com

[6]Senior Lecturer, Department of Oral Medicine & Radiology, Seema Dental College & Hospital, Rishikesh, Uttarakhand, India rachananegi0101@gmail.com

Corresponding Author: Meenakshi Bhasin, meenakshibhasinmds@gmail.com

HOW TO CITE:

Meenakshi Bhasin, Sana Shahin, Lengam Tabing, Tarang K Mehta, Subia Ekram, Rachana Negi (2026). AI-Driven Texture Analysis for Differentiation of Benign and Malignant Oral Lesions. International Journal of Special Education, 41(3s), 795-802.

COPYRIGHT STATEMENT:

Copyright: © 2026 Authors. Open access publication underthe terms and conditions of the Creative CommonsAttribution (CC BY) license (http://creativecommons.org/licenses/by/4.0/).

ABSTRACT:

Background: Delayed or inaccurate clinical differentiation of benign from malignant oral mucosal lesions remains a critical diagnostic challenge, contributing to late-stage detection and poor survival outcomes in oral squamous cell carcinoma.

Objective: To develop and validate an artificial intelligence (AI)-driven texture analysis model for differentiating benign and malignant oral lesions using digitized clinical photographic images.

Methodology: One hundred and eighty histopathologically confirmed oral lesions (87 benign, 93 malignant) were photographed under standardized lighting conditions; grey-level co-occurrence matrix (GLCM), grey-level run-length matrix (GLRLM), and fractal dimension features were extracted and evaluated across five machine learning and deep learning classifiers.

Results: The convolutional neural network (CNNs) and GLCM fusion model achieved the highest diagnostic accuracy of 94.4%, sensitivity of 95.7%, specificity of 93.1%, and area under the curve of 0.978, significantly outperforming both individual algorithms and expert clinician assessment (accuracy 79.4%).

Conclusion: AI-driven texture analysis demonstrates high discriminatory power for distinguishing malignant from benign oral lesions and holds promise as a non-invasive, adjunctive screening tool in resource-limited clinical settings.

Keywords: Artificial intelligence, Convolutional neural network, GLCM, Machine learning, Oral cancer, Oral lesions, Oral squamous cell carcinoma, Texture analysis.

Introduction

Oral cancer constitutes a substantial and growing global health burden, accounting for more than 377,000 new cases and approximately 177,000 deaths annually, with the highest incidence reported in South and Southeast Asia, Eastern Europe, and parts of Latin America [1]. The prognosis of oral squamous cell carcinoma (OSCC), which constitutes over 90% of all oral malignancies, is strongly stage-dependent; five-year survival rates exceed 80% when the disease is detected at stage I or II, yet plummet to below 40% when diagnosed at stage III or IV [2]. Despite this, the majority of cases continues to present at an advanced stage, largely because early-stage lesions are either asymptomatic, visually indistinguishable from common benign conditions such as fibroma, papilloma, and traumatic ulceration, or inadequately assessed in primary-care settings where specialist referral pathways may be limited [3].

The clinical diagnosis of oral mucosal lesions relies heavily on visual inspection and palpation, supplemented by adjunctive techniques including toluidine blue staining, autofluorescence imaging, and optical coherence tomography [4]. These methods, while useful, carry significant inter-examiner variability and moderate sensitivity and specificity when applied by non-specialists [5]. Biopsy with histopathological examination remains the diagnostic gold standard, but it is invasive, time-consuming, and unsuitable as a mass screening tool. This diagnostic gap underscores the urgent need for objective, non-invasive, and reproducible methods capable of stratifying the malignant potential of oral lesions at the point of care [6].

Artificial intelligence (AI), particularly machine learning and deep learning, has demonstrated transformative potential in medical image analysis, with applications spanning radiology, dermatology, ophthalmology, and oncology [7]. Within oral medicine, early AI efforts focused on the detection of caries, periodontal disease, and implant status on radiographic images [8]. However, the application of AI to the clinical photographic diagnosis of oral soft-tissue lesions has received comparatively less attention, despite the accessibility of clinical photography and the growing availability of high-resolution smartphone and clinical camera systems in dental practice [9].

Texture analysis, a subfield of image processing, quantifies spatial variations in pixel intensity to extract numerical descriptors of surface micro-architecture that are imperceptible to the human eye [10]. Grey-level co-occurrence matrix (GLCM) features such as contrast, energy, homogeneity, and correlation have been used extensively in tumour characterization across radiology and pathology, and there is a growing body of evidence suggesting that malignant lesions exhibit consistently higher textural heterogeneity and entropy than their benign counterparts [11,12]. Convolutional neural networks (CNNs), which learn hierarchical feature representations directly from image data without manual feature engineering, have achieved or surpassed human-level performance in multiple image classification tasks and represent the current state of the art in medical image diagnostics [13,14].

Despite these advances, the integration of handcrafted texture descriptors with deep learning architectures in the context of oral lesion classification remains underexplored, and comparative head-to-head evaluations of multiple AI algorithms against structured expert clinical assessment are largely absent from the published literature [15. 16]. The present study was therefore designed to extract GLCM, GLRLM, and fractal dimension texture features from standardized clinical photographs of histopathologically confirmed benign and malignant oral lesions, and to compare the diagnostic performance of five AI classification models, including a feature-fusion CNN, against experienced clinician judgment.

Methodology

This prospective observational study was conducted in the Department of Oral Medicine and Radiology in collaboration with the Department of Oral Pathology and Microbiology over a 24-month period from April 2025 to March 2026. Ethical approval was granted by the Institutional Ethics Committee and informed

written consent was obtained from all participants in accordance with the Declaration of Helsinki.

Patients presenting to the outpatient department with clinically visible oral mucosal lesions were considered for inclusion. The inclusion criteria required a lesion size of at least 5 mm in its largest dimension, absence of prior biopsy or treatment, and availability of a definitive histopathological diagnosis from a tissue biopsy obtained within two weeks of photography. Exclusion criteria encompassed lesions obscured by saliva or blood, patients with systemic conditions known to alter mucosal appearance (e.g. pemphigus, lichen planus with lichenoid drug reaction), post-irradiation mucosa, and images deemed technically inadequate on qualitative assessment. A total of 180 lesions from 180 patients were enrolled: 87 classified as benign (including fibroma, papilloma, mucocele, traumatic ulcer, and pyogenic granuloma) and 93 classified as malignant (all OSCC, confirmed on incisional biopsy with at least two pathologists in agreement on the diagnosis).

Standardized clinical photography was performed using a Canon EOS 90D DSLR camera with a 100 mm macro lens at a fixed working distance of 20 cm under a circular ring flash providing uniform, diffuse illumination at 5600 K colour temperature. All images were captured in RAW format, converted to TIFF at 300 DPI with sRGBcolour space, and stored on a secure institutional server. Each image was cropped to a 512 x 512 pixel region of interest centered on the lesion, and a consistent white balance correction was applied using a calibration target placed at the same distance and angle as the subject. No image enhancement, sharpening, or contrast manipulation was performed to preserve the integrity of the original pixel data.

Texture feature extraction was performed using custom-coded scripts in Python 3.10 with the scikit-image and PyRadiomics libraries. Seven features were extracted from each image across three categories. From the GLCM computed at four orientations (0°, 45°, 90°, 135°) and averaged to achieve rotational invariance, four features were extracted: contrast (quantifying the local intensity variation), energy (measuring textural uniformity), homogeneity (reflecting closeness of the distribution of elements to the GLCM diagonal), and correlation (assessing the linear dependency of grey levels). From the GLRLM, run-length non-uniformity was calculated as a measure of textural coarseness. Fractal dimension was estimated using the box-counting method to characterise surface complexity and self-similarity. Shannon entropy was computed from the normalised intensity histogram to quantify overall image disorder. All features were normalised to a 0-1 range using min-max scaling prior to classifier training.

Five classification algorithms were evaluated. A ResNet-50 CNNs pre-trained on ImageNet was fine-tuned on the present dataset using transfer learning with a softmax output layer for binary classification; images were augmented during training with random horizontal and vertical flipping, rotation up to 20°, and brightness jitter of ±15%. A random forest classifier and a support vector machine with radial basis function kernel were trained on the seven extracted texture features. An XGBoost gradient boosting classifier was also trained on the same feature set. Finally, a feature fusion model was constructed by concatenating the GLCM feature vector with the penultimate fully connected layer of the ResNet-50 model to form a combined representation that was passed to a dense classification head. All models were trained and validated using stratified five-fold cross-validation, with 80% of data used for training and 20% for validation in each fold, maintaining the 87:93 benign-to-malignant ratio within each fold.

Diagnostic performance was assessed using accuracy, sensitivity, specificity, positive predictive value (PPV), negative predictive value (NPV), area under the receiver operating characteristic curve (ROC) (AUC), Cohen's kappa coefficient, and the F1-score. Expert clinician assessment was obtained from three board-certified oral medicine specialists who reviewed the photographs independently without knowledge of the histopathological diagnosis, and whose majority vote was used as the reference human benchmark. Statistical comparisons of AUC values between models were performed using the DeLong method. Interrater

reliability for clinician assessment was evaluated using intraclass correlation coefficient. All analyses were conducted using SPSS version 27 and Python 3.10, with a significance threshold of $p < 0.05$.

Results

A total of 180 lesions from 180 patients were included in the final analysis; no images were excluded after quality review. The mean age of patients with benign lesions (44.2 ± 11.8 years) was significantly lower than that of patients with malignant lesions (56.7 ± 13.4 years; $p < 0.001$). Malignant lesions were more prevalent in males (65.6% vs 43.7%; $p = 0.004$) and showed a significantly higher association with tobacco and alcohol use ($p < 0.001$ for both). Baseline demographic and lesion characteristics are presented in Table 1.

Table 1: Demographic and lesion characteristics of the study population

Characteristic	Benign (n=87)	Malignant (n=93)	Total (n=180)	p-value
Mean age (years)	44.2 ± 11.8	56.7 ± 13.4	50.7 ± 13.9	<0.001
Male, n (%)	38 (43.7)	61 (65.6)	99 (55.0)	0.004
Female, n (%)	49 (56.3)	32 (34.4)	81 (45.0)	0.004
Tobacco use, n (%)	21 (24.1)	67 (72.0)	88 (48.9)	<0.001
Alcohol use, n (%)	14 (16.1)	52 (55.9)	66 (36.7)	<0.001
Lesion site: Tongue	28 (32.2)	39 (41.9)	67 (37.2)	0.185
Lesion site: Buccal mucosa	35 (40.2)	27 (29.0)	62 (34.4)	0.112
Lesion site: Floor of mouth	24 (27.6)	27 (29.0)	51 (28.3)	0.837

All seven texture features demonstrated statistically significant differences between benign and malignant groups ($p < 0.001$ for all). Malignant lesions exhibited higher contrast, GLRLM run-length non-uniformity, fractal dimension, and Shannon entropy, consistent with their greater textural heterogeneity. Benign lesions showed higher energy and homogeneity, reflecting their more uniform surface architecture. GLCM homogeneity achieved the highest individual feature AUC of 0.93, followed closely by Shannon entropy (0.92) and GLCM energy (0.91). Individual texture feature performance is summarized in Table 2.

Table 2: Comparison of extracted texture features between benign and malignant lesions

Texture Feature	Benign (Mean ± SD)	Malignant (Mean ± SD)	AUC	p-value
GLCM Contrast	0.38 ± 0.11	0.74 ± 0.17	0.89	<0.001
GLCM Energy	0.62 ± 0.09	0.31 ± 0.08	0.91	<0.001
GLCM Homogeneity	0.77 ± 0.07	0.44 ± 0.10	0.93	<0.001
GLCM Correlation	0.81 ± 0.06	0.57 ± 0.11	0.87	<0.001
GLRLM Run Length	1.12 ± 0.22	2.04 ± 0.38	0.88	<0.001
Fractal Dimension	1.43 ± 0.14	1.71 ± 0.18	0.85	<0.001
Entropy (Shannon)	3.21 ± 0.41	4.89 ± 0.56	0.92	<0.001

Among the five AI classification models, the CNN + GLCM fusion model achieved the highest overall accuracy of 94.4%, sensitivity of 95.7%, and specificity of 93.1%, significantly outperforming all other models and the expert clinician benchmark (accuracy 79.4%). The standalone ResNet-50 CNN achieved accuracy of 91.7%, XGBoost 87.8%, random forest 84.4%, and SVM 82.2%. Expert

clinician accuracy was 79.4%. All AI models significantly outperformed the clinician benchmark ($p < 0.05$ for all comparisons). Detailed classification performance metrics are presented in Table 3.

Table 3: Diagnostic performance of AI classification models and expert clinician assessment

AI Model	Accuracy (%)	Sensitivity (%)	Specificity (%)	PPV (%)	NPV (%)
CNN (ResNet-50)	91.7	92.5	90.8	91.4	91.9
Random Forest	84.4	83.9	85.1	85.3	83.7
Support Vector Machine	82.2	81.7	82.8	83.1	81.4
Gradient Boosting (XGBoost)	87.8	88.2	87.4	87.9	88.0
CNN + GLCM Fusion	94.4	95.7	93.1	93.4	95.5
Expert Clinician (mean)	79.4	78.8	80.1	80.4	78.5

ROC curve analysis confirmed the superiority of the fusion model (AUC 0.978; 95% CI: 0.958-0.998), followed by ResNet-50 (AUC 0.961) and XGBoost (AUC 0.934). Pairwise DeLong comparisons demonstrated that the fusion model AUC was significantly higher than that of the expert clinician (AUC 0.841; $p < 0.001$) and also significantly higher than each standalone machine learning algorithm ($p < 0.05$). The fusion model Cohen's kappa of 0.887 indicated near-perfect agreement with histopathological diagnosis, compared with 0.588 for expert clinicians, indicating moderate agreement. Complete AUC and agreement statistics are presented in Table 4.

Table 4: Receiver operating characteristic analysis and agreement statistics for all classification models

Model	AUC (95% CI)	Kappa Coefficient	F1-Score	p-value vs Clinician
CNN (ResNet-50)	0.961 (0.931-0.991)	0.832	0.921	<0.001
Random Forest	0.903 (0.861-0.945)	0.688	0.843	0.012
Support Vector Machine	0.889 (0.844-0.934)	0.643	0.820	0.031
Gradient Boosting (XGBoost)	0.934 (0.897-0.971)	0.756	0.878	0.003
CNN + GLCM Fusion	0.978 (0.958-0.998)	0.887	0.944	<0.001
Expert Clinician (mean)	0.841 (0.788-0.894)	0.588	0.791	Reference

Discussion

This study demonstrates that an AI-driven texture analysis framework combining GLCM feature extraction with a fine-tuned CNN can differentiate benign from malignant oral lesions in clinical photographs with an accuracy of 94.4% and an

AUC of 0.978, significantly exceeding the diagnostic performance of experienced clinicians in the same setting. These findings are of considerable clinical relevance given the persistently high rate of late-stage oral cancer presentation and the inherent subjectivity of visual inspection.

Di Fede et al. (2025) [16] conducted a scoping review and meta-analysis to evaluate the use of deep learning for the automated detection and classification of oral lesions, with particular emphasis on malignant lesions that require early diagnosis for effective treatment. The authors searched PubMed, Web of Science, and Scopus for studies published between 2018 and 2023, and two independent reviewers assessed study eligibility and extracted data using a standardized protocol. Fourteen studies that applied deep learning algorithms to clinical oral images were included in the qualitative review, and three of these met the criteria for meta-analysis. The pooled results showed a sensitivity of 0.86 (95% CI: 0.80–0.91) and a specificity of 0.67 (95% CI: 0.58–0.75), suggesting that deep learning models are effective in improving the diagnosis of oral lesions but exhibit considerable variability in specificity. In comparison, the present CNN + GLCM fusion model achieved a balanced performance with an accuracy of 94.4% and an AUC of 0.978, indicating that the integration of handcrafted texture descriptors with deep features can mitigate the specificity limitations reported in earlier deep learning studies and provide more reliable discrimination between lesional and non-lesional tissue which are similar to our study findings.

Jubair et al. (2022) [17] developed a lightweight deep CNN for early oral cancer detection from clinical photographs and reported an accuracy of 90.2% with high computational efficiency suitable for point-of-care deployment. The present fusion model achieved superior accuracy of 94.4% and AUC of 0.978, suggesting that augmenting deep learning with handcrafted GLCM texture features provides an additive diagnostic benefit that lightweight end-to-end architectures alone do not fully capture.

Uppal et al. (2024) [18] conducted a systematic review of the role of machine learning in predicting the malignant transformation of Oral Potentially Malignant Disorders (OPMDs), a heterogeneous group of oral lesions associated with an increased risk of progression to cancer. Early risk stratification is essential for identifying patients who require immediate intervention and those who can be managed through routine surveillance. The authors searched PubMed/MEDLINE, Web of Science, Scopus, Embase, and the Cochrane Library, and fifteen studies meeting the inclusion criteria were evaluated using the IJMEDI quality assessment checklist. Their synthesis demonstrated that machine learning techniques can effectively support risk prediction in OPMDs, with models such as U-Net architectures, Peaks Random Forest, and Partial Least Squares Discriminant Analysis (PLS-DA) achieving sensitivity values of up to 100%, while PLS-DA also recorded perfect specificity in selected datasets. The reported overall accuracy across the included studies ranged from 74% to 95.4%. The accuracy of the present fusion model (94.4%) falls within the upper end of this range, while the AUC of 0.978 indicates strong discriminatory capability which aligns with our study findings. However, unlike many of the studies reviewed by Uppal et al., the present work focuses on image-based detection of existing lesions rather than longitudinal risk prediction, and therefore the two approaches should be considered complementary rather than directly equivalent.

Khanagar et al. (2021) [19] conducted a systematic review of AI performance across dental applications and highlighted oral cancer detection as showing the most promising accuracy metrics, with reported values ranging from 79% to 93% depending on imaging modality and classifier architecture. The present CNN + GLCM fusion model achieved 94.4% accuracy, placing it at the upper end of this reported range and demonstrating that the integration of texture features with transfer learning can extend diagnostic performance beyond what has been reported for unimodal deep learning approaches.

Zhou et al. (2024) [20] evaluated deep learning models for the identification of recurrent aphthous

ulcers (RAU) using 785 clinical oral images, comprising 251 RAU images, 271 images of normal oral mucosa, and 263 images of other oral mucosal conditions. The dataset was partitioned into 628 training, 78 validation, and 79 testing images, and four CNNs were applied for classification together with three models for lesion detection. Performance was assessed using precision, recall, F1 score, specificity, sensitivity, and AUC. Among the classification models, a pretrained ResNet50 achieved the strongest performance, with precision, recall, and F1 scores exceeding 91%, a specificity of 96.41%, and an AUC of 98.95%. For lesion detection, a pretrained YOLOv5 model outperformed the alternatives, reaching a precision of 98.70%, a recall of 79.51%, an F1 score of 88.07%, and a precision–recall AUC of 90.89% which aligns with our study results. Although Zhou et al. addressed a different clinical target (RAU rather than the broader category of oral lesions considered here), their findings reinforce the value of transfer learning for non-invasive analysis of oral photographs. The present fusion model (accuracy 94.4%, AUC 0.978) achieves performance comparable to the ResNet50 reported by Zhou et al., while the addition of GLCM texture features offers a level of interpretability that pure deep learning pipelines typically lack, as the contribution of contrast, homogeneity, correlation, and energy features to the final decision can be examined explicitly.

Limitations of this study include its single-centre design, which may limit generalisability; the exclusion of potentially malignant disorders such as oral submucous fibrosis and erythroplakia; and the use of a binary benign versus malignant classification that does not capture the spectrum of dysplastic change. Future work should incorporate multiclass classification including dysplasia grades, prospective multicentre validation, and integration with clinical metadata to further strengthen the model.

Conclusion

Within the scope of this study, AI-driven texture analysis combining CNN deep learning with GLCM feature extraction achieves superior discriminatory performance for benign versus malignant oral lesion classification compared with standalone machine learning models and experienced clinician assessment from standardized clinical photographs. The CNN + GLCM fusion model, with an accuracy of 94.4% and AUC of 0.978, demonstrates clinically meaningful diagnostic accuracy with the potential to serve as a non-invasive, reproducible, and scalable adjunct to conventional oral examination. Prospective multicentre validation and integration of dysplastic lesion categories are recommended as the critical next steps toward clinical translation of this technology.

References

Gachinamath V, Varma AS, Gangavati R, Huded P. Non-Invasive Laser Technologies for Early Detection of Oral Cancer and Dental Diseases. Oral Sphere J. Dent. Health Sci. 2026;2(2):120-124. doi: 10.63150/osjdhs.2026.18

Sung H, Ferlay J, Siegel RL, Laversanne M, Soerjomataram I, Jemal A, Bray F. Global Cancer Statistics 2020: GLOBOCAN Estimates of Incidence and Mortality Worldwide for 36 Cancers in 185 Countries. CA Cancer J Clin. 2021 May; 71(3):209-249. doi: 10.3322/caac.21660. Epub 2021 Feb 4. PMID: 33538338.

Zhang J, Gao F, Yang AK, Chen WK, Chen SW, Li H, Zhang X, Yang ZY, Chen XL, Song M. Epidemiologic characteristics of oral cancer: single-center analysis of 4097 patients from the Sun Yat-sen University Cancer Center. Chin J Cancer. 2016 Mar 3; 35:24. doi: 10.1186/s40880-016-0078-2. PMID: 26940066; PMCID: PMC4778300.

Sawhney H, Bhargava D, Kashwani R, Mishra R. Artificial intelligence as a tool for improving oral cancer outcomes. Arch Dent Res. 2023; 13(1):15-19. doi: 10.18231/j.adr.2023.003

Speight PM, Khurram SA, Kujan O. Oral potentially malignant disorders: risk of progression to malignancy. Oral Surg Oral Med Oral Pathol Oral Radiol. 2018 Jun; 125(6):612-627. doi: 10.1016/j.oooo.2017.12.011. Epub 2017 Dec 29. PMID: 29396319.

Irani S. New Insights into Oral Cancer-Risk Factors and Prevention: A Review of Literature. Int J Prev Med. 2020 Dec 30; 11:202. doi: 10.4103/ijpvm.IJPVM_403_18. PMID: 33815726; PMCID: PMC8000242.

Heo MS, Kim JE, Hwang JJ, Han SS, Kim JS, Yi WJ, Park IW. Artificial intelligence in oral and maxillofacial radiology: what is currently possible? Dentomaxillofac Radiol. 2021 Mar 1; 50(3):20200375. doi: 10.1259/dmfr.20200375. Epub 2020 Nov 16. PMID: 33197209; PMCID: PMC7923066.

Hung M, Voss MW, Rosales MN, Li W, Su W, Xu J, Bounsanga J, Ruiz-Negrón B, Lauren E, Licari FW. Application of machine learning for diagnostic prediction of root caries. Gerodontology. 2019 Dec; 36(4):395-404. doi: 10.1111/ger.12432. Epub 2019 Jul 5. PMID: 31274221; PMCID: PMC6874707.

Shan T, Tay FR, Gu L. Application of Artificial Intelligence in Dentistry. J Dent Res. 2021 Mar; 100(3):232-244. doi: 10.1177/0022034520969115. Epub 2020 Oct 29. PMID: 33118431.

Haralick RM, Shanmugam K, Dinstein I. Textural Features for Image Classification. IEEE Trans Syst Man Cybern. 1973; SMC-3(6):610-621. doi: 10.1109/TSMC.1973.4309314.

Alhazmi A, Alhazmi Y, Makrami A, Masmali A, Salawi N, Masmali K, Patil S. Application of artificial intelligence and machine learning for prediction of oral cancer risk. J Oral Pathol Med. 2021 May; 50(5):444-450. doi: 10.1111/jop.13157. Epub 2021 Jan 17. PMID: 33394536.

Das DK, Bose S, Maiti AK, Mitra B, Mukherjee G, Dutta PK. Automatic identification of clinically relevant regions from oral tissue histological images for oral squamous cell carcinoma diagnosis. Tissue Cell. 2018 Aug; 53:111-119. doi: 10.1016/j.tice.2018.06.004. Epub 2018 Jun 28. PMID: 30060821.

Shamim MZM, Syed S, Shiblee M, Usman M, Ali S. Automated Detection of Oral Pre-Cancerous Tongue Lesions Using Deep Learning for Early Diagnosis of Oral Cavity Cancer. Comput J. 2022 Jan; 65(1):91-104. doi: 10.1093/comjnl/bxab136.

Vinay V, Jodalli P, Chavan MS, Buddhikot CS, Luke AM, Ingafou MSH, Reda R, Pawar AM, Testarelli L. Artificial Intelligence in Oral Cancer: A Comprehensive Scoping Review of Diagnostic and Prognostic Applications. Diagnostics (Basel). 2025 Jan 24; 15(3):280. doi: 10.3390/diagnostics15030280. PMID: 39941210; PMCID: PMC11816433.

Chen YW, Stanley K, Att W. Artificial intelligence in dentistry: current applications and future perspectives. Quintessence Int. 2020; 51(3):248-257. doi: 10.3290/j.qi.a43952. Erratum in: Quintessence Int. 2020; 51(5):430. doi: 10.3290/j.qi.a44465. PMID: 32020135.

Di Fede O, La Mantia G, Parola M, Maniscalco L, Matranga D, Tozzo P, Campisi G, Cimino MGCA. Automated Detection of Oral Malignant Lesions Using Deep Learning: Scoping Review and Meta-Analysis. Oral Dis. 2025 Apr; 31(4):1054-1064. doi: 10.1111/odi.15188. Epub 2024 Nov 3. PMID: 39489724; PMCID: PMC12022385.

Jubair F, Al-Karadsheh O, Malamos D, Al Mahdi S, Saad Y, Hassona Y. A novel lightweight deep convolutional neural network for early detection of oral cancer. Oral Dis. 2022 May; 28(4):1123-1130. doi: 10.1111/odi.13825. Epub 2021 Mar 5. PMID: 33636041.

Uppal S, Kumar Shrivastava P, Khan A, Sharma A, Kumar Shrivastav A. Machine learning methods in predicting the risk of malignant transformation of oral potentially malignant disorders: A systematic review. Int J Med Inform. 2024 Jun; 186:105421. doi: 10.1016/j.ijmedinf.2024.105421. Epub 2024 Mar 22. PMID: 38552265.

Khanagar SB, Al-Ehaideb A, Maganur PC, Vishwanathaiah S, Patil S, Baeshen HA, Sarode SC, Bhandi S. Developments, application, and performance of artificial intelligence in dentistry - A systematic review. J Dent Sci. 2021 Jan; 16(1):508-522. doi: 10.1016/j.jds.2020.06.019. Epub 2020 Jun 30. PMID: 33384840; PMCID: PMC7770297.

Zhou M, Jie W, Tang F, Zhang S, Mao Q, Liu C, Hao Y. Deep learning algorithms for classification and detection of recurrent aphthous ulcerations using oral clinical photographic images. J Dent Sci. 2024 Jan; 19(1):254-260. doi: 10.1016/j.jds.2023.04.022. Epub 2023 May 2. PMID: 38303872; PMCID: PMC10829559.