ISSN: 3079-6601

Progress in Early Diagnosis of Oral Squamous Cell Carcinoma: From Cytology to AI-Based Imaging Analysis

Jiayi Xu[1], Wenjing Zhou[1], Liping Gui[1], Hengrui Deng[1], Tianyi Zhou[2]

[1] Hunan University of Medicine, College of Stomatology, Huaihua 418000, China

[2] University of South China, School of Basic Medical Sciences, Hengyang 421001, China

Abstract

Oral squamous cell carcinoma (OSCC) is the most common type of oral malignancy, and its five-year survival rate is highly dependent on the timing of diagnosis. In recent years, advances in cytology, optical imaging technologies, and artificial intelligence (AI) based image analysis have enriched the available strategies for early detection of OSCC. This review summarizes traditional diagnostic approaches, including oral brush cytology, liquid-based cytology, tissue biopsy, and immunohistochemistry. It further highlights recent developments in optical techniques such as narrow-band imaging and autofluorescence-based methods, and provides an overview of emerging deep learning driven automated recognition technologies. Current evidence suggests that future OSCC early-diagnosis systems will evolve toward multimodal, intelligent, and non-invasive frameworks, offering greater potential for improving clinical outcomes.

Keywords

Oral squamous cell carcinoma (OSCC); Early diagnosis; Cytology; Optical imaging; Narrow-band imaging; Autofluorescence; Artificial intelligence (AI); Deep learning.

Introduction

Oral squamous cell carcinoma (OSCC) accounts for more than 90% of all oral malignancies and represents a significant global public health burden. Its development is closely associated with multiple risk factors, including tobacco use, alcohol consumption, betel-quid chewing, chronic irritation, and viral infections. Evidence indicates that early detection and intervention at the T1–T2 stage can markedly improve the five-year survival rate. However, in clinical practice, many cases are still diagnosed at advanced stages, leading to poor prognosis. Consequently, enhancing the capacity for early screening and early diagnosis of OSCC has become a major focus in both clinical and research settings.With continuous advancements in cytological techniques, rapid development of optical imaging technologies, and the growing application of artificial intelligence in medical image analysis, early detection strategies for OSCC are becoming increasingly diversified and precise. This review summarizes current progress in early diagnostic approaches, covering both traditional methods and emerging technologies.

ISSN: 3079-6601

2 Advances in Traditional Diagnostic Methods

2.1 Oral Brush Cytology and Liquid-Based Cytology

Oral brush cytology is widely used for the preliminary assessment of suspicious lesions due to its simplicity and non-invasive nature. Liquid-based cytology improves cell preservation and slide quality compared with conventional smears, enabling clearer visualization of cellular morphology and substantially enhancing the detection of abnormal cells. Studies have reported that the sensitivity of liquid-based cytology in identifying oral potentially malignant disorders (OPMDs) ranges from 70% to 90%, making it a practical option for routine screening in primary healthcare settings.

2.2 Tissue Biopsy

Tissue biopsy remains the “gold standard” for diagnosing OSCC. Histopathological examination allows direct evaluation of epithelial differentiation, invasion depth, and the presence of precancerous or malignant alterations. Nevertheless, biopsy is invasive, may cause discomfort, and requires considerable expertise from pathologists. These limitations restrict its feasibility for large-scale early screening, especially in general dental or community settings.

2.3 Role of Immunohistochemistry (IHC) in Molecular Diagnosis

Immunohistochemistry is frequently applied as an adjunct diagnostic tool and plays an important role in prognostic assessment. Common biomarkers include p53, Ki-67, and Cyclin D1. Aberrant p53 expression may indicate epithelial dysplasia, while elevated Ki-67 levels typically reflect increased cellular proliferation. These markers assist in distinguishing suspicious lesions from true malignant transformation, thereby improving overall diagnostic accuracy.

Table 2. 1 Comparison of Traditional Diagnostic Methods for Early Detection of OSCC

Diagnostic Method	Principle / Technical Basis	Advantages	Limitations	Clinical Application Scenarios
Oral Brush Cytology / Liquid-Based Cytology	Collection of superficial epithelial cells for morphological analysis; liquid-based preparation improves cell preservation and smear quality	Non-invasive and easy to perform; good cellular preservation; higher detection rate of abnormal cells; suitable for large-scale preliminary screening	Possible false-negative results; relies on cytologist’s expertise; cannot replace histopathological confirmation	Primary screening in community settings; follow-up of OPMDs; evaluation of suspicious lesions
Tissue Biopsy	Excision of lesion tissue followed by	Highest diagnostic	Invasive procedure with	Final diagnosis; pre-treatment

ISSN: 3079-6601

Diagnostic Method	Principle / Technical Basis	Advantages	Limitations	Clinical Application Scenarios
	histopathological examination of epithelial differentiation, basement membrane integrity, and invasion depth	accuracy; determines grade of differentiation and invasion; regarded as the "gold standard"	low patient compliance; requires experienced pathologists; unsuitable for large-scale screening	staging; sampling of heterogeneous lesions
Immunohistochemistry (IHC)	Antibody-based detection of specific proteins (e.g., p53, Ki-67, Cyclin D1) to assess cell proliferation and malignant transformation	Helps differentiate benign from malignant lesions; provides molecular-level evidence supporting histopathology; useful for prognostic evaluation	Requires tissue samples; relatively high cost; results influenced by antibody selection and technical conditions	Risk assessment of precancerous lesions; molecular classification after diagnosis; prognosis prediction

3. Optical Imaging Techniques for Early Diagnosis of OSCC

With advances in non-invasive diagnostic technologies, a range of optical imaging methods has been increasingly applied to the early detection of oral squamous cell carcinoma (OSCC) and its precancerous lesions. These techniques can capture subtle alterations in vascular patterns, fluorescence signals, or epithelial microstructures before they become clinically apparent, making them valuable tools for risk assessment and screening of oral mucosal lesions. This section summarizes the principles and recent clinical developments of narrow band imaging, autofluorescence imaging, Lugol's iodine staining, and optical coherence tomography.

3.1 Narrow Band Imaging (NBI)

NBI enhances visualization of superficial mucosal microvasculature by using specific narrow-band light wavelengths. Early malignant transformation is often accompanied by characteristic vascular changes, such as punctate dilation, irregular arrangement, or arborizing proliferation. NBI can reveal these alterations even when the lesion is not clearly visible under conventional white light. Clinical studies have reported a 20–40% improvement in diagnostic accuracy for suspicious lesions compared with white-light examination, along with markedly better delineation of lesion margins. NBI is particularly useful for risk stratification of oral potentially malignant disorders (OPMDs), such as leukoplakia and erythroplakia, as well as for preoperative assessment of lesion extent.

ISSN: 3079-6601

3.2 Autofluorescence Imaging (AFI)

AFI is based on the excitation and emission properties of endogenous fluorophores within tissues, including collagen and NADH. When epithelial dysplasia or early malignant transformation occurs, fluorescence intensity typically decreases, producing darkened areas on the image. AFI offers high sensitivity and is easy to perform, making it suitable as a rapid screening tool in outpatient settings. However, its specificity is relatively low, and it can be affected by inflammation, mechanical irritation, or epithelial thickening. As a result, AFI is often used in combination with other diagnostic methods to reduce the false-positive rate.

3.3 Lugol's Iodine Staining

Lugol's iodine staining is a low-cost, simple auxiliary imaging technique that relies on differences in glycogen content between normal and abnormal epithelium. Normal mucosa with higher glycogen content absorbs iodine and stains brown, whereas areas with reduced glycogen such as dysplastic or malignant epithelium—remain unstained, appearing as "iodine-negative" zones. Lugol staining is widely applied for risk assessment of OPMDs and is frequently used during OSCC surgery to identify tumor margins, helping reduce local recurrence and improve completeness of excision.

3.4 Optical Coherence Tomography (OCT)

OCT is an imaging technique that uses near-infrared light to generate high-resolution cross-sectional views of tissue, often described as an "optical biopsy." It provides real-time visualization of epithelial thickness, basement membrane integrity, and the epithelial–lamina propria interface, offering structural information comparable to histopathology. In OSCC diagnosis, OCT can detect key features such as epithelial irregular thickening and basement membrane disruption, aiding in differentiating benign lesions, precancerous changes, and early invasive carcinoma. Its non-invasive and repeatable nature makes OCT suitable for longitudinal monitoring and surveillance of high-risk areas.

4. Advances in AI-Based Imaging Analysis for Early Diagnosis of OSCC

With the rapid development of artificial intelligence, deep learning has become an important tool for identifying oral mucosal lesions. Its major advantage lies in the ability to automatically extract high-dimensional features from large image datasets, overcoming the limitations of manual feature engineering. Current mainstream models used for detecting OSCC and premalignant lesions include convolutional neural networks (CNNs), the ResNet family, MobileNet, VGGNet, InceptionNet, and more recently, Vision Transformers (ViT). CNNs are effective in capturing both low- and high-level structural features such as color variation, texture, and boundary irregularities. ResNet, with its residual architecture, improves the stability of deep networks and is frequently used for detecting lesions in complex clinical backgrounds. MobileNet, owing to its lightweight design, is suitable for mobile and portable devices. VGGNet and InceptionNet often serve as baseline architectures in multi-center studies, whereas Transformer-based models excel in capturing global contextual information and perform particularly well when assessing lesions with unclear margins or atypical presentations.These models not only demonstrate strong capabilities in classification and segmentation but can also be paired with interpretability tools such as Grad-CAM and SHAP to highlight areas of diagnostic interest, thereby

ISSN: 3079-6601

improving clinician trust and supporting clinical decision-making. Numerous studies have confirmed the effectiveness of artificial intelligence in the early detection of OSCC. Overall, AI models typically achieve diagnostic accuracies of 85–95%, with sensitivities of 80–92% and specificities of approximately 80–90%. Their performance approaches that of experienced specialists and may even surpass that of junior clinicians, particularly in primary care settings or large-scale screening programs. For instance, MobileNet has reached accuracies of about 92% in classifying leukoplakia and erythroplakia, while ResNet has shown consistent results in distinguishing precancerous lesions from early squamous cell carcinoma. Certain Transformer-based models perform especially well when interpreting atypical lesions or those with indistinct borders.However, AI performance remains influenced by factors such as image quality, variability in acquisition devices, challenges in annotation, and the scarcity of large multi-center datasets. The generalizability and stability of these models still require validation in broader clinical cohorts.

In recent years, the improvement of mobile device capabilities has stimulated growing interest in portable screening systems. By capturing mucosal images using smartphone cameras and analyzing them with lightweight models such as MobileNet, it is possible to achieve real-time “photo-based screening” at low cost and with high accessibility. This approach is particularly valuable for resource-limited settings, high-risk populations, and community-level screening. Some studies have further integrated external optical accessories such as miniature lenses and auxiliary lighting to enhance image quality and improve diagnostic reliability. Although mobile-based screening systems face limitations, including variability among devices, inconsistent image acquisition standards, and limited clinical validation, their potential for public health applications in regions with high OSCC incidence has been widely recognized. They are increasingly viewed as a promising direction for future early detection strategies

Table 4.1 Characteristics and Applications of Deep Learning Models for Early OSCC Diagnosis

Model Type	Technical Features	Advantages	Limitations	Typical Application Scenarios
CNN (Convolutional Neural Network)	**Extracts features from local textures and color gradients, gradually forming high-level representations**	**Strong feature extraction capability; applicable to various lesion types**	**Limited sensitivity to complex global structures**	**Classification and risk assessment of leukoplakia, erythroplakia, and mucosal ulcers**
ResNet Series	**Uses residual connections to prevent gradient vanishing and enable deeper network**	**Supports deeper architectures; performs well under low-light or complex backgrounds**	**Larger model size; higher hardware requirements**	**Differentiation between OPMDs and early squamous cell carcinoma**

ISSN: 3079-6601

Model Type	Technical Features	Advantages	Limitations	Typical Application Scenarios
	training			
MobileNet (Lightweight Model)	**Employs depthwise separable convolutions to reduce parameters significantly**	**Deployable on mobile devices; real-time inference; suitable for resource-limited settings**	**Feature representation slightly weaker than larger models**	**Mobile-based "photo-and-diagnose" systems; screening in underserved regions**
VGGNet	**Regular and straightforward layer structure**	**Easy to train and widely adopted as a baseline model**	**Large number of parameters; computationally expensive**	**Baseline architecture for multi-center datasets and comparative studies**
InceptionNet	**Multi-scale convolution branches extract diverse feature patterns**	**Effective for mucosal images with complex texture variations**	**Structural complexity; slower inference speed**	**Analysis of diverse mucosal lesions and optical imaging datasets**
Vision Transformer (ViT)	**Uses global attention mechanisms to model the overall image structure**	**High sensitivity for atypical lesions and blurred boundaries**	**Requires large datasets; prone to overfitting on small samples**	**Identification of atypical OSCC lesions; complex imaging conditions**
Explainability Modules (Grad-CAM, SHAP)	**Provides visual interpretation for deep learning models**	**Improves clinical trust and enhances model transparency**	**Only aids interpretation; does not increase classification accuracy**	**Clinician decision support and visualization of lesion attention regions**

5 Future Directions and Integrated Diagnostic Trends in OSCC

Overall, future approaches to the diagnosis of oral squamous cell carcinoma (OSCC) are expected to move toward integrated multimodal strategies rather than relying on a single testing method. Cytological assessment, optical imaging, molecular biomarkers, and computer-assisted analysis will likely be combined to support a tiered system of early screening and precise diagnosis. Such integration is particularly valuable for improving diagnostic specificity. The complementary strengths of optical techniques and computational analysis may help reduce false-positive findings and, in turn, unnecessary biopsies. At the same time, the development of low-cost, easy-to-use mobile screening tools is anticipated to enhance accessibility in primary care settings. Wider adoption of these technologies could expand early-detection coverage and lower the mortality associated with delayed diagnosis. Taken together, multimodal integration, improved specificity, and greater accessibility are likely to shape the future direction of OSCC screening and diagnostic practice.

ISSN: 3079-6601

6. Conclusion

Early diagnostic approaches for oral squamous cell carcinoma (OSCC) are shifting from traditional morphology-based assessment toward more multimodal, non-invasive, and technology-enhanced strategies. While cytology and biopsy remain fundamental tools, advances in optical imaging and artificial intelligence are expanding the efficiency and accessibility of early screening. As cross-disciplinary technologies continue to develop, the future diagnostic framework for OSCC is expected to become increasingly precise, offering new opportunities to reduce the overall disease burden associated with oral cancer.

REFERENCE

Romano, A., Di Stasio, D., Gentile, C., Petruzzi, M., Serpico, R., & Favia, G. (2021). Noninvasive imaging methods to improve the diagnosis of oral squamous cell carcinoma: A comprehensive review. Journal of Clinical Medicine, 10(11), 2345.

Lau, A., Li, K. Y., McGrath, C., & Yiu, C. K. (2024). Adjunctive aids for the detection of oral squamous cell carcinoma and oral potentially malignant disorders: A systematic review. Journal of Stomatology, Oral and Maxillofacial Surgery, 125(1), 101–110.

Thakuria, P., Sanjeevi, R., & Kumar, N. (2024). Deep learning for early diagnosis of oral cancer via smartphone and DSLR image analysis: A systematic review. Oral Oncology, 149, 106–120.

Lin, H., Chen, H., Jiang, S., Zhang, S., & Ye, J. (2021). Automatic detection of oral cancer in smartphone-based imaging using deep learning. Frontiers in Oncology, 11, 705–718.

van Schaik, P. M., Lakshmanan, R., Baatenburg de Jong, R. J., & Events, E. (2021). Optical imaging for early detection of oral and oropharyngeal cancers: Clinical status and future perspectives. Current Opinion in Oncology, 33(3), 234–241.

Kikuta, S., Iwai, H., Hanai, N., et al. (2018). Clinical application of the IllumiScan fluorescence visualization device in detecting oral mucosal lesions. Cureus, 10(11), e3552.

Carreras-Torras, C., & Gay-Escoda, C. (2015). Techniques for early diagnosis of oral squamous cell carcinoma: A systematic review. Medicina Oral, Patología Oral y Cirugía Bucal, 20(3), e305–e315.

Noorlag, R., van Kempen, P. M. W., & Smeele, L. E. (2022). Image-guided surgery in oral cancer: Toward improved surgical margins and earlier detection. Current Opinion in Otolaryngology & Head and Neck Surgery, 30(2), 85–92.

Tsai, M. T., Lee, C. K., Lee, H. C., et al. (2020). Optical coherence tomography for oral cancer diagnosis during clinical examination. Journal of Biomedical Optics, 25(6), 065002.

ISSN: 3079-6601

Feng, X., Lu, Z., Liu, X., et al. (2022). Deep learning–based classification of oral leukoplakia and squamous cell carcinoma using autofluorescence and white-light images. Photodiagnosis and Photodynamic Therapy, 38, 102806.