

Artificial Intelligence-Based Classification of Oral Pigmented Lesions in Adolescents

Kanad Chaudhuri¹, Vinaya Pawar², Gaurav Vishal³, Sumit Bhatt⁴, Razya Rahim Sheikh⁵, Mohit Zarekar⁶

¹Assistant Professor, Department of Oral Medicine & Radiology, Dr. R. Ahmed Dental College & Hospital, Kolkata, West Bengal, India kanad.chaudhuri@gmail.com

²Assistant Professor, Department of Oral Medicine & Radiology, TPCT's Terna Dental College, Navi Mumbai, Maharashtra, India dr.vinayapawar95@gmail.com

³Head & Neck Oncosurgeon, Department of Surgical Oncology, Prathima Cancer Institute, Warangal, Telangana, India vishal.gaurav512@gmail.com

⁴Assistant Professor & PhD Scholar, Department of Oral & Maxillofacial Surgery, Rajasthan Dental College & Hospital, Nirwan University, Jaipur, Rajasthan, India sumitbhatt22@gmail.com

⁵Senior Resident, Department of Pediatric & Preventive Dentistry, Maulana Azad Institute of Dental Sciences, New Delhi, India razyarahim1823@gmail.com

⁶Senior Resident, Department of Pediatric & Preventive Dentistry, Maulana Azad Institute of Dental Sciences, New Delhi, India zarekarmohit@gmail.com

Corresponding author: Kanad Chaudhuri email: kanad.chaudhuri@gmail.com

Abstract

Background:

Oral pigmented lesions in adolescents encompass a range of conditions with overlapping clinical features, making accurate recognition important for appropriate diagnosis and management. Artificial Intelligence (AI), particularly deep learning, may provide useful support for image-based classification of these lesions. This study evaluated the diagnostic performance of a ResNet50-based AI model in classifying oral pigmented lesions among adolescents.

Materials and Methods:

An observational diagnostic accuracy study was conducted at a tertiary care centre from January to March 2026. A total of 150 adolescents aged 12–19 years with clinically identifiable oral pigmented lesions were evaluated. Standardized intraoral photographs and relevant clinical records were assessed using a ResNet50 convolutional neural network (CNN). Lesions were classified into five diagnostic categories: melanotic macule, melanocytic nevus, physiological pigmentation, exogenous pigmentation, and melanoacanthoma. AI-generated classifications were compared with the reference diagnosis. Diagnostic performance was assessed using accuracy, sensitivity, specificity, positive predictive value, negative predictive value, F1-score, and Cohen's kappa (κ). AI performance was also compared with that of a clinical examiner.

Results:

The AI model correctly classified 139 of 150 cases, yielding an overall accuracy of 92.7%. Category-specific accuracy ranged from 90.0% to 96.7%. The overall sensitivity, specificity, positive predictive value (PPV), and negative predictive value (NPV) were 92.7%, 98.2%, 92.7%, and 98.2%, respectively. κ demonstrated excellent agreement between AI-based classification and the reference diagnosis ($\kappa = 0.909$; 95% confidence interval (CI): 0.846–0.972; probability (p) < 0.001). Classification accuracy remained comparable across gender and adolescent age groups. The AI model achieved higher observed accuracy than the clinical examiner, with accuracies of 92.7% and 87.3%, respectively.

Conclusion:

The ResNet50-based AI system demonstrated reliable classification performance and excellent concordance with the reference diagnosis for oral pigmented lesions in adolescents. Its consistent performance across demographic groups suggests potential utility as an adjunctive diagnostic tool. Larger multicentre studies with independent external validation and prospective clinical evaluation are warranted before broader clinical implementation.

Keywords: Artificial intelligence, Adolescents, Convolutional neural network, Deep learning, Diagnostic accuracy, Oral pigmented lesions, ResNet50.

Introduction

Oral pigmented lesions encompass a heterogeneous group of conditions characterized by abnormal pigmentation of the oral mucosa resulting from melanin deposition, vascular alterations, exogenous materials, or other pathological processes. These lesions may present as localized or diffuse discolorations, exhibiting variations in colour, size, distribution, and morphological characteristics. In adolescents, oral pigmentation may be associated with physiological changes, genetic predisposition, inflammatory responses, medication use, and melanocytic proliferations. Although most oral pigmented lesions are benign, their clinical resemblance to potentially significant pathological conditions presents diagnostic challenges, necessitating systematic clinical evaluation and appropriate investigations [1]. Accurate identification and differentiation of these lesions are essential for establishing appropriate treatment strategies, preventing unnecessary invasive procedures, and ensuring timely referral when suspicious features are identified [2].

Adolescence represents an important developmental period during which physiological, hormonal, and immunological changes may influence oral mucosal characteristics. Oral pigmented lesions observed in this age

group include melanotic macules, melanocytic nevi, post-inflammatory pigmentation, and other localized or generalized pigmentary alterations. Their clinical manifestations may overlap considerably; making differentiation based solely on visual examination difficult [3].

Conventional diagnostic approaches involve detailed patient history, clinical examination, assessment of lesion morphology, and histopathological evaluation when indicated. However, variations in clinical presentation, examiner experience, and the subjective interpretation of colour, borders, and pigmentation patterns may contribute to diagnostic uncertainty. Consequently, the identification of reliable, non-invasive adjunctive diagnostic methods is of particular relevance in adolescent patients, in whom minimizing unnecessary invasive investigations is an important consideration [4].

Artificial intelligence (AI) has increasingly gained attention in healthcare owing to its ability to analyse complex datasets, recognize patterns, and support diagnostic decision-making. Machine learning and deep learning techniques, particularly convolutional neural networks (CNNs), have demonstrated considerable potential in medical image analysis through automated feature extraction and image classification. These algorithms can identify subtle variations in colour, texture, shape, and spatial distribution that may be difficult to assess consistently through conventional visual interpretation [5]. In dentistry and oral medicine, AI-based applications have been explored for the detection and classification of oral mucosal abnormalities, dental diseases, and potentially malignant disorders. The integration of AI into oral diagnostic workflows may facilitate standardized image assessment, improve diagnostic reproducibility, and provide clinicians with additional information to support clinical decision-making [6].

The application of AI-based image classification to oral pigmented lesions may offer an objective approach to differentiating lesions with overlapping clinical characteristics. Deep learning algorithms trained on standardized intraoral photographs can potentially recognize pigmentation patterns, lesion margins, colour intensity, and anatomical distribution to categorize different types of oral pigmentation [7]. Such systems may assist clinicians in identifying lesions requiring closer observation, further investigation, or specialist referral. Nevertheless, the diagnostic performance of AI models is influenced by several factors, including image resolution, lighting conditions, variations in mucosal pigmentation, dataset diversity, sample size, and the accuracy of reference diagnoses. Therefore, rigorous model training, independent validation, and comprehensive evaluation using parameters such as accuracy, sensitivity, specificity, precision, and area under the receiver operating characteristic curve (AUC) (ROC) are necessary to establish their reliability and clinical applicability [8], [9].

Despite advancements in AI-assisted diagnostic imaging, the classification of oral pigmented lesions specifically among adolescents remains an area that warrants further investigation. The diversity of pigmentary conditions, overlapping clinical manifestations, and limited representation of age-specific cases in image datasets may restrict the generalizability of existing classification approaches. Developing and evaluating AI-based systems in this population may help address diagnostic variability and provide insights into the feasibility of integrating automated image analysis into routine oral medicine practice. Accordingly, the present study aims to evaluate the performance of an AI-based image classification model for identifying and differentiating oral pigmented lesions in adolescents, with particular emphasis on diagnostic accuracy, classification reliability, and its potential role as an adjunctive tool in clinical decision-making.

Materials & Methods

Study Design and Study Population

The present observational diagnostic accuracy study was conducted at a single tertiary care centre from January to March 2026 to evaluate the performance of an AI-based image classification system in identifying oral pigmented lesions among adolescents. A total of 150 adolescents aged 12–19 years with clinically identifiable oral pigmented lesions were included. The study population comprised 82 males and 68 females. The participants were categorized into three age groups: 12–14 years, 15–17 years, and 18–19 years. Clinical photographs and corresponding diagnostic records were used to assess the classification performance of the AI model against the reference diagnosis.

Inclusion Criteria

- Adolescents aged 12–19 years at the time of evaluation

- Participants presenting with clinically identifiable oral pigmented lesions
- Availability of clear, diagnostically adequate intraoral photographs
- Availability of relevant clinical information and a documented reference diagnosis
- Participants with sufficient records to permit assessment of AI-based diagnostic classification
- Availability of informed consent from parents or legal guardians and assent from the adolescent, as applicable

Exclusion Criteria

- Individuals younger than 12 years or older than 19 years
- Participants with oral lesions that did not exhibit identifiable pigmentation
- Photographs with inadequate illumination, poor resolution, blurring, or incomplete visualization of the lesion
- Cases with insufficient clinical information or an unavailable reference diagnosis
- Images showing extensive artefacts or obscured anatomical landmarks that could interfere with image interpretation
- Duplicate images of the same lesion, where these could result in repeated evaluation of an identical case.
- Participants whose records were incomplete or unsuitable for diagnostic assessment.

Clinical Examination and Image Acquisition

All participants underwent a clinical examination to document the presence, anatomical location, distribution, colour, size, shape, surface characteristics, and clinical appearance of oral pigmented lesions. Relevant demographic and clinical details, including age and gender, were recorded using a standardized data collection format. Intraoral photographs were obtained to document the clinical characteristics of the lesions. Images were reviewed for adequate focus, illumination, colour representation, and visualization of the affected anatomical site. Only photographs meeting the predefined image-quality requirements were considered suitable for AI-based analysis.

Reference Diagnosis and Classification of Lesions

The reference diagnosis was established using the available clinical and diagnostic records, supported by histopathological examination wherever indicated and available. The lesions were classified into five diagnostic categories: melanotic macule, melanocytic nevus, oral melanoacanthoma, physiologic pigmentation, and other relevant exogenous oral pigmented lesions, according to the study's diagnostic classification protocol. Each case was assigned a reference diagnosis for comparison with the output generated by the AI model.

AI-Based Image Classification

A deep learning-based convolutional neural network (CNN) using the ResNet50 architecture was employed for the classification of oral pigmented lesions. ResNet50 is a residual learning architecture designed to facilitate the training of deep neural networks through the use of shortcut connections, allowing effective extraction of complex image features. The clinical photographs were processed and provided to the AI model for image-based classification. The model analyzed image characteristics, including pigmentation patterns, colour distribution, lesion boundaries, and morphological features, to generate a predicted diagnostic category. The predicted classification for each case was compared with its corresponding reference diagnosis to determine the diagnostic performance of the system.

Assessment of Diagnostic Performance

The diagnostic performance of the AI model was evaluated using the following parameters:

- **Accuracy:** Proportion of cases correctly classified among all evaluated cases
- **Sensitivity:** Ability of the model to correctly identify cases belonging to a particular diagnostic category
- **Specificity:** Ability of the model to correctly exclude cases that did not belong to the respective category

- **Positive predictive value (PPV):** Proportion of correctly classified positive cases among all cases predicted as positive
- **Negative predictive value (NPV):** Proportion of correctly classified negative cases among all cases predicted as negative
- **F1-score:** Harmonic mean of sensitivity and PPV
- **AUC ROC:** Measure of the model's ability to distinguish between diagnostic categories.

The classification results were also analyzed according to the three adolescent age groups to evaluate variations in diagnostic accuracy across different age ranges.

Comparison with Clinical Examiner

The diagnostic performance of the AI model was compared with that of a clinical examiner using the same set of 150 cases. The clinical examiner's classification was evaluated against the reference diagnosis, and the number and proportion of correctly and incorrectly classified cases were determined. The overall classification accuracy of both approaches was compared to assess their relative diagnostic performance.

Inter-Rater Agreement

Agreement between the AI-based classification and the reference diagnosis was evaluated using Cohen's kappa coefficient (κ). The standard error and 95% CI of the kappa coefficient were calculated to quantify the precision of the agreement estimate. The statistical significance of the observed agreement was assessed, with a p-value below 0.05 considered statistically significant.

Ethical Considerations

The study was conducted in accordance with applicable ethical principles for research involving human participants. Participant confidentiality was maintained throughout data collection, image assessment, analysis, and reporting. Clinical photographs and associated participant information were handled exclusively for the purposes of the study, with identifying information excluded from the analytical dataset.

Statistical Analysis

All statistical analyses were performed using the Statistical Package for the Social Sciences (SPSS), IBM SPSS Statistics for Windows, version 29.0 (IBM Corp., Armonk, NY, USA). Categorical variables such as age group, gender, anatomical site, lesion type, colour, surface characteristics, border characteristics, and lesion distribution were expressed as frequency (n) and percentage (%). Continuous variables such as age were summarized using mean and standard deviation (SD). The study included 150 adolescents aged 12–19 years. For evaluation of the AI system, the reference diagnosis was compared with the AI-generated classification using a cross-tabulation/confusion matrix. The diagnostic performance of the AI system was assessed separately for each lesion category using sensitivity, specificity, PPV, NPV, and F1-score. Overall diagnostic accuracy was calculated as the proportion of correctly classified cases among the total number of evaluated cases. The AI correctly classified 139 of 150 cases, giving an overall accuracy of 92.7%. The level of agreement between the AI classification and the reference diagnosis was assessed using κ statistic, along with its standard error and 95% confidence interval (CI). The observed agreement was 92.7%, with a κ value of 0.909 (95% CI: 0.846–0.972; probability (p) < 0.001). For comparison between the AI model and clinical examiner, the number and percentage of correctly and incorrectly classified cases were calculated and the difference in observed accuracy was reported. The AI model showed 92.7% accuracy compared with 87.3% for the clinical examiner, corresponding to an absolute difference of 5.4 percentage points. A p-value <0.05 was considered statistically significant, and a 95% CI was used for inferential estimates.

Results

Table 1 & Figure 1 present the demographic characteristics of the 150 adolescents included in the study. The participants were between 12 and 19 years of age, with a mean age of 15.62 ± 2.14 years. The 15–17-year age group constituted the largest proportion of the study population, comprising 69 (46.0%) participants, followed by the 12–14-year age group with 43 (28.7%) participants and the 18–19-year age group with 38 (25.3%) participants.

Regarding gender distribution, 82 (54.7%) participants were male and 68 (45.3%) were female, indicating a slightly higher proportion of males in the study population.

Table 1: Demographic characteristics of the study participants

Characteristic	n	%
Age (years)		
12–14	43	28.7
15–17	69	46.0
18–19	38	25.3
Mean age $\pm$ SD	15.62 $\pm$ 2.14	
Gender		
Male	82	54.7
Female	68	45.3
Total	150	100.0

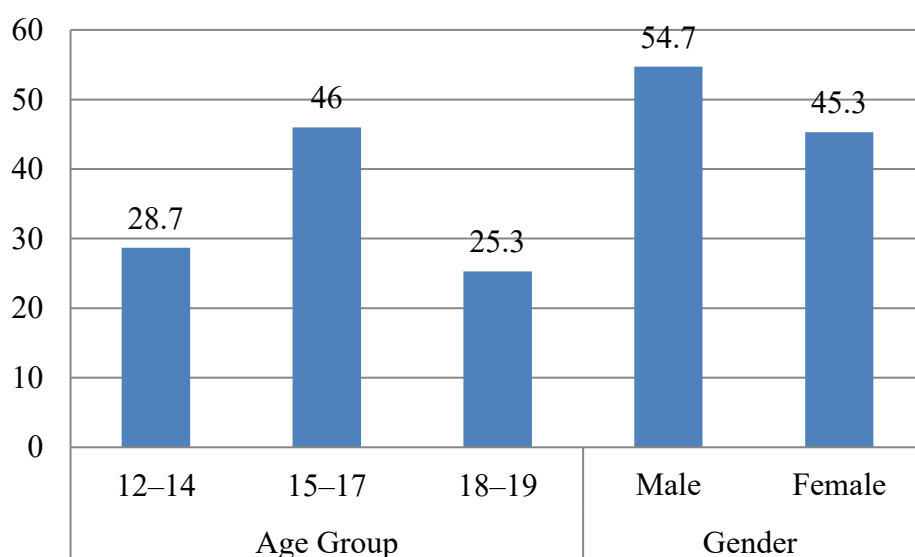


Figure 1: Age-group and gender distribution of the study participants

Table 2 & Figure 2 present the distribution of the oral pigmented lesions included in the study. An identical distribution was observed across all five diagnostic categories. Melanotic macule, melanocytic nevus, physiological pigmentation, exogenous pigmentation, and melanoacanthoma were each identified in 30 (20.0%) cases. Therefore, all five lesion types represented an equal proportion of the total study population, with no single diagnostic category showing a higher frequency than the others.

Table 2: Distribution of oral pigmented lesions

Lesion type	n	%
Melanotic macule	30	20.0
Melanocytic nevus	30	20.0
Physiological pigmentation	30	20.0
Exogenous pigmentation	30	20.0
Melanoacanthoma	30	20.0
Total	150	100.0

- Melanotic macule ■ Melanocytic nevus
- Physiological pigmentation ■ Exogenous pigmentation
- Melanoacanthoma

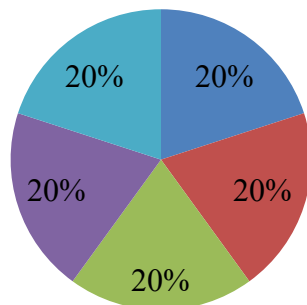


Figure 2: Distribution of the different diagnostic categories of oral pigmented lesions among the study participants. Table 3 & Figure 3 present the anatomical distribution of the 150 oral pigmented lesions evaluated in the study. The gingiva was the most frequently affected site, with 43 (28.7%) cases, followed by the lower lip with 32 (21.3%) cases and the buccal mucosa with 28 (18.7%) cases. The palate was involved in 22 (14.7%) cases, whereas 15 (10.0%) lesions were observed on the tongue. The remaining 10 (6.7%) lesions occurred at other anatomical sites. Overall, the gingiva represented the most common site of oral pigmented lesions in the study population.

Table 3: Distribution According to Anatomical Site

Anatomical site	n	%
Gingiva	43	28.7
Lower lip	32	21.3
Buccal mucosa	28	18.7
Palate	22	14.7
Tongue	15	10.0
Other sites	10	6.7
Total	150	100.0

- Gingiva ■ Lower lip ■ Buccal mucosa
- Palate ■ Tongue ■ Other sites

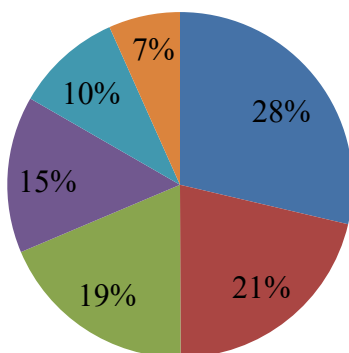


Figure 3: Anatomical site-wise distribution of oral pigmented lesions among the study participants

Table 4 & Figure 4 present the clinical characteristics of the 150 oral pigmented lesions evaluated. Brown was the predominant colour, observed in 74 (49.3%) cases, followed by dark brown/black pigmentation in 49 (32.7%) cases. Bluish-grey pigmentation was noted in 17 (11.3%) cases, while 10 (6.7%) lesions appeared black. With regard to surface morphology, most lesions were flat or macular, accounting for 128 (85.3%) cases, whereas 22 (14.7%) lesions were slightly raised. The majority of lesions had well-defined borders, observed in 119 (79.3%) cases, while 31 (20.7%) exhibited irregular borders. In terms of distribution, 121 (80.7%) lesions were solitary and 29 (19.3%) were multiple. Overall, the most frequently observed clinical features were brown pigmentation, a flat or macular surface, well-defined margins, and solitary presentation.

Table 4: Clinical characteristics of the lesions

	Clinical characteristic	n	%
Colour	Brown	74	49.3
	Dark brown/black	49	32.7
	Bluish-grey	17	11.3
	Black	10	6.7
Surface	Flat/macular	128	85.3
	Slightly raised	22	14.7
Border	Well-defined	119	79.3
	Irregular	31	20.7
Distribution	Solitary	121	80.7
	Multiple	29	19.3

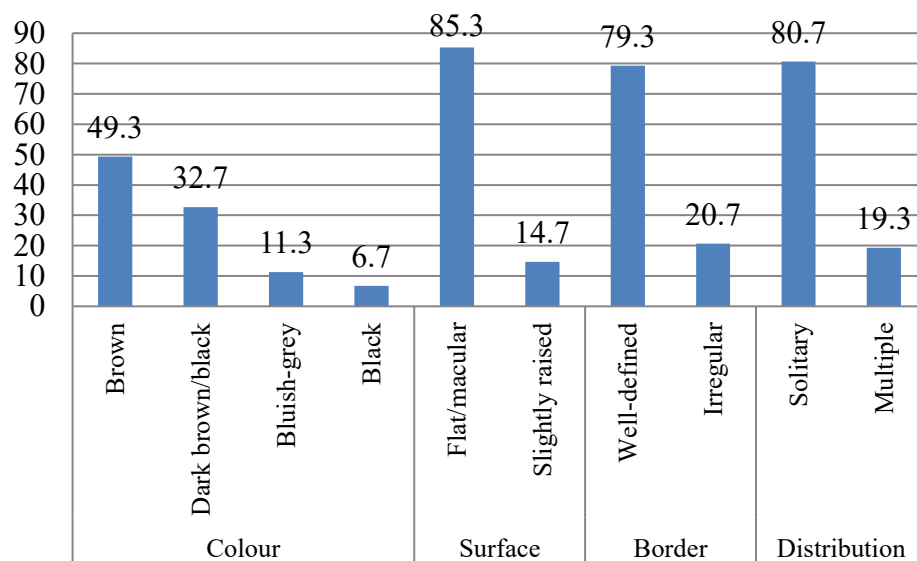


Figure 4: Distribution of oral pigmented lesions according to colour, surface morphology, border characteristics, and lesion distribution.

Table 5 & Figure 5 present the comparison between AI-based classification and the reference diagnosis of the 150 oral pigmented lesions. Among the 30 cases of melanotic macule, the AI system correctly identified 29 (96.7%), while 1 (3.3%) was misclassified as a melanocytic nevus. Of the 30 melanocytic nevus cases, 27 (90.0%) were accurately classified, whereas 1 (3.3%) case each was incorrectly categorized as melanotic macule, physiological pigmentation, and exogenous pigmentation. Similarly, among the 30 cases of physiological pigmentation, 28 (93.3%) were correctly identified, while 1 (3.3%) case each was misclassified as melanocytic nevus and exogenous pigmentation. Of the 30 cases of exogenous pigmentation, 27 (90.0%) were correctly classified, while the

remaining 3 (10.0%) were misidentified as melanocytic nevus, physiological pigmentation, and melanoacanthoma, with 1 (3.3%) case in each category. Among the 30 melanoacanthoma cases, 28 (93.3%) were accurately classified, while 1 (3.3%) case each was incorrectly identified as physiological pigmentation and exogenous pigmentation. Overall, the AI system achieved a classification accuracy of 92.7%, correctly identifying 139 of the 150 lesions, while 11 (7.3%) cases were misclassified. The class-specific accuracy was 96.7% for melanotic macule, 93.3% for physiological pigmentation and melanoacanthoma, and 90.0% for both melanocytic nevus and exogenous pigmentation.

Table 5: AI-based classification of oral pigmented lesions

Reference diagnosis	AI: Melanotic macule	AI: Nevus	AI: Physiological	AI: Exogenous	AI: Melanoacanthoma	Total
Melanotic macule	29 (96.7%)	1 (3.3%)	0 (0.0%)	0 (0.0%)	0 (0.0%)	30 (100.0%)
Melanocytic nevus	1 (3.3%)	27 (90.0%)	1 (3.3%)	1 (3.3%)	0 (0.0%)	30 (100.0%)
Physiological pigmentation	0 (0.0%)	1 (3.3%)	28 (93.3%)	1 (3.3%)	0 (0.0%)	30 (100.0%)
Exogenous pigmentation	0 (0.0%)	1 (3.3%)	1 (3.3%)	27 (90.0%)	1 (3.3%)	30 (100.0%)
Melanoacanthoma	0 (0.0%)	0 (0.0%)	1 (3.3%)	1 (3.3%)	28 (93.3%)	30 (100.0%)

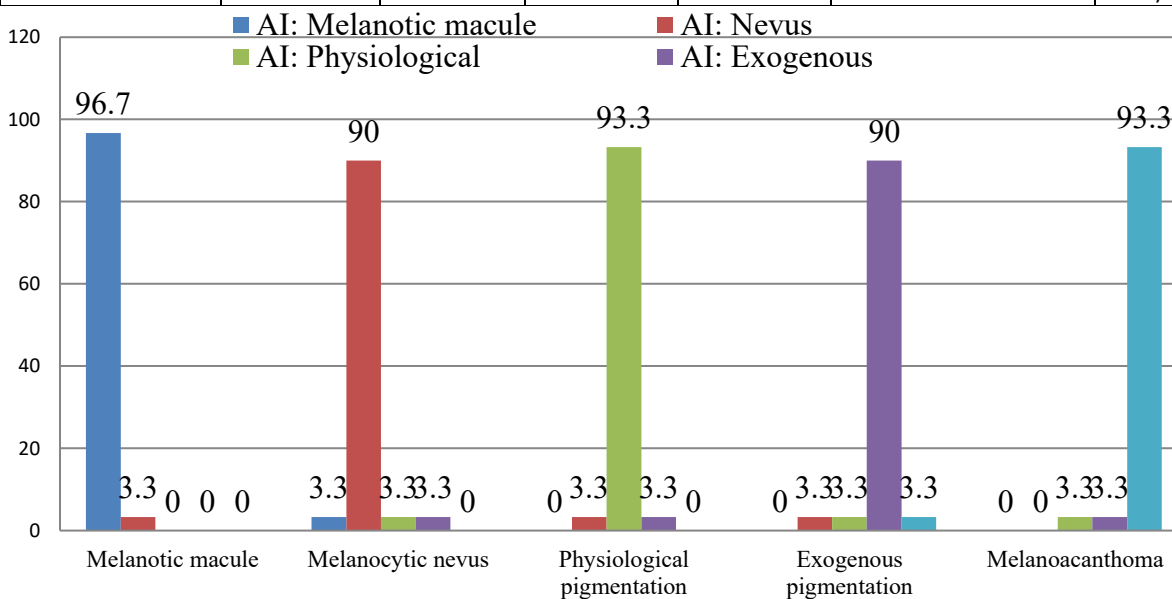


Figure 5: Confusion matrix showing the AI-based classification of oral pigmented lesions in comparison with the reference diagnosis.

Table 6 & Figure 6 present the diagnostic performance of the AI system across the five categories of oral pigmented lesions. Performance was assessed using sensitivity, specificity, PPV, NPV, and F1-score. The AI system demonstrated consistently high diagnostic performance across all evaluated categories. For melanotic macule, the AI system demonstrated a sensitivity of 96.7% and specificity of 99.2%, with a PPV of 96.7% and NPV of 99.2%. The

corresponding F1-score was also high, indicating a strong balance between sensitivity and precision. For melanocytic nevus, the sensitivity was 90.0%, while specificity was 97.5%. The PPV and NPV were 90.0% and 97.5%, respectively, with an F1-score of 90.0%. Similarly, exogenous pigmentation showed a sensitivity of 90.0% and specificity of 97.5%, with a PPV of 90.0% and NPV of 97.5%. The F1-score further indicated good classification performance for this category. For physiological pigmentation, the AI system achieved a sensitivity of 93.3% and specificity of 97.5%. The PPV was 90.3%, the NPV was 98.3%, and the F1-score was 91.8%. For melanoacanthoma, the system demonstrated a sensitivity of 93.3% and specificity of 99.2%, with high PPV, NPV, and F1-score values, indicating reliable classification of this lesion category. Overall, the AI system demonstrated a sensitivity of 92.7% and specificity of 98.2%, with a PPV of 92.7% and NPV of 98.2%. These findings indicate consistently strong diagnostic performance across the five evaluated categories, with specificity generally exceeding sensitivity, suggesting a particularly high ability of the system to correctly identify lesions that did not belong to a given diagnostic category.

Table 6: Diagnostic performance of the AI System

	Sensitivity (%)	Specificity (%)	PPV (%)	NPV (%)
Melanotic macule	96.7	99.2	96.7	99.2
Melanocytic nevus	90.0	97.5	90.0	97.5
Physiological pigmentation	93.3	97.5	90.3	98.3
Exogenous pigmentation	90.0	97.5	90.0	97.5
Melanoacanthoma	93.3	99.2	96.6	98.3
Overall	92.7	98.2	92.7	98.2

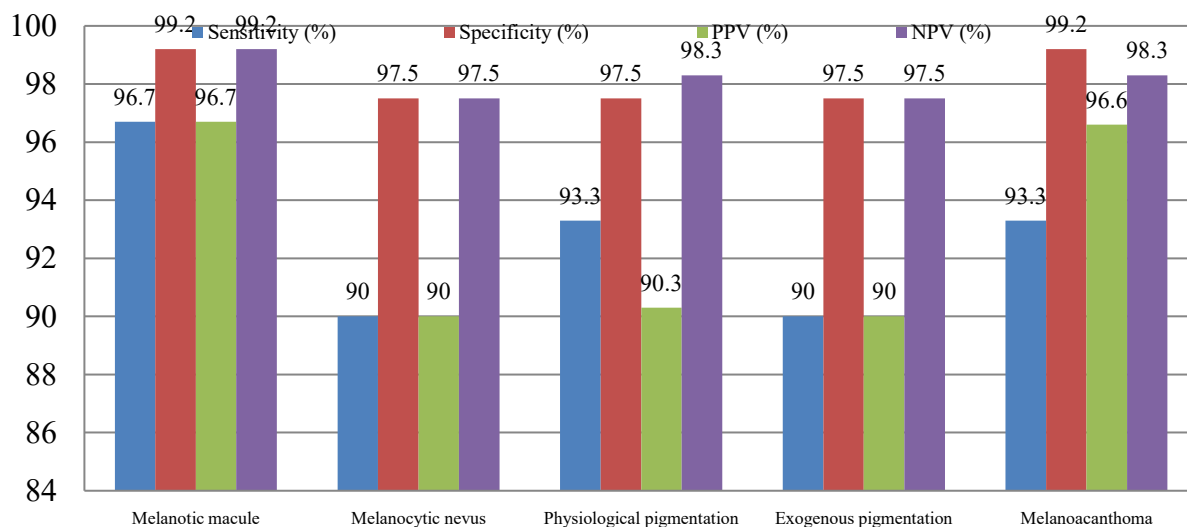


Figure 6: Diagnostic performance of the AI system across different categories of oral pigmented lesions

Table 7 & Figure 7 present the overall diagnostic performance of the AI system. Among the 150 images analyzed, 139 were correctly classified, whereas 11 were misclassified, corresponding to an overall diagnostic accuracy of 92.7%. The AI system demonstrated an overall sensitivity of 92.7% and specificity of 98.2%. The PPV and NPV were 92.7% and 98.2%, respectively. Collectively, these findings indicate that the AI system achieved high diagnostic performance in the classification of oral pigmented lesions within the study population.

Table 7: Overall Diagnostic Accuracy

Performance parameter	Result
Total images evaluated	150
Correct classifications	139
Incorrect classifications	11
Accuracy	92.7%
Sensitivity	92.7%
Specificity	98.2%
PPV	92.7%
NPV	98.2%

■ Accuracy ■ Sensitivity ■ Specificity ■ PPV ■ NPV

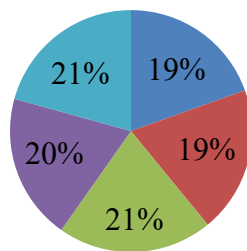


Figure 7: Overall diagnostic performance and classification accuracy of the AI system for oral pigmented lesion

Table 8 & Figure 8 present the gender-wise classification performance of the AI system. Among the 82 male participants, 76 cases (92.7%) were correctly classified, whereas 6 cases (7.3%) were misclassified. Among the 68 female participants, 63 cases (92.6%) were correctly classified and 5 cases (7.4%) were incorrectly classified. Overall, 139 of the 150 cases (92.7%) were correctly classified, while 11 cases (7.3%) were misclassified. The classification accuracy was therefore 92.7% among male participants and 92.6% among female participants, demonstrating comparable AI performance across the two gender groups.

Table 8: AI classification according to gender

Gender	Correct classification	Incorrect classification	Accuracy (%)
Male (n=82)	76	6	92.7
Female (n=68)	63	5	92.6
Total	139	11	92.7

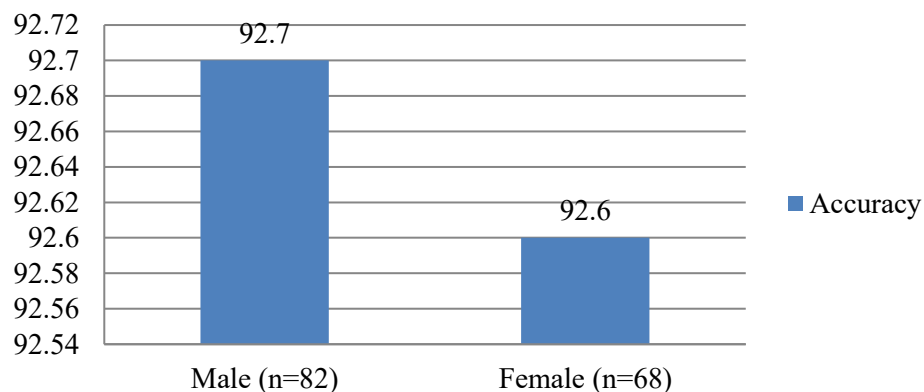


Figure 8: Gender-wise classification accuracy of the AI system for oral pigmented lesions.

Table 9 & Figure 9 illustrate the age-wise diagnostic performance of the AI system. Among participants aged 12–14 years, 40 of 43 cases were correctly classified, achieving an accuracy of 93.0%. In the 15–17 years age group, 64 of 69 cases were accurately classified, corresponding to 92.8% accuracy. Similarly, in the 18–19 years age group, 35 of 38 cases were correctly identified, yielding an accuracy of 92.1%. Overall, the AI system correctly classified 139 (92.7%) of the 150 cases, while 11 (7.3%) were misclassified. The accuracy remained relatively uniform across all age groups, with the highest value recorded among participants aged 12–14 years (93.0%), indicating consistent classification performance across the adolescent age range.

Table 9: AI classification according to age group

Age group	Correct	Incorrect	Accuracy (%)
12–14 years	40	3	93.0
15–17 years	64	5	92.8
18–19 years	35	3	92.1
Total	139	11	92.7

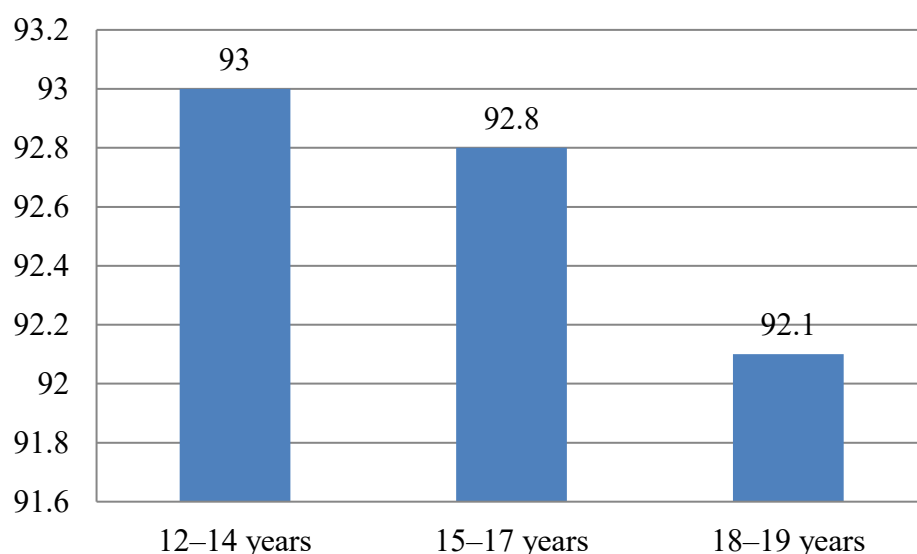


Figure 9: Age-wise comparison of AI-based classification accuracy across different adolescent age groups.

Table 10 presents the agreement between the AI-based classification system and the reference diagnosis. The observed agreement was 92.7%, reflecting a high degree of concordance between the two assessments. κ was 0.909 standard error (SE = 0.032), with a 95% CI ranging from 0.846 to 0.972. The agreement was statistically significant ($p < 0.001$), demonstrating excellent agreement between the AI system and the reference diagnosis.

Table 10: Agreement between AI and reference diagnosis

Measure	Value
Observed agreement	92.7%
κ	0.909
Standard error	0.032
95% CI	0.846–0.972
p-value	<0.001

Table 11 & Figure 10 illustrate the comparative diagnostic accuracy of the AI model and the clinical examiner across 150 cases. The AI model correctly identified 139 cases (92.7%), with 11 cases (7.3%) misclassified. In comparison, the clinical examiner accurately classified 131 cases (87.3%), while 19 cases (12.7%) were incorrectly

classified. The AI model achieved an accuracy of 92.7%, exceeding the clinical examiner's accuracy of 87.3% by 5.4 percentage points.

These findings indicate that the AI model demonstrated higher observed classification accuracy than the clinical examiner in the evaluated dataset.

Table 11: Comparison of AI with Clinical Diagnosis

Diagnostic method	Correct classifications	Incorrect classifications	Accuracy (%)
AI model	139/150	11/150	92.7
Clinical examiner	131/150	19/150	87.3
Difference	8 cases		5.4

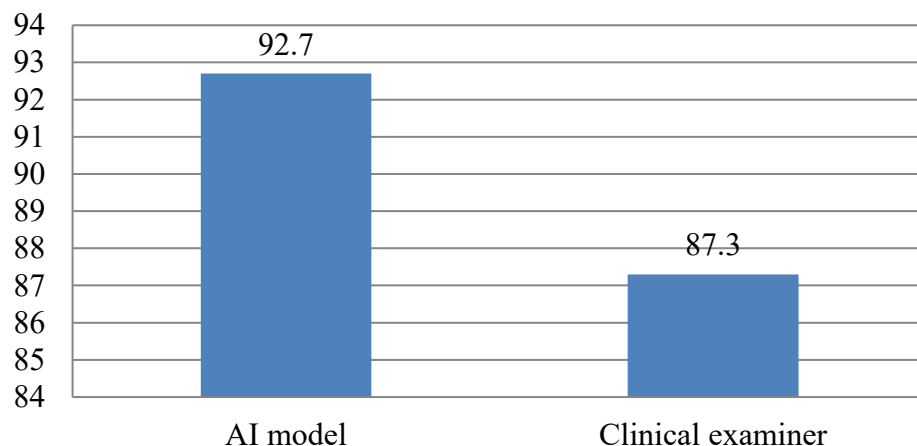


Figure 10: Comparative assessment of diagnostic accuracy between the AI model and the clinical examiner across 150 cases.

Discussion

Oral pigmented lesions comprise a heterogeneous group of conditions with considerable variation in their clinical appearance. Although many lesions are benign, similarities in colour, morphology, and distribution may make clinical differentiation challenging, particularly when assessment is based on photographic images. The present study evaluated the ability of an AI-based image classification system to distinguish five categories of oral pigmented lesions in adolescents and demonstrated strong overall classification performance.

Müller (2010) [10] described the broad clinical spectrum of melanin-associated pigmented lesions of the oral mucosa and emphasized the importance of careful clinical assessment and differential diagnosis. Oral pigmentation may arise from several physiological, reactive, melanocytic, exogenous, and neoplastic processes, and these conditions can demonstrate overlapping clinical characteristics. The predominance of brown and dark brown/black pigmentation observed in the present study is consistent with the importance of colour assessment in the initial characterization of oral pigmented lesions. However, the findings also indicate that colour alone may not be sufficient for reliable categorization, supporting the potential value of image-based analytical methods.

Buchner and Hansen (1979) [11] conducted a clinicopathological study of oral melanotic macules and highlighted the characteristic clinical presentation of this lesion while demonstrating the importance of distinguishing it from other causes of oral pigmentation. In the present investigation, melanotic macules showed the highest category-specific classification accuracy among the evaluated lesions. This relatively strong performance may reflect the presence of recognizable visual characteristics that can be captured by an image-based classification system. Nevertheless, clinical appearance should always be interpreted together with relevant clinical history and, where indicated, histopathological findings.

The application of deep learning to oral image interpretation has expanded considerably in recent years. Zhou et al. (2024) [12] demonstrated the feasibility of using deep-learning algorithms to classify and detect recurrent aphthous ulcerations from clinical oral photographs. Although their study investigated ulcerative rather than pigmented lesions, it provides evidence that deep-learning models can identify clinically relevant visual patterns

from oral photographs. The high classification performance observed in the present study similarly supports the potential usefulness of AI for recognizing visually distinct oral mucosal conditions.

Su et al. (2025) [13] evaluated a deep-learning system for the differential diagnosis of oral mucosal lesions using clinical photographic images. Their findings demonstrated the potential of automated image analysis to assist in distinguishing clinically overlapping oral conditions. This is particularly relevant to the present study because several pigmented lesions may exhibit similar colour and morphological characteristics. The ability of the AI system to maintain high classification performance across all five diagnostic categories suggests that automated image interpretation may provide an additional layer of diagnostic support during the preliminary assessment of oral pigmentation.

Ye et al. (2024) [14] investigated automated oral-lesion detection using deep learning in a multicentre setting and demonstrated the potential applicability of AI-based image analysis across clinical datasets. Their work is important in the context of the present findings because image acquisition conditions, patient characteristics, lesion morphology, and clinical settings can influence AI performance. The strong diagnostic performance observed in the current study indicates that deep learning may be useful for standardized image-based assessment; however, validation using larger and independently collected multicentre datasets would be necessary to determine whether similar performance can be achieved in broader clinical populations.

Keser et al. (2023) [15] assessed a deep-learning algorithm for classification of oral lichen planus lesions using photographic images and demonstrated that AI-based analysis could assist in identifying characteristic patterns of an oral mucosal disorder. Their findings further support the growing role of photographic image analysis in oral diagnosis. Although oral lichen planus differs clinically and pathologically from pigmented lesions, the underlying principle of extracting diagnostically relevant visual characteristics from clinical photographs is applicable to the present investigation.

In the present study, the AI system correctly classified the majority of evaluated lesions and demonstrated substantial agreement with the reference diagnosis. The relatively small differences in classification accuracy between male and female participants and among the different age groups suggest that the model's observed performance was relatively consistent across these demographic subgroups. The higher observed accuracy of the AI model compared with the clinical examiner also indicates that automated image analysis may offer useful diagnostic support. However, the observed difference should not be interpreted as evidence that AI can replace clinical expertise, particularly because clinical diagnosis incorporates information such as patient history, palpation, lesion evolution, and contextual findings that may not be represented in a photograph.

Overall, the findings of the present study extend the emerging literature on AI-assisted oral image analysis to the classification of oral pigmented lesions in adolescents. The earlier clinical and clinicopathological literature by above mentioned studies demonstrate the expanding application of deep learning to oral photographic diagnosis. Taken together, these findings suggest that AI may function as a valuable adjunct for preliminary classification and clinical decision support. Nevertheless, prospective validation with larger, diverse, and multicentre datasets, together with assessment against clinically and histopathologically confirmed diagnoses, is required before such systems can be incorporated routinely into clinical practice.

Limitations

The study has certain limitations that should be considered when interpreting the findings. First, the relatively limited sample size and single-centre design may restrict the generalizability of the results to broader adolescent populations. Second, the AI model was evaluated using a specific dataset, and its performance may vary when applied to images obtained from different institutions, imaging devices, photographic conditions, or patient populations. Third, variations in lesion appearance, image quality, pigmentation intensity, and anatomical location may influence AI classification. The study also focused on a defined set of oral pigmented lesions, which may not represent the full spectrum of pigmented conditions encountered in clinical practice. In addition, comparison with a single clinical examiner may not fully reflect variability among clinicians with different levels of experience. Finally, external validation and prospective clinical evaluation were not undertaken; therefore, the findings should be interpreted as preliminary evidence of diagnostic potential rather than confirmation of routine clinical applicability.

Conclusion

The ResNet50-based AI model demonstrated strong and consistent diagnostic performance in identifying oral pigmented lesions among adolescents, with excellent agreement with the reference diagnosis. Its stable performance across age groups and favourable diagnostic accuracy compared with the clinical examiner indicate its potential as a useful adjunct to clinical assessment. Further studies involving larger, multicentre cohorts and external validation are required to establish its reliability and clinical applicability.

Conflicts of Interest: Nil

Financial Support: None

References

1. Gondak RO, da Silva-Jorge R, Jorge J, Lopes MA, Vargas PA. Oral pigmented lesions: Clinicopathologic features and review of the literature. *Med Oral Patol Oral Cir Bucal*. 2012 Nov 1;17(6):e919-24. doi: 10.4317/medoral.17679. PMID: 22549672; PMCID: PMC3505710.
2. Meleti M, Vescovi P, Mooi WJ, van der Waal I. Pigmented lesions of the oral mucosa and perioral tissues: a flow-chart for the diagnosis and some recommendations for the management. *Oral Surg Oral Med Oral Pathol Oral Radiol Endod*. 2008 May; 105(5):606-16. doi: 10.1016/j.tripleo.2007.07.047. Epub 2008 Feb 21. PMID: 18206403.
3. Kauzman A, Pavone M, Blanas N, Bradley G. Pigmented lesions of the oral cavity: review, differential diagnosis, and case presentations. *J Can Dent Assoc*. 2004 Nov;70(10):682-3. PMID: 15530266.
4. Sreeja C, Ramakrishnan K, Vijayalakshmi D, Devi M, Aesha I, Vijayabanu B. Oral pigmentation: A review. *J Pharm Bioallied Sci*. 2015 Aug;7(Suppl 2):S403-8. doi: 10.4103/0975-7406.163471. PMID: 26538887; PMCID: PMC4606629.
5. Ko E, Panchal N. Pigmented Lesions. *Dermatol Clin*. 2020 Oct;38(4):485-494. doi: 10.1016/j.det.2020.05.009. PMID: 32892857.
6. Abati S, Sandri GF, Finotello L, Polizzi E. Differential Diagnosis of Pigmented Lesions in the Oral Mucosa: A Clinical Based Overview and Narrative Review. *Cancers (Basel)*. 2024 Jul 8;16(13):2487. doi: 10.3390/cancers16132487. PMID: 39001549; PMCID: PMC11240708.
7. 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.
8. Warin K, Suebnukarn S. Deep learning in oral cancer- a systematic review. *BMC Oral Health*. 2024 Feb 10;24(1):212. doi: 10.1186/s12903-024-03993-5. PMID: 38341571; PMCID: PMC10859022.
9. Rokhshad R, Mohammad-Rahimi H, Price JB, Shoorgashti R, Abbasiparashkouh Z, Esmaeili M, Sarfaraz B, Rokhshad A, Motamedian SR, Soltani P, Schwendicke F. Artificial intelligence for classification and detection of oral mucosa lesions on photographs: a systematic review and meta-analysis. *Clin Oral Investig*. 2024 Jan 13;28(1):88. doi: 10.1007/s00784-023-05475-4. PMID: 38217733.
10. Müller S. Melanin-associated pigmented lesions of the oral mucosa: presentation, differential diagnosis, and treatment. *Dermatol Ther*. 2010 May-Jun;23(3):220-9. doi: 10.1111/j.1529-8019.2010.01319.x. PMID: 20597941.
11. Buchner A, Hansen LS. Melanotic macule of the oral mucosa. A clinicopathologic study of 105 cases. *Oral Surg Oral Med Oral Pathol*. 1979 Sep;48(3):244-9. doi: 10.1016/0030-4220(79)90011-2. PMID: 289929.
12. 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.
13. Su AY, Wu ML, Wu YH. Deep learning system for the differential diagnosis of oral mucosal lesions through clinical photographic imaging. *J Dent Sci*. 2025 Jan;20(1):54-60. doi: 10.1016/j.jds.2024.10.019. Epub 2024 Oct 28. PMID: 39873061; PMCID: PMC11763237.
14. Ye YJ, Han Y, Liu Y, Guo ZL, Huang MW. Utilizing deep learning for automated detection of oral lesions: A multicenter study. *Oral Oncol*. 2024 Aug;155:106873. doi: 10.1016/j.oraloncology.2024.106873. Epub 2024 Jun 3. PMID: 38833826.
15. Keser G, Bayrakdar İŞ, Pekiner FN, Çelik Ö, Orhan K. A deep learning algorithm for classification of oral lichen planus lesions from photographic images: A retrospective study. *J Stomatol Oral Maxillofac Surg*. 2023 Feb;124(1):101264. doi: 10.1016/j.jormas.2022.08.007. Epub 2022 Aug 11. PMID: 35964938.