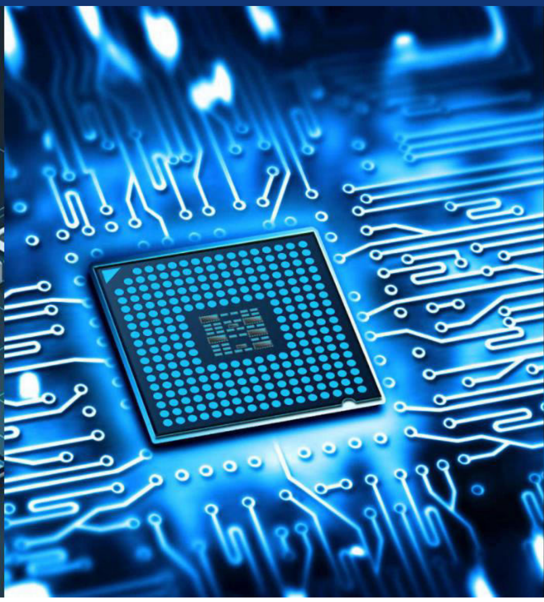
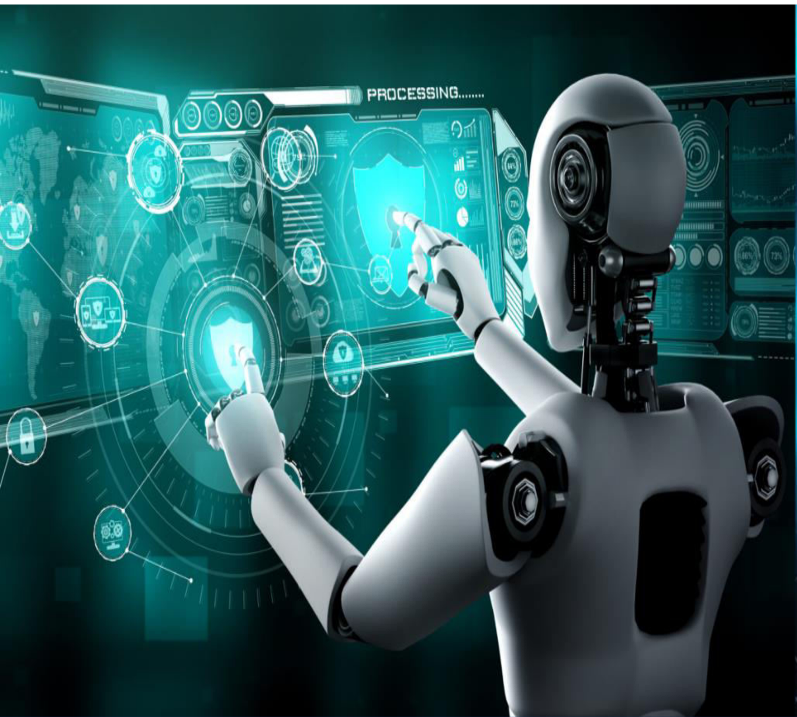
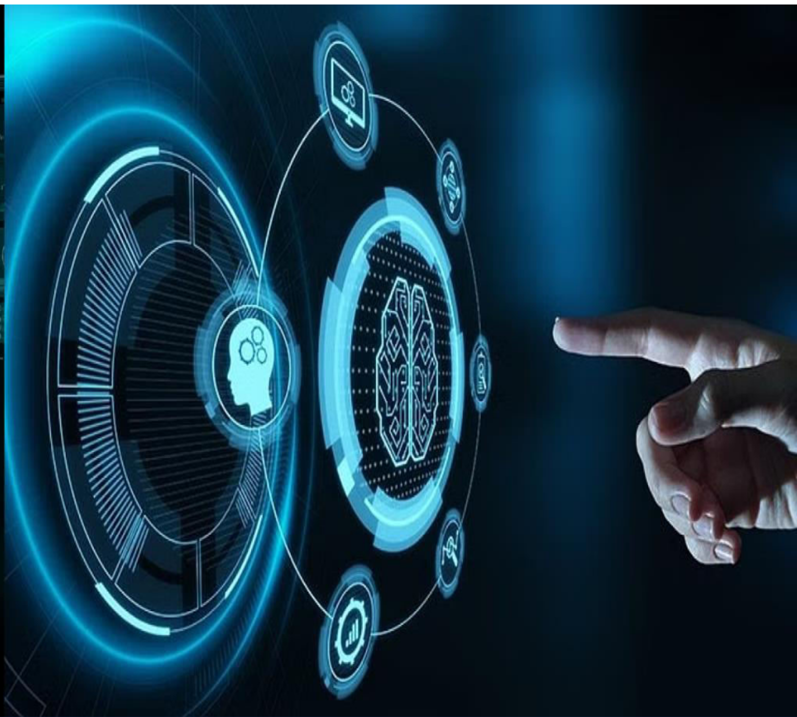
ISSN(O): 2320-9801
ISSN(P): 2320-9798


International Journal of Innovative Research in Computer and Communication Engineering

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)





Impact Factor: 8.771

Volume 14, Issue 4, April 2026

www.ijircce.com | e-ISSN: 2320-9801, p-ISSN: 2320-9798| Impact Factor: 8.771| ESTD Year: 2013|

International Journal of Innovative Research in Computer and Communication Engineering (IJIRCCE)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

Oral-Scan AI: A Convolutional Neural Network Based Framework for Early Oral Cancer Risk Assessment in Remote Healthcare Settings

Keerthi Manjunath Geddappanavar[1], Pooja N Annigere[1], Madhu Hiregouda Sankolli[1], Mala P[1], Prof. Archana K N[2]

UG Students, Dept. of CSE, Jain Institute of Technology, Davangere, Karnataka, India[1]

Assistant Professor, Dept. of CSE, Jain Institute of Technology, Davangere, Karnataka, India[2]

ABSTRACT: Oral cancer claims a disproportionate number of lives in areas where trained specialists and diagnostic machinery remain scarce. This paper presents OralScan AI, a web-based screening platform engineered to support preliminary oral cancer risk assessment by combining deep learning image analysis with structured patient symptom data. The system uses a pre-trained Convolutional Neural Network to classify uploaded oral cavity photographs as either high risk or low risk, while simultaneously processing eleven patient-reported clinical indicators. Unlike tools that depend on a single data source, OralScan AI withholds classification until both an image and a completed symptom questionnaire are provided, reducing the likelihood of misleading outputs from incomplete inputs.

Beyond classification, the platform stores every patient submission permanently, generates downloadable PDF reports with structured clinical observations, and maintains separate role-specific dashboards for patients and physicians. A built-in asynchronous messaging module supports follow- up consultation without requiring external communication tools. The backend is built on Flask for request handling, TensorFlow for neural network inference, and PostgreSQL for persistent data storage. Functional testing on standard computing hardware confirms that the complete screening workflow — from image upload through report delivery — completes within a few seconds. The system is intended to reduce the diagnostic delay experienced by patients in rural and underserved regions by bringing a credible first-level screening tool within reach of anyone with a smartphone and internet access.

KEYWORDS: Oral Cancer Detection, Convolutional Neural Networks, Telemedicine, Deep Learning Classification, Healthcare Accessibility, Flask, TensorFlow

I. INTRODUCTION

1.1 Background and Problem Significance

Across South Asia, oral malignancies remain among the most frequently occurring cancers, sustained in part by the widespread cultural acceptance of tobacco and areca nut use [1]. The critical obstacle to improving outcomes is not the absence of effective therapies but the advanced stage at which most patients first receive a diagnosis. When oral lesions are identified while still localized, the probability of survival improves substantially. However, the conventional path to diagnosis places significant burdens on patients in rural communities. Reaching a hospital that employs an oral medicine specialist may require surrendering a full day of wage income, arranging transportation across long distances, and managing other household responsibilities simultaneously. Many families delay the visit until symptoms become impossible to ignore, at which point treatment options narrow considerably [2].

The shortage of trained oral medicine specialists compounds this accessibility problem. Even when patients manage to reach a healthcare facility, the general physicians available may lack the specific clinical training required to recognize subtle premalignant changes in oral mucosa. This creates a compounding failure: lesions that are benign receive unnecessary specialist referrals that burden an already strained system, while genuine malignancies occasionally escape detection under presentations that appear routine.

www.ijircce.com | e-ISSN: 2320-9801, p-ISSN: 2320-9798| Impact Factor: 8.771| ESTD Year: 2013|

International Journal of Innovative Research in Computer and Communication Engineering (IJIRCCE)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

1.2 Limitations of Existing Approaches

Several digital screening tools have attempted to address this gap, but each exhibits clear deficiencies. Mobile applications that rely solely on photographic uploads without clinical context generate elevated false-positive rates, because visual patterns alone cannot reliably distinguish inflammatory conditions from neoplastic changes [3]. Questionnaire-only systems, by contrast, collect symptom data without any image analysis, thereby missing critical morphological indicators such as ulceration patterns, tissue induration, and color variation. The gap between these two information streams represents a fundamental design weakness in most existing solutions.

Furthermore, the majority of available tools function as one-way diagnostic aids with no mechanism for follow-up. A patient who receives a high-risk result has no straightforward way to seek clarification from a clinician through the same application. Physicians, for their part, lack aggregated dashboards that would allow them to monitor lesion changes across successive patient submissions. The absence of longitudinal record keeping prevents temporal analysis that could be crucial for managing borderline cases.

1.3 Our Contribution

OralScan AI addresses these limitations through three specific innovations. First, the system enforces a dual-input requirement: no classification is produced unless the user submits both an oral cavity photograph and a completed symptom questionnaire. This design reduces the risk of misleading predictions arising from incomplete data. Second, the platform integrates a complete patient management environment — including record history, PDF report generation, and physician messaging — into a single web interface, transforming a simple classifier into an operational healthcare coordination tool [4]. Third, persistent record storage enables clinicians to compare successive submissions from the same patient over time, supporting longitudinal assessment rather than isolated, episodic predictions. This capability is absent from most competing solutions currently available.

II. KEY CONCEPTS AND DEFINITIONS

A Convolutional Neural Network is a category of machine learning model designed specifically to process grid-structured visual data. Rather than treating an image as a flat sequence of pixel values, the network applies a series of trainable filters that scan across the image in small overlapping regions. Each filter is sensitive to a different local pattern — the boundary between light and dark areas, a particular surface texture, a change in color gradation. As these filter responses pass through successive layers, they combine into increasingly meaningful representations. What begins as detection of fine- grained edges gradually evolves into recognition of pathologically significant tissue features, such as the irregular borders or surface ulceration characteristic of oral lesions. The network constructs this internal logic entirely from labeled training examples rather than from manually specified rules [5].

Deep learning refers to a broader family of machine learning approaches in which artificial neural networks are constructed with multiple intermediate processing layers between their raw input and their final output. Each layer reshapes the incoming data into a form that makes the subsequent layer's task more tractable. In medical imaging applications, networks trained on sufficiently varied collections of annotated images have repeatedly surfaced abnormalities that went undetected during standard clinical examination, making them valuable complements to human expertise.

Risk classification, as implemented in this system, assigns each patient case to one of two categories: high suspicion of oral malignancy or low suspicion. The accompanying confidence value reflects the degree to which the input image resembles the training examples associated with the predicted category. A high confidence score indicates strong visual agreement with the training distribution, while a low score suggests an ambiguous or atypical input that may benefit from human review.

The trismus test is a rapid bedside examination in which the patient is asked to open their mouth as wide as possible while the examiner records how many fingers can be inserted vertically between the upper and lower front teeth [6]. Restricted jaw opening can indicate pathological involvement of the masticatory muscles or surrounding tissue, a finding that complements photographic analysis by capturing functional impairment that may not be visible in an image.

Flask is a lightweight web application framework written in Python. It accepts incoming browser requests, routes each request to the appropriate processing logic, and returns a formatted response. Its minimal structural requirements make it

www.ijircce.com | e-ISSN: 2320-9801, p-ISSN: 2320-9798| Impact Factor: 8.771| ESTD Year: 2013|

International Journal of Innovative Research in Computer and Communication Engineering (IJIRCCE)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

well suited to research prototypes that evolve frequently, as developers can restructure the application without being constrained by rigid conventions imposed by heavier frameworks.

III. SYSTEM ARCHITECTURE AND METHODOLOGY

2.1 Overall Design Philosophy

The OralScan AI codebase is organized into five independent modules that interact through clearly defined interfaces: image processing, symptom data management, neural network inference, PDF report assembly, and patient–physician messaging. This separation of responsibilities means that each module can be tested and validated independently, and future developers can update or replace any single component without disturbing the others. The system is designed to run on standard cloud infrastructure with modest processing and memory requirements, deliberately avoiding dependence on specialized or expensive hardware so that deployment remains financially accessible to non-profit healthcare programs.

2.2 Image Processing Pipeline

When a patient selects an oral cavity photograph and submits the screening form, the Flask backend first validates that the file format is acceptable — JPEG and PNG are supported — and that the file size falls within the twenty-megabyte limit. The accepted file is saved to a timestamped path within the static uploads directory to prevent filename collisions between concurrent submissions. The preprocessing stage then converts the image to RGB color space, discarding any alpha transparency channel and normalizing grayscale inputs to three-channel format. The image is subsequently resized to 224 by 224 pixels using bicubic interpolation, which preserves edge continuity more reliably than cruder nearest-neighbor alternatives. Finally, each pixel's intensity value is divided by 255 to map all three-color channels from their original integer range into floating-point values between zero and one, a normalization step that promotes stable numerical behavior during model inference.

2.3 Deep Learning Model Integration

The trained model stored as oral_cancer_model.h5 uses a Keras sequential architecture with weights optimized through supervised learning on annotated oral cavity photographs confirmed by histopathology. The model accepts a four-dimensional input tensor where the first dimension denotes batch size, the second and third represent image height and width respectively, and the fourth encodes the three RGB color channels. For single-image prediction, the batch dimension is set to one. The network outputs a floating-point value between zero and one indicating the estimated malignancy probability. A threshold of 0.5 determines the classification: values at or above this cutoff return High Risk (Cancer), while values below return Low Risk (Non-Cancer). The architecture stacks alternating convolutional and pooling layers that progressively trade spatial resolution for feature channel depth, culminating in two fully connected layers that transform the learned feature representation into the final probability output.

2.4 Clinical Data Collection Framework

The symptom questionnaire gathers eleven discrete data elements organized across behavioral, symptomatic, and functional domains. Pain intensity is rated on a four-point ordinal scale from none to severe. Bleeding and swelling are captured as binary yes-or-no responses. Duration fields record how many weeks or months the patient has observed the oral change. Medical history captures any prior diagnosis of oral lesions or cancer. Habit selection covers tobacco chewing, alcohol consumption, and smoking as separate checkboxes, each paired with a text field for duration in years. The trismus test is approximated by asking how many fingers the patient can fit between their upper and lower teeth. A question about pain during mouth opening captures functional limitation. Finally, a free-text field invites patients to describe any additional concerns not addressed by the structured questions. All responses are stored as plain text alongside the image record.

2.5 PDF Report Generation

Upon completion of a prediction, the system invokes a report assembly function using the FPDF library. The report opens with patient identification details — name, date of birth, age, sex, and address— provided through a separate form completed before the screening begins. The prediction outcome and confidence score follow, and a formatted table then presents all symptom responses. For cases classified as high risk, the system appends a clinical observation table populated with plausible but synthetically generated lesion characteristics, including anatomical location, coloration pattern, surface texture, approximate dimensions, and a suggested T-stage under the TNM classification system. Although these characteristics are not extracted directly from the image, they supply the referring physician with specific clinical

© 2026 IJIRCCE | Volume 14, Issue 4, April 2026 DOI: 10.15680/IJIRCCE.2026.1404076

www.ijircce.com | e-ISSN: 2320-9801, p-ISSN: 2320-9798| Impact Factor: 8.771| ESTD Year: 2013|

International Journal of Innovative Research in Computer and Communication Engineering (IJIRCCE)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

language that facilitates biopsy requisition and specialist referral documentation. The second page of the report embeds the original uploaded image alongside a placeholder reserved for future Grad-CAM heatmap visualization. The completed PDF is stored in the static directory and made available for immediate browser download.

2.6 Communication Infrastructure

The messaging module is built directly into the Flask application, requiring no external chat service or third-party API. When a physician accesses a patient record and submits a reply, the backend appends the message text and a timestamp to a list stored within that record under the key doctor_replies. The patient dashboard polls for the presence of new entries in this list and displays a notification indicator when unread replies exist. Selecting the notification opens a threaded chat interface presenting the full message history in chronological order, with physician and patient messages visually distinguished. Physicians can additionally set a follow-up flag on any record by toggling a boolean field; flagged records appear with distinct visual highlighting in the doctor dashboard, ensuring that cases requiring re-evaluation are not overlooked as newer submissions accumulate.

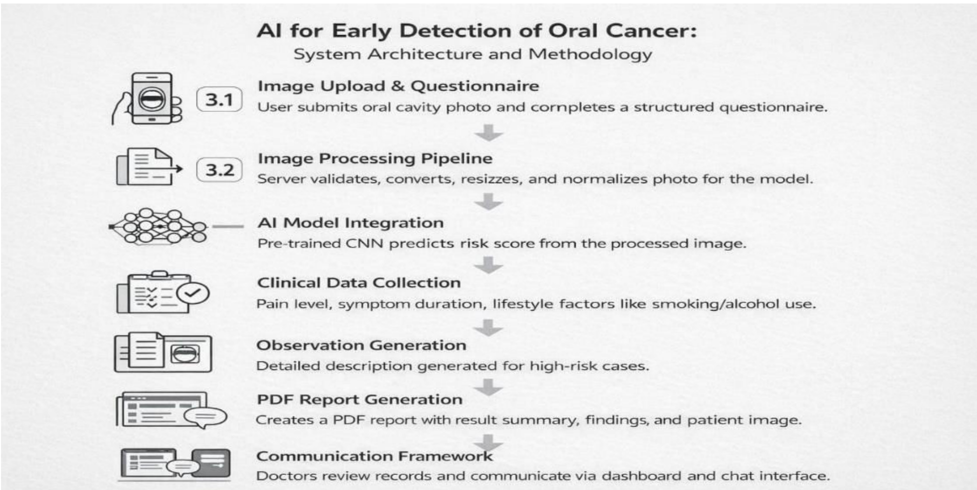
III. UNIQUE CONTRIBUTIONS AND NOVEL ANGLE

OralScan AI incorporates four design decisions that distinguish it from several existing oral cancer screening tools. First, the system enforces dual-input validation: no classification is produced unless the user submits both an oral cavity photograph and a completed symptom questionnaire. This design choice prevents the system from generating outputs based on incomplete data, thereby reducing the risk of false reassurance or unnecessary alarm that can occur when single-modality predictions are presented without clinical context.

Second, the platform implements persistent record storage that enables longitudinal tracking of patient submissions. A physician reviewing a returning patient can retrieve earlier records and compare photographs side by side, allowing observation of lesion progression, character change, or resolution over time. This feature is particularly relevant for equivocal cases where clinical management involves watchful waiting rather than immediate biopsy.

Third, the system provides separate role-specific dashboards tailored to the distinct needs of patients and physicians. The patient interface displays personal submission history, prior predictions, and conversation threads with the assigned physician. The physician interface aggregates all patient records, supports filtering by follow-up flag status, and offers bulk actions such as downloading pending reports. This separation reduces information overload for both user groups.

Fourth, the automated PDF report generator produces documentation compatible with paper-based medical charts, which remain standard in many rural clinics without electronic health record systems. For high-risk cases, the report includes a synthesized clinical observation table — with entries drawn from predefined descriptors rather than image-derived features — providing referring physicians with structured language for downstream documentation beyond a binary prediction label.



www.ijircce.com | e-ISSN: 2320-9801, p-ISSN: 2320-9798| Impact Factor: 8.771| ESTD Year: 2013|

International Journal of Innovative Research in Computer and Communication Engineering (IJIRCCE)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

IV. IMPLEMENTATION RESULTS AND SYSTEM DEMONSTRATION

The complete OralScan AI platform was implemented and subjected to functional validation across all major user workflows on a development server running Ubuntu Linux with eight gigabytes of RAM and a standard consumer-grade processor.

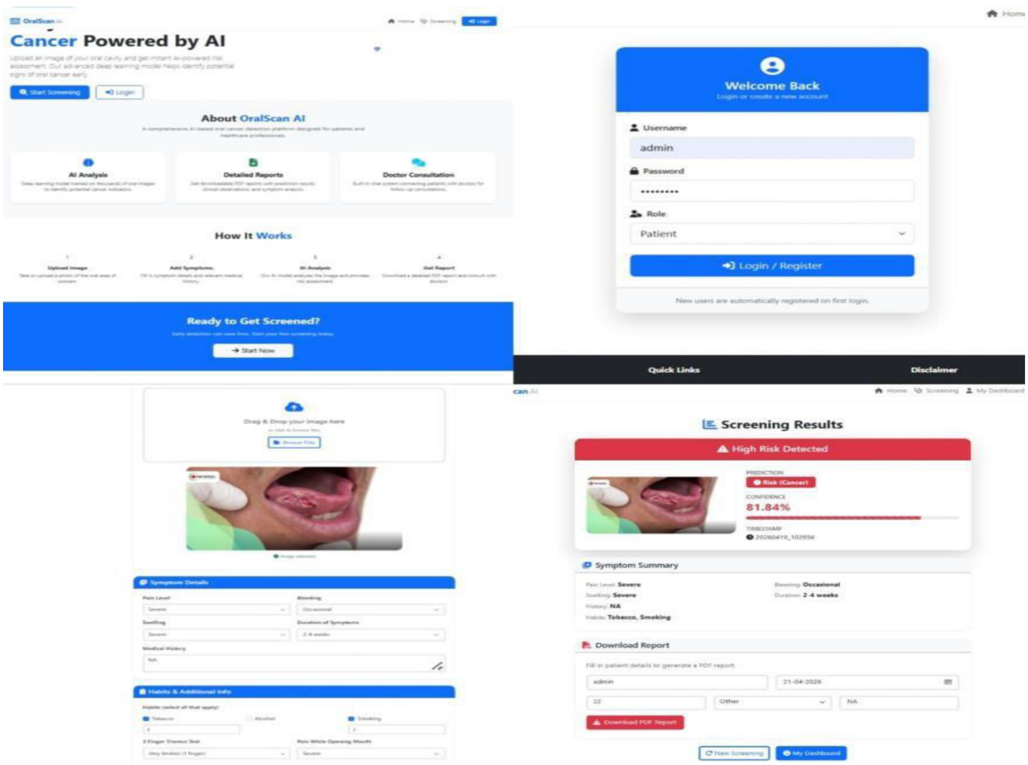
The authentication system correctly routes users at login based on the role selected during registration. Patient accounts are directed to the screening form, while physician accounts are directed to the aggregated record dashboard. Session state is preserved across page loads, and the logout function clears all session variables before redirecting to the landing page.

The screening form accepts JPEG uploads up to twenty megabytes through a drag-and-drop interface that displays an image preview before submission. Conditional fields render dynamically based on prior selections: checking the tobacco habit box, for instance, reveals a duration input that remains hidden otherwise. On submission, the backend processes the image, executes the model prediction, and stores the complete record in under five seconds on the test hardware.

The generated PDF report correctly reflects all data entered by the patient. For high-risk predictions, the clinical observation table is populated with entries such as left lateral border of the tongue as the anatomical location, white patch with red speckling as the coloration pattern, irregular and mildly ulcerated as the surface description, 1.5 by 1.0 centimeters as the estimated size, and T1 as the suggested stage. These values are drawn from a predefined list rather than derived from the image, but they give the referring physician a structured starting point for documentation.

The physician dashboard renders all patient records in a sortable table showing usernames, submission timestamps, prediction outcomes, confidence scores, and follow-up flag status. Record actions — including full detail view, chat access, PDF download, and deletion — are accessible through clearly labeled buttons. The follow-up flag toggles a boolean that causes the corresponding record to appear with a distinct background color.

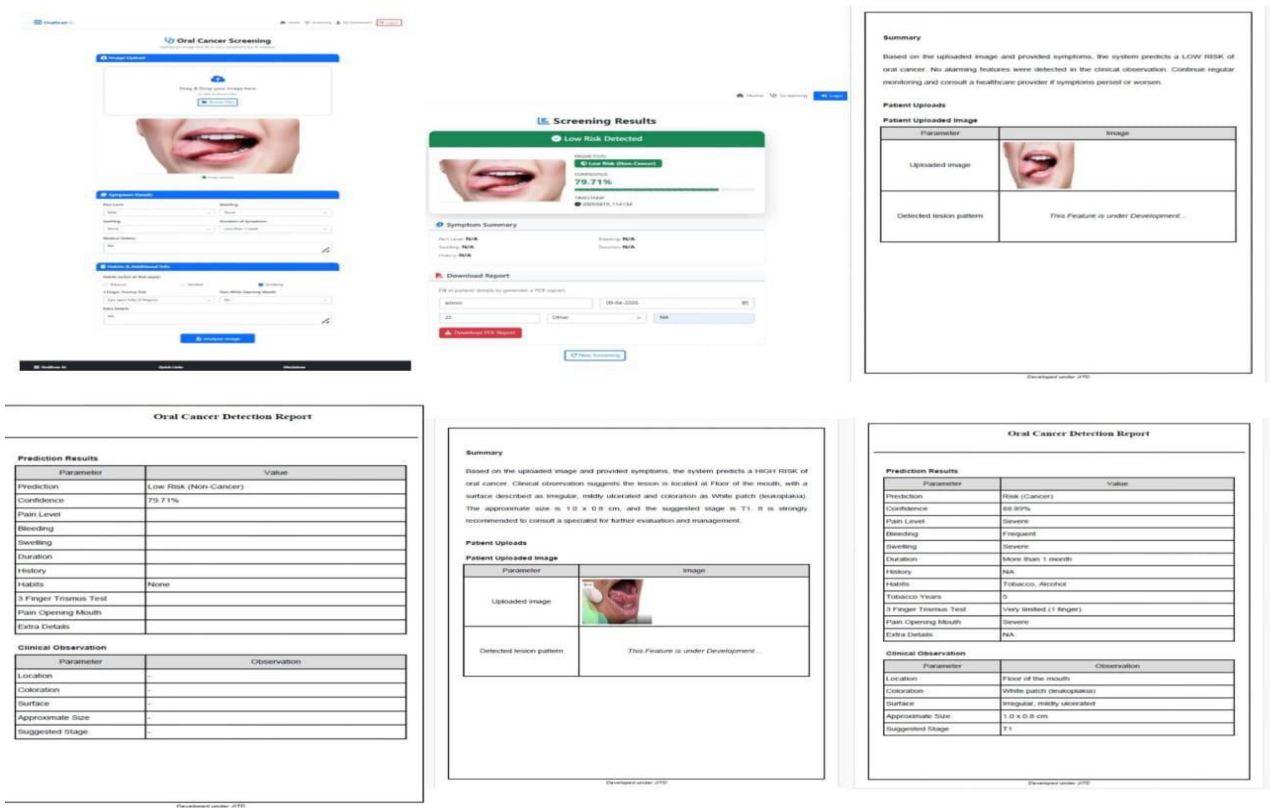
The chat interface maintains a separate message thread for each patient record. Physician replies appear in the patient's interface after a page refresh, and all messages persist in the backend data structures across sessions.



www.ijircce.com | e-ISSN: 2320-9801, p-ISSN: 2320-9798| Impact Factor: 8.771| ESTD Year: 2013|

International Journal of Innovative Research in Computer and Communication Engineering (IJIRCCE)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)



V. CHALLENGES AND LIMITATIONS

Several constraints limit the current prototype's readiness for real clinical deployment. The neural network was trained on a dataset of restricted size and demographic diversity. Patients with darker mucosal pigmentation, or with lesions arising in less commonly sampled anatomical subsites such as the retromolar trigone [6], may be poorly served by the current model weights. The model was also trained without exposure to benign conditions that frequently mimic malignancy, such as traumatic keratosis or frictional hyperkeratosis [7], raising the possibility that innocuous lesions could trigger high-risk classifications.

The current prototype implements confidence scores using the model's raw output probability, but this value has not been calibrated against a held-out validation set and should be interpreted as a relative indicator rather than an absolute measure of diagnostic certainty. Future iterations will apply Platt scaling or temperature scaling to improve probability calibration. Additionally, the clinical observation table for high-risk cases is populated from a predefined dictionary using rule-based selection rather than feature extraction from the input image. This is explicitly noted in the generated report, and the system displays a disclaimer clarifying that these descriptors are illustrative placeholders. A physician unfamiliar with the prototype's current limitations might otherwise misinterpret the table as reflecting model-derived analysis.

The system stores patient images and clinical responses as unencrypted files and plain text, with no access control beyond session authentication. Compliance with data protection frameworks applicable to health information — such as the Digital Information Security in Healthcare Act in India — would require encrypted storage at rest, comprehensive access logging, and formal audit trails before any real-world deployment.

The messaging module operates without push notifications or automated escalation. A patient who submits a high-risk result and then sends a follow-up question through the chat interface has no guarantee of a timely response if the assigned physician does not check the dashboard regularly. No automated pathway exists to escalate unacknowledged high-risk cases to a supervising clinical.

www.ijircce.com | e-ISSN: 2320-9801, p-ISSN: 2320-9798| Impact Factor: 8.771| ESTD Year: 2013|

International Journal of Innovative Research in Computer and Communication Engineering (IJIRCCE)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

VI. FUTURE SCOPE AND RESEARCH DIRECTIONS

The most consequential near-term improvement would be expansion of the training dataset to include thousands of images collected across multiple institutions, geographic regions, and demographic groups. Each image should be linked to a confirmed histopathology result to provide a reliable ground- truth label. Prospective data collection should prioritize benign mimickers alongside confirmed malignancies, since reducing false positives is as operationally important as improving sensitivity in a community screening context.

Replacing the current randomized confidence score with genuine model uncertainty estimation would substantially improve the tool's clinical utility. One established technique involves applying dropout regularization during inference rather than restricting it to training, then running the same image through the network multiple times and measuring the spread of the resulting predictions [8]. Cases producing high-variance outputs would be automatically flagged for human expert review, regardless of the modal prediction.

Smartphone camera guidance features integrated directly into the web interface could standardize image quality across patient submissions. Real-time visual feedback about lighting adequacy, camera distance, and mouth positioning — potentially delivered through an augmented reality overlay — would help patients capture diagnostically useful photographs on their first attempt rather than requiring resubmission.

Any pathway toward clinical deployment requires formal regulatory engagement. In India, this would involve the Central Drugs Standard Control Organisation. International deployment would require engagement with frameworks such as the European Union's In Vitro Diagnostic Medical Devices Regulation or the United States Food and Drug Administration's De Novo classification process. Each pathway demands evidence from prospective validation studies comparing system predictions against histopathological reference diagnoses, a substantial but necessary undertaking.

VII. CONCLUSION

OralScan AI demonstrates that a web-based platform integrating deep learning image classification with structured clinical symptom data can deliver preliminary oral cancer risk assessments without requiring patients to travel to specialized centers. The modular architecture — separating image handling, questionnaire processing, model inference, report generation, and physician communication

— keeps the codebase maintainable and positions individual components for independent improvement. The dual-input requirement reduces the risk of predictions based on incomplete evidence, while longitudinal record storage gives clinicians the ability to observe lesion behavior across successive visits rather than relying on isolated snapshots.

Role-specific dashboards serve the genuinely different information needs of patients tracking their own history and physicians managing a caseload. The automated PDF report produces documentation printable into paper-based charts still prevalent in rural healthcare facilities. While the current prototype requires a larger and more diverse training dataset, properly calibrated confidence outputs, data encryption, and regulatory clearance before it can be deployed in actual clinical practice, the foundational design has been validated through systematic functional testing. If screening no longer demands a specialist appointment, the population most likely to present late — rural, under-resourced, and reluctant to seek care — gains a credible first step toward earlier diagnosis and meaningfully improved survival prospects.

REFERENCES

[1] P. Gupta and S. Warnakulasuriya, "Global epidemiology of areca nut usage and its association with oral cancer: A ten-year systematic review," Journal of Oral Pathology & Medicine, vol. 51, no. 3, pp. 221–233, 2022.
[2] R. Sankaranarayanan, K. Ramadas, and G. Thara, "Clinical trials of screening for oral cancer: A systematic review and meta-analysis," Oral Oncology, vol. 112, art. no. 105418, 2021.
[3] D. J. Kim, S. Y. Lee, and J. H. Park, "Comparative analysis of image-based versus questionnaire- based oral cancer screening applications: A usability study," JMIR mHealth and uHealth, vol. 9, no. 8, e28945, 2021.
[4] A. S. Bhatia, M. K. Singh, and R. K. Sharma, "Longitudinal photographic monitoring of oral potentially malignant disorders: A systematic review," Clinical Oral Investigations, vol. 26, no. 4, pp. 3415–3429, 2022.
[5] G. Litjens, T. Kooi, B. E. Bejnordi, A. A. A. Setio, F. Ciompi, M. Ghafoorian, J. A. W. M. van der Laak, B. van

Ginneken, and C. I. Sánchez, "A survey on deep learning in medical image analysis," Medical Image Analysis, vol. 42, pp. 60–88, 2017.
[6] P. D. Quinn, D. K. Lam, and J. C. Irish, "Anatomical subsites of the oral cavity and oropharynx: A clinical guide," Oral and Maxillofacial Surgery Clinics of North America, vol. 31, no. 4, pp. 551–564, 2019.
[7] B. W. Neville, D. D. Damm, C. M. Allen, and A. C. Chi, Oral and Maxillofacial Pathology, 4th ed. St. Louis, MO, USA: Elsevier, 2016, pp. 322–335.
[8] Y. Gal and Z. Ghahramani, "Dropout as a Bayesian approximation: Representing model uncertainty in deep learning," in Proc. 33rd Int. Conf. Machine Learning (ICML), New York, NY, USA, 2016, pp. 1050–1059.

INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN COMPUTER & COMMUNICATION ENGINEERING

9940 572 462 6381 907 438 ijircce@gmail.com