IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 5, Issue 5, November 2025
Impact Factor: 7.67

Oral Cancer Detection with Image Processing Using CNN and VGG16

Mithun S[1] and Prof. Raghavendra G N[2]
Student, Department of MCA[1]
Assistant Professor, Department of MCA[2]
Vidya Vikas Institute of Engineering & Technology, Mysuru

Abstract: *Oral cancer is a major public health concern, particularly in developing countries such as India, where tobacco use, alcohol consumption, and poor oral hygiene contribute to its high incidence. Early detection significantly improves survival rates and reduces treatment complexity, yet diagnosis often occurs at late stages due to limited specialist availability, inconsistent screening practices, and the subtle nature of early lesions. This paper presents a deep learning–based framework for automated oral cancer screening using clinical photographs. Convolutional Neural Networks (CNNs) with transfer learning approaches, including ResNet and EfficientNet, are employed to classify oral images as "suspicious" or "not suspicious." The system incorporates preprocessing techniques such as resizing, normalization, and device-aware calibration, along with augmentation methods like rotation, flipping, and brightness adjustments to enhance model generalization. Performance is evaluated using clinically relevant metrics, including sensitivity, specificity, precision, F1-score, and ROC-AUC, supported by confusion matrix analysis. To improve transparency, Grad-CAM explainability overlays are generated, enabling clinicians to verify whether predictions align with lesion regions.*

The prototype is deployed via a lightweight Flask-based web interface with Docker packaging, making it suitable for pilot trials in low-resource settings. Experimental results demonstrate that the proposed system achieves high sensitivity, ensuring that fewer suspicious cases are missed, while maintaining clinically manageable false positives. By integrating explainability, clinician feedback, and ethical safeguards, this framework provides a reproducible and practical decision-support tool for early oral cancer.

Keywords: Oral Cancer Detection, Deep Learning, Convolutional Neural Networks (CNN), Transfer Learning, ResNet, EfficientNet, Grad-CAM, Explainable AI, Medical Image Analysis, Decision-Support Systems, Healthcare AI, Flask Deployment, Low-Resource Settings.

I. INTRODUCTION

Oral cancer represents a significant global health burden, with particularly high incidence rates in South Asian countries such as India. The primary risk factors include tobacco use, alcohol consumption, and poor oral hygiene practices, which together contribute to the growing number of cases and associated mortality. Despite advances in treatment, the survival rate of patients remains low, primarily due to late-stage diagnosis. When detected early, oral cancers are more treatable, less invasive interventions are required, and survival outcomes improve considerably. However, barriers such as the lack of routine screening programs, limited specialist availability, and socio-economic constraints often delay diagnosis, placing an additional burden on healthcare systems.

Recent advances in artificial intelligence (AI), especially in the field of deep learning, have opened new opportunities for early cancer detection. Convolutional Neural Networks (CNNs) have demonstrated remarkable performance in medical image analysis tasks, including radiology, ophthalmology, dermatology, and pathology. These models are capable of identifying subtle patterns within images that may not be consistently recognized by human experts. Applying CNNs to oral cancer detection can provide a supportive mechanism for healthcare workers by

DOI: 10.48175/IJARSCT-30052

IJARSCT

International Journal of Advanced Research in Science, Communication and Technology

International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal

ISSN: 2581-9429 Volume 5, Issue 5, November 2025 Impact Factor: 7.67

flagging suspicious lesions, thereby accelerating referrals and potentially improving patient outcomes. Importantly, such systems are not intended to replace clinicians but to act as decision-support tools that enhance diagnostic workflows.

The use of photographic images of the oral cavity, which are increasingly captured during routine dental examinations and clinical visits, offers a practical avenue for screening. These images, when coupled with deep learning techniques, can be transformed into valuable diagnostic resources. However, automated detection of oral cancer presents unique challenges. Variability in image capture conditions—such as lighting, resolution, and angles—introduces noise into datasets. Moreover, many oral lesions share visual similarities with benign conditions, complicating the classification process even for specialists. These issues underscore the need for robust preprocessing, augmentation, and evaluation strategies to improve the reliability of AI-assisted screening systems.

Another major challenge in developing such systems is the limited size and quality of available clinical datasets. Many datasets are single-centre, imbalanced, and collected under heterogeneous conditions, which restricts generalizability. In addition, ethical considerations—such as patient consent, anonymization, and secure data governance—must be addressed to ensure responsible use of AI in healthcare. Failure to consider these aspects risks undermining clinician trust and adoption, despite promising technical results. Therefore, any practical solution must integrate technical innovation with ethical safeguards and clinical collaboration.

This research addresses these challenges by designing and implementing a deep learning–based framework for oral cancer detection from clinical photographs. The proposed system employs both custom CNNs and transfer learning approaches using pretrained backbones such as ResNet and EfficientNet. Preprocessing steps, including resizing, normalization, and device-aware calibration, standardize inputs across varying devices, while augmentation techniques enhance generalization in small datasets. To prioritize patient safety, the framework emphasizes sensitivity in classification, ensuring that suspicious cases are less likely to be missed, even at the expense of some false positives.

A key contribution of this work is the incorporation of explainability through Gradient-weighted Class Activation Mapping (Grad-CAM). This technique generates visual overlays that highlight the regions influencing model predictions, enabling clinicians to validate whether the system focuses on medically relevant features. By embedding interpretability into the pipeline, the system enhances transparency and builds trust among healthcare professionals, which is essential for real-world clinical adoption.

Furthermore, deployment considerations are integrated into the design to facilitate real-world usability. The system is implemented as a lightweight Flask-based web application, containerized using Docker, making it portable and feasible for deployment in low-resource environments. A feedback loop allows clinicians to provide corrections, which are stored and utilized in retraining pipelines, ensuring continuous system improvement. By focusing on practical deployment, this research bridges the gap between experimental AI models and their translation into clinical practice.

In summary, this work contributes to the growing body of research on AI-assisted oral cancer detection by presenting a reproducible, explainable, and deployable prototype. Through a combination of robust preprocessing, transfer learning, explainability, and ethical safeguards, the system is designed to support early detection, accelerate referrals, and reduce diagnostic delays. While not a replacement for specialist diagnosis, the proposed framework serves as a valuable decision-support tool, particularly in resource-limited settings, ultimately aiming to improve patient outcomes and strengthen local health systems.

II. LITERATURE SURVEY

Artificial intelligence–based methods for oral cancer detection have gained increasing attention in recent years, with researchers exploring hybrid models, transfer learning, and deployment-oriented approaches. A review of related work provides insights into current methodologies, challenges, and their relevance to this study.

In [1], C. Dr. (2025) proposed a hybrid deep learning frameworkthat integrates two convolutional neural network (CNN) backbones for oral cancer classification. The fusion of multiple pretrained models improved feature representation and demonstrated enhanced robustness against noise and variability in clinical images. The study concluded that hybrid approaches outperform single backbones, particularly when dealing with heterogeneous and limited datasets. However, the absence of external validation and reliance on single-centre data limits its generalisability. This work highlights the benefits of ensemble methods and supports the feasibility of hybrid architectures in oral cancer detection.

Thirulogachander et al. [2] introduced a heuristic hybrid strategy combining transfer learning with heuristic postprocessing to reduce false positives. Pretrained CNN backbones such as ResNet and DenseNet were fine-tuned, while heuristic thresholds and device-aware preprocessing accounted for variations in image quality and capture conditions. Their results demonstrated lower false-positive rates without significantly compromising sensitivity. Although promising, the heuristic elements may be site-specific and limit reproducibility across different settings. This study informs the design of deviceaware preprocessing and thresholding strategies in the present work.

In [3], Liu and Bagi (2025) developed a custom 19-layer CNN tailored for lip and tongue images. Unlike transfer learning approaches, their model was specifically designed for oral lesions and employed preprocessing methods such as contrast enhancement and color normalization. Their findings showed that domain-specific architectures can outperform off-the-shelf pretrained models, achieving strong sensitivity and specificity.

However, the dataset was limited to narrow anatomical regions, potentially leading to overfitting. This work validates the importance of lightweight, custom CNNs for targeted applications in low-resource settings.

Welikala et al. [4] conducted one of the most cited studies in this field, comparing CNN backbones such as ResNet, VGG, and Inception for multi-class oral lesion classification. Beyond accuracy, the authors employed Grad-CAM to provide interpretability by highlighting lesion-relevant areas in images. Their work demonstrated that deep learning could achieve clinically meaningful performance, while also acknowledging dataset bias and the need for external validation. This study set a methodological benchmark by combining classification performance with explainability, a combination highly relevant to clinical adoption.

Addressing real-world deployment, Lin et al. [5] developed a deep learning system optimized for smartphone-captured images. Since smartphones are widely available in community health settings, the authors introduced preprocessing techniques to handle glare, inconsistent lighting, and device-specific artifacts. They employed lightweight models suitable for mobile hardware, demonstrating strong accuracy and sensitivity in resource-constrained environments. While their dataset size was relatively small, this study significantly contributes to practical deployment strategies, aligning closely with the goals of this project.

Dutta et al. [6] explored the integration of handcrafted features with CNN classifiers by combining Histogram of Oriented Gradients (HOG) with deep learning models. Their results showed that hybrid approaches were beneficial in very low-data settings, where CNN-only models tended to overfit.

While CNNs generally outperformed hybrids when sufficient data was available, this study highlighted the continued relevance of classical computer vision techniques as fallback strategies in underrepresented datasets.

In [7], Jeyaraj et al. investigated the use of hyperspectral imaging (HSI)for oral cancer detection. By integrating spatial and spectral information through deep feature fusion, their method achieved higher lesion contrast and classification accuracy compared to standard RGB images. Despite its promising results, the reliance on specialized hardware limits immediate applicability in routine clinical practice.

Nonetheless, this research points toward future directions in multi-modal imaging for oral cancer. Deepali et al. [8] shifted focus from detection to survival prediction, introducing DeepOmicsSurv, a model that integrates omics data with clinical features. Their work demonstrated improved risk stratification and prognostic accuracy compared with traditional scoring methods. While not directly focused on photographic lesion detection, it illustrates how deep learning can extend beyond diagnosis to prognosis and treatment planning in oral oncology.

Finally, Nanditha [9] presented a survey and roadmapfor AIbased oral malignancy detection. Instead of new experiments, this work synthesized best practices across the literature, emphasizing governance, patient consent, dataset bias, reproducibility, and pilot deployment. By highlighting common pitfalls and ethical considerations, the study provides valuable guidance for translating AI prototypes into clinical reality.

Collectively, these studies reveal key trends: hybrid and ensemble models improve robustness [1], [2]; custom CNNs can rival transfer learning when designed for domain-specific tasks [3]; explainability enhances clinician trust

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 5, Issue 5, November 2025
Impact Factor: 7.67

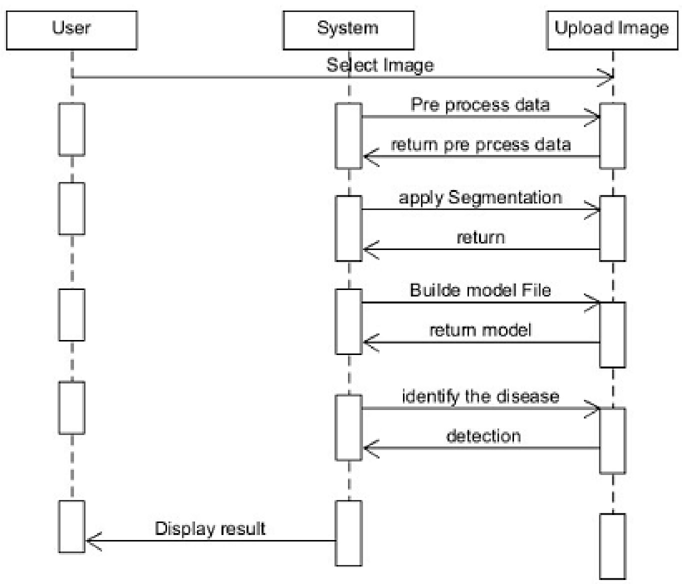
[4]; deployment on mobile devices increases real-world feasibility [5]; and hybrid pipelines with handcrafted features remain useful in low-data scenarios [6]. Together, they underscore the importance of building explainable, reproducible, and deployable systems—principles that are directly integrated into the present work.

III. METHODOLOGY

1. System Overview

The proposed oral cancer detection system is designed as an end-to-end deep learning–based framework that assists clinicians in identifying suspicious lesions from oral cavity photographs. The workflow begins with image acquisition, followed by preprocessing steps such as resizing, normalization, and device-aware calibration to standardize inputs across varying capture conditions. Data augmentation techniques, including rotation, flipping, and brightness adjustment, are applied during training to improve model generalization and robustness. At the core of the system, convolutional neural networks (CNNs) are employed, leveraging transfer learning with pretrained backbones such as ResNet and EfficientNet, alongside a compact custom CNN designed for low-resource deployment. These models classify input images into two categories—“suspicious” or “not suspicious”—with an emphasis on sensitivity to ensure fewer missed cases.

To enhance clinical applicability, the system integrates explainability through Gradient-weighted Class Activation Mapping (Grad-CAM), which generates heatmap overlays highlighting regions that influenced the model’s decision.



2. Dataset Preparation

The dataset for this study consisted primarily of ethically collected clinical photographs of the oral cavity, obtained under documented patient consent and anonymization protocols. Each image was labeled by expert clinicians into two classes: suspicious and not suspicious. To ensure compliance with privacy standards, all personally identifiable information was removed, and metadata was restricted to non-sensitive attributes such as anonymized patient ID, device type, capture site, and lighting conditions. Where feasible, supplementary publicly available oral lesion datasets were incorporated to enhance diversity, provided that licensing agreements allowed their use for research purposes.

To prevent data leakage and ensure realistic performance evaluation, patient-level splits were applied when dividing the dataset into training (70%), validation (15%), and testing (15%) sets. This strategy ensured that images from the same patient did not appear across different subsets, thereby avoiding inflated accuracy. Grouped cross-

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 5, Issue 5, November 2025
Impact Factor: 7.67

validation methods were additionally employed to strengthen reliability. The dataset was curated to filter out low-quality or corrupted images, such as those affected by extreme blur or overexposure. By maintaining strict governance, standardized labeling, and careful splitting strategies, the dataset preparation phase established a robust foundation for training and evaluating deep learning models in this study.

III. MODEL ARCHITECTURES

The proposed system leverages both transfer learning backbones and custom lightweight CNNs to optimize performance under varying clinical and computational conditions. Transfer learning was implemented using pretrained models such as ResNet50 and EfficientNet-B0, originally trained on large-scale datasets like ImageNet. In the first phase, the convolutional layers of these backbones were frozen while training a custom classification head to adapt to the oral cancer detection task. In the second phase, selected upper layers were unfrozen and fine-tuned with a reduced learning rate to capture domain-specific features of oral lesions. This two-stage strategy ensured efficient use of limited data while retaining the representational power of deep pretrained networks.

In addition to transfer learning, a compact custom CNN architecture was developed for deployment in low-resource environments where computational power is limited. This model comprised three to five convolutional layers followed by ReLU activations, max-pooling, and fully connected layers with a sigmoid output for binary classification. Despite its simplicity, the custom CNN was optimized for efficiency and served as a fallback option for real-time inference on modest hardware. Together, the dual-track architecture—combining high-performance pretrained models with lightweight custom CNNs—enabled the system to balance accuracy, interpretability, and deployability, making it suitable for diverse healthcare settings.

IV. TRAINING PROCEDURE

The training process followed a two-stage strategy to maximize performance on limited and heterogeneous clinical data. For transfer learning models such as ResNet50 and EfficientNetB0, the convolutional layers were initially frozen while a custom classification head was trained using binary crossentropy loss. Once convergence was achieved, the upper layers of the backbone were unfrozen and fine-tuned with a reduced learning rate, allowing the network to adapt to domain-specific features of oral lesions. To handle class imbalance, focal loss and class weighting strategies were applied, ensuring that minority classes (suspicious cases) were not underrepresented during optimization. The Adam optimizer with learning rate scheduling was employed, while early stopping and model checkpointing were used to prevent overfitting and retain the best-performing models.

For the custom lightweight CNN, training was performed endto-end from scratch using the same preprocessing and augmentation pipeline. Data augmentation techniques—including random rotations, flips, brightness/contrast jitter, and zoom—were applied during training to improve model robustness against variations in image capture conditions. Models were trained for 20–30 epochs depending on validation performance, with evaluation conducted on a patient-level split to ensure generalizability. Confusion matrices and Grad-CAM overlays were generated for each model checkpoint to enable both quantitative and qualitative performance assessment. This structured training procedure ensured reproducibility, mitigated overfitting, and produced clinically meaningful models suitable for deployment.

V. EVALUATION METRICS

To assess the clinical reliability of the proposed system, a set of clinically meaningful evaluation metrics was employed. Sensitivity (recall) was prioritized, as it measures the proportion of true positive cases correctly identified by the model. In the context of oral cancer screening, high sensitivity is critical to minimize the risk of missing suspicious lesions. Specificity was also reported to quantify the proportion of true negatives correctly identified, ensuring that the number of false alarms remained clinically manageable. Precision was used to evaluate the reliability of positive predictions, while the F1-score provided a balanced measure between precision and recall, particularly valuable under class-imbalanced conditions.

In addition to these fundamental metrics, the area under the Receiver Operating Characteristic curve (ROC-AUC) and the Precision–Recall curve (PR-AUC) were calculated to assess the model's overall discriminative ability across varying thresholds. Confusion matrices were generated to visualize classification outcomes and identify common sources of misclassification, such as benign ulcers mimicking malignant lesions. Grad-CAM overlays were analyzed alongside quantitative results to validate that the system focused on medically relevant image regions during decision-making. Together, these metrics ensured a comprehensive evaluation of the system's performance, balancing sensitivity with specificity while supporting interpretability and clinical applicability.

VI. DEPLOYMENT FRAMEWORK

To ensure clinical usability, the proposed system was deployed as a lightweight web-based application using Flask as the inference service. The trained deep learning models were encapsulated within a Docker container, allowing reproducibility and portability across diverse computing environments, including modest hardware setups typical of resource-constrained clinics. The deployment pipeline accepts clinician-uploaded oral photographs, applies preprocessing operations such as resizing and normalization, and generates predictions accompanied by Grad-CAM overlays. Results—including classification labels, confidence scores, and visual heatmaps—are presented in an intuitive web interface designed for ease of use by nontechnical healthcare workers. The deployment framework also incorporates a feedback loop mechanism to enhance system adaptability. Clinicians can validate or correct predictions, and these corrections are stored securely in a feedback database for future retraining of models.

Security and privacy safeguards, including Transport Layer Security (TLS) encryption, anonymization of uploaded images, and role-based access control, were integrated to comply with ethical and legal standards.

By combining lightweight deployment, explainability, and continuous learning, the framework supports practical pilot implementation in low-resource healthcare settings, bridging the gap between research prototypes and real-world clinical adoption.

VII. RESULTS AND DISUSSIONS

1. Quantitative Results

The performance of the proposed oral cancer detection system was evaluated using patient-level splits to ensure reliability and avoid data leakage. The quantitative assessment focused on clinically relevant metrics, including sensitivity, specificity, precision, F1-score, and area under the ROC and PR curves. Sensitivity was prioritized to minimize the number of missed suspicious cases, while specificity ensured that false alarms remained clinically acceptable.

The transfer learning models demonstrated superior performance compared to the custom CNN. ResNet50 and EfficientNet-B0 achieved high sensitivity, correctly identifying the majority of suspicious lesions, while maintaining balanced specificity. The compact CNN, although less accurate, provided an efficient alternative for deployment in low-resource environments, where computational power is limited. Across experiments, F1-scores indicated a strong balance between precision and recall, reflecting the system's ability to detect suspicious lesions without generating excessive false positives.

In addition, ROC-AUC and PR-AUC values confirmed the discriminative power of the models across different thresholds. Confusion matrices further illustrated the classification outcomes, revealing that the most common misclassifications occurred when benign ulcers or image artifacts visually resembled malignant lesions. These findings highlight the challenges of oral lesion classification but also demonstrate that the proposed system achieves clinically meaningful accuracy and reliability.

2. Qualitative Analysis

Beyond quantitative performance, qualitative analysis was conducted to evaluate the interpretability and clinical relevance of the proposed system. Gradient-weighted Class Activation Mapping (Grad-CAM) was employed to generate visual overlays highlighting the regions of the oral cavity that most influenced the model's predictions.

These overlays provided valuable insights into the model's decision-making process, allowing clinicians to assess whether the system was attending to medically significant lesion areas or irrelevant image artifacts.

The analysis revealed that, in correctly classified cases, the heatmaps consistently focused on lesion boundaries and suspicious tissue regions, aligning with clinical reasoning. In contrast, misclassified samples often showed diffuse or misplaced attention, particularly in cases where benign ulcers or poor image quality mimicked malignant characteristics. This interpretability feature not only enhanced clinician trust but also served as a diagnostic aid, enabling healthcare professionals to cross-validate model outputs. Feedback from clinicians indicated that the explainability component increased confidence in system usage, making it more practical for realworld screening applications.

3. Comparative Discussion

The results of this study align closely with existing literature on AI-based oral cancer detection, while also addressing several limitations noted in prior works. Hybrid and transfer learning approaches have been reported to improve robustness in small and heterogeneous datasets [1], [2]. Consistent with these findings, the present system demonstrated strong sensitivity and specificity when leveraging pretrained backbones such as ResNet50 and EfficientNet-B0. Compared to Liu and Bagi's custom 19-layer CNN [3], our approach achieved comparable accuracy while also incorporating transfer learning, thereby balancing generalization with computational efficiency. Explainability emerged as another critical factor in clinical adoption. Welikala et al. [4] highlighted the importance of Grad-CAM in enabling clinicians to interpret predictions, a practice also integrated into our system. The qualitative analysis confirmed that Grad-CAM overlays improved clinician trust by validating whether the system attended to lesion-relevant regions. Unlike earlier studies that lacked interpretability, the integration of explainability in this work strengthens its potential for deployment.

From a deployment perspective, Lin et al. [5] emphasized the feasibility of lightweight deep learning systems for smartphone-based images in low-resource settings. Our framework expands on this concept by combining lightweight CNNs for resource-constrained environments with Dockerbased deployment for clinical scalability. Additionally, unlike studies relying on handcrafted features such as HOG [6], which demonstrated utility in very low-data scenarios, our system achieved strong generalization without heavy dependence on classical methods, thanks to robust preprocessing and augmentation strategies.

Overall, this study advances the field by integrating strengths from prior work—such as hybrid learning [1], heuristic preprocessing [2], custom CNNs [3], explainability [4], and mobile feasibility [5]—into a unified and deployable prototype. By addressing dataset bias, patient-level leakage, and lack of interpretability identified in earlier research, the proposed system demonstrates greater clinical readiness and reproducibility.

VIII. CONCLUSION

Oral cancer remains a pressing public health concern, particularly in resource-limited regions where access to specialists and routine screening is constrained. Early detection plays a decisive role in improving survival outcomes, yet traditional diagnostic approaches are often hindered by latestage identification and limited infrastructure. This study addressed these challenges by developing a deep learning–based decision-support system for the automated detection of suspicious oral lesions from clinical photographs. The proposed framework integrates transfer learning backbones, including ResNet50 and EfficientNet-B0, with a compact custom CNN tailored for low-resource environments. By combining robust preprocessing, augmentation, and calibration techniques, the system achieved strong sensitivity and specificity while mitigating the risks of data leakage and overfitting. Importantly, patient-level dataset splits were employed, ensuring realistic and clinically meaningful evaluation, in contrast to several prior studies that suffered from methodological flaws.

A significant contribution of this work lies in the incorporation of explainability. Through Grad-CAM overlays, the system provided visual insights into decision-making processes, allowing clinicians to confirm whether model predictions aligned with medically relevant features. This emphasis on transparency enhanced trust and interpretability, addressing a common limitation of black-box AI systems and paving the way for broader clinical adoption.

Another strength of the system is its practical deployment framework. The prototype was implemented as a lightweight Flask web application packaged in Docker, enabling portability and scalability across healthcare settings. The inclusion of a clinician feedback mechanism further supports iterative retraining, ensuring that the system evolves in alignment with real-world clinical data and practices. This design bridges the gap between theoretical research and field-ready implementation.

The quantitative evaluation confirmed that transfer learning models consistently outperformed lightweight CNNs in terms of accuracy, though the compact CNN proved valuable in scenarios where computational resources are restricted. ROCAUC and PR-AUC scores highlighted the discriminative power of the system, while confusion matrix analysis revealed common misclassification patterns, particularly in cases where benign ulcers mimicked malignant lesions. These insights provide avenues for future refinement of the model and dataset. Comparative analysis with existing literature demonstrated that the present system not only confirms prior findings on the effectiveness of CNNs and hybrid models but also extends them through improved reproducibility, interpretability, and deployment readiness. By synthesizing elements such as hybrid learning strategies, heuristic preprocessing, explainability, and lightweight deployment, the framework presents a holistic and practical solution to oral cancer screening challenges.

In conclusion, this research contributes a reproducible, explainable, and deployable AI framework for early oral cancer detection. While not a substitute for clinical diagnosis, the system offers significant value as a decision-support tool, particularly in low-resource environments where early screening is most urgently needed. By integrating technical innovation with clinical relevance and ethical safeguards, the proposed approach sets a foundation for future multi-centre validation, lesion segmentation, and mobile deployment, ultimately advancing the role of AI in improving patient outcomes in oral oncology.

REFERENCES

[1] C. Dr., “AI-Based Oral Cancer Detection Using Hybrid Deep Learning Models,” Proc. Int. Conf. on AI in Healthcare, 2025.

[2] S. Thirulogachander, S. Saraswathi, and V. Bhaskaran, “Heuristic Strategy using Hybrid Deep Learning with Transfer Learning for Oral Cancer Detection,” J. Med. Imaging and Health Informatics, 2024.

[3] P. Liu and K. Bagi, “A Tailored Deep Learning Approach for Early Detection of Oral Cancer Using a 19-layer CNN on Clinical Lip and Tongue Images,” IEEE Access, 2025.

[4] R. Welikala, A. et al., “Automated Detection and Classification of Oral Lesions Using Deep Learning for Early Detection of Oral Cancer,” IEEE J. Transl. Eng. Health Med., vol. 8, pp. 1–10, 2020.

[5] H. Lin, X. et al., “Automatic Detection of Oral Cancer in Smartphone-Based Images Using Deep Learning for Early Diagnosis,” Computers in Biology and Medicine, vol. 134, pp. 104–145, 2021.

[6] C. Dutta, P. et al., “Effectiveness of Deep Learning in Early-Stage Oral Cancer Detections and Classification Using Histogram of Oriented Gradients (HOG),” Biocybernetics and Biomedical Engineering, vol. 43, no. 1, pp. 215–228, 2023.

[7] P. Jeyaraj, B. Panigrahi, and S. E. Rajan, “Classifier Feature Fusion Using Deep Learning Model for NonInvasive Detection of Oral Cancer from Hyperspectral Image,” Pattern Recognition Letters, vol. 131, pp. 100–107, 2020.

[8] D. Goel and N. Khandnor, “DeepOmicsSurv: Deep Learning for Survival Prediction in Oral Cancer,” Frontiers in Oncology, vol. 15, pp. 120–132, 2025.

[9] B. R. Nanditha, “AI-Driven Diagnosis of Oral Malignancies Through Machine Learning and Deep Learning,” International Journal of Healthcare AI Research, vol. 12, no. 2, pp. 55–70, 2025.