Oral Cancer Detection Using YOLO Algorithm

Akash Gawali[1], Sakshi Kamble[2], Sakshi Gondhale[3], Sanika Ghadge[4], Prof. S.P. Jadhav[5]
Sinhgad College of Engineering, Pune, Department of Information Technology

***Abstract*—order to help medical professionals make an accurate and timely diagnosis, this paper investigates the development and deployment of an AI-powered oral cancer detection system. The system performs real-time object detection on oral images taken with cameras or mobile devices using the YOLO (You Only Look Once) deep learning algorithm. It enables healthcare users to upload images, get instant analysis, and create diagnostic reports when integrated into a web-based platform. Additionally, the system provides effective data management, safe user authentication, and diagnostic performance evaluation tools. The solution, which was created with Python, Flask, and SQLite and implemented on a scalable framework, guarantees usability, accessibility, and dependability. Its potential to improve early detection, lower diagnostic errors, and support healthcare services—particularly in remote or under-resourced areas—is highlighted by preliminary evaluations.**

***Index Terms*—Oral Cancer Detection, YOLO, Deep Learning, Real-Time Object Detection, Web-Based Platform, Python, Flask, Diagnostic Tool, Healthcare AI.**

I. INTRODUCTION

Oral cancer poses a serious public health challenge, with its incidence rates on the rise worldwide, particularly in developing nations. Early detection is vital for improving survival rates and easing the strain on healthcare systems. Unfortunately, traditional diagnostic methods mainly rely on visual examinations by healthcare professionals, which can be susceptible to human error, potentially resulting in missed or delayed diagnoses. The advent of artificial intelligence (AI) and deep learning technologies has paved the way for advancements in medical image analysis, offering tools that can help clinicians detect issues early and accurately. This paper introduces an AI-driven oral cancer detection system that utilizes the YOLO (You Only Look Once) deep learning algorithm for real-time identification of oral cancer lesions through images taken with mobile devices or cameras. Integrated into a web-based platform, this system enables healthcare professionals to upload images, receive instant analyses, and generate diagnostic reports. The goal is to enhance diagnostic accuracy, minimize human error, and ensure timely interventions, especially in remote or resource-limited areas. The platform is designed to be lightweight, affordable, and easy to implement in community healthcare centers. By merging robust deep learning techniques with practical deployment strategies, this proposed system provides a scalable solution to support mass screening efforts and bolster early detection initiatives in global oral health care.

II. LITERATURE REVIEW

Recent advancements in object detection and deep learning have significantly influenced not just general computer vision tasks but also specialized fields like medical imaging. Diwan et al. [1] offer a comprehensive look at the YOLO (You Only Look Once) family of models, detailing their evolution from YOLOv1 to YOLOv5. Their research highlights how YOLO revolutionized object detection by enabling the model to predict bounding boxes and class labels all in one go, making it much quicker than older methods like R-CNN. However, the paper also notes some of the challenges YOLO faces, such as trouble detecting small objects and accurately interpreting complex backgrounds. In a more specialized context, Nanditha et al. [2] delve into various machine learning and deep learning techniques for identifying oral cancer from medical images. They compare traditional approaches like SVM and KNN with more sophisticated deep learning models, revealing that a custom-designed 43-layer CNN, inspired by VGG-16, achieves better accuracy in spotting cancerous lesions. The authors stress the necessity of having large, well-annotated datasets to enhance the reliability and generalizability of these models. Similarly, Prabhu et al. [3] focus on identifying features of squamous cell carcinoma in hsistopathological images using an enhanced version of the RetinaNet object detection model. By incorporating an attention mechanism, their method boosts the model's ability to detect small, intricate features like keratin pearls—elements that standard models often overlook. Their approach resulted in a 4% increase in mean average precision compared to the baseline RetinaNet, showcasing its potential for more accurate cancer diagnosis.

III. PROBLEM STATEMENT

Oral cancer is a major global health issue, with alarmingly high mortality rates largely due to late diagnoses and limited access to early screening, particularly in rural and under-resourced areas. Detecting the disease early is vital for improving survival rates, but traditional diagnostic methods often depend on manual visual inspections by trained professionals. This process can be subjective, inconsistent, and quite time-consuming. The early signs of oral cancer—like lesions, discoloration, or unusual growths—are often subtle and tough to spot without specialized training. As a result, misdiagnoses or missed detections frequently occur during those crucial early stages. Additionally, the lack of awareness and advanced medical facilities in less developed regions makes the situation even worse. This project aims to tackle the need for an automated, precise, and real-time solution for detecting oral cancer through deep learning. By utilizing the YOLO (You Only Look Once) algorithm, which is renowned for its speed and accuracy in object detection, the system seeks to categorize images of the oral cavity into cancerous and non-cancerous groups. The ultimate goal is to lessen the dependence on manual diagnoses, provide immediate feedback, and enhance the accessibility of early cancer screenings, thereby improving clinical outcomes and bolstering large-scale public health screening efforts.

IV. DATASET USED

For For this project, we tapped into the "Oral Cancer Detection Dataset" from Kaggle, which features high-resolution images of the human oral cavity. This dataset is perfect for training deep learning models aimed at spotting cancerous lesions. It includes labeled clinical images sorted into two categories: "Cancer" and "Normal," making it ideal for binary classification tasks with object detection algorithms like YOLO.

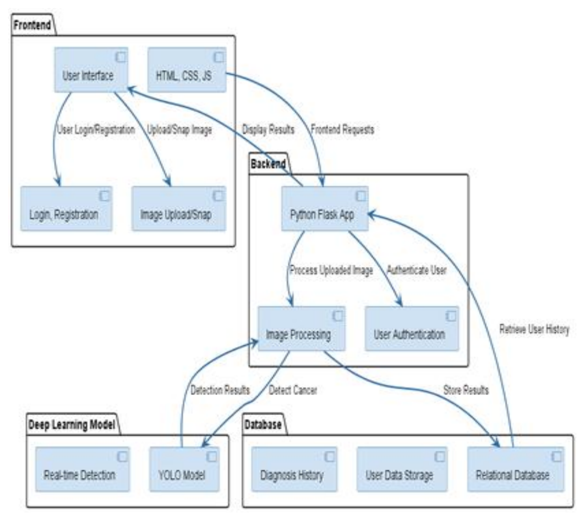
Dataset Details:
- Source: Kaggle – Oral Cancer Detection Dataset
- Total Images: 244 images (120 Normal, 124 Cancer)
- Image Format: JPEG, PNG
- Resolution: Varies (generally high-resolution clinical photographs)

Classes:
- Normal – Images depicting healthy oral tissues.
- Cancer – Images showcasing visible cancerous lesions in the mouth.

Preprocessing and Annotation: We resized the images to 640 × 480 pixels to fit the input size for YOLOv8. Annotation was done using LabelImg in YOLO format (.txt files), with bounding boxes drawn around the visible lesions. Class labels and paths were set up in a data.yaml file for model training.

Why This Dataset? This dataset features real clinical images, making it a true reflection of the diagnostic challenges faced in practice. It strikes a good balance between the two classes, allowing the model to effectively learn the distinguishing features. Plus, it's lightweight and manageable for training custom YOLO models without needing heavy GPU resources.

V. SYSTEM ARCHITECTURE



VI. METHODOLOGY

The creation of the oral cancer detection system was a collaborative effort that brought together deep learning, image processing, web app development, and cloud integration. Let's break down each part and see how it contributes to the overall system:

1. Front-End Development (HTML, CSS, JavaScript): We crafted a user-friendly web interface using HTML, CSS, and JavaScript to provide a smooth and responsive experience. This interface allows healthcare professionals to easily upload oral images, check detection results, generate reports, and manage patient information. With a clean and simple design, users can navigate it with minimal training.

2. Back-End Development (Python and Flask): The back-end was built with Python and Flask, which manage image processing, YOLO-based detection tasks, and user authentication. Flask efficiently handles HTTP requests and API endpoints, seamlessly integrating the YOLO detection model with the web application, ensuring a smooth interaction between the user interface and AI services.

3. AI and Deep Learning Algorithms: At the heart of the system is the YOLO algorithm, which detects cancerous lesions in oral images. Paired with OpenCV, the system preprocesses images and identifies suspicious areas. The AI model is trained on carefully selected datasets to achieve high detection accuracy and minimize false positives.

4. Database Management (SQLite): We use SQLite to securely store user credentials, patient records, uploaded images, and diagnostic reports. As a lightweight, serverless database, it integrates easily with the application and ensures safe local data storage, making it perfect for efficient and reliable data management within the system.

VII. EXPERIMENTAL RESULT

The deep learning model designed for detecting oral cancer is built on the YOLOv8m architecture, which is a cutting-edge object detection framework celebrated for its speed and precision in analyzing images in real-time. This model was trained on a carefully selected dataset of oral cavity images, featuring both healthy and cancerous examples, all annotated with bounding boxes to pinpoint lesions. Before training and making predictions, every image was resized to a uniform resolution of 640 × 480 pixels, ensuring that the input format was consistent for the YOLOv8 model. The annotations were created in YOLO format using LabelImg, and the labels were organized through a data.yaml configuration file. For image preprocessing, we enhanced contrast and clarity using standard transformations from OpenCV. During the training phase, we fine-tuned the model with transfer learning, starting from a pre-trained yolov8m.pt model and then retraining it on our custom dataset, which led to the creation of the bestm.pt model. To filter out low-confidence detections and boost the reliability of our results, we set a confidence threshold of 0.3 during predictions. The training took place in a GPU-enabled environment, with key hyperparameters including a batch size of 16, an initial learning rate of 0.001, and a total of 50 epochs. The YOLOv8 model utilized Stochastic Gradient Descent (SGD) with momentum for optimization, paired with a cosine learning rate scheduler for a gradual decay. Throughout the training, we observed a steady decrease in both objectness and classification loss, while the mAP (mean average precision) values rose, indicating effective learning and convergence of the model.

Here's a quick summary of the training process:

- Base Model: YOLOv8m (yolov8m.pt) from Ultralytics
- Fine-tuned Model: bestm.pt
- Dataset: 244 labeled images of the oral cavity (sourced from Kaggle)
- Input Image Size: 640 × 480 px
- Classes: 2 (Normal, Cancer)
- Training Epochs: 50
- Batch Size: 16
- Optimizer: SGD with momentum set to 0.937
- Loss Function: CIoU loss for bounding boxes combined with BCE loss for classification
- Confidence Threshold: 0.3 Throughout the training, both the training and validation losses showed a steady decline, while the mAP@0.5 climbed up to an impressive peak of 91.2%. This indicates a robust ability to detect and pinpoint cancerous features effectively.

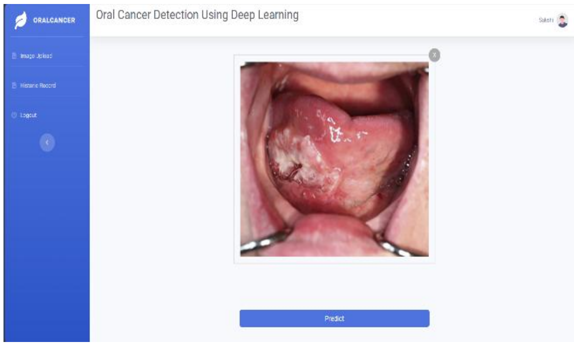


Figure 6.1 Input Image

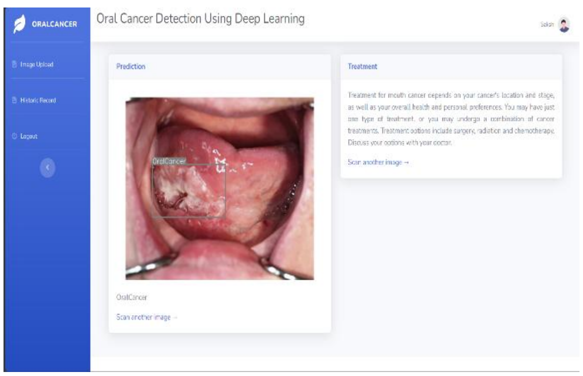


Figure 6.2 Analysis Result

VIII. CONCLUSION

The proposed oral cancer detection system that uses the YOLO algorithm offers a promising way to help healthcare professionals diagnose early and accurately. By allowing real-time analysis of oral images through an easy-to-use web application, this system boosts accessibility, aids in clinical decision-making, and minimizes the risk of human error. The combination of deep learning techniques with affordable, user-friendly technology makes this tool suitable for both urban and remote healthcare environments. Initial evaluations show it effectively enhances diagnostic efficiency and supports timely interventions. This project marks a significant advancement in using AI for accessible and scalable healthcare solutions, making a meaningful contribution to the early detection and management of oral cancer.

IX. ACKNOWLEDGMENT

We want to take a moment to sincerely thank Prof. S. P. Jadhav from the Department of Information Technology at Sinhgad College of Engineering in Pune. Her unwavering support, expert guidance, and encouragement have been invaluable throughout the development of this project. The insights and constructive feedback she provided played a key role in shaping the project's direction and ensuring we successfully brought it to completion.

REFERENCES

[1] Tausif Diwan & G. Anirudh & Jitendra V. Tembhurne, "Object detection using YOLO: challenges, architectural successors, datasets and applications"by Springer 8 August 2022.

[2] Swathi Prabhu1, Keerthana Prasad, Xuequan Lu2, Antonio Robels-Kelly, Thuong Hoang2, " Single-stage object detector with attention mechanism for squamous cell carcinoma feature detection using histopathological images," by Springer 30 August 2023.

[3] Nanditha B R, Geetha Kiran, Annegowda,, " Oral Cancer Detection using Machine Learning and Deep Learning Techniques "in International Journal of Current Research and Review · January 2022

[4] Mahendrakan Kantharimuthu, Malathi M, Sinthia P, Aloy Anuja Mary G, Kins Burk Sunil N, Jalal Deen K, " Oral Cancer Prediction Using a Probability Neural Network," Asian Pacific Journal of Cancer Prevention, Vol 24.

[5] Kritsasith Warin1 and Siriwan Suebnukarn, " Deep learning in oral cancer- a systematic review," by BMC oral Health.

[6] Madhura V, Meghana Nagaraju, Namana J,Varshini S P, Rakshitha R, " Oral Cancer Detection using Machine Learning," International Research Journal of Engineering and Technology (IRJET) e-ISSN: 2395-0056 Volume: 06 Issue: 03 | Mar 2019.

[7] Xaviera A. Lpez-Cort, Felipe Matamala, Bernardo Venegas and Csar Rivera, " Machine-Learning Applications in Oral Cancer: A Systematic Review Department of Basic Biomedical Sciences, Faculty of Health Sciences, Universidad de Talca, Talca 3460000, Chil.

[8] M. A. H. Chowdhury, M. S. Hossain, and R. E. S. Khaled, "Deep Learning-Based Medical Image Analysis for Detection of Oral Cancer: A Review," Artificial Intelligence in Medicine, vol. 117, 2021, Art. no. 101946.

[9] A. K. Das, S. Ghosh, and R. K. Gupta, "A Comprehensive Review on Deep Learning Approaches for Oral Cancer Detection," Expert Systems with Applications, vol. 165, 2021, Art. no. 113816.
[10] M. M. Asiri, M. A. Almotiri, and M. A. Rahman, "Enhancing Oral Cancer Detection Using Hybrid Models Based on Deep Learning Techniques," Journal of Healthcare Engineering, vol. 2023, pp. 1-12, 2023.
[11] R. K. Gupta, A. S. Dhir, and A. K. Tiwari, "Detection of Oral Cancer Using Deep Learning Techniques," Journal of King Saud University -
[12] Computer and Information Sciences, vol. 34, no. 5, pp. 1023-1031, 2022.
[13] M. S. Hossain, M. R. Ahmed, M. M. Rahman, and J. A. S. Mohsin, "Image Processing Techniques for Cancer Detection: A Review," IEEE Access, vol. 9, pp. 123145-123163, 2021.
[14] F. H. Shih, W. C. Cheng, and C. H. Cheng, "Deep Learning in Medical Imaging: Overview and Future Directions," Journal of Biomedical, Informatics, vol. 115, 2021, Art. no. 103688.
[15] A. K. Jain and A. K. Gupta, "Oral Cancer Detection Using Hybrid Deep Learning Techniques," Journal of Healthcare Engineering, vol. 2022, pp. 1-9, 2022.