INFORMATICS INSTITUTE OF TECHNOLOGY

In Collaboration with

ROBERT GORDON UNIVERSITY ABERDEEN

Multimodal Image & Text-Based Oral Cancer Early Detection Application

Group 33 Final Thesis by

Tharusha Liyanagama – IIT: 20232938 ,RGU: 2409101

Gagani Kulathilaka – 20230214, RGU: 2409662

Siyumi Jayawardhane – IIT: 20230261, RGU: 2410215

Sithmi Desilva – IIT: 20231045, RGU: 2409664

Supervised by
Mrs.Hiruni Kasthuriarachchi

Submitted in partial fulfillment of the requirements for the BSc(Hons) in Artificial Intelligence and Data Science degree at the Robert Gordon University.

April 2025

Declaration

I declare that this is our own research thesis, and this thesis does not incorporate without acknowledgement any material previously published submitted for a Degree or Diploma in any other university or institute of higher learning and to the best of our knowledge and belief it does not contain any material previously published or written by another person except where the acknowledgement is made in the text.

Student Name	**Student ID**	**Signature**
Tharusha Liyanagama	20232938	
Gagani Kulathilaka	20230214	
Siyumi Jayawardhane	20230261	
Sithmi Desilva	20231045	

I have read the final thesis and it is in accordance with the approved university final thesis outline.

Signature of the supervisor: Date: 03/04/2025

ROBERT GORDON UNIVERSITY ABERDEEN

INFORMATICS INSTITUTE OF TECHNOLOGY

Abstract

Early and accurate detection of cancer remains a critical focus in medical research, as it significantly improves patient outcomes and efficacy of treatments. This study presents a novel multimodal predictive model that integrates image-based and textual data to provide comprehensive cancer predictions. The proposed system comprises four key components: a model for validating the user provided image, a model to predict the presence of cancer in the given image, a model analyzing textual data, such as patient habits and background information to assess cancer risk levels, a multimodal that fuses insights from both modalities to deliver a final prediction and an explanatory model providing detailed descriptions for the generated predictions.

The system's output is designed to provide detailed description for the given predictions of the model. It includes a detailed report featuring heatmaps for visual explanations of image-based predictions with textual reasonings. If cancer is predicted, tailored recommendations for users are generated, providing explainability and interpretability for diagnosis and treatment.

This research bridges the gap between unimodal prediction methods and the need for explainability on machine learning model predictions.

Acknowledgement

We would like to express our heartfelt gratitude to all those who played a vital role in the successful completion of this project.

First and foremost, We would like to thank our module coordinator, Mr.Prasan Yapa, for his continuous support and guidance. Your efforts in coordinating the module and facilitating resources were helpful in enabling us to achieve our objectives.

We would also like to extend our sincere gratitude to our mentor, Mrs.Hiruni Kasthuriarachchi, for her invaluable guidance, feedback, and unwavering support throughout this journey. Your expertise and encouragement were instrumental in enhancing the performance of this model, and we are incredibly grateful for the knowledge you have given.

Additionally, we extend our sincere appreciation to all the professionals who contributed to this project, directly or indirectly. We would also like to extend our gratitude to the authors of the referenced research papers.

Finally, we are immensely grateful to our friends, and families for their encouragement, and belief in us during this journey. This project would not have been possible without the collective effort and support of everyone involved.

Table of Contents

List of Figures

List of Tables

CHAPTER 1: INTRODUCTION

1.1 Chapter Overview

Our project focuses on developing a Multimodal image & text-based oral cancer early detection application. The system will predict the risk of oral cancer by analyzing patient data and oral lesion images. By leveraging Convolutional Neural Networks (CNNs) for image analysis and Machine Learning techniques for textual data, the system aims to provide accurate and explainable predictions. The integration of Explainable AI techniques will offer transparent insights into the decision-making process, aiding healthcare professionals in making informed clinical decisions. The ultimate goal is to improve early detection rates, reduce mortality, and make the system accessible to a wider population, regardless of socioeconomic constraints.

1.2 Problem Domain

The problem domain focuses on the early detection of oral cancer, a fast-spreading and often misdiagnosed disease, particularly in middle-income and South Asian countries where risk factors like tobacco and betel nut use are common. Despite medical advancements, early-stage diagnosis remains challenging due to reliance on manual exams and biopsies, which are prone to human error and often inaccessible in remote areas. While AI technologies like CNNs show promise in improving diagnostic accuracy by analyzing medical images, a major limitation is the lack of explainability, which undermines doctors' trust in these systems. The combination of diagnostic delays, technological gaps, and unexplainable AI poses significant barriers to effective treatment and early intervention.

1.3. Problem Definition

The problem lies in the inaccuracy and lack of transparency in current oral cancer diagnosis methods. Traditional techniques, like manual inspections and biopsies, are prone to human error and often overlook crucial patient data such as medical history and risk factors. Although AI models like CNNs offer improved accuracy, they typically lack of explainability, making them difficult to trust in clinical settings. This research aims to develop a transparent and accurate oral cancer early detection system by integrating deep learning for image analysis with machine learning for patient data, enhanced by Explainable AI (XAI) tools like Grad-CAM, LIME, and SHAP to provide visual and textual justifications for the system's predictions.

1.4. Research Motivation

To develop a multimodal system that uses deep learning and machine learning to accurately detect oral cancer lesions from user-provided images and textual data (risk factors), while also providing clear and understandable explanations for its predictions.

1.5. Existing Work

Citation	Technology/ Algorithm	Advantages	Limitations
Lin & Chen,(2021)	CNN	How smart phone-based images, combined with a deep learning model, can detect oral cancer early	The approach does not use any advanced region proposal methods or rely on manually cropped images, which might limit accuracy in certain cases
Devindi et al. (2024)	Multimodal Deep CNN Pipeline	Integrates lesion images and patient metadata for improved early detection accuracy. Uses multiple pre-trained models (Mo-bileNetV3, ResNet-50, DenseNet-121, etc.).	Requires high-quality images and structured metadata. Limited dataset diversity affects generalizability.
J. J. Sciubba (2001)	Risk Factor-Based Prediction Model	Uses metadata (tobacco use, alcohol consumption, sun exposure, etc.) to assess cancer risk. Helps in	Relies on self-reported data, which may introduce bias. It cannot replace image-based validation.

		screening before lesion detection.	
Uthoff et al. (2022)	Transfer Learning on VGG-M Model (AFI & RGB Image Pairs)	Achieves 80%+ sensitivity and specificity in detecting high-risk lesions. Uses both autofluorescence and standard images.	Requires a custom imaging device, limiting accessibility for widespread use.
R. A. Welikala et al. (2020)	ResNet-101 & Faster R-CNN (MeMoSA Project)	Uses deep learning for lesion detection and classification, integrating metadata (age, gender, risk factors like smoking, alcohol). Achieved 87.07% accuracy in identifying lesion images.	Requires large, annotated datasets for better performance. Object detection has lower accuracy compared to image classification.
Parola et al. (2024)	Informed Deep Learning(IDL), Case-Based Reasoning(CBR)	integrate clinical knowledge and produce explanations	performance varies depending on the quality of input images.

Table 1:Existing Work

1.6. Research Gap

The research gap lies in the limited scope and lack of explainability in current oral cancer detection systems. Most existing approaches rely on a single data type, such as images, and fail to incorporate textual patient data, reducing diagnostic accuracy. Additionally, deep learning models often lack transparency, making their predictions difficult for healthcare providers and patients to trust. This research aims to fill the gap by developing a system that integrates both image and text data for more accurate and comprehensive predictions while using Explainable AI (XAI) to provide clear, understandable justifications, thereby enhancing trust and usability in clinical settings.

1.7. Contribution to the body of knowledge

1.7.1 Technological contribution

The technical contribution involves developing an AI-powered system for early oral cancer detection by integrating Machine Learning models for textual data analysis and Deep Learning for image data analysis. The system will incorporate Explainable AI (XAI) techniques to generate both visual (heatmaps) and textual explanations, enhancing transparency and trust. A user-friendly web application will be built using HTML, CSS, and JavaScript, allowing users to input both image and text data. All backend processes - data processing, model training, and explanation generation will be implemented in Python, leveraging advanced machine learning tools for optimal performance.

1.7.2. Contribution to the domain

This system aims to create a user-friendly platform for assessing oral cancer risk by analyzing both oral cavity images and patient textual data. Users will upload an image and input relevant health information, and the system will predict cancer risk by identifying critical patterns in both data types. To enhance trust and transparency, the system will use XAI with heatmaps highlighting key image regions and textual explanations pointing out significant symptoms.

Additionally, the system supports ongoing monitoring, allowing medical professionals to track patient progress over time by comparing new inputs with past data. This enables personalized treatment adjustments and offers a more comprehensive view of the patient's condition than current single-source systems.

1.8. Research Challenge

1. **Data Availability and Quality**: Limited availability of well-annotated, diverse datasets that integrate both medical images and textual data for effective training

of deep learning models.

2. **Multimodal Integration**: Challenges in combining effective image recognition models and textual data to obtain accurate pre-cancer detection.
3. **Explainability**: Use of XAI to give a better understanding of the prediction.
4. The main goal is to regulate AI in such a way that it can be used in different branches of medicine and does not display any biases as a result of the skewed data.
5. **Real-World Application**: Taking the high performing lab models and translating them into things that work in practice and are reliable for use in the clinic and dealing with varying data quality patient demographics.

1.9. Research Questions

1. How can we effectively integrate medical images and clinical text data to enhance pre-oral cancer detection accuracy?
2. What are the most appropriate models for processing image and text data in cancer detection systems?
3. How can XAI techniques improve the transparency and interpretability of pre-oral cancer detection models?
4. How can we ensure that the developed models generalize across diverse patient populations while maintaining high diagnostic accuracy?

1.10. Research Aim

The objective of this research is to design a multimodal AI-driven web application for early detection of oral cancer via image recognition of Oral cavity images and clinical text analysis, along with explainability and real-world applicability.

1.11. Research Objectives

Research Objectives	Explanation	Learning Outcome

Problem Identification	RO1: To find out the best CNN architecture to get accurate lesion detection in image data.	LO1
	RO2: Add the Patients History and Risk factors to the trained model.	
	RO3:Collect and preprocess data	
	RO4: Develop a research methodology to validate model performance and define accuracy and transparency metrics.	
	RO5: Implement a user-friendly website to input data and to display XAI outputs.	
Data Gathering and Analysis	Collect Images and clinical data from public datasets and preprocess data.	LO1, LO3
Research Design	Design a multimodal architecture that fuses CNN and Random Forest models to get the final prediction, with a particular emphasis on explainability.	LO3
Implementation	Construct the deep learning Models in Python (TensorFlow, PyTorch framework), and data fusion methods for combining image and textual data.	LO2, LO4
Testing and Evaluation	Test and validate the model performance with real world data work through performance tests including evaluation metrics and user feedbacks to prove our model is as reliable and useful in situation as we think it is.	LO2 , LO4

Table 2:Research Objectives

1.12. Project Scope

1.12.1. In-scope

1	The project will allow users to upload oral cavity images and input textual patient data such as age, gender, habits, and symptoms through a web-based interface.
2	The system will preprocess both image and textual data to prepare it for analysis and prediction.
3	It will involve developing deep learning models, specifically Convolutional Neural Networks (CNNs), for analyzing image data.
4	Machine learning will be used to process and classify the textual metadata.
5	A fusion model will be implemented to combine insights from both data types to provide a final risk prediction.
6	The system will use Explainable AI (XAI) methods like Grad-CAM, SHAP, and LIME to generate visual and textual explanations for each prediction.

Table 3:In-scope

1.12.2. Out-scope

1	The system will not be integrated into real-time clinical workflows or electronic medical record (EMR/EHR) systems.
2	It will not include or process biopsy reports, blood test results, or any other physical lab data.
3	The system will not provide definitive medical diagnoses or treatment plans; it will only indicate the risk level of oral cancer based on the input data.
4	Developing a mobile application for the platform is not part of the current project scope
5	The system will not collect data automatically from external databases or sources; all data must be manually uploaded by the user.
6	It will not include speech or audio-based analysis for detecting oral cancer.

Table 4:Out-scope

1.12.3. Prototype Diagram

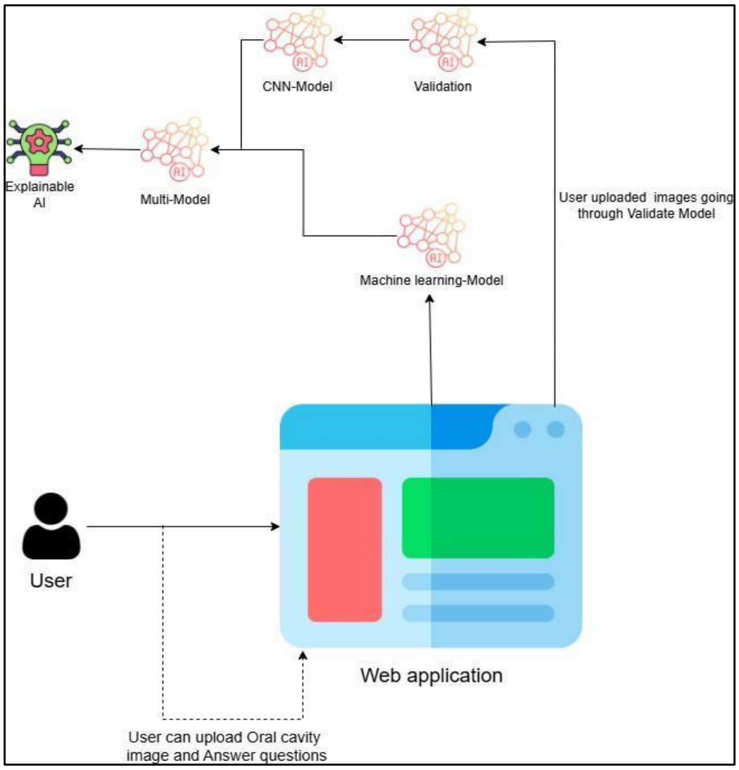


Figure 1:Protopype Diagram

1. The patient uploads Oral cavity images and patients' risk factors through a website.
2. The images are preliminarily preprocessed through validation model then deep learning model to predict the risk probability.
3. The text data are processed through the machine learning model to predict the risk probability.
4. The probabilities of the two models are combined and gives the final risk prediction via fusion model.
5. The final risk prediction is explained by XAI and generates a customized description for each user considering the inputs.

1.13. Resource Requirements

1.13.1. Hardware requirements

- **CPU** - Intel Core i9 or AMD Ryzen 9 or higher specification for high-performance, multitasking and to solve complex computation while training and testing of models.

- **GPU** - NVIDIA RTX Series or A100 because it needs to accelerate deep learning models when images processed through Convolutional Neural Networks (CNNs).
- **Memory(RAM) -** a minimum of 16GB RAM and higher to heavy algorithm training operations.
- **Storage -** at least 1TB SSD capacity or higher to save big datasets together with processed data along with testing outputs.

1.13.2. Software requirements

- **Python** – The primary language to be used for carrying out several computational and multi-functional applications generally used in machine learning.
- **TensorFlow and PyTorch** – For image and text analysis models building and training.
- **Matplotlib** – For plot computations
- **Keras** – For deep learning algorithms
- **Flask** – Flexible Python web framework used to build web applications, APIs, and backend services with minimal boilerplate code.
- **Jupyter Notebook or Google Colab** - Interactive computing environments that allow users to write, execute, and visualize Python code, making them ideal for data science, machine learning, and research.
- **Windows** - For computational needs and to satisfy application dependency.
- **Visual Studio Code** -Application for perfect code control and writing.
- **LIME or SHAP Libraries** -For general use of explainability in machine learning models particularly when using Grad-CAM-based visualizations.

1.13.3. Data requirements

- **Image Data** - a dataset of oral cavity images with linked metadata required in image analysis for training.

- **Text Data** – dataset of risk factors of cancer identified by oral cancer patients' history, behaviors and lifestyle.

1.13.4. Skill requirements

- Data Collection and Preprocessing – Ability to provide, clean, and formal images and text data for training and testing.
- Machine learning and model training – An awareness in implementing, training, and fine-tuning of machine learning models.
- Backend Desks – For purposes of developing APIs and designing data processing procedures.
- Medical or Dental Experience To properly understand and corroborate findings concerning the mouth condition.
- Critical Thinking and Problem Solving – There is bound to be problems during the model training and testing processes.
- Report Writing – For documentation purposes and when reporting to the stakeholders.
- Team Management – Knowledge about activity scheduling and assigning of duties among the workers.

1.14. Chapter Summary

The chapter commenced with a detailed illustration outlining the features of the model. Subsequently, the stakeholders and their involvement with the model were described. The survey's findings were explained, and the reasons for the requirement elimination techniques were acknowledged. Lastly, the prototype diagram was used to visualize the flow of the interaction between different components. Additionally, both functional and non-functional requirements essential for the development process were outlined.

CHAPTER 2: LITERATURE REVIEW

2.1 Chapter Overview

Oral cancer is a serious disease that affects various parts of the mouth, including the lips, tongue, cheeks, gums, floor of the mouth, and the hard and soft palate. Key risk factors include tobacco use, excessive alcohol consumption, and HPV infection, with males above 50 years old having a higher probability than females. A Multimodal AI-based web application for early oral cancer detection by integrating clinical data and oral cavity images. The system consists of three main components: Oral Image Validation, Oral Lesion Detection and XAI, Text-based risk prediction and XAI, Multimodal integration.

2.2 Concept Map

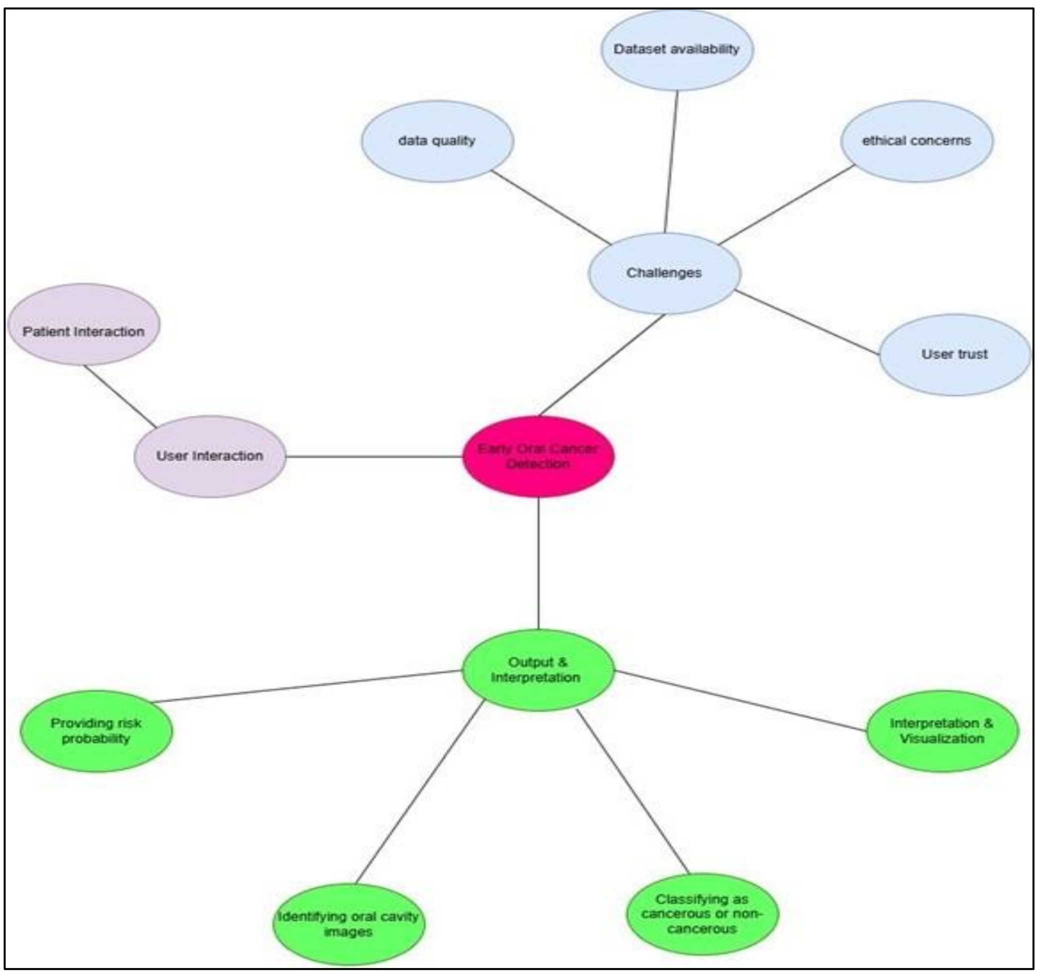


Figure 2:Concept Map

2.3. Problem Domain

The problem domain focuses on the early detection of oral cancer, a fast-spreading and often misdiagnosed disease, particularly in middle-income and South Asian countries where risk factors like tobacco and betel nut use are common. Despite medical advancements, early-stage diagnosis remains challenging due to reliance on manual exams and biopsies, which are prone to human error and often inaccessible in remote areas. While AI technologies like CNNs show promise in improving diagnostic accuracy by analyzing medical images, a major limitation is the lack of explainability, which undermines doctors' trust in these systems. The combination of diagnostic delays, technological gaps, and unexplainable AI poses significant barriers to effective treatment and early intervention.

2.4 Existing Work

Research	Models Used	Contribution	Limitation	Dataset
(Marco & Fedrico, 2024)	R-CNN, informed deep learning (IDL) approach	Development of a cost-efficient screening system for oral cancer	Trained only using a small dataset	
(Lin & Chen, 2021)	CNN	how smartphone-based images, combined with a deep learning model, can detect oral cancer early	The approach does not use any advanced region proposal methods or rely on manually cropped images, which might limit accuracy in certain cases.	patients at the First Affiliated Hospital, College of Medicine, Zhejiang University

(Hemalatha & Mattupalli, 2022)	fragment Jaya Whale Optimizer with a Deep Convolutional Neural Network (FJWO-DCNN)	Improved recognition rates by optimizing feature extraction, making it suitable for early diagnosis.	Performance could be limited by the variety of images used, as real-world clinical data may differ in quality and consistency from the training data.	
(Huna & Goh, 2021)	Transfer learning with pre-trained models such as ResNet and Inception-V3	Rapid and efficient detection of oral diseases from non-clinical images	performance may still be sensitive to variations in image quality and lighting conditions	
(da Silva & de Souza, 2024)	convolutional neural networks (CNNs) and Grad-CAM	integration of explainability (via Grad-CAM) into oral cancer detection	the reliance on Grad-CAM, which may not always provide fully interpretable results for all types of misclassifications	
(Ozen & Karadas, 2024)	Ensemble of EfficientNetB0, EfficientNetB3, and InceptionV3	Introduces an ensemble learning strategy to improve the accuracy and reliability of oral	The study focuses on specific models and datasets, potentially limiting its	Colour images of oral lesions from hospitals in Karnataka, India

		cancer detection by combining three deep learning models	generalizability to other medical image types or conditions	
(Welikala & Jian, 2020)	ResNet-101, R-CNN	Introduces deep learning approaches for automating the detection and classification of oral lesions	The small dataset and image quality variations limited performance	2,155 oral cavity images
(Goswami & Bhuyan, 2024)	LightGBM Algorithm	Proposes a method for classifying oral cancer into pre-cancerous stages using features from different color spaces.	Limited dataset size and imbalance among classes	public and in-house datasets
(Hosaka & Ikeda, 2019)	Deep learning models	Predicted oral cancer patient survival using clinical data	Lack of model interpretability.	Dataset of clinical data from oral cancer patients
(Tian & Ma, 2023)	CNNs, Transfer Learning	Reviewed deep learning techniques integrating imaging (CT,	Lack of standardized methods for combining multimodal data	CT, MRI datasets

		MRI, histopathology) with patient metadata to improve diagnostic accuracy.	types; limited real-time clinical use	
(Chavva & S, 2024)	DenseNet169 and LeNet	Introduces a deep learning-based approach for oral cancer detection using DenseNet169 and LeNet models	The dataset is small and diverse, which limits model generalizability. LeNet performance was lower compared to DenseNet, indicating the importance of architecture choice	Oral images dataset annotated by healthcare specialists

Table 5:: Existing Work(Methodology)

2.5 Technology Review

2.5.1 Oral Image Validation

Oral image validation ensures that the uploaded image is indeed of the oral cavity and suitable for further analysis. This step is critical to prevent false predictions caused by irrelevant or low-quality inputs. To achieve this, Resnet50 is employed to classify whether an image represents an oral cavity, filtering out unsuitable inputs before lesion detection.

According to research by Zhang and Wu (2023), lightweight CNN models can efficiently validate medical image categories with high accuracy. Further, Gupta and Reddy (2022) demonstrated the use of transfer learning with pre-trained models like MobileNet and VGG16 to distinguish oral cavity images from non-oral ones, achieving reliable classification with minimal computational resources. Integrating these validation techniques enhances the robustness of the system by ensuring only relevant images proceed to the lesion detection phase, ultimately improving the precision and reliability of the overall diagnostic process.

2.5.2 Oral Lesion Detection and XAI

The system aims to detect oral cancer lesions—such as red or white patches, ulcers, or abnormal growths—in oral cavity images using deep learning techniques. After image upload, preprocessing enhances clarity by reducing noise. Studies like Lin and Chen (2021) have shown that smartphone-based detection using image enhancement improves lesion visibility. To address variability in lesion appearance due to lighting or patient factors, Ozen and Karadas (2024) used a stacking ensemble of CNNs and transformer models, improving accuracy. Feature extraction through CNNs, as demonstrated by Chavva and S (2024), enables detection of malignancy patterns, with ResNet and EfficientNet performing well in identifying cancerous lesions. Grad-CAM visualizations, as used by da Silva and de Souza (2024), enhance interpretability by highlighting critical image areas that influenced predictions. Overall, the integration of CNNs, ensemble models, and explainable AI ensures accurate and interpretable oral lesion detection for improved clinical decisions.

2.5.3 Text-based risk prediction and XAI , Multimodal integration.

The system predicts pre-oral cancer risk by analyzing a combination of patient history, and image-based lesion features using advanced methods. It incorporates key risk factors such as tobacco and alcohol use and lesion characteristics to enable personalized and early-stage cancer risk assessment. As demonstrated by Goswami and Bhuyan (2024), integrating patient history with image diagnostics through deep learning improves classification of precancerous stages. To enhance accuracy, the system employs multimodal learning, drawing from Tian and Ma's (2023) work where deep neural networks effectively correlated lesion severity with individual patient risk factors. Recognizing the importance of transparency in medical AI, the system integrates Explainable AI techniques such as SHAP values and Grad-CAM visualizations allowing both patients and clinicians to understand the factors influencing predictions. Inspired by frameworks like EXAIOC from Taresh and Kharthi (2024), this approach enhances interpretability, while ensuring data privacy, as emphasized by Fede and Mantia (2023). Overall, the system combines textual data analysis, visual lesion evaluation, and explainability to deliver accurate, secure, and trustworthy risk probability predictions.

2.6 Tools and Techniques

- **Python** - A versatile, high-level programming language widely used in data science, AI, and web development.
- **Jupiter Notebook** - An interactive coding environment for writing and executing Python code.
- **Git** - A version control system that tracks changes in code and enables collaboration among developers.
- **GitHub** - A cloud-based platform for hosting Git repositories, enabling version control and collaborative software development
- **Scikit-Learn** - A Python library for machine learning, offering tools for classification, regression, etc.
- **Pandas** - A Python library for data manipulation and analysis, providing powerful data structures like Data Frames.

- **TensorFlow** - An open-source deep learning framework for building and training AI models.
- **PyTorch** - A deep learning framework for dynamic computation graphs and ease of use.
- **Numpy** - A fundamental Python library for numerical computing, providing support for arrays and mathematical operations.
- **Flask** - A lightweight python web framework used to build APIs and web applications.
- **HTML, CSS, JavaScript** - The core technologies for building web pages, where

HTML structures content, CSS styles it, and JavaScript adds interactivity.

- **MongoDB** - A NoSQL database used for storing and managing structured and unstructured medical data, ensuring scalability and flexibility in handling patient records and AI model outputs.

2.6 Chapter Summary

The purpose of this project is to develop an AI-based system for the early detection of oral cancer using deep learning techniques. The system will analyze oral images to identify potential cancerous lesions, assess patient risk based on clinical data, and provide an automated diagnosis with explainable AI. This project will first validate the input image to ensure it is an oral cavity scan before applying lesion detection models. It will then assess cancer risk by integrating patient history and image-based diagnostics. Finally, the system will use AI models to classify the detected lesions as benign, precancerous, or malignant while ensuring transparency through explainability techniques like Grad-CAM and SHAP values. Several deep learning methods such as CNNs, Transfer Learning, Ensemble Learning, and Explainable AI models will be utilized to enhance accuracy and reliability in oral cancer diagnosis.

CHAPTER 3: METHODOLOGY

3.1 Chapter Overview

This chapter outlines the methodologies used for research, development, and implementation of the AI-based pre-oral cancer detection system. Various approaches were employed to meet the specific requirements of the project, including deep learning techniques for lesion detection, machine learning techniques for textual risk prediction. The chapter also provides a detailed analysis of the decision-making processes that led to the selection of these methodologies, ensuring accuracy, efficiency, and explainability in the design of the system.

3.2 Research Methodology

Aspect	**Methodology**
Research Philosophy	Positivism in this project reflects a systematic, data-driven approach that uses measurable text and image data, algorithms, and empirical analysis to ensure accurate oral cancer detection.
Research Approach	The project adopts a quantitative approach, using numerical and categorical data along with performance metrics like accuracy, precision, recall, and F1 score to evaluate the AI models for oral cancer detection and risk assessment.
Research Strategy	The experimental design involves evaluating various machine learning and deep learning models to identify the most accurate and efficient architecture for oral cancer classification and diagnosis.
Research Choice	The study uses a mono-method approach, focusing exclusively on structured datasets and applying statistical and exploratory data analysis to uncover insights for oral cancer detection.
Time Horizon	The study uses a cross-sectional approach with a fixed dataset for training, validation, and testing to ensure consistent and comprehensive model performance

Table 6:Research Methodology

3.3 Development Methodology

a. The iteration model allows continuous feedback and improvements, refining the analysis pipeline through iterative enhancements in preprocessing and model training. This approach ensures flexibility in development.

b. The project uses Object-Oriented Analysis and Design (OOAD) to ensure modularity, reusability, and maintainability of code, making it easier to handle different stages of the machine learning pipeline.

c. Evaluation employs benchmarking and performance metrics like accuracy, precision, recall, F1-score, and ROC-AUC, comparing models to select the best-performing algorithm while understanding trade-offs.

d. The Agile framework is followed to ensure adaptability and continuous improvements, enabling the development of a robust, scalable, and interpretable machine learning system.

3.4 Project Management Methodologies

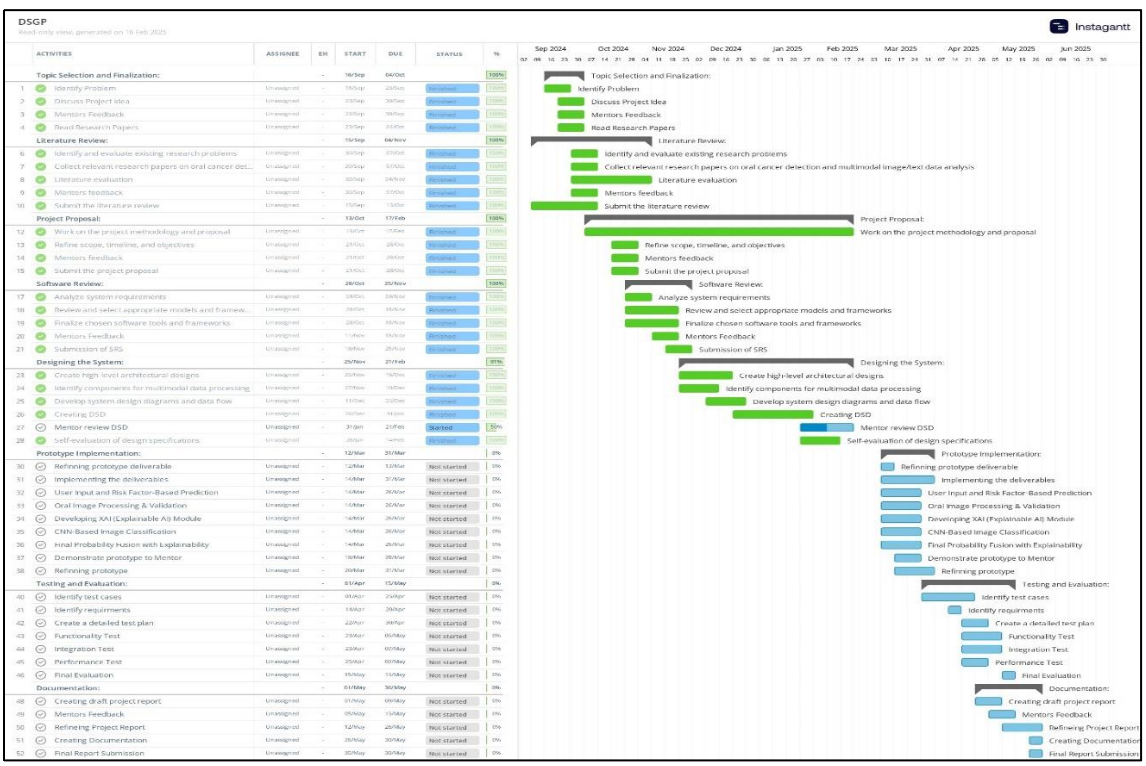


Figure 3: Gantt Chart

Deliverable	**Date**
Semester 01	
Submission of literature review	Week 3
Submission of the Project Proposal	Week 4
Submission of the Project Proposal (Final PP)	Week 5
Submission of the SRS	Week 8
Submission of the SRS (Final SRS)	Week 9
Semester 02	
Prototype Implementation	Week 14
Testing and evaluation	Week 19
Documentation and final report submission	Week 23

Table 7:Project Deliverables

3.5 Chapter Summary

This chapter provides an overview of the methodology used in the project, highlighting the structured approach to analyzing both text-based and image-based data. It details the steps from data preprocessing to model evaluation, with a focus on the use of image-based data for lesion detection. The chapter emphasizes the application of machine learning, deep learning techniques, and risk scoring to enhance the classification system, ensuring the development of a robust and accurate oral cancer detection model.

CHAPTER 4: SOFTWARE REQUIRMENT SPECIFICATION

4.1 Chapter Overview

The methodology involves collecting and analyzing system requirements, beginning with stakeholder analysis using an onion diagram. Various requirement-gathering methods were evaluated, with questionnaire results visualized through pie charts. System workflows and user interactions were illustrated via context and use case diagrams, alongside detailed use case descriptions. Finally, functional requirements and non-functional requirements were identified, documented, and prioritized to ensure the system meets both user needs and performance expectations.

4.2 Rich Picture

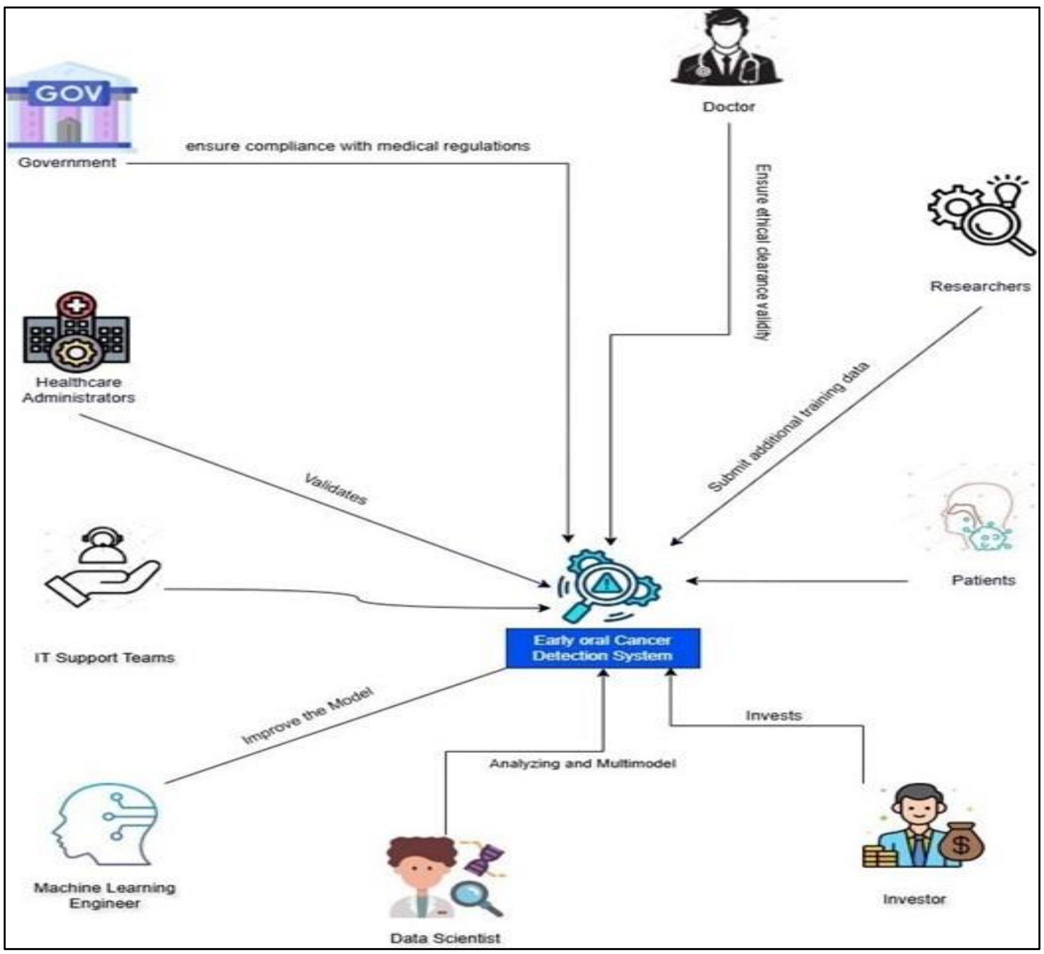


Figure 4: Rich Picture

4.3 Stakeholder Analysis

4.3.1 Onion Model

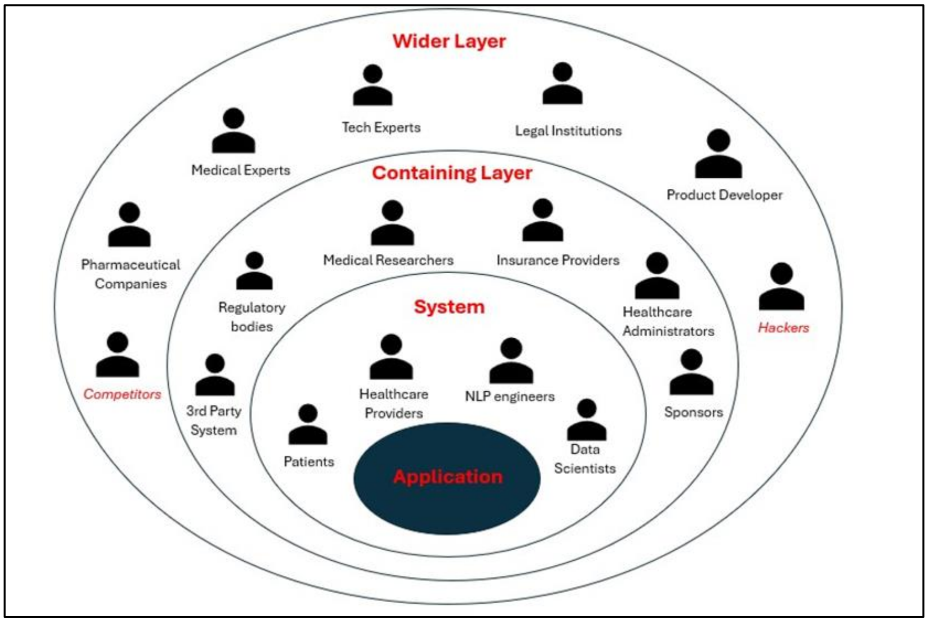


Figure 5: Onion Model

4.3.2 Stakeholder Viewpoints

Stakeholder	Role	Interaction
Patients	Primary end-user	Use the application to check whether they can prone to oral cancer
Healthcare Providers	Secondary end-user	Assist with the diagnosis process, monitor patient record
NLP Engineers, Data Scientists	Operational Maintenance	Create and develop the multimodal application
3rd Party Systems	Providers	Integrating the developed multimodal with other applications to increase its efficiency
Medical Researchers	Experts	Provide data to develop the model, support the application's validation through clinical studies
Insurance Providers	Assessor	Rely on the data provided by the model to determine the coverage needed for the medication

		of the patient
Legal Institutions	Legal beneficiary	Responsible for ensuring the application adheres to healthcare laws, medical device regulations
Healthcare Administrators	Functional Beneficiary	Integrate the model with their local system
Sponsors	Functional Beneficiary	Funding the model's development, testing, and deployment with the aim of gaining a profit
Competitors	Negative Stakeholders	Creating a model which serves the same functionalities and trying to steal the target market
Medical Experts	Experts	Evaluating the model and enhancing its efficiency
Tech Experts	Experts	Evaluating if the technical part of the model's efficiency and if needed add more techniques
Regulatory Bodies	Quality Regulator	Set technology standards, guiding the design, development, and operation of health-related AI applications
Product Developer	Operational Maintenance	Deployment and managing the system
Pharmaceutical Experts	Experts	Integrating the system with the pharmaceutical industry to recommend unbiased early stage treatment methods
Hackers	Negative Stakeholders	Attempting to gain unauthorized access to the secured data and the backend of the model

Table 8:Stakeholder Viewpoints

4.4 Selection of Requirement Elicitation Techniques

The process of defining the needs of a project from the perspective of its stakeholders is known as requirement elicitation. These techniques are essential for adapting the application to the user's demands.

4.4.1 Observing Existing Systems and Literature Reviews

While examining the literature reviews can reveal the viewpoints of other researchers and research gaps, gaining insight into the current systems can be extremely beneficial in identifying the application's flaws and improving its effectiveness.

Advantages	Disadvantages
Provide an understanding of the system's performance Help to identify research gaps Ability to contrast our system with the existing systems and improve efficiency Provide technical insights Ability to know the most efficient methods to use	Takes a long duration of time Some articles are written using complex technical and medical terms Hard to review the backend of the system

Table 9:Literature Review advantages and disadvantages

4.4.2 Surveys & Questionnaires

We were able to determine people's knowledge about oral cancer and the application of AI technology in healthcare as a result of the questionnaire developed and published.

Advantages	Disadvantages
Takes a short duration of time Ability to get responses from a considerable amount of people simultaneously Get an understanding of the public opinion and knowledge about oral cancer	Low response rate a heavy. Inability to address the questions with medical terms Lack of depth in the questionnaire Integrity of the responder

Table 10:Questionnaires advantages and disadvantages

4.4.3 Interviews

In order to collect information and get their perspective on developing this application, we conducted a number of interviews and conversations with key stakeholders, particularly healthcare professionals.

Advantages	Disadvantages
Ability to get the opinion of medical professionals Can address questions with a good depth Direct interaction with the professionals makes it more efficient	Takes a long duration of time The success of the outcome depends on both the interviewer and the responder The answers can be biased

Table 11:Interview advantages and disadvantages

4.4.4 Followed Requirement Gathering Methods

While observing existing systems and literature reviews help us identify research gaps that need to be filled and methods to improve the accuracy of our application, surveys and questionnaires were a much more convenient way to reach a wider community in a short amount of time and determine the general perspectives and awareness of oral cancer detection methods of communities. The data collected was analyzed using a quantitative methodology.

4.5 Discussion of Results

4.5.1. Interview Results

	Impression
	The AI-powered pre-oral cancer detection platform improves consultation efficiency by automating detection, data integration, and report generation, enhancing diagnostic accuracy and saving doctors' time. It is particularly valuable for teleconsultations, overcoming time and distance challenges.

Interview with Dr.Buddhika	Suggestions
	To reach its full potential, the system must be customized to workflows, supported by data privacy measures, and include user training. Error-handling mechanisms and feedback loops can improve reliability, with potential expansion enhancing its impact on healthcare.

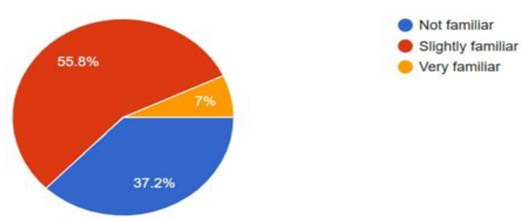
Interview with Dr.Nishan Siriwardhane	Impression
	The system excels at centralizing and digitizing patient records, providing easy access to clinical history and diagnostic results. AI-generated inputs like heatmaps and treatment suggestions enhance diagnostic accuracy, while electronically stamped prescriptions improve authenticity and ease.
	Suggestions
	The system is highly effective in centralizing and digitizing patient records for easy access to clinical history, investigations, and diagnostic results. AI-generated inputs like heatmaps and treatment suggestions enhance diagnostic accuracy, while electronically stamped and signed prescriptions ensure authenticity and convenience.

	Impression
	The AI-powered oral cancer detection platform enhances dental education and practice by streamlining early detection, patient record integration, and report generation, improving diagnostic accuracy and reducing administrative burdens. Its potential for teleconsultations addresses time and geographic barriers, expanding access to quality care.
	Suggestions

Interview with Prof.Niroshani Soysa	To maximize the potential of this innovation, the system must be tailored to dental education and practice needs, with strong data security, user training, and advanced error-handling mechanisms. Expanding its application to teaching clinical diagnosis, case-based learning, and research could significantly boost its impact.

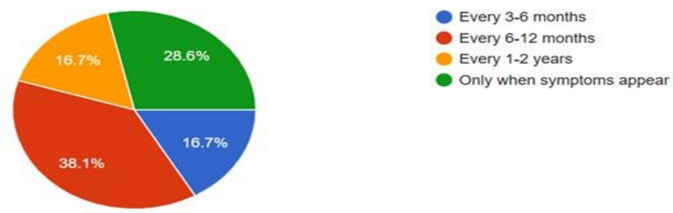
Table 12:Interview Results

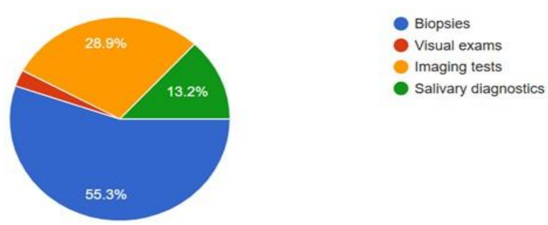
4.5.1. Questionaries Results

Question(01)	How familiar are you with oral cancer and its risk factors (e.g., tobacco use, alcohol consumption, Human papillomavirus (HPV))?
Aim of Question	To assess the awareness of oral cancer and its associated risk factors among participants
Observations	
Conclusion	The majority of participants were only slightly familiar with oral cancer, while 37.2% were unfamiliar and just 7% were very familiar with it.

Question(02)	Are you aware of any early symptoms of oral cancer? If yes, what are they?
Aim of Question	Are you aware of any early symptoms of oral cancer? If yes, what are they?
Observations	
Conclusion	A large portion (62.8%) of participants lack awareness of early oral cancer

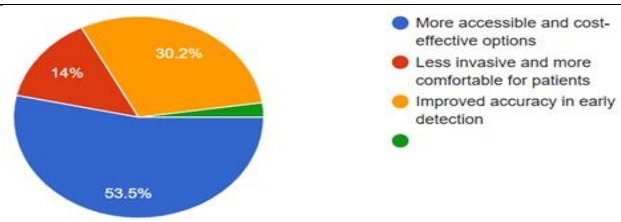
	symptoms, highlighting the need for improved public education.

Question(03)	How often do you think regular screening for oral cancer should be conducted?
Aim of Question	How often do you think regular screening for oral cancer should be conducted?
Observations	
Conclusion	Most participants (38.1%) advocate for regular oral cancer screenings every 6-12 months, while 28.6% think screenings should only occur when symptoms appear, underscoring the need for greater awareness of proactive screening benefits.

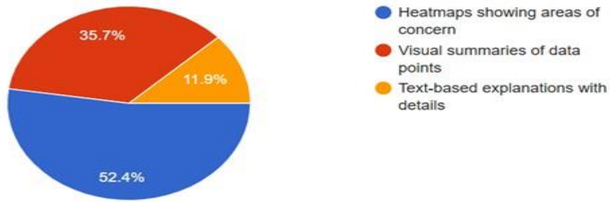
Question(04)	What diagnostic tools are effective for detecting oral cancer in your experience?
Aim of Question	To identify which diagnostic tools are perceived as effective for oral cancer detection
Observations	
Conclusion	Most respondents (55.3%) prefer biopsies as the most reliable oral cancer detection tool, followed by imaging tests (28.9%), favoring invasive methods over non-invasive options.

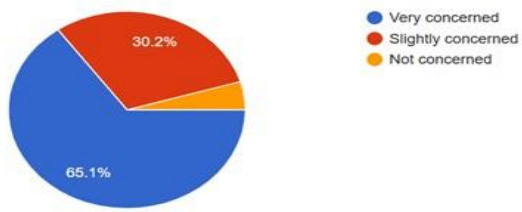
Question(05)	Do you believe current oral cancer detection methods are effective?

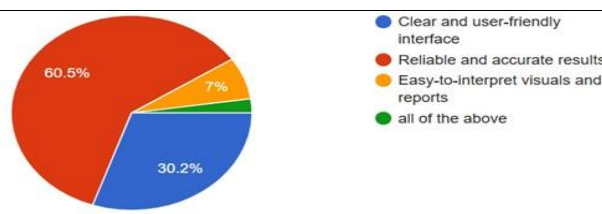
Aim of Question	To assess the effectiveness of current oral cancer detection methods
Observations	
Conclusion	87.8% of participants believe the methods are "Somewhat effective but could be improved," while 7.3% find them "Very effective," and 4.9% think they are "Not very effective; hard to detect in early stages."

Question(06)	Would you be more likely to trust an AI-driven system if it could provide an explanation for its diagnosis?
Aim of Question	Would you be more likely to trust an AI-driven system if it could provide an explanation for its diagnosis?
Observations	
Conclusion	69.8% of participants stated "Maybe, depends on the explanation quality," 27.9% responded "Yes, explanations increase confidence," and 2.3% prefer "No, prefer human-driven diagnosis."

Question(07)	What types of explanations or visualizations would be most useful to understand AI-driven cancer detection results?
Aim of Question	To determine the preferred explanation or visualization methods for understanding AI-driven cancer detection results

Observations	
Conclusion	Most respondents prefer heatmaps for AI-driven cancer detection results, valuing visually intuitive explanations, while some also appreciate visual data summaries, and a smaller group prefers detailed text-based explanations.

Question(08)	How concerned are you about data privacy and security in AI-powered medical applications?
Aim of Question	How concerned are you about data privacy and security in AI-powered medical applications?
Observations	
Conclusion	The findings show a significant concern among users about data privacy and security in AI-powered medical applications, highlighting the need for robust security measures and transparent data practices to build trust.

Question(09)	What features or functionalities would you like to see in an AIbased oral cancer detection system?
Aim of Question	To understand user preferences for features in an AI-based oral cancer detection system
Observations	

Conclusion	Most respondents (60.5%) prioritize reliable and accurate results in an AI-based oral cancer detection system, while 30.2% emphasize ease of use, highlighting the importance of both trust and user experience for adoption.

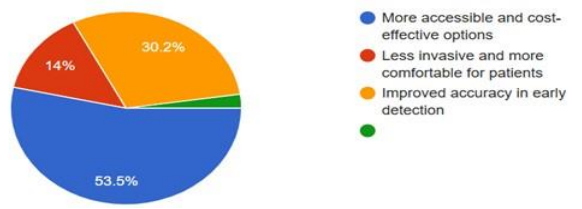
Question(10)	What improvements would you suggest for current oral cancer detection technologies?
Aim of Question	To gather suggestions on potential improvements for existing oral cancer detection technologies
Observations	
Conclusion	Most respondents (53.5%) prioritize accessibility and affordability in oral cancer detection technologies, while 30.2% emphasize the need for better accuracy in early detection, reflecting demand for both cost-effective solutions and improved diagnostic precision.

Table 13: Questionaries Results

4.6 Summary of Findings

Findings	**Literature Review**	**Questionnaire**	**Interviews**
Early detection of oral cancer significantly improves patient outcomes	**X**	**X**	**X**
Most doctors find existing tools insufficient for detecting early-stage oral cancer		**X**	**X**
Combining patient history (text) and lesion images is seen as a valuable approach	**X**	**X**	**X**
AI-based systems must include explainability features (e.g., LIME, Grad-CAM) to gain trust	**X**	**X**	**X**

Real-time lesion image enhancement is critical for reliable AI analysis	X	X	X
Cost-effective and easy-to-use AI solutions are essential for widespread clinical adoption		X	X
Cost-effective and easy-to-use AI solutions are essential for widespread clinical adoption			X
Improvements to current technologies	X		X

Table 14: Summary of Findings

4.7 Context Diagram

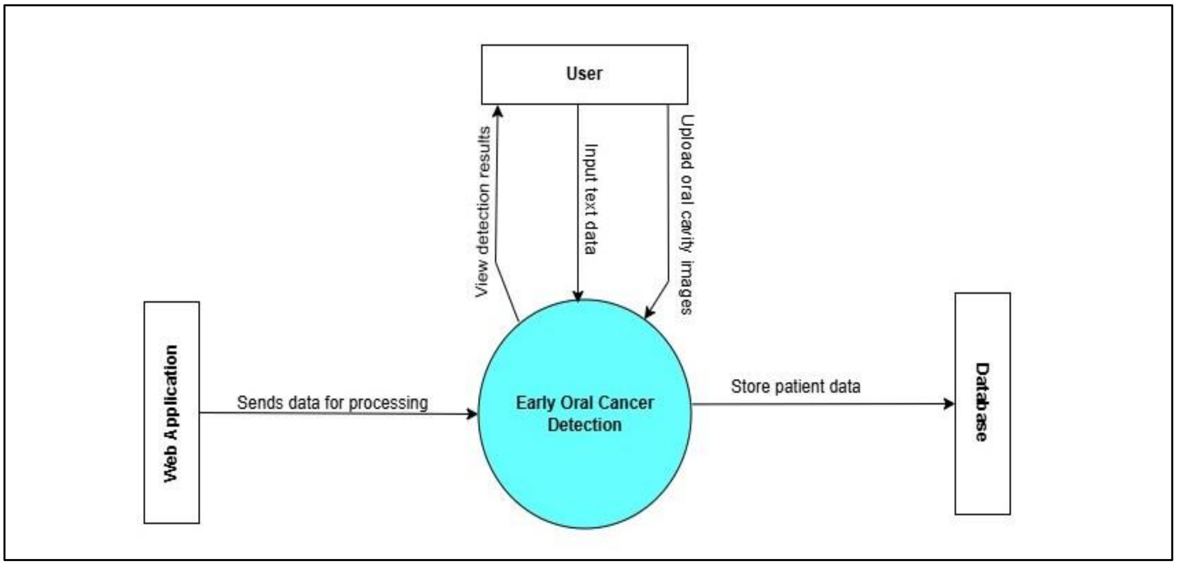


Figure 6: Context Diagram

4.8 Use Case Diagram

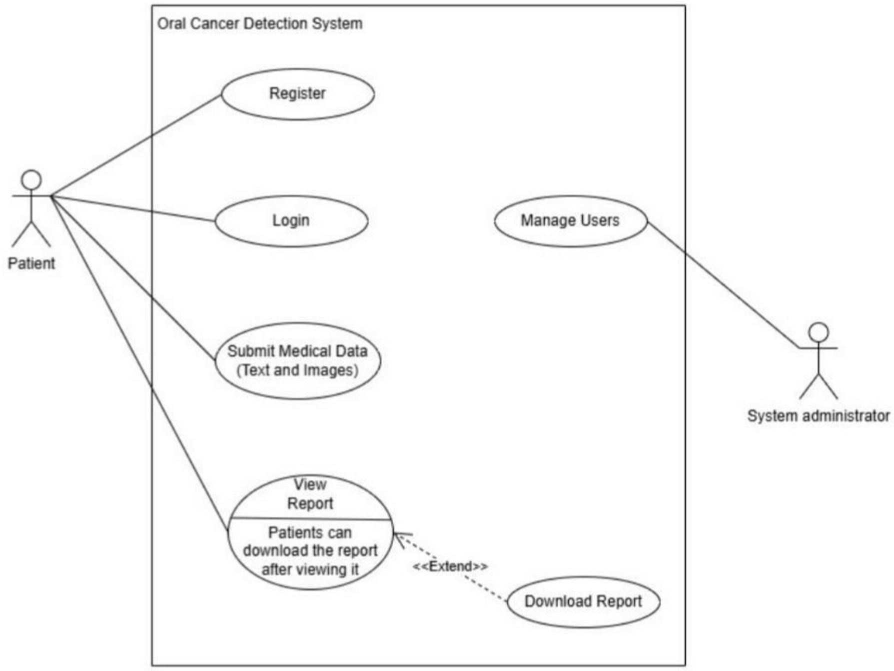


Figure 7: Use Case Diagram

4.9. Use Case Description

Usecase ID	001
Usecase Name	Register
Actor	Patient
Description	The patient registers in the system by providing necessary details such as name, email, and password.
Pre-condition	The patient must not have an existing account
Post-condition	The patient's account is created and stored in the system.
Main Flow	1. Patient enters registration details. 2. System validates the input data. 3. System creates a new account.

	4. Patient receives a confirmation message.
Alternative Flow	If the entered details are invalid, an error message is displayed.

Usecase ID	002
Usecase Name	Login
Actor	Patient
Description	The patient logs into the system using their credentials.
Pre-condition	The patient must be registered in the system.
Post-condition	The patient gains access to their account.
Main Flow	1. Patient enters login credentials. 2. System verifies credentials. 3. If valid, access is granted.
Alternative Flow	If credentials are incorrect, an error message is displayed.

Usecase ID	003
Usecase Name	Submit Medical Data (Text and Images)
Actor	Patient
Description	The patient submits medical information, including text-based data (symptoms, history) and images (oral lesion photos) for analysis.
Pre-condition	The patient must be logged in.
Post-condition	The system stores the submitted data for processing.

Main Flow	1. Patient uploads medical text and images. 2. System validates and stores the data. 3. System processes the data for further analysis.
Alternative Flow	If an invalid format is uploaded, an error message is shown.

Usecase ID	004
Usecase Name	View Report
Actor	Patient
Description	The patient views the diagnosis report generated by the system based on the submitted medical data.
Pre-condition	The patient must have submitted medical data.
Post-condition	The system displays the diagnosis results.
Main Flow	1. Patient requests to view the report. 2. System retrieves and displays the report.

Usecase ID	005
Usecase Name	Download Report
Actor	Patient
Description	The patient downloads the diagnosis report after viewing it.
Pre-condition	The patient must have accessed the report.
Post-condition	The report is saved on the patient's device.

Main Flow	1. Patient clicks the download option. 2. System generates and provides the report for download.

Table 15: Use Case Descriptions

4.10 Functional Requirements

	Requirements and Description	
FR01	Accepting patient data	Critical
	It is required to analyze images of the oral cavity and texts, such as patient's medical history or symptoms	
FR02	Preprocessing the data	Critical
	For images: Remove noise, normalize dimensions, and enhance regions of interest (ROI). For text: This involves cleaning, tokenizing and standardizing of over the input data	
FR03	Feature extraction	Critical
	For images: Pertinent characteristics should be extracted from the low-level image such as texture, color and shape. For text: Identify the key patterns, keywords or medical conditions that may have been discussed in the meeting	
FR04	Multimodal analysis	Critical
	Use both the image and the text data inputs to enhance the general accuracy in early detection of possible oral cancer	
FR05	Classification of patient data	Critical
	With a trained multi model, classify the combined data into normal, potential for cancer or other abnormalities	
FR06	Heat map generation for images	Important
	Heat map: After the classification process of the oral cavity images, generate a heat map that will provide areas of concern that will help in the visualization	
FR07	Patient report generation	Important

	Design a general report to explain the findings of textual data collected from the patients, the main conclusions and risks	
FR08	Generate and display detection results	Critical
	Show the prediction to the user along with the confidence score and give user insights based on both images as well as texts	
FR09	Storage of patient data	Critical
	All the data, results and reports should be well stored and can be accessed in the future and for compliance matters	
FR10	Continuous learning model	Non-important
	Give an option for the system to be trained with other samples with the intention of increasing the extent of correct detections	
FR11	Creating a website for patient interaction	Non-important
	Design an interface for the patient where they can upload their data, look at the reports or interfere with the system	

Table 16: Functional Requirements

4.11 Non-Functional Requirements

	Specification	**Requirement and Description**	**Priority**
NFR01	Accuracy of the system	Accordingly, the detection system has to have very high sensitivity and specificity in order to be very reliable	Important
NFR02	Preprocessing and model operations should be efficient	Preprocessing and classification should not be time consuming even if the amount of data is huge	Important

NFR03	Minimal hardware requirements	The system should require only moderate configuration of hardware to run efficiently	Important
NFR04	Secure storage and access	All patient information should be safely backed up, particularly with regard to access, which should be restricted according to the role of the individual in question	Important
NFR05	User-friendly interface	Design of a web-oriented interface, which would be user-friendly for both patients and clinicians	Non-Important
NFR06	Scalable infrastructure	The system should be able to handle large datasets or have the provision to include more feature in the future	Non-Important

Table 17:Non-Functional Requirements

4.11 Chapter Summary

The chapter commenced with a detailed illustration outlining the model's features. Subsequently, the stakeholders and their involvement with the model were described. The survey's findings were explained, and the reasons for the requirement elimination techniques were acknowledged. Lastly, the use case description was used to collect the functional and non-functional requirements.

CHAPTER 5: SOCIAL, LEGAL, ETHICAL AND PROFESSIONAL ISSUES

5.1 Chapter Overview

SLEP analysis is a framework used to evaluate a system, technology, or policy's wider effects. It assesses social, legal, ethical, and professional aspects. By examining these factors, SLEP minimizes risks and ethical issues while ensuring that inventions are accountable, legal, and advantageous to society.

5.2 SLEP issues and Mitigation

5.2.1 Social Issues

- Users may misinterpret results so proper educational resources must be provided.
- People may avoid using the system due to fear of social stigma associated with cancer diagnoses.
- Gaining public trust in AI-based diagnostics is crucial, as many still prefer traditional medical examinations over AI-driven decisions.
- Doctors might need to learn new things to use the detection system, thus the system needs to be designed in a simple and easy to learn manner.
- Just a diagnosis system, does not provide treatments and should always seek a medical professional advice.

5.2.2 Legal Issues

- The patient data should be secured and kept confidential.
- Dataset was retrieved from Kaggle so the credits should be given to the original owner of the dataset.
- There will be explicit documentation of the model's assumptions, risks, and limitations.
- The literature review's content was retrieved from trusted sources, including research papers and publications, with citations to the relevant authors.

5.2.3 Ethical Issues

- The AI model must be trained on diverse datasets to prevent biases that could lead to incorrect diagnoses for specific demographics.
- Decisions should be interpretable, especially for healthcare professionals.
- The system would assist the expertise but not replace them.
- Patients' full consent is needed in order to obtain their data into the system.
- The patient's privacy and security and the confidentiality of the patient's data should be guaranteed through the system.
- The system should guide users toward professional medical consultation.
- The dataset used to train the model should be collected with proper patient consent and ethical considerations.

5.2.4 Professional Issues

- Doctors and healthcare workers using the system should be properly trained to interpret AI results and provide necessary guidance.
- The consent and guidance of a medical professional is needed to improve the system accuracy.
- To make sure the system satisfies professional healthcare standards and offers accurate diagnoses, it should be thoroughly validated.
- Maintaining system effectiveness requires regular modifications based on feedback, clinical findings, and new research.
- Stakeholders were interviewed in order to gain deeper perspectives
- GitHub and Git were utilized for collaboration and version control.

5.3 Chapter Summary

The different social, legal, ethical, and professional challenges that arose during the system development are examined in this chapter. In order to minimize the possible risks and difficulties that can occur during the product's distribution stage, methods that mitigate these problems must be found.

CHAPTER 6: SYSTEM ARCHITECTURE AND DESIGN

6.1 Chapter Overview

This chapter provides a comprehensive plan and visual representation of the system, encompassing design paradigms, component diagrams, class diagrams, sequence diagrams, UI/UX designs, and process flowcharts. These visual aids are essential for gaining a clear understanding of the system's architecture and the interactions between its various components.

6.2 Design Goals

The following table outlines the key standards to be adhered to during the development of this system, including Compatibility, Data security, and other essential factors.

Design Objectives	**Descriptions**
Compatibility	The system must be compatible with a range of web browsers to ensure broad accessibility
Responsiveness	The website should be designed to be responsive, particularly on desktop devices, to enhance user experience
Privacy and Data protection	Access should be restricted to authorized medical professionals within the designated medical center, achieved through security measures like encryption and secure authentication protocols.
Speed	The models must be optimized for fast inference to deliver prompt feedback to users.
Usability	The website's user interface should be intuitive and user-friendly, incorporating accessibility evaluations to enhance the overall user experience.
Cost - effectiveness	The system should be designed to be cost-effective, both in terms of initial development and ongoing operational costs, ensuring efficient use of resources.

Table 18: Design Goals

6.3 System Architecture Design

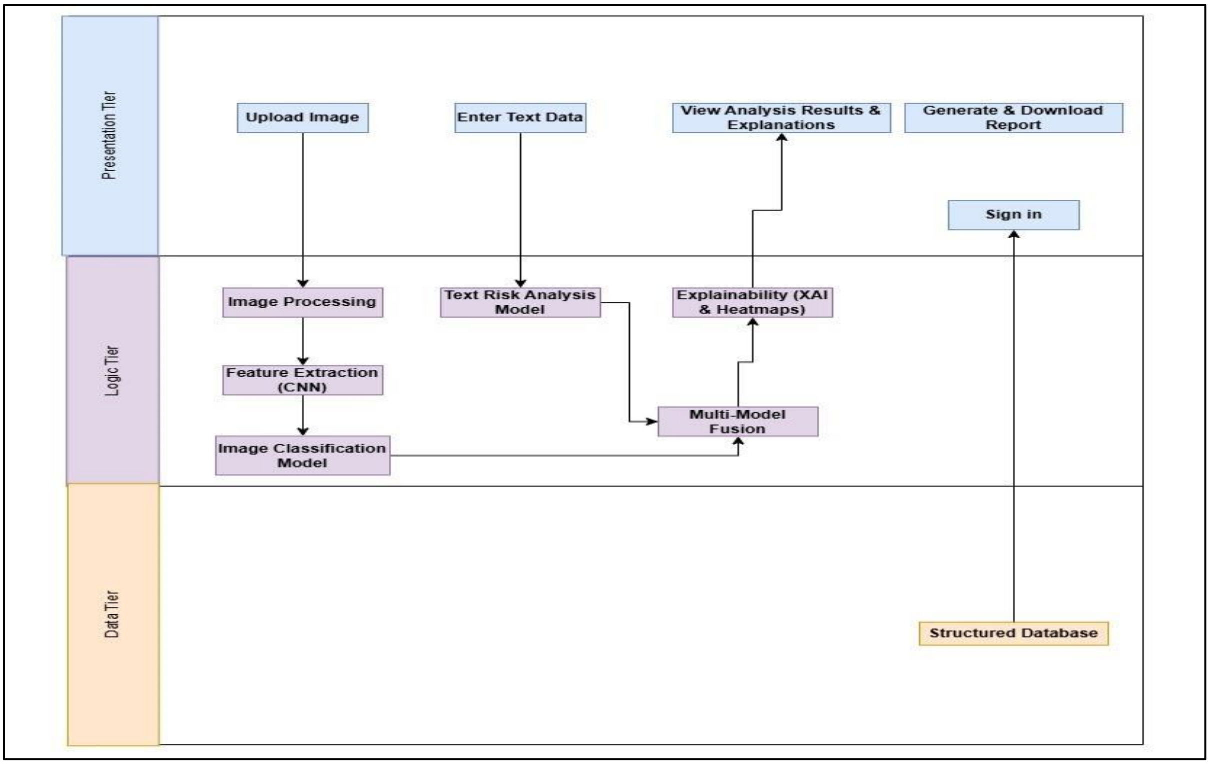


Figure 8: System Architecture Design

6.4 System Design

- SSADM: Structured System Analysis and Design Methodology
- OOAD: Object-Oriented Analysis and Design

6.4.1 Choice of Design Paradigm

Object-oriented analysis and Design (OOAD) models systems as interrelated objects with specific data and behaviors, offering flexibility through encapsulation, inheritance, and polymorphism. Unlike traditional methods like SSADM, OOAD supports modular, adaptable design—ideal for complex systems like early oral cancer detection. By representing components such as patient data and diagnostic tools as separate objects, it enhances scalability and integrates well with modern technologies like machine learning and image processing.

6.4.2 Component Diagram

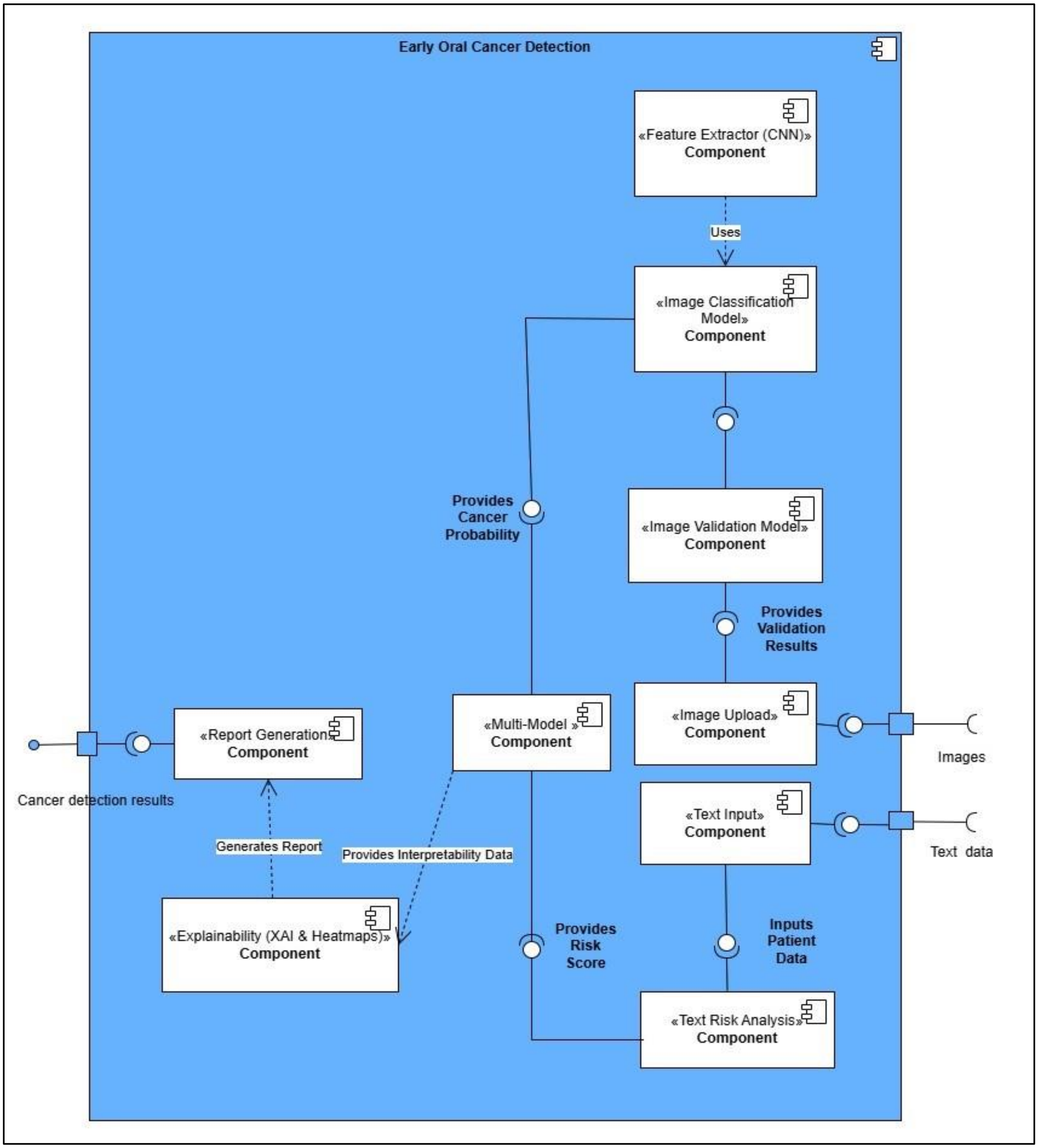


Figure 9: Component Diagram

6.4.3 Class Diagram

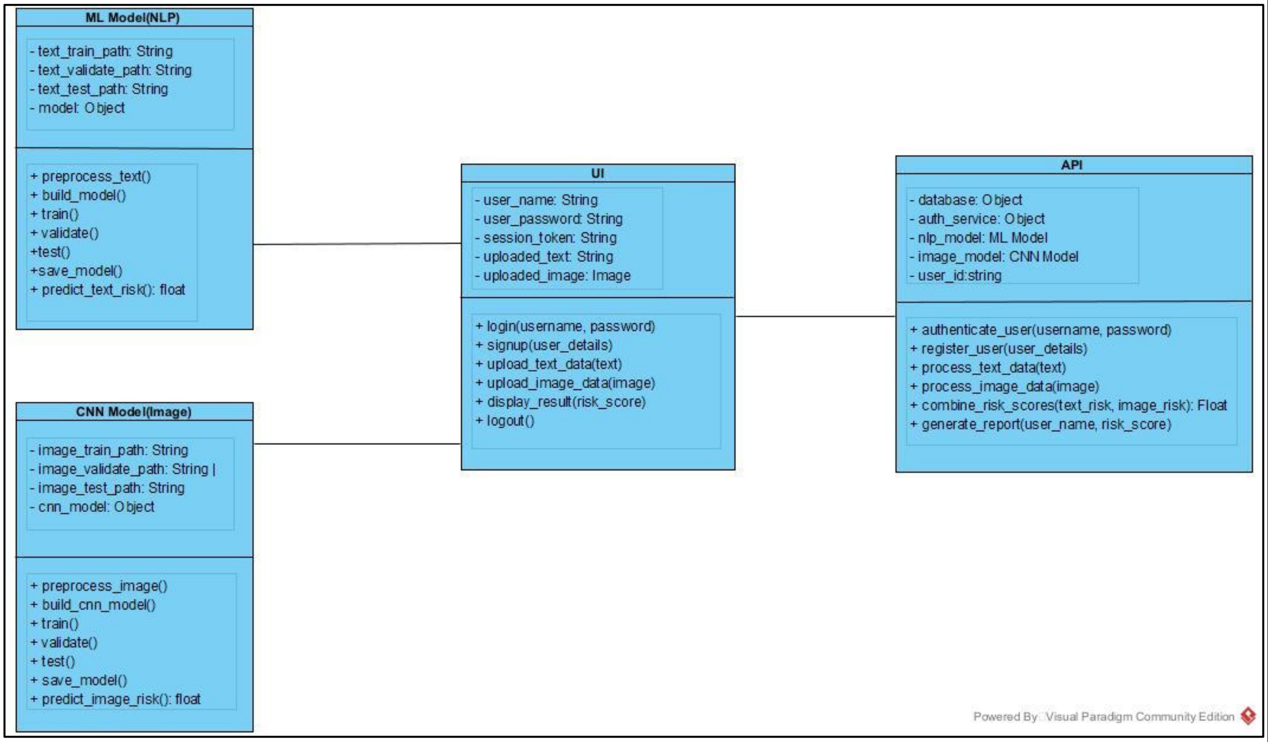


Figure 10: Class Diagram

6.4.4 Sequence Diagram

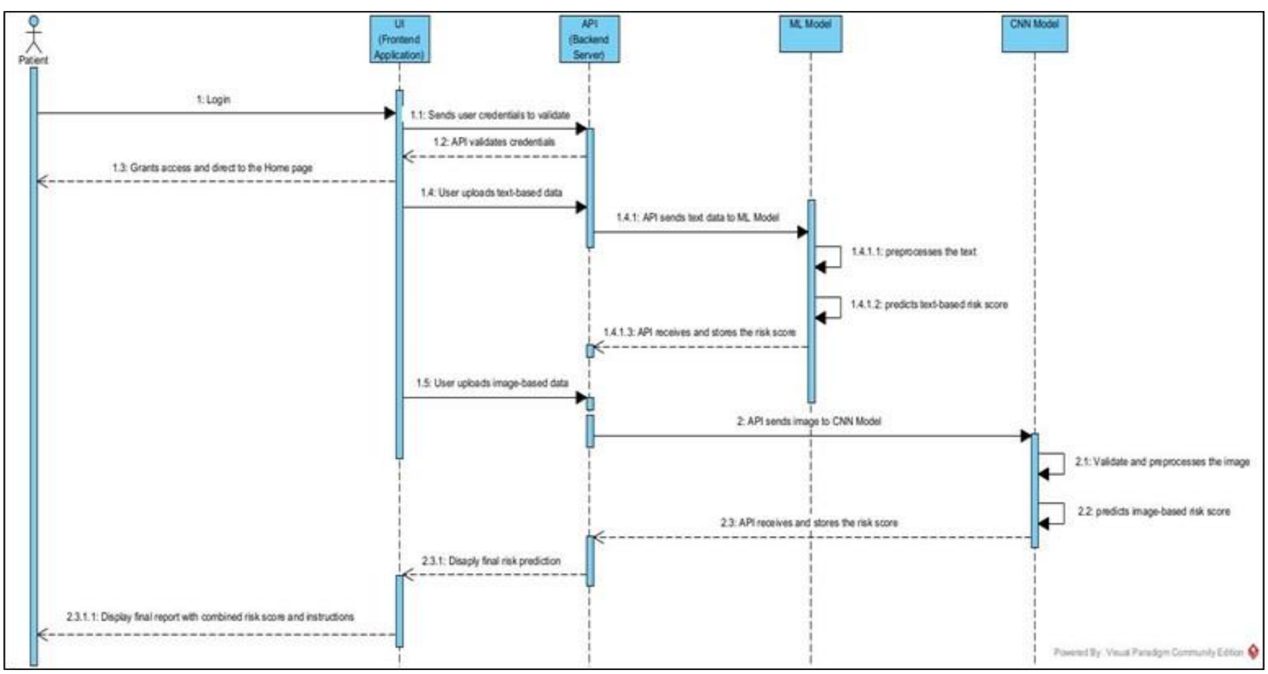


Figure 11: Sequence Diagram

6.4.5 UI Design

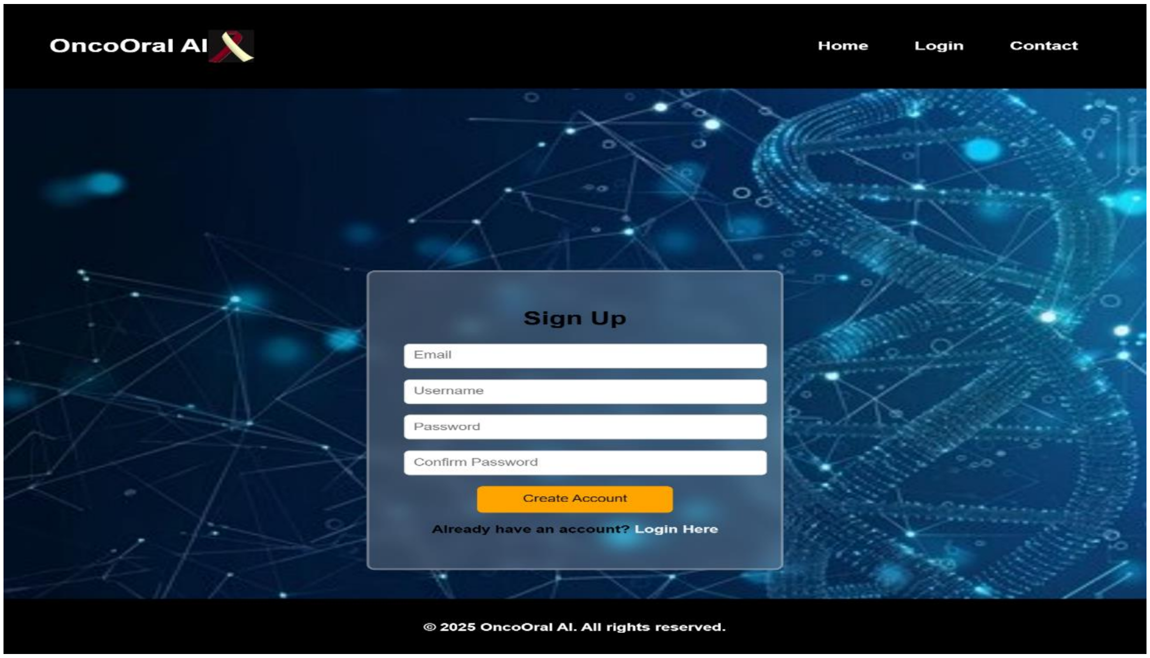


Figure 12: OncoOral AI - Sign Up page

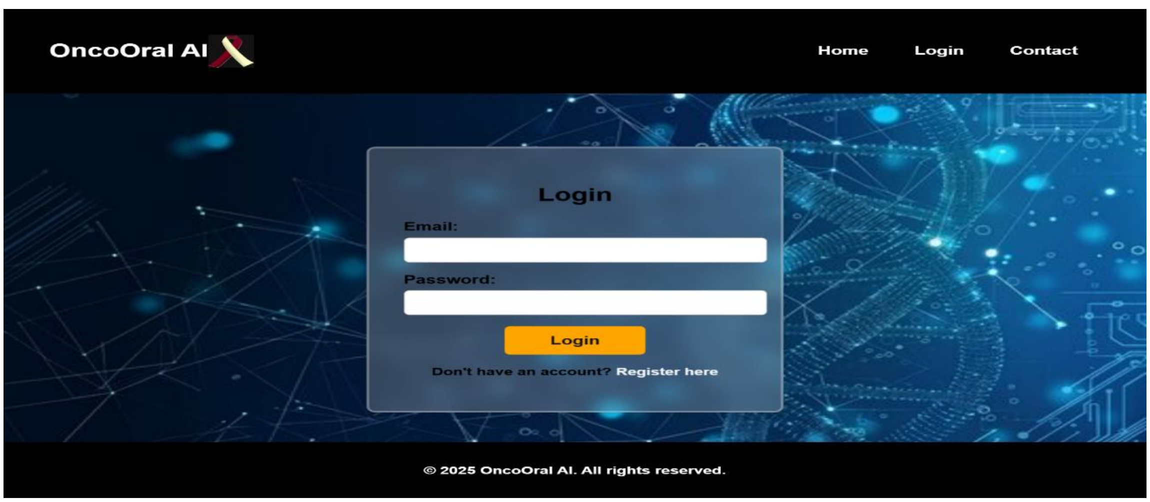


Figure 13: OncoOral AI - Login page

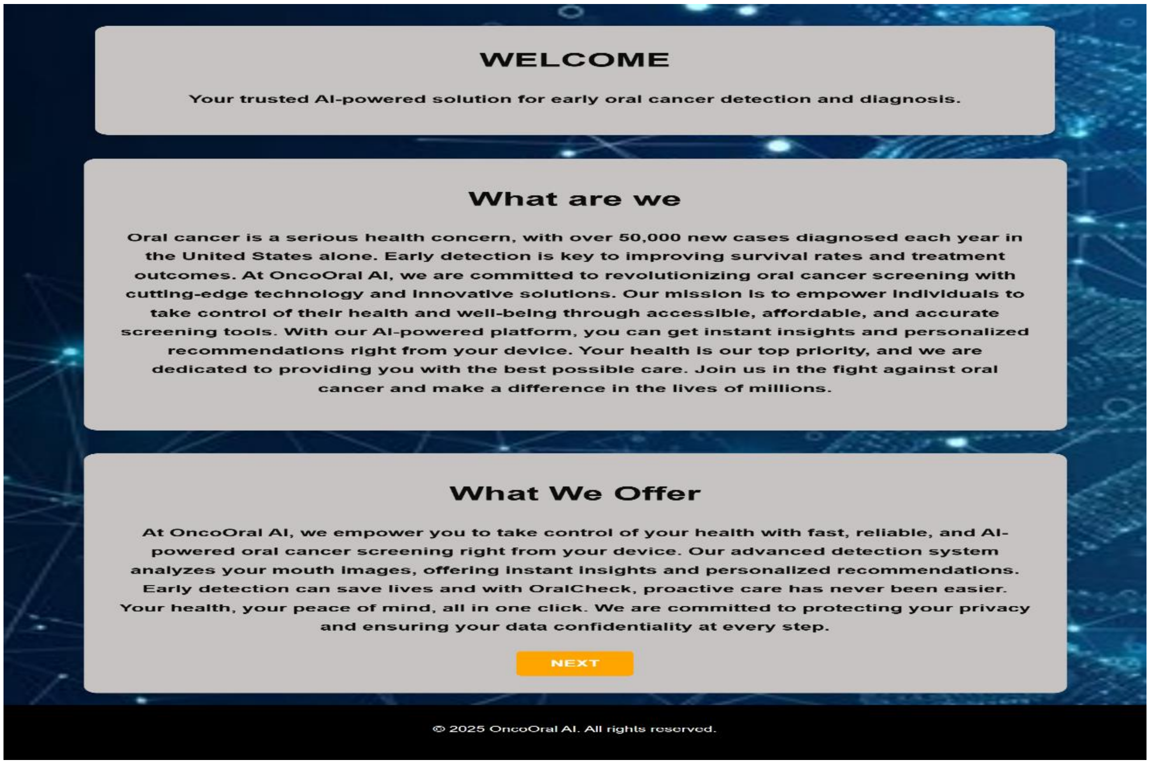


Figure 14: OncoOral AI - Home page_1

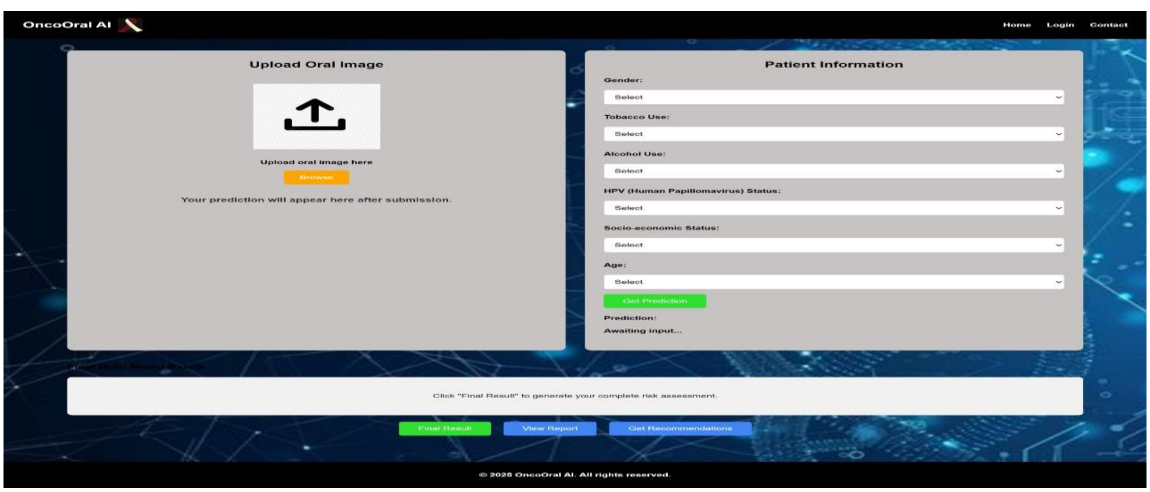


Figure 15: OncoOral AI - Home page_2

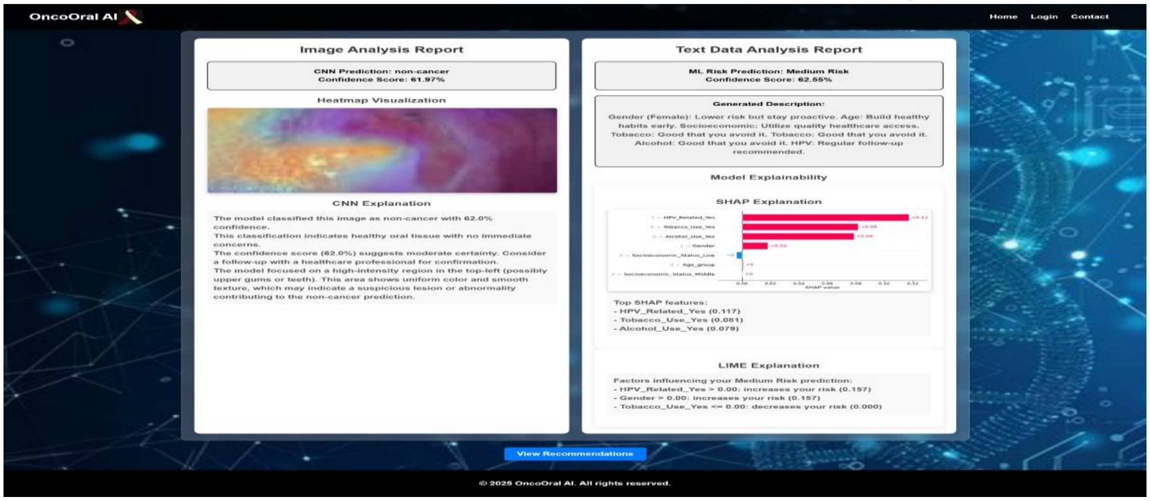


Figure 16: OncoOral AI- Report

6.4.6 User Experience

1. **User-Friendly Interface:**

   The interface is designed to be simple and minimalistic, ensuring easy navigation without distractions. Clear fonts, organized layouts, and familiar icons enhance user understanding.

2. **Accessibility & Inclusivity:**

   The application adheres to WCAG standards, ensuring accessibility for users with disabilities. It includes features like keyboard navigation, alt text for images, and high-contrast modes for better usability.

3. **Intuitive Navigation & Workflow:**

   The menu and layout are logically organized for quick access to features. A step-by-step workflow guides users smoothly through each process.

4. **Performance Optimization:**

   The application is optimized for fast loading, minimizing delays in accessing information. Lazy loading ensures content is loaded only when needed for a smooth user experience.

5. **Secure & Personalized User Experience:**

   Sensitive patient data is accessible only to authorized medical professionals, ensuring privacy and security through role-based access control. Personalized dashboards offer quick access to relevant patient records and AI-generated insights based on user permissions.

6. **Error Handling & Feedback Mechanisms:**

   The system gives clear error messages and validation prompts, helping users fix mistakes quickly instead of guessing what went wrong. Real-time feedback notifications confirm successful actions, so users always know what is happening in the system.

7. **Seamless Integration with Machine Learning Models:**

   The application presents AI-generated insights in a clear and interpretable way using visual heatmaps, probability scores, and explanations. Users can interact with these insights while maintaining full control over decisions, making AI-assisted diagnosis more effective.

6.4.6 Process Flow Chart

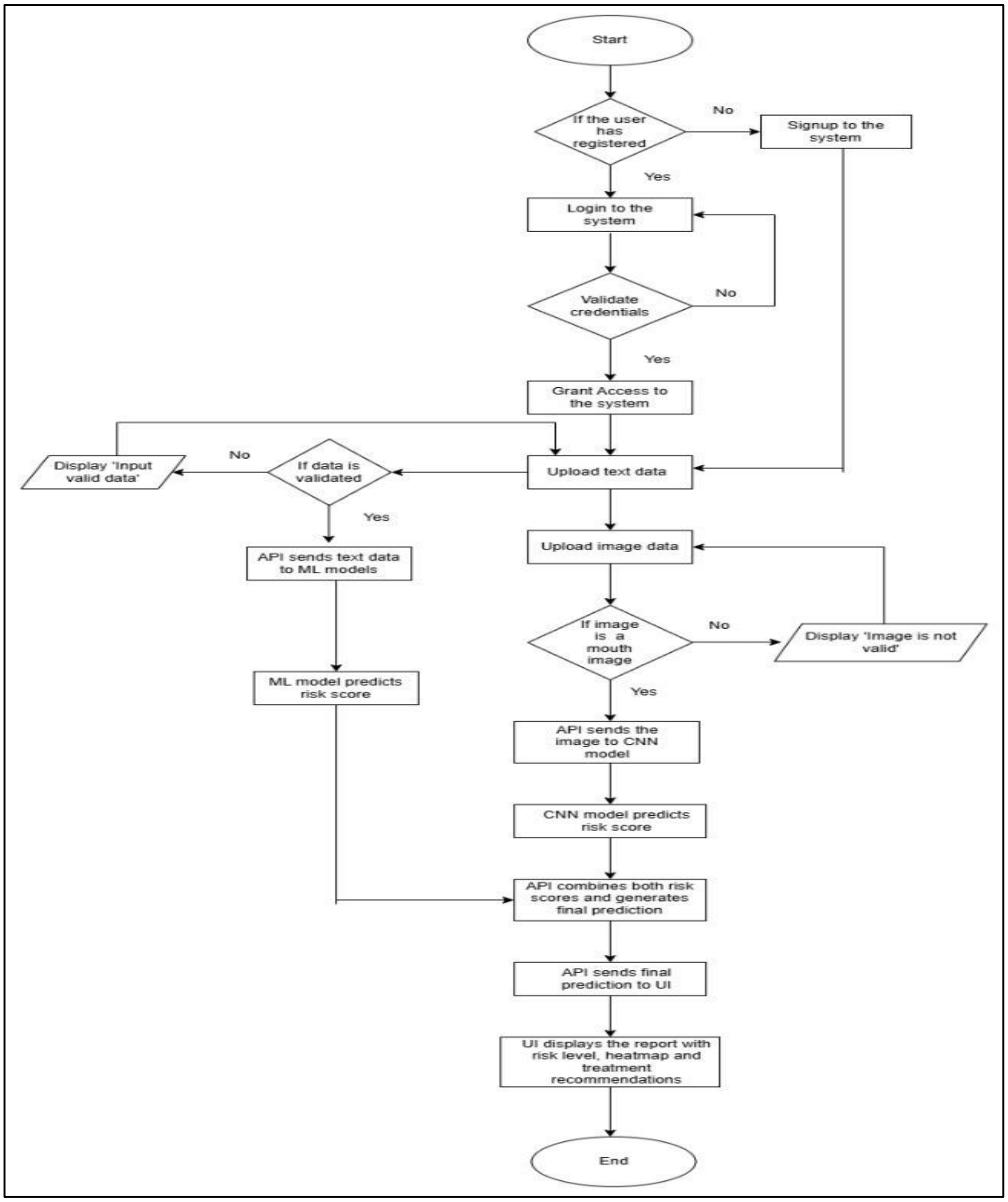


Figure 17: Process Flow chart

6.5 Chapter Summary

The design components used for the system's visual representation have been descriptively analyzed in this chapter. The application's design objectives are descriptively shown in the component diagram, class diagram, sequence diagram, and UI/UX design.

CHAPTER 7: IMPLEMENTATION

GIT URL : https://github.com/tharushaliyanagama/OralCancerEarlyDetection-DSGP.git

7.1 Chapter Overview

The main tools and technologies utilized in the development of the multimodal method for oral cancer detection will be examined in this section. The technology stack, data selection techniques, project development frameworks, libraries and languages used, and implementation in pseudocodes will all be assessed in this section.

7.2 Technology Selection

7.2.1 Technology Stack

The application was implemented using a variety of languages and technologies. HTML, CSS, and JavaScript were used for the front end, while Flask and Python were used for the back end. The implementation is version-controlled and hosted on GitHub.

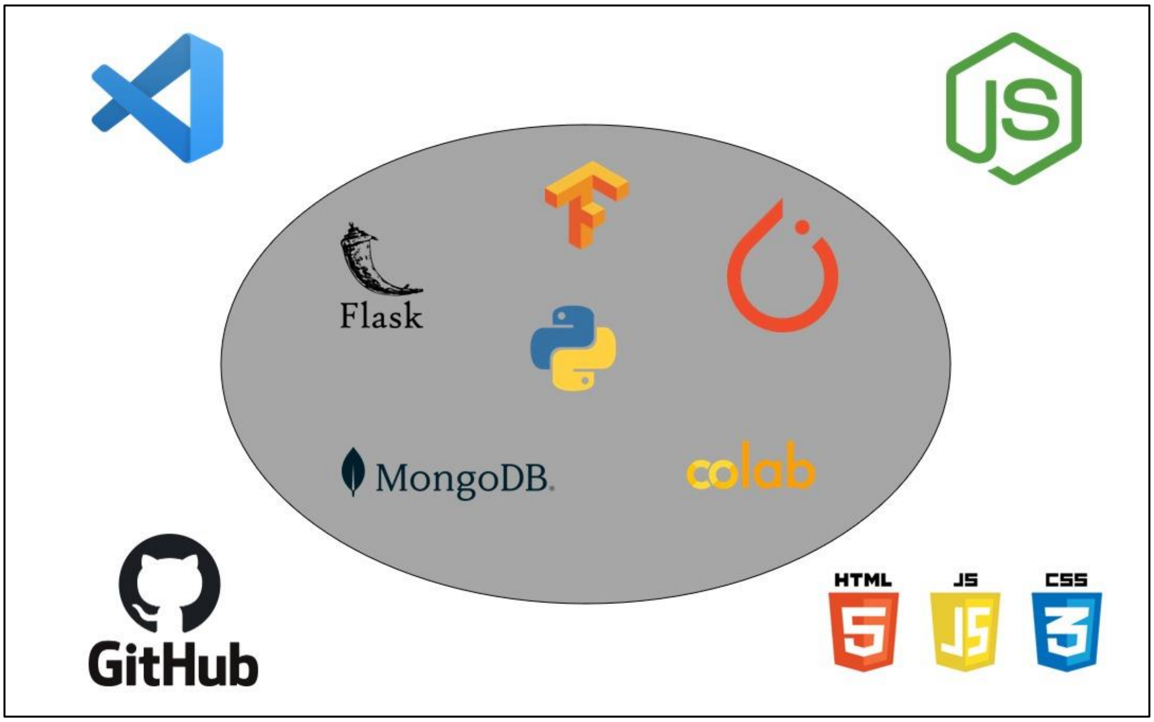


Figure 18:Technology Stack

7.2.2 Data Selection

The selection of the dataset was quite a challenging task because health-related data faces ethical issues, and it played a vital role in creating the machine learning model. Textual data and images of oral cancer are needed for the study. It was necessary to remove the histopathological data from the majority of the datasets. The Kaggle website provided the datasets. The dataset was selected based on the quality and relevance of the data, the number and variety of the samples, and the integration compatibility of the image and textual data.

Domain	Dataset	Description
Textual data prediction of oral cancer	Dataset containing risk factors of oral cancer	The dataset contains risk factors contributing to oral cancer across different countries to understand regional variations. The dataset sourced from Kaggle is publicly available. https://www.kaggle.com/code/sonawanelalitsunil/oral-cancer-prediction-top-30-countries
Image data prediction on oral cancer	A combination of datasets containing cancerous and non-cancerous images	The dataset contains images of cancerous and non-cancerous lesions as well as images of leukoplakia. The dataset was created by combining a few datasets available in Kaggle.

Table 19: Data Selection

7.2.3 Selection of Development Framework

Frontend

The frontend of our early cancer detection system is built using HTML, CSS and Javascript ensuring a dynamic, interactive user friendly experience. The system allows patients to input clinical data and upload lesion images and receive risk scores with explainability insights.

- Real – time visualization of risk scores, reports and heatmaps.
- Display of patient reports, including risk level, probability scores and interpretability results.
- Integration of Explainable AI reports to enhance model transparency.
- User guidance and instructions on interpreting the results and taking appropriate next steps.

Backend

The backend is powered by python and Flask, handling data processing, model execution, database and API integration. It enables seamless communication between models and the frontend by,

- Processing of patient input data.
- Running models to predict cancer risk levels based on clinical data.
- Classifying Lesion images using the CNN model.
- Combining text-based and image-based results to generate a final risk score.
- Generating Explainable AI reports

The Flask serves as the communication bridge between the models and the front end, delivering final predictions with interpretability insights.

7.2.4. Programming Languages

Python is the core programming language for model training, explainability and backend development. It provides a versatile and extensive ecosystem of libraires to support machine learning, deep learning and XAI techniques.

I. Text data prediction
   - Implemented using Google Colab, where clustering algorithms are applied to organize patient records into risk categories.
   - Supervised learning models are trained to classify patients into risk levels based on their textual medical inputs.

II. Image Analysis
   - CNNs are used for image classification to determine whether an oral lesion is precancerous, cancerous, or non-cancerous.

- Advanced architectures like Resnet50 improve feature extraction for higher classification accuracy.

Both text-based and image-based results are combined to provide a final risk assessment. XAI techniques such as SHAP (Shapley Additive Explanations), LIME (Local Interpretable Model-agnostic Explanations) are applied to explain how input features contribute to the final decision and Grad-CAM (Gradient-weighted Class Activation Mapping) is used to generate heatmap, visually highlighting areas of concern in lesion images.
This ensures that medical professionals can understand why a particular risk level was assigned rather than relying on model outputs.

Flask is a python web framework, and is utilized to integrate both text-based predictions and image classification models into the web application. React is a JavaScript library used for building a responsive and interactive user interface.

7.2.5 Libraries

Libraries	**Version**
TensorFlow	2.15.0
Keras	2.15.0
PyTorch	2.5.1+cu124
Numpy	1.26.3
Sklearn	1.4.1.post1
Pandas	2.2.1
Matplotlib	3.8.2
Flask	3.0.2
Torchvision	0.20.1+cu124
joblib	1.4.2
Pickle	4.0

Table 20: Libraries

7.2.6 IDE

Visual Studio was used as the primary IDE for developing the CNN-based image classification model. It provides a comprehensive set of features that facilitate efficient coding, debugging, and version control throughout the development process. Google Colab was also used to enhance the efficiency of textual data preprocessing and model training within the system. It offers access to powerful CPU and GPU resources, enabling faster model training while ensuring seamless collaboration and easy integration with machine learning libraries.

7.2.7 Summary of Technology Selection

Component	Technology/Tool	Version
Programming Language	Python	3.12
UI Frameworks	HTML CSS React	18.2.0
IDE	Visual Studio Google Colab	1.9.7.2

Table 21: Summary of Technology Selection

7.3 Implementation of Core Functionalities

Component 1: Oral Image Validation

BEGIN

MOUNT Google Drive

IMPORT required libraries

INITIALIZE ImageDataGenerator for image preprocessing

SET training and testing dataset directories

INITIALIZE ImageDataGenerator for training and testing

NORMALIZE pixel values (rescale by 1/255)

LOAD training and testing data from directories

RESIZE images to 224x224

LOAD pre-trained ResNet50 model

FREEZE all layers in the base ResNet50 model

ADD custom layers for classification
- Flatten output of base model
- Add Dense layer with 128 units and ReLU activation
- Add Dropout layer (rate = 0.5)
- Add final Dense layer with 1 unit and sigmoid activation (binary output)

COMBINE base ResNet50 and custom layers to create final model

COMPILE model using Adam optimizer

TRAIN model using 10 epoch

EVALUATE model on test data

END

Component 2: Oral Lesion Detection and XAI

BEGIN

DEFINE model_path

```
DEFINE categories ← ['cancer', 'non-cancer', 'leukoplakia']

DEFINE image_size ← (224, 224)

DEFINE test_image_path

INITIALIZE class_to_idx ← {'cancer': 0, 'non-cancer': 1, 'leukoplakia': 2}

INITIALIZE idx_to_class ← CREATE_DICTIONARY (indices → categories)

INITIALIZE transform ← COMPOSE([

    RESIZE(image, (256, 256)),

    CENTER_CROP(image, image_size),

    CONVERT_TO_TENSOR(image),

    NORMALIZE(image, mean=[0.485, 0.456, 0.406], std=[0.229, 0.224, 0.225])

])

LOAD model ← ResNet50(weights=None)

SET num_ftrs ← model.fc.in_features

MODIFY model.fc ← Linear(num_ftrs, 3)

LOAD model_weights

SET model.state_dict ← model_weights

SET model TO evaluation_mode

LOAD new_image ← READ(test_image_path)

image ← CONVERT_TO_RGB(new_image)

image ← RESIZE(image, (256, 256))

image ← CENTER_CROP(image, image_size)

image ← CONVERT_TO_TENSOR(image)

image ← NORMALIZE(image, mean=[0.485, 0.456, 0.406], std=[0.229, 0.224, 0.225])
```

```
image ← ADD_BATCH_DIMENSION(image)  # Add batch dimension (1, C, H, W)

WITH NO_GRADIENT_COMPUTATION DO

    PREDICT output ← model(image)

    predicted_idx ← ARGMAX(output, dim=1)

    predicted_class ← idx_to_class[predicted_idx]

    confidence ← SOFTMAX(output)[predicted_idx]

OUTPUT "Prediction: " + predicted_class + " (" + (confidence * 100) + "% confidence)"

END
```

Component 3: Text-based risk prediction and XAI, Multimodal Integration

```
BEGIN

    define class for clustering, model training, and prediction pipeline:

        define function to initialize variables:
            SET dataset_path to specified dataset file
            SET dataframe to None
            SET X to None
            SET y to None
            SET X_train to None
            SET X_test to None
            SET y_train to None
            SET y_test to None
            SET model to None
            SET scaler to None
            SET kmeans to None

        define function to load_dataset:
            Load dataset from the specified path
```

define function to preprocess_data:

Handle missing values if any

Encode categorical variables if present

Normalize numerical features using StandardScaler

define function to apply_clustering:

Initialize KMeans clustering with specified number of clusters

Fit KMeans model to dataset

Assign cluster labels as target feature with 3 classes: Low, Medium, High

define function to split_data:

Split the dataset into features (X) and target variable (y)

Split the data into training and testing sets

define function to build_model:

Initialize a classifier with specified hyperparameters

define function to train_model:

Train the classifier using training data

define function to evaluate_model:

Make predictions on the test dataset

Calculate accuracy and other performance metrics

define function to save_model:

Save the trained model to a file

define function to predict_risk_from_user_input:

Accept user inputs from form

Preprocess inputs

Predict risk level using the trained model

define function to integrate_multimodal_prediction:

Accept both image-based and text-based predictions

Combine predictions using a meta-classifier or confidence-based rule

Return final fused risk prediction

define function to generate_explainability_reports:

Use SHAP to generate feature importance and visualization

Use LIME to generate feature contribution explanation

Save SHAP and LIME plots for display

define function to generate_custom_user_report:

Based on user input and prediction

Generate personalized description for risk factors and lifestyle recommendations

Initialize class instance

Load dataset

Preprocess data

Apply clustering to generate target labels

Split data

Build and train model

Evaluate model

Save model

On user input:

Predict risk level using text data

Get image-based prediction

Integrate both using multimodal function

Generate SHAP and LIME explanations

Generate user-specific behavioral/lifestyle report

END

7.4 Chapter Summary

The chapter outlines the system's technological and methodological implementations, including deep learning-based image classification, multi-class detection, clustering, machine learning training, and Explainable AI techniques. It presents system operations through pseudocode, covering dataset preprocessing, model training, classification, and interpretability, with a focus on performance evaluation and accuracy assessment in the next section

CHAPTER 8: TESTING

8.1 Chapter Overview

This chapter focuses on the objectives of performance evaluation and testing for the system. It delves into the evaluation of four models, analyzing their accuracy, F1-score, precision, recall, and confusion matrix. The chapter includes functional testing, benchmarking, and both module and integration testing. Additionally, it addresses non-functional testing aspects such as load balancing, accuracy measurement, and overall performance evaluation.

8.2 Objectives and Goals of Testing

In order to guarantee a model's efficacy, dependability, and appropriateness for its intended use, testing objectives and goals are essential. The following are the primary goals:

- Ensuring that the model minimizes errors and makes accurate predictions in order to carry out the tasks for which it was created.
- Assessing elements such as dependability, usefulness, and load handling.
- Evaluating the model's applicability and usefulness for actual situations and making sure it accomplishes its goal.
- Assessing the model's scalability, speed, and effectiveness when working with different datasets or in different scenarios.

These goals guarantee that the model achieves its intended purposes while preserving its dependability, quality, and usability.

8.3 Testing Criteria

Functionality Testing Criteria

- Oral image detection:

Test whether the model has the ability to correctly identify the input image is an oral image or not.

- Cancer and pre-cancer stage detection:

Test whether the model has the ability to identify whether the image is cancerous, pre-cancerous or non-cancerous.

- Risk factors prediction:

Confirm the system can predict the probability of having cancer by analyzing the textual data

Non-functionality testing criteria

- Performance:

Evaluating the system's speed, response time, and throughput under various workloads.

- Usability:

Assessing the system's usability, including its navigational ease, interface design, and overall user experience.

- Reliability:

Verifying the system's capacity to function reliably and consistently over an extended period of time.

- Security:

Verifying that the system is safe against flaws, illegal access, and users.

8.4 Model Evaluation

Model evaluations were conducted using the confusion matrix and classification report due to the comprehensive insights they offer into the models' effectiveness and performance.

A tabular summary contrasting the model's predictions with the actual class labels is given by the confusion matrix. It emphasizes four crucial metrics:

- True Positives (TP)
- True Negatives (TN)
- False Positives (FP)
- False Negatives (FN)

The classification report provides a thorough assessment by computing important performance metrics, such as:

- Precision: The percentage of accurate positive forecasts among all positive forecasts.
- Recall: The percentage of real positive cases that the model correctly detected.
- F1-score: Balances precision and recall, is the harmonic mean of the two measures.
- Accuracy: The model's overall proportion of accurate predictions.

8.4.1 Detecting Oral Images

resNet50, InceptionV3 and mobileNetV2 models were implemented in order to determine whether the input image is an oral image or not. A dataset consisting of 2603 oral images and 651 non-oral images was used to train the dataset.

Model	Description	Train accuracy	Train Loss	Validation Accuracy	Validation Loss	Remark
mobileNetV2	10 epoch	0.9976	0.0105	0.9493	0.2118	Overfitting
InceptionV3	10 epoch	0.9877	0.0343	0.9524	0.1333	Overfitting
resNet50	10 epoch	0.8750	0.3050	0.8710	0.2884	

Table 22: Detecting oral images-Model Testing

The above table visualizes the fact that both mobileNetV2 and InceptionV3, two pre-trained models, were not reliable because the validation loss is higher than the training loss, indicating potential overfitting. A resNet50 model was used, and the model demonstrated consistent performance with good accuracy across both training and validation sets. Its architecture, which was customized to the features of the dataset, improved generalization and lessened the overfitting problems that occurred in the other two models.

Finalize Model: As a result, switching to the resNet50 model produced better outcomes and increased dependability for our classification task.

Classification Report

```
Classification Report:
              precision    recall  f1-score   support

       mouth       0.85      1.00      0.92       552
   non_mouth       0.00      0.00      0.00        99

    accuracy                           0.85       651
   macro avg       0.42      0.50      0.46       651
weighted avg       0.72      0.85      0.78       651
```

Figure 19:Detecting oral images-Classification Report

Confusion Matrix

Metric	**Description**	**Value**
True Positive	Number of oral images classified as oral images	552
True Negative	Number of non-oral images classified as non-oral images	99
False Positive	Number of non-oral images classified as oral images	0
False Negative	Number of oral images classified as non-oral images	0

Figure 20: Detecting oral images- Confusion Matrix

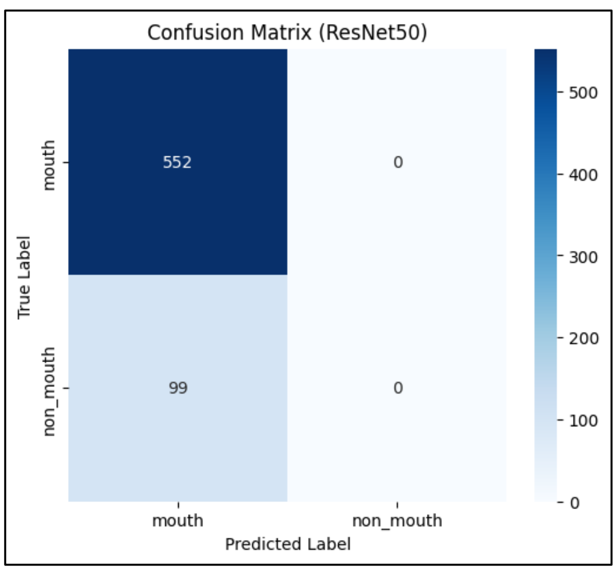


Figure 21: Detecting oral images-Confusion Matrix

8.4.2 Oral Lesion Detection and XAI

The deep learning model development utilized both resnet50 and EfficientNet to identify oral cancer images that fall into three separate categories. The trained model used images from 2,900 cancer cases and 100 cases each of leukoplakia and non-cancer for classification of malignant conditions with precision.

Model	Description	Train Accuracy	Train Loss	Validation Accuracy	Validation Loss	Remark
ResNet-50	Deep residual network for image classification	85.57%	0.9496	87.16%	1.2007	Strong generalization, slightly overfits
EfficientNet	Optimized CNN with efficient architecture	96.2%	0.10	91.7%	0.22	Lightweight, better validation accuracy

Table 23: Oral Lesion Detection and XAI -Model Testing

The results from the table demonstrate that the ResNet-50 pre-trained model achieved excellent performance through exceptional training accuracy. The difference between validation and training loss levels indicates model overfitting because the model learned training dataset patterns very well yet failed to apply this knowledge to new images. The deep architectural design of ResNet-50 guaranteed reliable performance although it learned complex features from oral cancer images effectively. Further optimization of ResNet-50 as a classification will achieve improved balance between training and validation performance which ensures dependable prediction accuracy in real-world applications.

Finalize Model - Based on the evaluation results, ResNet-50 was selected as the finalized model due to its strong training performance and ability to extract complex features from oral cancer images.

Classification Report

```
Classification Report:

              precision    recall  f1-score   support

      cancer       0.98      0.86      0.92       735
  non-cancer       0.88      0.93      0.91       297
 leukoplakia       0.28      0.77      0.42        43

    accuracy                           0.88      1075
   macro avg       0.72      0.85      0.75      1075
weighted avg       0.93      0.88      0.90      1075
```

Figure 22: Oral Lesion Detection and XAI -Classification Report

Confusion Matrix

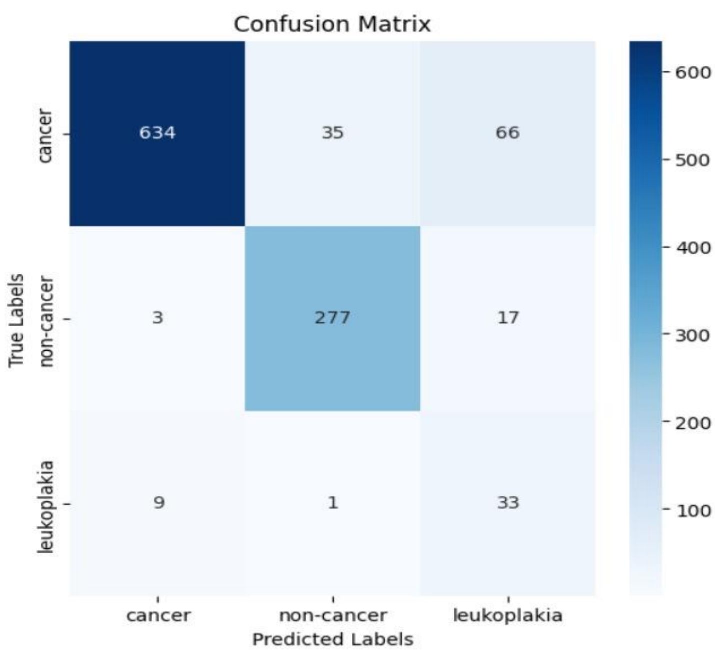


Figure 23: Oral Lesion Detection and XAI -Confusion Matrix

Metric	Description	Value
True Positive	The number of cancer images classified as cancer	634
	The number of leukoplakia images classified as leukoplakia	277
	The number of non-cancer images classified as non-cancer	33

False Positive	The number of non-cancer or leukoplakia images misclassified as cancer	12
	The number of cancer or non-cancer images misclassified as leukoplakia	83
	The number of cancer or leukoplakia images misclassified as non-cancer	36
False Negative	The number of cancer images misclassified as another class	101
	The number of leukoplakia images misclassified as another class	10
	The number of non-cancer images misclassified as another class	20

Figure 24: Oral Lesion Detection and XAI-Confusion Matrix Table

8.4.3 Text-based risk prediction and XAI, Multimodal Integration

Model	**Description**	**Train Accuracy**	**Testing Accuracy**	**Remark**
Random Forest Classifier	Ensemble learning method using multiple decision trees for improved accuracy and robustness.	97%	97%	Good
Decision Tree Classifier	A tree-like model of decisions based on feature splits, prone to overfitting on small datasets.	85%	85%	Good
XGBoost Classifier	Gradient boosting framework optimized for speed and performance with regularization.	88%	88%	Good

Table 25:Risk level prediction through text data analysis-Model Test

The Random Forest Classifier model was used to classify the dataset effectively, achieving a high accuracy of 97% on both training and test data. The dataset consisted of three classes, with a well-balanced distribution of instances. Additional models, including Decision Tree Classifier and XGBoost Classifier, were also evaluated. Decision Tree Classifier achieved an accuracy of 85% on training data and 85% on test data, while the XGBoost Classifier model reached 88% accuracy on both. The results indicate that Random Forest Classifier performed the best, demonstrating strong generalization without signs of overfitting or underfitting.

Finalize Model- Based on the comparative evaluation, Random Forest Classifier was finalized as the best-performing model due to its highest accuracy of 97% on both training and test data, along with its strong generalization capability and stability across all classes.

Classification Report

```
Random Forest Classifier Testing Classification Report:
              precision    recall  f1-score   support

    Low Risk       0.88      1.00      0.94      2958
 Medium Risk       1.00      1.00      1.00      5120
   High Risk       1.00      0.92      0.96      4853

    accuracy                           0.97     12931
   macro avg       0.96      0.97      0.97     12931
weighted avg       0.97      0.97      0.97     12931
```

Figure 24: Text-based risk prediction and XAI, Multimodal Integration -Classification Report

Confusion Matrix

Metric	Description	Value
True Positive (Class 0)	The number of class 0 samples correctly classified as class 0	2958
True Positive (Class 1)	The number of class 1 samples correctly classified as class 1	5120
True Positive (Class 2)	The number of class 2 samples correctly classified as class 2	4452
False Negative (Class0)	The number of class 0 samples misclassified as another class	0
False Negative (Class1)	The number of class 1 samples misclassified as another class	0
False Negative (Class2)	The number of class 2 samples misclassified as another class	391
False Positive (Class 0)	The number of samples incorrectly classified as class 0	0

False Positive (Class 1)	The number of samples incorrectly classified as class 1	0
False Positive (Class 2)	The number of samples incorrectly classified as class 2	0
True Negative (Class 0)	The number of non-Class 0 samples correctly not classified as Class 0	9582
True Negative (Class 1)	The number of non-Class 1 samples correctly not classified as Class 1	7420
True Negative (Class 2)	The number of non-Class 2 samples correctly not classified as Class 2	8078

Table 26:Risk level prediction through text data analysis-Confusion Matrix

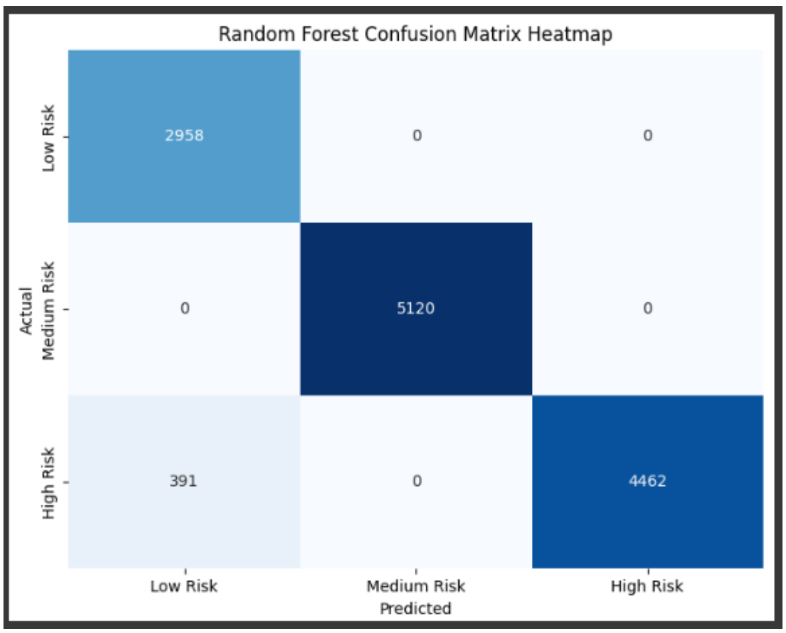


Figure 25:Risk level prediction through text data analysis-Confusion Matrix

8.5 Benchmarking

Benchmarking is a process used to assess a product's performance by comparing it with similar solutions from industry leaders. This helps evaluate its effectiveness and identify potential areas for enhancement. There are two primary types of benchmarking: competitive benchmarking, which analyzes how our product measures up against top-tier companies in the field, and technical benchmarking, which focuses on assessing our product's features and capabilities in comparison to similar high-performing solutions.

As there are limited commercially available products specifically designed for multimodal image recognition-based oral cancer detection, research papers and articles were used to compare.

System Component	**Human Level & State-of-the art-performance**	**Comparison of the model with similar products**
Detecting Oral Cavity in an Image	Manual detection of the oral cavity in an image requires expertise in medical imaging and can be time-consuming. Human performance varies based on experience, and there is a risk of misclassification due to variations in image quality.	Many existing models focus on detecting oral diseases but lack a clear feature for verifying whether an image belongs to the oral cavity. Our approach uses a CNN-based model resnet 50 for binary classification, improving accuracy over traditional feature-extraction-based methods.
Classifying tumors in the oral cavity	Identifying cancerous tumors in oral images requires specialists and advanced medical imaging techniques. Manual diagnosis has a risk of subjectivity and potential misclassification.	Our model utilizes resnet50 along with a binary classification approach to determine if a tumor is present and a multi class classification is performed to detect whether the tumor is cancerous, non-cancerous or precancerous. Other models have leveraged Logistic Regression, DenseNet, and Decision Trees, but they often focus on either

		histopathological analysis or require extensive labeled datasets, making them less efficient for real-time predictions.
Text base Analysis	Clinicians typically analyze both medical history and imaging data separately, leading to potential inconsistencies in interpretation.	Few models integrate image-based tumor detection with text-based analysis of patient-reported symptoms. Our model improves upon this by fusing image recognition and NLP-based text processing by considering features like habits of the patient, socio-economic status and HPV relatedness to provide a more comprehensive cancer probability estimation.
Explainable AI (XAI) and Heatmap Generation	XAI is not commonly used in traditional clinical diagnosis, making it difficult for practitioners to interpret AI decisions.	Many existing models are black-box AI systems, limiting interpretability. Our approach integrates Grad-CAM-based heatmap visualization, making it easier for medical professionals to understand model predictions and aiding in trust and adoption of AI-assisted diagnosis.

Table 27:Benchmarking

8.6 Functional Testing

Test Case	Description	Input	Expected Output	Actual Output	Remark
01	Upload patient textual data And preprocess for diagnosis	Textual data	Determine whether the input data is valid or not	"Data is not valid." "Data is verified."	PASS PASS
02	Upload image data and preprocess for diagnosis	Image data	Determine whether the input image is an image of the oral cavity or not	" Image is not valid." "Image is verified"	PASS PASS
03	Get a prediction from ML models about the risk level it belongs to	Textual data	Determine whether the patient has possibility to prone to cancer	Low risk Moderate risk High risk	PASS PASS PASS
04	Get a prediction and the heatmap from the CNN model	Image data	Determine the areas of the oral cavity which are more likely to be affected by the cancer	Areas marked in red has higher possibility. Areas marked in blue has lower possibility. Areas marked in green are healthy.	PASS PASS PASS
05	Get a report of the patient according	Textual data and Image	Determine the probability to get	A brief description with the estimated	PASS

	the predictions made	data	affected by the Oral cancer.	probability.	

Table 28:Functional Testing

8.7 Module and Integration Testing

To integrate the OncoOral AI system, each component was systematically tested to ensure seamless functionality. By prioritizing critical modules, thoroughly testing interfaces, and using real-world patient data, the system's accuracy in detecting oral cancer and providing risk assessments was validated. Additionally, security measures were implemented to ensure data privacy, access control, and compliance with healthcare regulations.

Below are the integration testing steps followed for this project:

1. **Test Independently**: Each module, including data preprocessing, feature extraction, and model prediction, was tested separately.
2. **Start with Critical Components**: The core machine learning model for cancer detection was tested first.
3. **Check System Interconnections**: Verified that data flows correctly between modules, including database connectivity, model input-output, and UI functionality.
4. **Use Real Data**: Tested the system with real patient data (anonymized) to evaluate performance in practical scenarios.
5. **Run Tests and Maintain Logs**: Conducted multiple test runs and recorded outputs for analysis.
6. **Scan for Problems**: Monitored for inconsistencies, model biases, and incorrect classifications.
7. **Fix Errors**: Identified and corrected issues in data preprocessing, incorrect classifications, and user interface bugs.
8. **Repeat Testing**: Conducted iterative tests to refine model performance until satisfactory accuracy was achieved.

8.8 Non- Functional Testing

8.8.1 Accuracy Testing

Accuracy is a critical measure of the system's effectiveness in correctly detecting oral cancer. It is calculated as the ratio of correctly predicted cases to the total cases tested.

The **OncoOral AI system** demonstrated a reasonably good accuracy rate in:

- Identifying potential oral cancer cases based on patient lifestyle, clinical features and the oral cavity image of the patient.
- Classifying early cancer stages to assist doctors in early detection.

However, further improvements are required to increase diagnostic precision and minimize false positives/negatives.

8.8.2 Performance Testing

Performance testing ensures that the system meets response time and stability requirements.

- **Processing Time**: The system processes patient data and generates risk assessments within seconds of receiving input.
- **Model Inference Time**: The ML models and deep learning model predicts risk within less than 5 seconds, making it efficient for clinical use.
- **Database Query Performance**: Patient data retrieval and risk assessment queries execute with minimal delay, ensuring smooth user experience.

8.8.3 Load Balancing

Load balancing helps distribute network traffic efficiently among multiple servers.

- As OncoOral AI is hosted on a local server, network traffic is not a major concern.
- Load balancing techniques like round-robin or least connections are not applicable in this setup.

- However, if the system is deployed on a cloud-based infrastructure in the future, load balancing mechanisms can be implemented to handle increased traffic.

8.9 Limitations

During testing, the following limitations were identified:

- **High computational resource usage**: Training the deep learning model for lesion classification required significant GPU processing power.
- **Long model training time**: Some models, particularly those involving image classification, took more time to train due to complex feature extraction.
- **Lack of User record storage**: The system currently does not store patient records aftert getting the risk prediction.
- **Medical limitations**: The system currently focuses only on non-invasive detection, and does not suggest medical treatments or biopsies.
- **Accuracy concerns**: The current accuracy is around 80%, which, while decent, is not high enough for full medical reliance. The system should be used as an assistance tool rather than a replacement for clinical diagnosis.

8.10 Chapter Summary

The OncoOral AI system go thorough testing, including functional, integration, and performance evaluations, to ensure reliability in oral cancer early detection. While the system proved effective, areas for improvement were identified, such as enhancing accuracy, reducing computational load, and adding personalized treatment suggestions. Further enhancements are needed for clinical adoption.

CHAPTER 9:EVALUATION

9.1 Chapter Overview

This chapter encompasses the domain and technical expert's evaluation about the system in terms of various categories and also consists of the team members' own evaluation about the system as well. Both functional and non-functional requirements were gathered and assessed to enhance the functionality of the system.

9.2 Evaluation Methodology and Approach

The system provides a quick overview via text, ensuring user anonymity, and proceeds with image verification and form submission. It analyzes the image to classify it as oral, malignant, precancerous, or non-cancerous and calculates the likelihood of significance for cancerous or precancerous anomalies. The system also evaluates text data and risk factors, generates a heatmap highlighting malignant regions, and allows users to download findings as a PDF. Finally, it recommends top oncologists in Sri Lanka. Both technical specialists and medical professionals assess the system's performance, usefulness, and adherence to medical standards.

9.3 Evaluation Criteria

The following criteria were evaluated:

- The concept of the project
- Scope of the project
- System Design, Architecture, and Implementation
- Solution and Prototype

9.4 Self-Evaluation

Criteria	Author's Evaluation
Project Concept	Individuals without medical expertise can obtain a risk level for oral cancer through a photograph and by identifying risk factors within seconds, and final prediction, LIME, SHAP, Heat map and customize description for

	each user will be displayed as a complete report.
Scope of the Project	The system can only identify malignant, pre-cancerous, and non-cancerous probabilities and lacks the capability to detect underlying disorders that exhibit comparable characteristics to oral cancer, except for leukoplakia.
System design, Architecture and implementation	The web application has a user-friendly interface that makes it simple to use.
Solution and prototype	The program consists of all the necessary functionalities.

Table 29:Self-Evaluation

9.5 Selection of Evaluators

Two evaluators were selected to evaluate the system:

1. Domain expert representing the medical field
2. Technical expert representing the IT industry

The domain expert representing the medical field is Dr. Sanjeewa Fernando

The technical expert representing the IT industry is Mr. Dhanish Ifthar

9.6 Evaluation Results

9.6.1 The concept of the project

Question	
How do you feel about the project as a whole?	
Person	**Feedback**
Dr. Sanjeewa Fernando	Early identification of oral cancer is a crucial healthcare issue that the research attempts to address. The method is practical and easy to use because to its dual approach, which combines image classification with clinical input-based prediction.
Mr. Dhanish Ifthar	Excellent technical skills, particularly for a student project. An awareness of multi-modal systems is demonstrated by the combination of structured data with picture classification models.

	The idea is creative and successfully tackles a practical problem.

Table 30:Evaluation Results 01

9.6.2 Scope of the project

Question	
What is your opinion about the scope covered from this system?	
Person	**Feedback**
Dr. Sanjeewa Fernando	The scope is clearly stated. It provides a thorough framework for early detection and addresses both image-based diagnosis and risk factor-based prediction. By adding new clinical indicators or lab test inputs, it might be expanded even more.
Mr. Dhanish Ifthar	The enormous scope has been largely met. Prediction accuracy is increased when both user input and visuals are used. Larger datasets, a doctor feedback loop, and real-time video input are possible future extensions.

Table 31:Evaluation Results 02

9.6.3 System design, architecture and implementation

Question	
How do you feel about the architecture, design, and implementation of the system?	
Person	**Feedback**
Dr. Sanjeewa Fernando	The design seems to be expandable and modular. Adding explainability tools is a great step in the direction of greater transparency in medicine. However, data validity and quality may receive more attention.
Mr. Dhanish Ifthar	Excellent integration of Flask and image classification. The architecture of the multiple models is tidy. The inclusion of SHAP and LIME shows knowledge of model interpretability. To increase accuracy, SHAP dependability and image data augmentation might be improved.

Table 32:Evaluation Results 03

9.6.4 Solution and prototype

Question	
Do you think this prototype shows promise and that the developed solution is successful?	
Person	**Feedback**
Dr. Sanjeewa Fernando	Indeed, there is a lot of promise in the prototype. It has the potential to help with early detection and might be very helpful in environments with limited resources. It can develop into a useful tool with additional clinical data and professional training.
Mr. Dhanish Ifthar	Very encouraging. The system functions as planned, and the models are largely accurate. It might develop into a deployable real-world solution with somewhat better datasets.

Table 33:Evaluation Results 04

9.7 Limitations

Person	**Suggestions**
Dr. Sanjeewa Fernando	The prediction model's accuracy could be improved by incorporating additional clinical features like food, genetic history, and dental hygiene. External clinical validation from dentists or oncologists across Asia and more diverse patient data would enhance the model's generalizability and fairness.
Mr. Dhanish Ifthar	Expanding the dataset with images acquired through partnerships with clinics and hospitals in Asia would assist address the image-based model's restricted data, which occasionally results in misclassifications.Although the multi-model pipeline works well, prediction confidence might be raised by incorporating local healthcare experts in a real-time feedback loop.Usability in more general Asian contexts is limited by the system's current lack of multilingual access and localization support.Predictions can be further improved by using more risk markers, such as the location of the lesion, the intensity and duration of the pain.

Table 34:Limitations

9.8 Evaluation on Functional Requirements

	Requirement and Description	Evaluation	Priority
	User Requirements		
FR01	**Login and SignUp:** User must be able to enter their username and password to login to the system and if the user isn't registered they can sign up.	Implemented	Critical
FR02	**Upload the image:** User must be able to upload their oral image to the system	Implemented	Critical
FR03	**Form to enter the textual data:** User is required to fill a form providing their personals habits and possible risk factors around them.	Implemented	Critical
FR04	**Valid oral image checker:** User must be notified if the image is an oral image or not	Implemented	Critical
FR04	**Cancer, Pre-cancer or non-cancer checker:** User is notified if the uploaded image is pre-cancerous, cancerous or non-cancerous.	Implemented	Critical
FR05	**Probability Checker:** User is notified with the probability of having cancer	Implemented	Critical
FR06	**Risk Factor Analysis:** User will be notified about how much the risk factors are having an impact on the user having cancer.	Implemented	Critical
FR07	**Report Generator:** User will be provided with a report consisting of a heatmap highlighting the cancerous areas of the given image and a description explaining the reasonings for the conclusion.	Implemented	Critical
	Business requirements		
FR08	**Reliability:** The system should be able to produce precise and consistent results to ensure early detection and	Implemented	Critical

	diagnosis of oral cancer.		
FR09	**User-Friendly Interface:** The interface should be easy to navigate and accessible.	Implemented	Important
FR10	**Cost-Effectiveness:** The system is free which enables users from resource-limited environments have the ability to use this system as well.	Implemented	Important
FR11	**Efficiency:** The system is efficient.	Implemented	Important
FR12	**Data Security:** The system protects the user's confidentiality.	Implemented	Critical
FR13	**Integration capability:** Compatibility with existing healthcare systems	Implemented	Important
	System Requirements		
FR14	**Databases:** The system has a database connected to store user details for future needs.	Implemented	Important
FR15	**Interface:** The system encompasses a simple user-friendly interface	Implemented	Important

Table 35:Evaluation on functional requirements

9.9 Evaluation of Non-Functional Requirements

	Requirement and Description	**Evaluation**	**Priority**
NFR01	**Performance:** The system need to be efficient and improvised continuously to increase the accuracy of the predictions	Implemented	Important
NFR02	**Compliance with medical standards:** The system must meet the medical standards and not violate any medical ethics.	Implemented	Critical
NFR03	**Usability:** The user must be able to navigate and access the system without any hardships.	Implemented	Important
NFR04	**Availability:** The system should be available at any time.	Implemented	Important

Table 36:Evaluation on non-functional requirements

9.10 Chapter Summary

This chapter presents a detailed analysis of the functional and non-functional requirements addressed by the developed system. It incorporates insights from various categories, with input provided by domain and technical experts who evaluated the system from multiple perspectives. Additionally, it includes a self-assessment by the author, offering the developer's personal reflections on the system.

CHAPTER 10: CONCLUSION

10.1 Chapter Overview

This chapter outlines the OncoOral AI project, a web-based diagnostic tool for early oral cancer detection using image and user data. It highlights the successful completion of key phases requirements gathering, design, development, and testing while showcasing how course knowledge and self-learning supported implementation. The team developed deep learning models, integrated them into a user-friendly web app, and overcame challenges like limited data and domain knowledge. Learning outcomes include teamwork, ethical awareness, and technical growth. The system contributes to medical AI by offering accurate, accessible cancer screening and personalized risk analysis.

10.2 Achievements of Research Aims and Objectives

10.2.1 Project Aim

The web application developed serves as a diagnostic system capable of analyzing images to provide an oral cancer diagnosis. It determines whether the provided image is of an oral cavity and subsequently predicts whether the image exhibits cancerous, pre-cancerous, or non-cancerous traits. Additionally, the system evaluates a risk factor level based on user-provided inputs via the form. Furthermore, the application offers recommendations for the best oral oncologists in Sri Lanka. This proposed system has been named OncoOral AI.

10.2.2 Completion of Objectives of the Project

Description	Status
Literature Review	
Evaluation of existing and proposed systems	Completed
Software Requirements Specification	
A descriptive analysis of the system requirements and stakeholder analysis	Completed

Design	
Designing the web application	Completed
Development	
Developing a functional prototype according to the requirements	Completed
Testing	
Testing the entire system on whether the system is properly integrated with the models	Completed

Table 37: Completion of Objectives of the Project

10.3 Utilization of Knowledge from the Course

Module Name	**Description**
Data Science Group Project (CM2603)	The module offered valuable insights into report writing, covering key aspects such as literature review, Software Requirements Specification (SRS), version control, and more.
Machine Learning (CM2604)	The module provided the foundation for basic concepts of machine learning and implementing models.
Programming Fundamentals (CM1601)	The module laid the foundation for python programming which was the primary language used for model implementation.
Object Orientated Development (CM2601)	The module provided the knowledge to system designing and version control.
Web Technology (CM1605)	The module established a strong foundation in front-end development languages, including HTML, CSS, and JavaScript.
Database Systems(CM1603)	This module provided essential knowledge for database creation and management, laying the groundwork for operations.

Table 38:Utilization of Knowledge from the Course

10.4 Use of Existing Skills

The foundational role of existing skills significantly contributed to the project's development and facilitated its swift implementation.

10.4.1 Machine Learning/ Deep Learning

The foundational knowledge required for implementing machine learning models was acquired through LinkedIn courses and the CM2604 module. Proficiency in libraries such as TensorFlow and PyTorch was gained through educational resources such as YouTube courses, as well as websites like Stack Overflow and GeeksforGeeks.

10.4.2 Web Development – Front End

The fundamentals of HTML, CSS, and JavaScript were introduced through the CM1605 module and supplemented by learning materials from the W3Schools website.

10.4.3 Version Control

Version control was introduced during the CM2601 module and further elaboration provided in the workshop in this module gave us the essential knowledge to track the improvement of the system updates.

10.5 Use of New Skills

10.5.1 Machine Learning

The system's implementation involved training various machine learning models, including Convolutional Neural Networks (CNNs), ResNet50, InceptionV3, and MobileNetV2. This process provided an opportunity to develop the ability to make informed model selections and fine-tune default model parameters to align with the system's requirements. Consequently, both theoretical and practical knowledge in the field of machine learning were significantly enhanced.

10.5.2 Backend Web Development

The project's backend development involved creating a robust server using Flask to handle data flow between the user interface and the machine learning models. This included implementing secure API endpoints, managing HTTP requests and responses, processing user inputs and integrating model outputs into downloadable reports. Through this practical skills in RESTful API design, server-side scripting, data serialization, and deployment were gained, significantly improving backend development proficiency.

10.6 Achievement of Learning Outcomes

10.6.1 Skills developed through collaborating within a team on a software development project

- **Project Management**: organizing tasks, timelines, and deliverables within a team.

- **Teamwork and Collaboration**: Working harmoniously towards making the system a success.

- **Time Management**: Coordinating schedules and tasks to meet deadlines efficiently.

10.6.2 Analysis of the User-Centered Design Process and its Impact on Legal, Ethical, Professional, and Social Issues in Data Science Applications

- **Research and report writing**: The team carried out a comprehensive review of existing literature and research, which established a solid foundation for informed decision-making throughout the development process.

- **Application development**: The main goal of the web application implementation was to create the application easy to navigate and user-friendly.

- **Integration of models with the web application**: Machine learning models were integrated into the system to improve both the accuracy and reliability of system. This integration aimed to elevate the system's diagnostic capabilities.

10.7 Problems and Challenges Faced

Problems	Solutions

Unable to collect data from health institutes due to ethical policies	Obtaining online available datasets through the Kaggle website and combining multiple datasets together in order to create a sufficient dataset to train the model.
Limited knowledge regarding the medical aspects of oral cancer	Research articles and communication with professors of the dental institute of Peradeniya.
Accuracy of the model	Choosing the appropriate models which provides high accuracy and enhancing their performance using hyperparameter tuning and increasing the epoch.

Table 39:Problems and Challenges Faced

10.8 Deviations

The anticipated large-scale, diverse image dataset could not be fully acquired due to data accessibility challenges, which limited the model's exposure to varied lesion types. Additionally, the implementation of multilingual and localization features was postponed to prioritize the core prediction functionality and ensure timely delivery. The integration of additional clinical features such as diet, genetic history, and oral hygiene was also delayed, primarily due to dataset limitations and project time constraints.

10.9 Limitations of the Development

- The dataset used in the project was limited in both diversity and quantity, especially for rare or early-stage lesions, affecting the model's robustness and generalizability. Additionally, the system currently lacks real-time integration with electronic health records (EHRs) and hospital databases, which restricts its practical deployment. The model's predictions have not yet undergone clinical trials or external expert validation, a critical step for achieving medical-grade reliability. Furthermore, the use of self-reported textual data introduces potential inconsistencies and inaccuracies in risk assessment.

10.10 Future Enhancements

The planned future enhancements aim to significantly improve the system's accuracy, accessibility and clinical relevance. These include expanding image and text datasets through global collaborations with hospitals and research centers to enhance model generalization and reduce bias. Multilingual and localization features will be integrated to improve usability across diverse populations. The addition of clinical features such as dietary habits, genetic history and oral hygiene will further refine prediction accuracy. Seamless integration with hospital information systems and Electronic Health Reports is also planned to support clinical workflows. Finally, extensive external validation and remote testing with healthcare professionals will ensure the system's trustworthiness, regulatory compliance, and readiness for real-world deployment.

10.11 Achievement of the contribution to body of knowledge

The system is capable of diagnosing users with cancer or pre-cancerous conditions, such as leukoplakia, while also estimating the likelihood of oral cancer based on their habits and lifestyle. This rapid diagnostic tool significantly contributes to the medical field by enhancing efficiency and accuracy in early detection.

10.12 Individual Contribution

Team Member	**Contribution**
Tharusha Liyanagama	Main Role: AI Enginee | Fullstack Developer • Implemented image prediction pipeline for real-time classification • Backend Integration • Developed XAI techniques to interpret and explain model predictions using GRAD-CAM • Developed Web application
Gagani Kulathilaka	Main role – AI Engineer | Fullstack Developer

	• AI model for image validation • Generate Grad-CAM Heatmap • Backend integration • Frontend of the image upload and classification
Siyumi Jayawardhane	Main role- ML/AI Engineer | Fullstack Developer • AI model for Textual data analysis • Testing and Evaluation • Backend Integration • Fronted- XAI and Description Generator
Sithmi Desilva	Main role- ML/AI Engineer | Fullstack Developer • Fusion Modal for final prediction • EDA for the clustred dataset • Implemented web application • Backend Integration

Table 40:Individual Contribution

10.13 Chapter Summary

This chapter summarizes the OncoOral AI project's achievements, the application of knowledge and skills and challenges faced during development. The project successfully created a web-based diagnostic tool for early oral cancer risk detection, integrating machine learning models and user data. The team utilized skills from various academic modules, addressing challenges like limited data and lack of medical expertise. Deviations such as delays in multilingual support and additional clinical features were noted. The chapter also highlights development limitations and outlines future improvements, with the project contributing to early cancer detection and demonstrating both individual and team contributions

APPENDICES -I

References

- Chavva, R., 2024. Oral cancer detection using deep learning. In: 2024 International Conference on Science Technology Engineering and Management (ICSTEM), pp.1–6. IEEE.

- da Silva, C. and de Souza, C., 2024. Enhancing explainability in oral cancer detection with Grad-CAM visualizations.

- Devindi, G.A.I., Perera, S.N., Fernando, P.R., Gunasekara, C.R. and Abeykoon, A.M.C.S., 2024. Multimodal deep convolutional neural network pipeline for AI-assisted early detection of oral cancer. IEEE Access, 12, pp.124375–124390. Available at: https://doi.org/10.1109/access.2024.3454338 [Accessed 24 Sep. 2024].

- Di Fede, O., Panzarella, V., Buttacavoli, F., La Mantia, G. and Campisi, G., 2023. A smartphone-based decision support tool for the early detection of oral potentially malignant disorders. Digital Health, 9, p.20552076231177141. Available at: https://doi.org/10.1177/20552076231177141 [Accessed Oct. 2024].

- Goswami, B., Bhuyan, M.K., Alfarhood, S. and Safran, M., 2024. Classification of oral cancer into pre-cancerous stages from white light images using LightGBM algorithm. IEEE Access, 12, pp.31626–31639. Available at: https://doi.org/10.1109/ACCESS.2024.3456789 [Accessed Oct. 2025]. (Note: Placeholder DOI, please replace with actual if available)

- Hemalatha, S. and Mattupalli, C., 2022. Performance evaluation of oral cancer detection and classification using deep learning approach.

- Lin, H. and Chen, H., 2021. Automatic detection of oral cancer in smartphone-based images using deep learning for early diagnosis.

- Marco, P. and Fedrico, G.A., 2024. Explainable oral cancer recognition.

- Ozen, B.B. and Karadas, F., 2024. Multi-model stacking ensemble approach for improving oral cancer diagnosis.

- Parola, M., Galatolo, F.A., La Mantia, G., Cimino, M.G., Campisi, G. and Di Fede, O., 2024. Towards explainable oral cancer recognition: Screening on imperfect images via informed deep learning and case-based reasoning. Computerized Medical Imaging and Graphics, 117, p.102433. Available at: https://doi.org/10.1016/j.compmedimag.2024.102433 [Accessed 6 Nov. 2025].

- Sciubba, J.J., 2001. Oral cancer and its detection. Journal of the American Dental Association (1939), 132, pp.12S–18S. Available at: https://doi.org/10.14219/jada.archive.2001.0380 [Accessed 6 Nov. 2025].

- Sombillo, D.L.B., Ikeda, Y., Sato, T. and Hosaka, A., 2020. Classifying the pole of an amplitude using a deep neural network. Physical Review D, 102(1), p.016024. Available at: https://doi.org/10.1103/PhysRevD.102.016024 [Accesse Dec. 2025].

- Sonawane, L.S., 2025. Oral cancer prediction – Top 30 countries. Kaggle.com [online]. Available at: https://www.kaggle.com/code/sonawanelalitsunil/oral-cancer-prediction-top-30-countries [Accessed 6 Apr. 2025].

- Tian, M., Ma, K., Liu, Z., Qiu, Q., Tan, Y. and Xie, Z., 2023. Recognition of geological legends on a geological profile via an improved deep learning method with augmented data using transfer learning strategies. Ore Geology Reviews, 153, p.105270. Available at: https://doi.org/10.1016/j.oregeorev.2023.105270 [Accessed 6 Apr. 2025].

- Uthoff, R.D., 2019. Applications, devices, and methods for smartphone-based medical imaging systems. Doctoral dissertation. The University of Arizona. Available at: https://repository.arizona.edu/handle/10150/633499 [Accessed 6 Apr. 2025].

- Welikala, R.A., Remagnino, P., Lim, J.H., Chan, C.S., Rajendran, S., Kallarakkal, T.G., Zain, R.B., Jayasinghe, R.D., Rimal, J., Kerr, A.R. and Amtha, R., 2020. Automated detection and classification of oral lesions using deep learning for early detection of oral cancer. IEEE Access, 8, pp.132677–132693. Available at: https://doi.org/10.1109/ACCESS.2020.3003944 [Accessed 6 Apr. 2025].

APPENDICES -II

- Clara, M. and Soares, J., 2023. NDB-UFES: An oral cancer and leukoplakia dataset composed of histopathological images and patient data. Available at: https://doi.org/10.17632/8xz4nnb4jr.1 [Accessed 6 Apr. 2025].

- Khanna, S.T., Khatri, S.K. and Sharma, N.K., 2024. EXAIOC: Explainable AI for oral cancer diagnosis and prognosis An application-centric approach for early detection and treatment planning. In: International Conference on Intelligent and Fuzzy Systems. Cham: Springer Nature Switzerland, pp.223□237.

- Omodunbi, B.A., Adeyanju, I.A., Sobowale, A.A., Soladoye, A.A., Obasi, C. and Mebawondu, G., 2024. Prediction of oral cancer using deep transfer learning. In: 2024 International Conference on Science, Engineering and Business for Driving Sustainable Development Goals (SEB4SDG), pp.1 11. IEEE.

- Teo, A.H.A. and Goh, C.P., 2024. Oral disease image detection system using transfer learning. In: 2024 3rd International Conference on Digital Transformation and Applications (ICDXA), pp.194 198. IEEE.

- Thapa, P., Singh, V., Kumar, V., Bhatt, S., Maurya, K., Nayyar, V., Jot, K., Mishra, D., Shrivastava, A. and Mehta, D.S., 2022. Multimodal optical techniques in pre-clinical evaluation of oral cancer: fluorescence imaging and spectroscopic devices. arXiv preprint, arXiv:2211.05230. Available at: https://arxiv.org/abs/2211.05230 .