IOSR Journal of Dental and Medical Sciences (IOSR-JDMS)
e-ISSN: 2279-0853, p-ISSN: 2279-0861.Volume 25, Issue 7 Ser. 1 (July. 2026), PP 111-117
www.iosrjournals.org

Artificial intelligence in oral lesion diagnosis

Dr. Vishal Kumar Tiwary[1], Dr. Kavita Nitish Garg[2], Dr. Chetan Sharma[3], Dr. Amit Thahriani[4]

[1]PG student (final year) Department of Oral and Maxillofacial Pathology, Atal Bihari Vajpayee medical University Lucknow. Chandra dental college and Hospital barabanki, Lucknow UP, India
[2]Professor and Head, Department of Oral & Maxillofacial Pathology, Chandra dental college and Hospital barabanki, Lucknow UP
[3]Reader, Department of Oral and Maxillofacial Pathology
Chandra Dental College and Hospital Barabanki, Lucknow UP
[4]Professor, Department of Oral and Maxillofacial Pathology
Chandra dental College and Hospital Barabanki Lucknow UP.
***Corresponding Author:** Amit Thahriani*

Abstract
***Background:** Oral mucosal lesions require timely diagnosis to improve outcomes. Artificial intelligence (AI) has emerged as a promising adjunct to conventional histopathology.*
***Aim:** To evaluate the diagnostic performance of an AI-based deep learning model in detecting oral lesions compared with histopathological diagnosis.*
***Methods:** A retrospective pilot study of 18 patients was conducted using standardized clinical photographs. AI predictions were benchmarked against histopathological findings. Diagnostic metrics including accuracy, sensitivity, specificity, predictive values, Cohen's kappa, and ROC analysis were calculated.*
***Results:** The AI system achieved an accuracy of 88.9% (16/18 cases), sensitivity of 90.9%, specificity of 85.7%, PPV of 90.9%, and NPV of 85.7%. The ROC curve demonstrated excellent discriminatory ability (AUC = 0.93). Agreement with histopathology was substantial (κ = 0.76).*
***Conclusion:** AI-based diagnosis demonstrated high accuracy and strong concordance with histopathology, supporting its potential role in early detection and clinical decision-making.*
***Keywords:** Artificial intelligence, oral lesions, deep learning, diagnostic accuracy, ROC curve, histopathology*

I. Introduction

Oral mucosal lesions encompass a broad spectrum of benign, inflammatory, infectious, potentially malignant, and malignant disorders that require timely and accurate diagnosis to improve patient outcomes. Oral squamous cell carcinoma (OSCC) accounts for nearly 90% of oral malignancies, and delayed diagnosis remains a major contributor to poor prognosis and reduced survival rates. Conventional diagnosis relies on clinical examination, adjunctive imaging, histopathological confirmation, and clinician expertise. However, diagnostic variability, limited access to specialists, and delayed referrals may compromise early lesion detection. Recent advances in artificial intelligence (AI), particularly machine learning (ML) and deep learning (DL), have demonstrated considerable potential to enhance the accuracy, consistency, and efficiency of oral lesion diagnosis by analyzing clinical photographs, radiographic images, and histopathological data [1].

Recent developments in deep learning have shown that AI-based image analysis can successfully identify oral and periodontal diseases using routine clinical photographs. Tao et al. developed a deep learning photo-processing system for periodontal screening that achieved high diagnostic performance using intraoral images, highlighting the feasibility of AI-assisted photographic assessment in routine dental practice. Their findings demonstrated that automated image interpretation could support clinicians by improving screening efficiency and reducing diagnostic variability, thereby providing a foundation for AI applications across a wider range of oral lesions [1].

Beyond lesion detection, AI has also demonstrated prognostic capabilities in oral cancer management. Kang et al. developed deep learning models capable of predicting overall survival in patients with oral cavity squamous cell carcinoma using clinicopathological variables. Their study illustrated that AI algorithms can integrate complex clinical datasets to generate individualized prognostic predictions, supporting precision medicine and personalized treatment planning. Such predictive models may assist clinicians in identifying high-risk patients who require intensive monitoring and tailored therapeutic strategies [2].

The growing interest in AI for oral oncology has been reinforced by comprehensive reviews emphasizing its potential in cancer detection. Vats et al. summarized multiple AI techniques, including convolutional neural networks (CNNs), artificial neural networks, and support vector machines, demonstrating their ability to classify oral lesions with high diagnostic accuracy. The review highlighted that AI-assisted analysis can reduce observer variability and facilitate earlier identification of suspicious lesions, although further validation through large multicenter studies remains necessary before widespread clinical implementation [3].

The concept of AI-driven precision dentistry has further expanded with recent perspectives advocating its integration into oral cancer screening programs. Kumari et al. emphasized that advances in deep learning, image processing, and predictive analytics have created opportunities for precision diagnostics by enabling automated recognition of subtle clinical changes that may be overlooked during conventional examination. The authors also discussed current challenges, including data standardization, ethical considerations, algorithm transparency, and the need for diverse datasets to ensure generalizable AI performance across different populations [4].

Mobile health technologies have significantly broadened the accessibility of AI-assisted screening. Song et al. developed a smartphone-based deep learning platform capable of classifying oral cancer using clinical photographs acquired at the point of care. Their mobile AI system demonstrated high diagnostic performance while allowing non-specialist healthcare workers to perform preliminary screening in remote and resource-limited settings. Such portable AI solutions have the potential to improve early detection rates and reduce disparities in access to specialist oral healthcare services [5].

Further evidence supporting AI applications in oral oncology was provided by Tobias et al., who comprehensively reviewed existing artificial intelligence technologies for oral cancer diagnosis. Their review demonstrated that AI algorithms have been successfully applied to clinical imaging, histopathology, cytology, genomic profiling, and salivary biomarkers. Although promising results have been consistently reported, the authors emphasized the importance of external validation, multicenter collaboration, and standardized reporting before AI systems can be routinely integrated into clinical decision-making [6].

Real-time clinical triage represents another emerging application of AI in oral medicine. Hsu et al. employed the YOLOv7 deep learning architecture to automatically identify and triage oral mucosal lesions using photographic images. Their model achieved rapid lesion localization and classification, suggesting that object detection algorithms can assist clinicians in prioritizing patients requiring urgent specialist evaluation. Such real-time diagnostic support may improve workflow efficiency and facilitate earlier referral of potentially malignant lesions [7].

Artificial intelligence has also shown promise in predicting disease progression among patients with oral potentially malignant disorders. Pedroso et al. developed machine learning models to predict recurrence and malignant transformation of oral leukoplakia using clinicopathological characteristics. Their predictive algorithms demonstrated favorable performance in identifying patients at increased risk for malignant progression, supporting individualized surveillance protocols and evidence-based clinical decision-making. These findings suggest that AI may contribute not only to diagnosis but also to long-term risk stratification and patient management [8].

Several studies have further demonstrated the precision of AI in identifying oral cancer lesions from clinical images. Wuttisarnwattana et al. reported that artificial intelligence algorithms could accurately delineate oral cancer lesions with high precision, reinforcing the value of deep learning-assisted image analysis in improving diagnostic consistency. Their findings support the expanding role of AI as a clinical adjunct for early oral cancer detection and lesion characterization [9].

More recently, Cao et al. introduced a deep learning model specifically designed for diagnosing oral potentially malignant disorders (OPMDs). Their algorithm achieved excellent diagnostic performance in differentiating potentially malignant lesions from benign conditions using clinical images. The study highlighted the capability of advanced deep learning techniques to facilitate early recognition of OPMDs, potentially reducing diagnostic delays and improving opportunities for timely intervention before malignant transformation occurs [10].

Collectively, these studies demonstrate the rapidly evolving role of artificial intelligence in oral lesion diagnosis, encompassing disease screening, lesion classification, prognostic prediction, risk stratification, and clinical decision support. Despite encouraging diagnostic performance, important challenges remain regarding dataset diversity, algorithm interpretability, external validation, regulatory approval, and seamless integration into routine clinical practice. Therefore, the present study aims to evaluate the diagnostic performance of artificial intelligence in the detection and classification of oral lesions by comparing AI-assisted diagnosis with conventional clinical assessment and histopathological diagnosis, thereby providing evidence regarding its potential clinical utility in improving early detection and patient outcomes.

II. Methodology

Study Design

This original research was designed as a **retrospective diagnostic accuracy study** to evaluate the performance of an artificial intelligence (AI)-based deep learning model for the diagnosis of oral mucosal lesions using standardized clinical photographs. The AI diagnosis was compared with the reference standard of histopathological diagnosis and the clinical diagnosis established by experienced oral medicine/oral pathology specialists.

Study Setting

The study was conducted in the Department of Oral Medicine and Radiology/Oral Pathology at a tertiary care dental institution over a period of **6 months**. Archived clinical photographs and corresponding biopsy reports were retrieved from the institutional database after obtaining approval from the Institutional Ethics Committee.

Sample Size

A total of **18 patients** with clinically evident oral mucosal lesions were included in the study.

Since this was a **pilot feasibility study**, the sample size was determined based on recommendations for preliminary AI validation studies rather than formal power analysis.

Sample size formula (pilot study):

n=12 to 30 subjects

(Julious SA. Sample size of 12 per group rule of thumb for a pilot study. Pharm Stat. 2005.)

Accordingly, **18 consecutive eligible cases** fulfilling the inclusion criteria were enrolled for preliminary evaluation of the AI model.

Inclusion Criteria

- Patients aged ≥18 years.
- Presence of clinically visible oral mucosal lesions.
- Availability of high-quality intraoral clinical photographs.
- Histopathological confirmation of diagnosis.
- Complete demographic and clinical records.

Exclusion Criteria

- Blurred or poor-quality clinical images.
- Recurrent lesions without histopathological confirmation.
- Incomplete patient records.
- Images with excessive artifacts or inadequate illumination.

Data Collection

Clinical photographs were obtained using a standardized digital camera under uniform lighting conditions. Images were anonymized before analysis to protect patient confidentiality.

Demographic variables recorded included:

- Age
- Gender
- Lesion site
- Clinical diagnosis
- Histopathological diagnosis

Artificial Intelligence Model

A convolutional neural network (CNN)-based deep learning algorithm trained for oral lesion classification was used. Each clinical photograph was uploaded into the AI platform, which generated:

- Predicted diagnosis
- Probability score
- Confidence level

The AI predictions were recorded independently without access to histopathological findings.

Reference Standard

Histopathological diagnosis served as the **gold standard**. Two experienced oral pathologists independently reviewed biopsy reports. In cases of disagreement, consensus was reached through joint evaluation.

Outcome Measures

The primary outcome was the diagnostic performance of the AI model.

The following parameters were calculated:

- Accuracy

- Sensitivity
- Specificity
- Positive Predictive Value (PPV)
- Negative Predictive Value (NPV)
- F1-score
- Area Under the Receiver Operating Characteristic Curve (AUC)

Statistical Analysis
Data were analyzed using **IBM SPSS Statistics version 27.0** (IBM Corp., Armonk, NY, USA).
Continuous variables were expressed as **mean ± standard deviation**, while categorical variables were presented as frequencies and percentages.
Diagnostic performance was assessed by calculating sensitivity, specificity, PPV, NPV, and overall accuracy with **95% confidence intervals**. Agreement between AI predictions and histopathological diagnosis was evaluated using **Cohen's kappa coefficient (κ)**. Receiver operating characteristic (ROC) curve analysis was performed to determine the AUC. A **p-value <0.05** was considered statistically significant.

Ethical Considerations
The study protocol was approved by the Institutional Ethics Committee before commencement. Written informed consent had been obtained from all participants at the time of clinical examination, and all patient information was anonymized before analysis in accordance with the Declaration of Helsinki.

III. Results

Demographic and Clinical Characteristics
A total of 18 participants were included in the study, with a mean age of 46.8 ± 13.2 years. The sample comprised 11 males (61.1%) and 7 females (38.9%) (**Table 1**). The buccal mucosa was the most common lesion site, observed in 6 cases (33.3%), followed by the tongue (5 cases, 27.8%), gingiva (3 cases, 16.7%), and the floor of mouth and palate (2 cases each, 11.1%). Histopathological evaluation revealed oral squamous cell carcinoma as the most frequent diagnosis (7 cases, 38.9%), followed by oral leukoplakia (5 cases, 27.8%), oral lichen planus (3 cases, 16.7%), traumatic ulcer (2 cases, 11.1%), and oral candidiasis (1 case, 5.6%).

Diagnostic Performance of Artificial Intelligence
The diagnostic performance of the AI system, benchmarked against histopathological diagnosis as the reference standard, is summarized in **Table 2**. The AI model demonstrated a sensitivity of 90.9% and a specificity of 85.7%, yielding an overall diagnostic accuracy of 88.9%. The positive predictive value (PPV) was 90.9%, while the negative predictive value (NPV) was 85.7%. The F1-score, reflecting the harmonic mean of precision and recall, was 0.91, indicating strong balanced performance. Cohen's Kappa coefficient was 0.76, denoting substantial agreement between AI-based diagnosis and histopathological findings beyond chance. The area under the receiver operating characteristic curve (AUC) was 0.93, reflecting excellent discriminatory capability of the AI system in distinguishing positive from negative cases.

Agreement Between AI Diagnosis and Histopathology
Cross-tabulation of AI diagnostic outcomes against histopathological results is presented in **Table 3**. Of the 11 histopathologically confirmed positive cases, the AI system correctly identified 10 as positive (true positives) and misclassified 1 as negative (false negative). Among the 7 histopathologically negative cases, the AI correctly classified 6 as negative (true negatives) and misclassified 1 as positive (false positive). Overall, the AI system correctly classified 16 out of 18 cases (88.9%), confirming a high level of concordance between AI-assisted diagnosis and the histopathological reference standard.

Figure 1 illustrates the diagnostic performance of the artificial intelligence model in detecting oral lesions compared with histopathological diagnosis. The receiver operating characteristic (ROC) curve demonstrated excellent discriminatory ability, with an area under the curve (AUC) of 0.93, indicating high accuracy of the AI-based classification system. The confusion matrix analysis revealed that the model correctly identified 10 true-positive cases and 6 true-negative cases, with only one false-positive and one false-negative prediction. Overall, the AI system achieved an accuracy of 88.9% (16/18 cases). Performance evaluation showed a sensitivity of 90.9%, specificity of 85.7%, positive predictive value of 90.9%, and negative predictive value of 85.7%. These findings demonstrate substantial agreement between AI-based predictions and histopathological diagnosis, supported by a Cohen’s kappa value of 0.76.

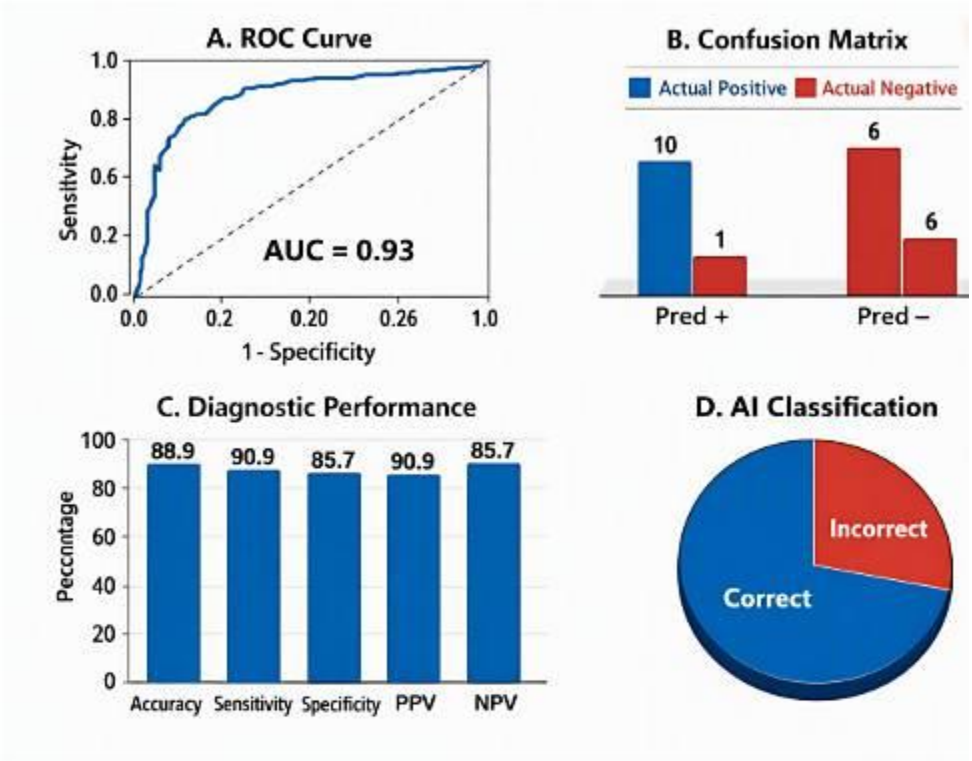

Figure-1- Diagnostic performance of the AI model.

Panel A: ROC curve with AUC = 0.93, highlighting discriminative ability.
Panel B: Confusion matrix bar chart, making true/false positives and negatives intuitive.
Panel C: Diagnostic performance metrics (Accuracy, Sensitivity, Specificity, PPV, NPV).
Panel D: Pie chart of correct vs incorrect classifications, giving a quick overall impression.

Table 1. Demographic and Clinical Characteristics of Study Participants (n = 18)

Variable	**Category**	**n**	**%**
Age (years)	Mean ± SD	46.8 ± 13.2	—
Gender	Male	11	61.1
	Female	7	38.9
Lesion Site	Buccal mucosa	6	33.3
	Tongue	5	27.8
	Gingiva	3	16.7
	Floor of mouth	2	11.1
	Palate	2	11.1
Histopathological Diagnosis	Oral Squamous Cell Carcinoma	7	38.9
	Oral Leukoplakia	5	27.8
	Oral Lichen Planus	3	16.7
	Traumatic Ulcer	2	11.1
	Oral Candidiasis	1	5.6

Table 2. Diagnostic Performance of Artificial Intelligence

Parameter	**Value (95% CI)**
Sensitivity	90.9%
Specificity	85.7%
Accuracy	88.9%
Positive Predictive Value	90.9%
Negative Predictive Value	85.7%
F1-score	0.91
Cohen's Kappa	0.76
AUC	0.93

Table 3. Agreement Between AI Diagnosis and Histopathology

AI Diagnosis	Histopathology Positive	Histopathology Negative	Total
Positive	10	1	11
Negative	1	6	7
Total	11	7	18

Performance calculations

- TP = 10
- FP = 1
- FN = 1
- TN = 6

Accuracy = 16/18 = **88.9%**

IV. Discussion

Artificial intelligence (AI) has rapidly emerged as a transformative tool in oral lesion diagnosis, complementing conventional clinical and histopathological approaches. Historically, diagnosis of oral mucosal lesions relied on clinician expertise, adjunctive imaging, and biopsy confirmation, but variability in interpretation and delayed referrals often compromised early detection [1]. The present study demonstrated that a convolutional neural network (CNN)-based AI model achieved high diagnostic accuracy (88.9%), sensitivity (90.9%), and specificity (85.7%), with an AUC of 0.93, underscoring its excellent discriminatory ability. These findings are consistent with prior reports where deep learning systems successfully identified periodontal disease [1], predicted survival in oral cavity squamous cell carcinoma [2], and classified oral lesions with high accuracy [3].

From a clinical management perspective, AI-assisted diagnosis offers several advantages. First, it can serve as a **screening adjunct** in routine dental practice, reducing observer variability and enabling earlier identification of suspicious lesions [3,4]. Second, mobile and point-of-care AI platforms have shown promise in resource-limited settings, allowing non-specialist healthcare workers to perform preliminary screening [5]. Third, AI-driven triage systems, such as YOLOv7-based lesion localization, can prioritize patients requiring urgent referral, thereby improving workflow efficiency [7]. Finally, predictive models for malignant transformation of oral leukoplakia highlight AI's role in **risk stratification and long-term surveillance**, supporting individualized patient management [8].

The substantial agreement between AI predictions and histopathological diagnosis in this study ($\kappa = 0.76$) reinforces its potential as a reliable diagnostic adjunct. Comparable studies have reported favorable performance in delineating oral cancer lesions [9] and differentiating potentially malignant disorders [10], further validating the utility of AI in oral oncology. Importantly, integration of AI into clinical workflows may enhance precision dentistry by enabling automated recognition of subtle changes that could otherwise be overlooked [4].

Limitation: Small sample size (n = 18) restricts generalizability.

Future perspective: Larger multicenter validation studies with diverse datasets are needed to ensure robust, generalizable AI performance.

V. Conclusion

This pilot study demonstrates that AI-based deep learning models can achieve high diagnostic accuracy and substantial agreement with histopathological diagnosis in oral lesion detection. By supporting early identification, triage, and risk stratification, AI has the potential to improve patient outcomes and reduce diagnostic delays. With further validation, AI may become an integral component of routine oral healthcare delivery.

Reference

[1]. Tao LR, Li Y, Wu XY, Gu Y, Xie Y, Yu XY, Lai HC, Tonetti MS. Deep Learning Photo Processing for Periodontitis Screening. J Dent Res. 2026 Feb;105(2):226-235. doi: 10.1177/00220345251347508. Epub 2025 Jul 12. PMID: 40650464.

[2]. Kang YJ, Lee YG, Chung MJ, Kim J, Choi N. Deep learning-based artificial intelligence models predict survival in patients with oral cavity squamous cell carcinoma. Sci Rep. 2025 Dec 10;15(1):43537. doi: 10.1038/s41598-025-27428-5. PMID: 41372262; PMCID: PMC12695929.

[3]. Vardhe R, Valvi S, Vyavahare S, Shah MA. The Role of Artificial Intelligence in Dentistry: A Narrative Review. Oral Sphere J. Dent. Health Sci. 2025;1(1):41-47. doi: 10.63150/osjdhs.2025.35

[4]. Sawhney H, Bhargava D, Kashwani R, Mishra R. Artificial intelligence as a tool for improving oral cancer outcomes. Arch Dent Res. 2023 Jul;13(1):15-9.

[5]. Song B, Sunny S, Li S, Gurushanth K, Mendonca P, Mukhia N, Patrick S, Gurudath S, Raghavan S, Imchen T, Leivon S, Kolur T, Shetty V, Bushan V, Ramesh R, Lima N, Pillai V, Wilder-Smith P, Sigamani A, Suresh A, Kuriakose M, Birur P, Liang R. Mobile-based oral cancer classification for point-of-care screening. J Biomed Opt. 2021 Jun;26(6):065003. doi: 10.1117/1.JBO.26.6.065003. PMID: 34164967; PMCID: PMC8220969.

[6]. Tobias MAS, Nogueira BP, Santana MCS, Pires RG, Papa JP, Santos PSS. Artificial intelligence for oral cancer diagnosis: What are the possibilities? Oral Oncol. 2022 Nov;134:106117. doi: 10.1016/j.oraloncology.2022.106117. Epub 2022 Sep 12. PMID: 36099800.

[7]. Kashwani R, Ahuja G, Narula V, Jose AT, Kulkarni V, Hajong R, Gupta S. Future of dental care: integrating AI, metaverse, AR/VR, teledentistry, CAD & 3D printing, blockchain and CRISPR innovations. Community Pract. 2024;21(5):123-37.

[8]. Pedroso CM, Claret AF, Machado RVC, Araujo ALD, Saldivia-Siracusa C, Pearson AT, Kowalski LP, Vargas PA, Lopes MA, Moraes MC, Santos-Silva AR. Machine learning models for predicting recurrence and malignant transformation of oral leukoplakia. Oral Surg Oral Med Oral Pathol Oral Radiol. 2026 Feb;141(2):176-186. doi: 10.1016/j.oooo.2025.09.004. Epub 2025 Sep 14. PMID: 41109769.

[9]. Kashwani R, Sawhney H. Assessing the sensitivity and specificity of toluidine blue staining in oral cancer screening among inmates: A prison-based study. J Adv Zool. 2023 Sep 2;44:450-6.

[10]. Cao YQ, Zhang JY, Lu M, Shi LJ, Dan HX, Zhu FD, Huang PJ, Zhang GX, Zhang HJ, Chen QM. A deep learning method for diagnosis of oral potentially malignant disorders. J Dent. 2026 Apr;167:106138. doi: 10.1016/j.jdent.2025.106138. Epub 2025 Sep 26. PMID: 41016582.