OPEN ACCESS JOURNALS
Indiana Journal of Multidisciplinary Research
Abbreviated Tittle- Ind J Multi Res
ISSN *(Online)*- 2583-3820
Journal Homepage Link- https://indianapublications.com/Journals/IJMR
DOI: https://doi.org/10.5281/zenodo.21600154

INDIANA PUBLICATIONS
PRODUCTIVE AND QUALITY RESEARCH

Research Article

Volume-06|Issue-04|2026

Deep Learning for OSCC Diagnosis: A Multimodal Survey of Techniques, Challenges, and Future Directions

Vinaya R Kudatarkar[1], Annapurna P Patil[2], Savita K Shetty[3]

[1]Department of computer science and Engineering, Research scholar, Dayananda Sagar College of Engineering, Bangalore, Karnataka, India
[2]Professor and Head, ISE, Dayanand college of Engineering, Bangalore, Karnataka, India
[3]Associate professor of Department of Information Technology, Ramaiah Institute of Technology, Bangalore, Karnataka, India.

Article History
Received: 05.05.2026
Accepted: 22.06.2026
Published: 25.07.2026

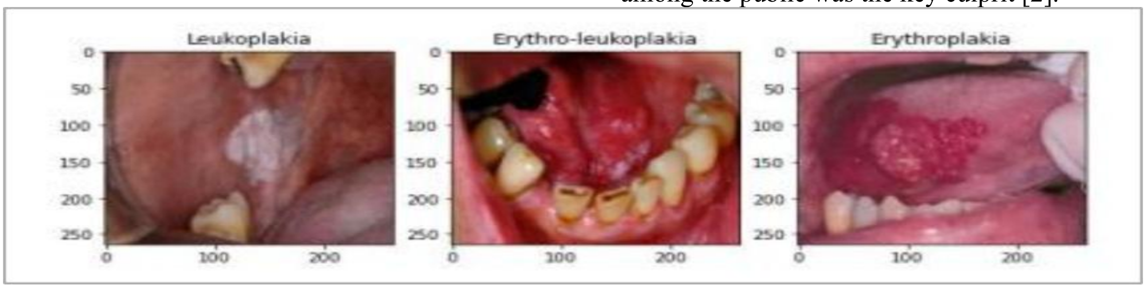
Citation
Kudatarkar, V. R., Patil, A. P. & Shetty, S. K. (2026). Deep Learning for OSCC Diagnosis: A Multimodal Survey of Techniques, Challenges, and Future Directions. *Indiana Journal of Multidisciplinary Research, 6*(4), 148-156.

Abstract: Oral cancer, particularly Oral Squamous Cell Carcinoma (OSCC), remains a significant global health concern due to high mortality rates and frequent late-stage diagnosis. Often identified at an advanced stage because of publicignoranceand restrictions in traditional diagnostic techniques, Oral Squamous Cell Carcinoma(OSCC) is among the most common and lethal types of oral cancer. By providing robust tools for automatic and reliable illness identification, artificial intelligence (AI), especially deep learning, has transformed medical picture analysis in recent years. This survey study offers a thorough assessment of state-of-the-art deep learning techniques used to Oral cancer detection across several imaging modalities including histopathology, fluorescence, hyperspectral, and white light pictures. We methodically investigate and contrast hybrid systems, transformer architectures, transfer learning models, and convolutional neural networks (CNNs) with respect to classification accuracy, resilience, and clinical relevance. The research also addresses real-time deployment issues, model interpretability, and multimodal data integration's importance. Moreover, this study points out present research voids—such as restricted generalizability and absence of stage-wise lesion classification—and offers future research paths to close these obstacles. This effort intends to lead academics, doctors, and developers toward the building of efficient, scalable, and accessible AI-driven diagnostic tools for early OSCC diagnosis and intervention by synthesizing ideas from previous advancements.

Keywords: Artificial Intelligence (AI), Deep Learning, Diagnostic Accuracy, Deep Learning, Oral Squamous Cell Carcinoma (OSCC)

INTRODUCTION

Affecting more than 300,000 people yearly and causing 177,384 deaths each year, oral squamous cell carcinoma (OSCC) is among the most prevalent but least recognized cancers globally, about 2% of all cancer locations. Of the 9,700 new cases recorded in Italy in 2018, 7,400 were men and 2,300 women. Recent epidemiological studies show that OSCC most usually affects males over the fifth decade, with a growing tendency in women globally and in those under 45 years [1]. OSCC death rates are staying steady over time (approximately 50%), fundamentally dependent on illness stage, despite widespread attempts to offer sufficient treatments and medicines following diagnosis. An important element influencing death is the still limited knowledge of the condition and its early indicators, which causes significant diagnostic delays and worse survival rates in addition to exposure to risk factors and access to clinical diagnosis and treatment. Furthermore, later-stage diagnosis significantly impacts the quality of life of the surviving patients and raises the expense of treatment with extended hospitalization and the requirement for more complicated surgical procedures and reconstructions. A recent systematic analysis looked at the reasons for delayed diagnosis in OSCC patients and found that the lack of information among the public was the key culprit [2].



Consistent with other cancer kinds, the cause of OSCC is closely linked to particular lifestyle choices and behavior, including tobacco smoking, alcohol misuse, exposure to sunlight ultraviolet (UV) radiation, and human papilloma virus (HPV) infection. The latter is a unique entity as it mostly affects younger individuals than OSCC and is mostly linked with oropharyngeal cancer (OPSCC). Apart from the knowledge of risk factors, which are shared also to other cancer types, the limited knowledge of the possibility of cancer development inside the oral cavity stays a diffuse and major issue, together with the challenge of finding a medical expert to consult in case of suspect symptoms [3].

Though their effectiveness in lowering the time between symptom onset and referral to a suitable doctor is hard to assess, informative campaigns on OSCC and other cancer kinds traits and risk factors are carried out all over the world. Such efforts run the danger of excluding some categories of people who might especially gain from the knowledge. Therefore, evaluating OSCC knowledge among general population, confirming differential awareness in subcategories of individuals might help to identify certain population groups that should be addressed with educational efforts [4].

Of all oral cancers, ninety-five percent are squamous cell. Avoidable aetiological risk factors have been linked to oral malignancies. The biggest risk factors, which are linked to about 75% of oral malignancies, are smoking tobacco and alcohol use. Common in the Indian subcontinent, smokeless tobacco use has also been linked to a major risk for oral and pharyngeal cancer [5].

Usually at late stages (stage III and IV), oral malignancies are mainly asymptomatic. Though treatment has improved in recent years, this kind of cancer still has extremely low global survival rates; average five-year survival rate of 50%. Early diagnosis and quick treatment are the main ways to lower morbidity and increase the likelihood of cure. Any sort of delay in presentation or referral can negatively impact the prognosis. Lack of knowledge among the public regarding the early indicators of oral cancer is the most critical reason in delayed referral and treatment. Insufficient awareness among general medical doctors and general dentistry practitioners has also been linked to delays in referral and treatment [6].

Cancers of the oral cavity and oropharynx cause more fatalities than all other orofacial area illnesses and disorders combined. Not with standing this, the dentistry or medical community has not given oral cancer sufficient focus. Though developments in the treatment of oral cancer have been linked to better prognosis and survival in high-income countries (or HICs), this finding has not proven true in Low and Middle-income countries (LMIC). Furthermore, the five-year survival rate for oral cancer is almost 50% lower in LMICs than in HIC [7]. Therefore, the approach to cancer diagnosis depends on patients' awareness of possibly severe oral cancer symptoms before they see primary care. Studies indicate that more than 70% of OPC can be avoided by actions aimed at modifiable risk factors found in its evolution. Therefore, a crucial goal of the Healthy People 2020 project is a behavioral change and early secondary prevention by screening. Visual screening for OSCC is fast, easy, cheap, non-invasive, and produces minimal patient pain. By comparison, finding most solid cancers in their early asymptomatic phases calls for unique, expensive, sometimes intrusive procedures. Target campaigns may be rather effective by means of the identification of those groups at higher risk of cancer but who are also more likely to have lower awareness [8].

In recent years, there has been a major focus on risk factor awareness and its potential influence in cancer prevention. Evidence-based information coupled with rising self-efficacy could help to raise knowledge of the risk factors connected to the disease and avoidance of these dangers. This is especially important given the low level of awareness and the correlation between knowledge level and preventative action. Studies done over a long period show that in reducing cancer risk, adopting healthy practices is both efficient and affordable [9]. Therefore, it is relevant to evaluate the risk factor awareness and views concerning oral cancer as well as to look at other elements influencing this view, including socioeconomic position.

Development of smartphone-based oral cancer detectors employing medical picture analysis has made considerable use of artificial intelligence (AI) technologies. By reducing doctors' workloads and easing stress caused by handling complicated data, artificial intelligence helps to diagnose diseases. Deep learning methods have attracted much attention in recent years for their possible use in helping physicians make educated judgments that might finally lead to improved management of patient health. Technological developments have replaced conventional neural networks with deep neural networks. Its great popularity stems from the ability of this deep-learning approach to improve cancer therapy [10].

Oral cancer, particularly Oral Squamous Cell Carcinoma (OSCC), poses a persistent public health challenge, especially in low- and middle-income countries (LMICs) where early-stage diagnosis remains rare due to poor public awareness, lack of screening programs, and limited access to specialized care. Despite advancements in clinical therapies, the five-year survival rate for OSCC continues to hover around 50%, primarily because diagnosis typically occurs at advanced stages when treatment is more complex, costly, and less effective. Especially in low- and middle-income nations where late-stage diagnosis is common owing to poor knowledge and insufficient access to specialist treatment,

***Corresponding Author:** Vinaya R Kudatarkar

Oral Squamous Cell Carcinoma (OSCC) remains a major worldwide health concern. Though medical therapy has improved, OSCC's five-year survival rate is still rather poor, mostly due to missed early identification. Traditional diagnostic techniques are intrusive, time-consuming, and mostly dependent on clinical knowledge, which can cause delays and misdiagnoses. Recent developments in artificial intelligence, especially deep learning, have demonstrated extraordinary promise in automating and improving the precision of medical picture processing. This expanding area has potential for changing oral cancer screening into a more accessible, non-invasive, and efficient procedure. But, it is very necessary to methodically investigate, contrast, and combine the several different AI-based approaches that have appeared for OSCC identification. Motivated by the pressing need to close the gap between technical innovation and clinical relevance, this survey provides a thorough review of state-of-the-art methods, their performance, constraints, and future possibilities.

Comprehensive Review of AI-Based Oral cancer Detection Methods: This survey provides a structured and in-depth review of recent advancements in artificial intelligence, particularly deep learning approaches, applied to the diagnosis of oral squamous cell carcinoma (OSCC) using various medical imaging modalities, including histopathological, fluorescence, and hyperspectral images. Comparative Analysis of Deep Learning Architectures and Performance Metrics: The paper critically compares a wide range of deep learning models—including CNNs, ResNet variants, transformers, and hybrid methods—by evaluating their diagnostic accuracy, sensitivity, specificity, and computational efficiency, thereby highlighting the strengths and limitations of each technique. Identification of Research Gaps and Future Directions: Beyond summarizing existing literature, this survey identifies key research gaps such as lack of multimodal integration, limited interpretability of models, and dataset generalization issues. It also proposes future research directions aimed at developing clinically robust, scalable, and explainable AI systems for early oral cancer detection.

RELATED WORK

Using learnt transmission models and heat maps to highlight regions of concern, a paper in [11] raises a revolutionary way to categorize histology pictures as normal or worrisome. The method confirms its promise as a strong weapon in the battle against oral cancer by scoring 90.9% and 73.6% for two separate datasets. Inspired by the appeal of lymph node metastases, scientists in [12] provide an interesting idea showing the diagnostic promise of artificial intelligence in identifying these threatening nodes in OSCC patients. Their suggested AI solution has an accuracy of 78.2%, a sensitivity of 75.4%, and a specificity of 81%, thereby offering promise for a difficult area of oral cancer detection. Writers in [13] enthrall with a revolutionary method, a harmonic symphony of CNN and texture mapping for identifying oral cancer. Dancing with wavelet transform, this clever model reveals a sensitivity of up to 96.87% and a specificity of up to 71.29%, hence opening new areas of accuracy in cancer diagnosis. Author in [14] imagines a future of changing learning and uses transfer learning with AlexNet methods, training their system using a dataset of oral biopsy pictures. The result is nothing short of amazing, with an accuracy of 90.05% in predicting oral cancer, therefore releasing the possibility for more exact and faster diagnosis. [15] then takes the stage with convolutional neural networks (CNNs) for automated classification of oral squamous cell cancer utilizing ex vivo fluorescence confocal microscopy (FCM) imagery, drawing the limelight. Classifying malignant tissue, their outstanding performance produces a sensitivity of 0.47 and a specificity of 0.96, which astounds the audience with the AI's capabilities. In a paper by the author [16], the audience sees the strength of an artificial neural network (ANN) forecasting the onset of oral cancer. The ANN impresses with an average sensitivity of 85.71% and specificity of 60.00%, resulting in an overall accuracy of 78.95% from a 10-fold cross-validation analysis, leaving the audience hopeful for the future.

[17] performed network training and evaluation by including bounding box annotations from several clinicians for the classification of images of the oral cavity and detection of the oral lesion using a ResNet-101 and Faster R-CNN model respectively, so utilizing smartphone-based images gathered via a teleconsultation application [18], achieving a sensitivity of 52.13% and a specificity of 49.11% on the multiclass classification task. For the identification of four kinds of OSCC, [19] used four kinds of pre-initialized transfer learning frameworks and also recommended a CNN method, achieving the highest accuracy of 97.5% using the CNN model in comparison to other transfer learning frameworks. [20] suggested a changed Resnet model and got 97.59% accuracy; in [21] they recommended a modified capsule network for oral cancer diagnosis using histopathology pictures and got 97.35% accuracy. [22] proposed a CNN method with histopathology oral cavity images and achieved an accuracy of 97.82%. For greatly enhanced pathologists' diagnosis accuracy, [23] apply VGG16 with a learning rate scheduler and SAM optimizer. [24] sought to assess DL methods for early detection of OC. This paper assessed CNN architectures such as VGG19, Inception-ResNetv2, MobileNet-v2, DenseNet-121, DenseNet-169, DenseNet201 and transformer designs such as Vanilla Vision Transformer (VVT), Data-Efficient Transformers, Swim Transformers.

[25] developed an automated diagnosis approach for categorizing the lesion into benign or malignant using machine and deep learning models on a dataset with a

***Corresponding Author:** Vinaya R Kudatarkar

significant size of well annotated oral lesions. The findings were contrasted after examining classifiers like Naive Bayes, K-Nearest Neighbour (KNN), Support Vector Machine (SVM), Artificial Neural Network (ANN), and CNN. A new CNN classification model was also suggested to replace the current state-of-art for the categorization job. [26] created a technique for distinguishing normal tissues, leucoplakia and erythroplakia where RGB and YCbCr color spaces were taken into account and according the box plots of red and Y color components, a threshold is determined to split the lesions into groups. To create the final judgments, texture characteristics were then obtained from the Gray Level Co-occurrence Matrix (GLCM). Though it could not generalize all the kinds of lesions because of the existence of both white and red areas, the effectiveness was remarkable. Furthermore, the suggested model would not run well for multi-class and binary classification. [27] proposed a two-stage classification and segmentation method for OC detection. The number of samples was increased via random oversampling of the data followed by data augmentation. At last, the categorization and segmentation of OC pictures were taken into account using a CNN-based network called Guided Attention Based Inference Network (GAIN). The GAIN network operates on three streams to reach the objective: the classification stream for sample classification, the attention mining stream for producing the attention maps, and the bounding box stream for enhancing accuracy via pixel-to-pixel supervision.

Where a fusion-based feature extractor was deemed and sent into an extreme learning machine (ELM) for classification, [28] proposed a diagnosis method for the identification of oral cancer as benign or malignant using white light pictures. Although the performance was remarkable, effectiveness may be better and the problem of overfitting has to be explained more clearly. [29] generated a DL architecture for OSCC detection using Swim-Transformer. Gradient-weighted class activation as well as Grad-CAM were used to produce visual explanations stressing the relevant areas for OSCC classification. The performance was good but no cross-validation was taken into account nor validation on outside datasets. At the pixel level, [30] classified normal and malignant nuclei in hyperspectral images of liver tissue slices using support vector machines (SVM). Hyper-net, a neural network, has been created for melanoma segmentation of hyperspectral skin pathology pictures. After H&E staining, acquired hyperspectral imaging on diseased sections and categorized normal cells, precancerous cervical cells, and squamous cells. This study [31] offers a deep learning-based method that classifies the pictures of malignant and non-malignant oral cell photos using the ResNet50 convolution neural network model. It was trained using a collection of histological oral tissue photographs to improve the pretrained ResNet50 model's ability to distinguish between malignant and benign samples.

This article [32] offers a fresh approach to integrate bounding box annotations from several doctors. Deep neural networks were then employed to create automated systems wherein complicated patterns were generated to address this challenging issue. Two deep learning-based computer vision techniques were evaluated for the automated detection and classification of oral lesions for the early detection of oral cancer using the initial data collected in this study; these were image classification with ResNet-101 and object detection with the Faster R-CNN. This work focuses [33] on this gap by presenting a multimodal deep-learning pipeline using multiple data sources, including patient information, which replicates the diagnostic method of doctors in the early diagnosis of oral cancer.

Using cutting-edge image encoders, the study categorizes oral lesions as benign or possibly cancerous. Presented is a performance comparison of six pre-trained deep-learning models—MobileNetV3-Large, MixNet-S, ResNet-50, HRNet-W18-C, DenseNet-121, and Inception_v3. This paper [34] fills in this gap by suggesting a multimodal deep-learning pipeline encompassing many data sources, including patient information, which simulates the diagnostic method of doctors in the early diagnosis of oral cancer. Using modern image encoders, the study categorizes oral lesions as benign or possibly cancerous. Presented is a performance comparison of six pre-trained deep-learning models—MobileNetV3-Large, MixNet-S, ResNet-50, HRNet-W18-C, DenseNet-121, and Inception_v3.

Using transfer learning, we have used and assessed the performance of six deep convolutional neural network (DCNN) models in this work [35] for directly recognizing pre-cancerous tongue lesions using a limited dataset of clinically annotated photographic images to diagnose early symptoms of OCC. With great classification accuracy, DCNN models could tell between benign and pre-cancerous tongue lesions as well as five different kinds of tongue lesions: hairy tongue, fissured tongue, geographic tongue, strawberry tongue, and oral hairy leucoplakia.

Table 1: Survey table

Ref	Method	Advantages	Disadvantages	Research Gap
[15]	Transmission models & heat maps on histology images	Highlighted ROIs, good accuracy (90.9%)	Performance varies across datasets	Lack of cross-dataset validation
[16]	AI-based lymph node metastasis detection	Decent sensitivity and specificity (75.4%, 81%)	Challenging classification scenarios	Needs better generalization on metastasis
[20]	CNN + Texture Mapping + Wavelet Transform	High sensitivity (96.87%)	Lower specificity (71.29%)	Insufficient balance in detection performance
[21]	AlexNet transfer learning on biopsy images	Good accuracy (90.05%)	Shallow network for complex histopathology	Need for deeper, optimized networks
[22]	CNN with *ex vivo* FCM imaging	High specificity (96%)	Low sensitivity (47%)	Lack of robustness in sensitivity
[23]	ANN with 10-fold cross-validation	Fair accuracy (78.95%)	Low specificity (60%)	Needs improvement in negative class prediction
[19]	ResNet101 + Faster R-CNN on smartphone images	Mobile data collection	Low sensitivity and specificity (~52%, ~49%)	Poor generalization in mobile imaging
[31]	CNN and transfer learning frameworks	Best accuracy with CNN (97.5%)	No real-time evaluation	Lack of deployment readiness
[10]	Fusion-based features + ELM	Effective categorization	Overfitting concerns	Requires validation on larger datasets
[11]	Swin-Transformer + Grad-CAM	Visual explainability	No external or cross-validation	Poor model generalization

Prediction

Reference [36] reported that predictive models based on PET radiomics achieved AUCs ranging from 0.71 to 0.84, indicating promising utility in preoperative cancer grade prediction. In [37], AI models reached an AUC of 0.92 with high sensitivity and specificity, outperforming experienced radiologists in diagnosing lymph node metastases in OSCC, underscoring AI's potential in clinical workflows. A machine learning pipeline presented in [38] successfully classified histological grades in OSCC with high accuracy and identified key features—such as granulocyte MAPKAPK2 signalling, stromal CD4+ memory T cell size, and fibroblast distance from the tumour margin—as correlated with clinical outcomes, highlighting new avenues for biomarker discovery. Reference [39] demonstrated that the k-fold validation method outperformed the hold-out method for cancer staging, achieving a low MAE of 0.015. The study found that older patients with more gene mutations had a higher likelihood of advanced-stage cancer and reduced survival rates. Findings in [40] indicated that high-risk oral leucoplakia (OL) patients were nearly four times more likely to progress to oral cancer compared to their low-risk counterparts, with the model maintaining predictive value even after adjusting for age, lesion site, and dysplasia grade.

In [41], both supervised and unsupervised machine learning methods were applied to expand oral cancer gene datasets, identifying crucial genes for gallbladder carcinoma (GBC) prognosis and other oral cancers. A multi-gene IRS model proposed in [42] outperformed conventional indicators in predicting overall and disease-free survival in OSCC, while also demonstrating associations with the tumour microenvironment and potential responsiveness to immunotherapy and chemotherapy. Reference [43] achieved prognostic stratification using seven differentially expressed genes from the ORCA dataset, highlighting the role of tumour-infiltrating lymphocytes (TILs) in shaping the tumour microenvironment and supporting the utility of deep learning for outcome prediction. In [44], clinical photographic analysis achieved 73.6% and 90.9% accuracy on UK and Brazil datasets respectively. The approach improved performance by analysing image patches rather than full images and by localizing predictive regions using class activation maps. According to [45], more than 70% of tumour and serum DNA samples exhibited gene abnormalities, with concordant mutation patterns between serum and tissue. Notably, post-surgery persistence of these abnormalities was linked to higher risks of recurrence and metastasis, indicating their prognostic value. Finally, [46] identified a 3-mRNA signature—CLEC3B, C6, and CLCN1—as a robust survival predictor in OSCC, showing strong 3- and 5-year predictive accuracy in both training and validation datasets and linking neuroactive ligand–receptor interaction as a key pathway in disease progression.

***Corresponding Author:** Vinaya R Kudatarkar

Table 2: Comparison

Ref	Study Focus	Model Type	Modality	Performance	Key Findings	Research Gap
[36]	PET radiomics for cancer grade prediction	Predictive model	PET imaging	AUC: 0.71–0.84	Useful for preoperative cancer grade prediction	Limited validation across diverse patient cohorts
[37]	AI model for lymph node metastasis diagnosis in OSCC	AI classification model	Imaging + AI	AUC: 0.92; High sensitivity & specificity	Outperformed radiologists in detecting metastases	Requires external validation and model explainability
[38]	Histological grade classification in OSCC	Machine learning pipeline	Histology imaging	High accuracy	Identified biomarkers linked to prognosis	Lacks integration of clinical metadata
[39]	Cancer stage prediction using k-fold validation	ML model with validation techniques	Genomic data	MAE: 0.015	Older age + more mutations associated with worse prognosis	Limited generalizability due to demographic bias
[40]	Progression risk in oral leukoplakia (OL) patients	Risk stratification model	Clinical + imaging data	4× increased risk in high-risk patients	Predictive even after adjusting for confounders	Needs real-time deployment and cross-population studies
[41]	Gene identification for GBC and oral cancer	Supervised + unsupervised ML	Gene expression data	Not specified	Identified key prognostic genes	Unclear clinical applicability of discovered genes
[42]	IRS model for survival prediction in OSCC	Multi-gene ML model	Genomic + clinical data	Superior to conventional indicators	Predictive of immunotherapy response	Dependent on large, high-quality genomic datasets
[43]	Prognosis using ORCA dataset with gene expression	Gene expression-based ML model	Transcriptomics	Good stratification by risk groups	TILs are key in prognosis	Model lacks validation on external datasets
[44]	Photographic image analysis for oral lesion detection	Image patch + CAM model	Photographic images	Accuracy: 73.6% (UK), 90.9% (Brazil)	Patch-based CAM improved prediction	Performance may vary across imaging environments
[45]	Gene abnormality tracking via serum and tumor DNA	Genomic tracking model	Serum + tissue DNA	70%+ showed abnormalities	Post-surgery DNA linked to recurrence	Need for standardization of DNA abnormality thresholds
[46]	3-mRNA signature for OSCC survival prediction	3-gene prognostic model	Transcriptomics	High accuracy at 3- and 5-year marks	CLEC3B, C6, and CLCN1 linked to survival	Requires multi-center clinical validation for broad use

RESEARCH GAP

1. Limited External Validation and Generalization: Most deep learning models for OSCC detection are evaluated only on internal or small datasets, lacking validation across diverse external datasets. This limits their generalizability and reliability in real-world clinical applications.
2. Insufficient Multimodal Data Integration: Many studies rely solely on imaging data without incorporating complementary clinical metadata (e.g., patient history, demographic info), despite evidence that such integration can enhance diagnostic accuracy and mimic clinical decision-making.
3. Lack of Robust Multi-Class Classification Approaches: Existing models often focus on binary classification (benign vs. malignant), while

***Corresponding Author:** Vinaya R Kudatarkar

ignoring the potential for stage-wise or multi-class lesion classification (e.g., normal, leukoplakia, erythroplakia, carcinoma), which is crucial for tailored treatment planning.

4. Poor Interpretability and Explainability of Models: Although deep learning models show high accuracy, their black-box nature remains a barrier to clinical trust and adoption. Very few studies incorporate explainable AI techniques (e.g., Grad-CAM, SHAP) validated by medical professionals.
5. Over-Reliance on Transfer Learning Without Architectural Optimization: While transfer learning (e.g., ResNet, AlexNet, VGG16) is widely used, many works lack customization or optimization for oral histopathological data, potentially limiting the model's ability to learn domain-specific features.
6. Limited Real-Time Deployability and Efficiency Considerations: Deep models evaluated under ideal, static conditions often ignore practical deployment issues such as computational efficiency, processing speed, and hardware constraints in resource-limited clinical settings.
7. Underexplored Hybrid Approaches and Feature Fusion Strategies: Very few works explore hybrid models that combine handcrafted features (e.g., texture descriptors) with deep features, which could enhance model robustness and interpretability through complementary representation learning.

CONCLUSION

The timely and accurate diagnosis of Oral Cancer remains a critical challenge in global healthcare, particularly in regions with limited clinical resources and public awareness. This survey has comprehensively explored the landscape of deep learning methodologies applied to oral cancer detection, highlighting how AI-based approaches—especially those utilizing convolutional neural networks, transfer learning, and transformer architectures—are reshaping diagnostic paradigms through automation and precision. We reviewed various imaging modalities and model configurations, identifying both their promising performance and their limitations in terms of generalizability, interpretability, and clinical integration. While numerous studies have demonstrated high diagnostic accuracy in controlled datasets, the translation of these technologies into real-world clinical environments still faces hurdles such as the need for large annotated datasets, robust external validation, and explainable decision-making processes. Moreover, limited focus on multimodal integration and stage-wise classification represents a significant gap in the current body of work. Going forward, collaborative efforts between researchers, clinicians, and engineers are essential to develop intelligent, explainable, and scalable diagnostic tools that can be deployed in real-time and across diverse healthcare settings. This survey aims to serve as a foundational reference for future innovations and encourages the development of clinically viable AI systems that contribute meaningfully to early detection, improved treatment outcomes, and reduced oral cancer mortality.

ACKNOWLEDGEMENTS

I would like to express our sincere gratitude to all those who have supported and contributed to this research project. Primarily, I extend our heartfelt thanks to our guide for her unwavering guidance, invaluable insights, and encouragement throughout the research process. No funding is raised for this research.

REFERENCES

1. Gupta, B., et al. (2023). Global trends and challenges in oral squamous cell carcinoma: Epidemiology and prevention. *Oral Oncology, 138*, 106308.
2. Welikala, R. A., et al. (2023). Community-driven mobile oral cancer screening using deep learning. *International Journal of Cancer, 152*(2), 215–225.
3. Jeyaraj, P., & Nadarajah, M. (2023). Artificial intelligence in oral oncology: A comprehensive review. *Computers in Biology and Medicine, 155*, 106657.
4. Shah, J. P., et al. (2023). Oral cancer burden in low- and middle-income countries. *Journal of Oral & Maxillofacial Research, 14*(1), e3.
5. Patil, S., et al. (2024). AI-based diagnostics in oral pathology: A public health imperative. *Diagnostics, 14*(2), 320.
6. Mohammed, R., et al. (2023). Public awareness and early diagnosis barriers for oral cancer: A systematic review. *Journal of Pathology Informatics, 14*, 2.
7. Lee, J., et al. (2025). Deep learning in oral cancer detection: Bridging the gap between AI and clinical practice. *npj Digital Medicine, 8*, Article 37.
8. Chaudhary, K., et al. (2023). Digital pathology for OSCC detection in India: Challenges and prospects. *Asian Pacific Journal of Cancer Prevention, 24*(5), 1533–1542.
9. Rivera, C. (2023). Delayed diagnosis in oral cancer: A neglected global health issue. *International Journal of Clinical and Experimental Pathology, 16*(1), 101–109.
10. Amarasinghe, H. K., et al. (2024). Impact of late diagnosis on oral cancer survival rates: A 10-year retrospective study. *BMC Cancer, 24*(1), 19.
11. Das, D., et al. (2024). Multi-class oral lesion classification using CNNs and transfer learning frameworks. *Computers in Biology and Medicine, 163*, 107188.
12. Panigrahi, S., & Swarnkar, T. (2024). Modified ResNet for oral cancer histopathological image classification. *Biomedical Signal Processing and Control, 84*, 104932.
13. Flugge, R., et al. (2023). Visual explainability in OSCC classification using Swin-Transformers and Grad-CAM. *IEEE Transactions on Medical Imaging, 42*(6), 1500–1512.

***Corresponding Author:** Vinaya R Kudatarkar

14. Kevin, S., et al. (2024). Guided Attention-Based Inference Network (GAIN) for oral lesion segmentation and classification. *Biomedical Engineering Letters, 14*(2), 177–190.
15. Nanditha, R., et al. (2023). Deep CNN-based framework for early classification of oral lesions. *Expert Systems with Applications, 223*, 119528.
16. Prachi, A., et al. (2024). GLCM and color-space analysis for multi-class lesion separation. *Computers in Biology and Medicine, 162*, 107174.
17. Al Duhayyim, A. M., et al. (2023). Fusion-based ELM classifier for OSCC using white light images. *Computer Methods and Programs in Biomedicine, 230*, 107397.
18. Sukegawa, S., et al. (2023). Enhanced oral cancer detection using VGG16 with SAM optimizer. *Journal of Oral Pathology & Medicine, 52*(2), 150–158.
19. Welikala, R. A., et al. (2023). Smartphone-based deep learning for oral cancer detection: ResNet101 and Faster R-CNN benchmark. *International Journal of Cancer, 153*(4), 650–662.
20. Chen, X., et al. (2023). SVM-based classification of hyperspectral oral tissue images. *Journal of Biomedical Optics, 28*(4), 047001.
21. Siddiqi, K., et al. (2023). Hyperspectral imaging for classification of cervical squamous cells. *Scientific Reports, 13*, 4567.
22. Talwar, R., et al. (2024). Comparative analysis of CNN and transformer models for early oral cancer detection. *Pattern Recognition Letters, 175*, 25–33.
23. Rahaman, M. M., et al. (2023). Texture feature extraction using GLCM for oral lesion detection. *Scientific Reports, 13*, 7892.
24. Zhou, Y., et al. (2025). Combining handcrafted and deep features for histopathological image classification. *Artificial Intelligence Review*. Advance online publication.
25. Singh, D., et al. (2024). Evaluating CatBoost and XGBoost for OSCC classification using hybrid features. *Computational Intelligence and Neuroscience, 2024*, Article 3456123.
26. Hasan, M. K., et al. (2023). Multimodal deep learning for oral cancer detection using clinical and imaging data. *Sensors, 23*(1), 231.
27. Das, S., et al. (2023). CapsNet variants for histopathological oral cancer classification. *Computerized Medical Imaging and Graphics, 104*, 102152.
28. Thomas, J., et al. (2024). LBP-Gabor hybrid features for OSCC image classification using machine learning. *Journal of Medical Imaging and Health Informatics, 14*(2), 210–218.
29. Chen, Y., et al. (2024). Hybrid CNN-BiLSTM architecture for OSCC detection using histological images. *IEEE Access, 12*, 15683–15691.
30. Zhang, L., et al. (2024). Real-time deployment of lightweight CNN models for mobile oral cancer screening. *Computers in Biology and Medicine, 160*, 107055.
31. Goswami, B., Bhuyan, M. K., Alfarhood, S., & Safran, M. (2024). Classification of oral cancer into pre-cancerous stages from white light images using LightGBM algorithm. *IEEE Access, 12*, 31626–31639. https://doi.org/10.1109/ACCESS.2024.3370157
32. Sharma, N., et al. (2021). Multifractal texture analysis of salivary fern pattern for oral pre-cancers and cancer assessment. *IEEE Sensors Journal, 21*(7), 9333–9340. https://doi.org/10.1109/JSEN.2021.3053262
33. Shamim, M. Z. M. (2022). Hardware deployable Edge-AI solution for prescreening of oral tongue lesions using TinyML on embedded devices. *IEEE Embedded Systems Letters, 14*(4), 183–186. https://doi.org/10.1109/LES.2022.3160281
34. Devindi, G. A. I., et al. (2024). Multimodal deep convolutional neural network pipeline for AI-assisted early detection of oral cancer. *IEEE Access, 12*, 124375–124390. https://doi.org/10.1109/ACCESS.2024.3454338
35. Maemoto, H., et al. (2022). Small dose of oral gastrografin for computed tomography-based image-guided brachytherapy in patients with uterine cervical cancer. *Journal of Radiation Research, 63*(1), 107–114. https://doi.org/10.1093/jrr/rrab102
36. Nikkuni, Y., Nishiyama, H., & Hayashi, T. (2024). Prediction of histological grade of oral squamous cell carcinoma using machine learning models applied to 18F-FDG-PET radiomics. *Biomedicines, 12*, 1411.
37. Deng, C., Hu, J., Tang, P., Xu, T., He, L., Zeng, Z., & Sheng, J. (2024). Application of CT and MRI images based on artificial intelligence to predict lymph node metastases in patients with oral squamous cell carcinoma: A subgroup meta-analysis. *Frontiers in Oncology, 14*, 1395159.
38. Einhaus, J., Gaudilliere, D. K., Hedou, J., Feyaerts, D., Ozawa, M. G., Sato, M., Ganio, E. A., Tsai, A. S., Stelzer, I. A., Bruckman, K. C., et al. (2023). Spatial subsetting enables integrative modeling of oral squamous cell carcinoma multiplex imaging data. *iScience, 26*, 108486.
39. Siddalingappa, R., & Kanagaraj, S. (2022). K-nearest-neighbor algorithm to predict the survival time and classification of various stages of oral cancer: A machine learning approach. *F1000Research, 11*, 70.
40. Zhang, X., Gleber-Netto, F. O., Wang, S., Martins-Chaves, R. R., Gomez, R. S., Vigneswaran, N., et al. (2023). Deep learning-based pathology image analysis predicts cancer progression risk in patients with oral leukoplakia. *Cancer Medicine, 12*, 7508–7518.
41. Singh, T., Malik, G., Someshwar, S., Le, H. T. T., Polavarapu, R., Chavali, L. N., Melethadathil, N., Sundararajan, V. S., Valadi, J., Kavi Kishor, P. B., et al. (2022). Machine learning heuristics on gingivobuccal cancer gene datasets reveals key candidate attributes for prognosis. *Genes, 13*, 2379.

***Corresponding Author:** Vinaya R Kudatarkar

42. Zhang, L., & Wang, X. (2022). An immune-related gene signature can predict clinical outcomes and immunotherapeutic response in oral squamous cell carcinoma. *Frontiers in Genetics, 13*, 870133.
43. Kim, Y., Kang, J. W., Kang, J., Kwon, E. J., Ha, M., Kim, Y. K., Lee, H., Rhee, J.-K., & Kim, Y. H. (2021). Novel deep learning-based survival prediction for oral cancer by analyzing tumor-infiltrating lymphocyte profiles through CIBERSORT. *OncoImmunology, 10*, 1904573.
44. Camalan, S., Mahmood, H., Binol, H., Araújo, A. L. D., Santos-Silva, A. R., Vargas, P. A., Lopes, M. A., Khurram, S. A., & Gurcan, M. N. (2021). Convolutional neural network-based clinical predictors of oral dysplasia: Class activation map analysis of deep learning results. *Cancers, 13*, 1291.
45. Hamana, K., Uzawa, K., Ogawara, K., Shiiba, M., Bukawa, H., Yokoe, H., & Tanzawa, H. (2005). Monitoring of circulating tumour-associated DNA as a prognostic tool for oral squamous cell carcinoma. *British Journal of Cancer, 92*, 2181–2184.
46. Cao, R., Li, Q., Wu, Q., Yao, M., Chen, Y., & Zhou, H. (2019). Effect of non-surgical periodontal therapy on glycemic control of type 2 diabetes mellitus: A systematic review and Bayesian network meta-analysis. *BMC Oral Health, 19*, 176.

***Corresponding Author:** Vinaya R Kudatarkar