Self-Supervised Vision Transformers for Dental Caries Detection from Radiographs

Bill Chakairu

151339

ICS 4D

Supervisor Name

Kevin Ochieng' Omondi

An Informatics and Computer Science Project Documentation Submitted to the School of Computing and Engineering Sciences in Partial Fulfilment of the Requirements for the Award of a Degree in Bachelor of Science in Informatics and Computer Science

School of Computing and Engineering Sciences

Strathmore University

Nairobi, Kenya

October 2025

Declaration and Approval

I declare that this work has not been previously submitted and approved for the award of a degree by this or any other University. To the best of my knowledge and belief, this research documentation contains no material previously published or written by another person except where due reference is made in the research documentation itself.

Student Name: Bill Chakairu

Admission Number: 151339

Student Signature: ___________________________ Date:__________________________

The Documentation of **Bill Chakairu** has been reviewed and approved by **Mr. Kevin Ochieng' Omondi**

Supervisor Signature: ___________________________ Date:__________________________

Acknowledgement

I would like to express my gratitude to God for granting me the opportunity, strength, and clarity to undertake this project. I am especially thankful to my supervisor, Mr. Kevin Ochieng' Omondi, for his invaluable guidance, and consistent support throughout the course of this project.

I also wish to extend my gratitude to Strathmore University for providing a conducive academic environment and resources necessary to work on this project. Finally, I am thankful to my family and friends for the support throughout the duration of the project.

Abstract

Early detection of dental caries helps prevent tooth decay and reduce the burden of restorative treatments. However, identifying small radiographic signs in bitewing X-rays is a diagnostic challenge. The adoption of Artificial Intelligence models such as CNN to help in diagnosis of dental caries has really helped bridge this gap. However, CNN models use many annotated datasets for training. Such datasets may not only be difficult to acquire but also expensive to annotate. This documentation outlines a self-supervised learning framework using vision transformers to detect dental caries from bitewing radiographs. This will minimize reliance on extensive manual annotations which are required for CNN-based models.

The research adopts Masked Autoencoders as the backbone for pretraining the vision transformer model on unlabelled radiographic images. A small, labelled subset was used to further tune the model for binary classification (presence or absence of caries). The experimental results show that this approach competes favourably alongside existing CNN models on standard evaluation metrics such as accuracy, precision, recall, and F1-score. Furthermore, the system offers clinical advice to the dentists based on the outcome of the prediction. This makes it a valuable tool for healthcare providers seeking ai-assisted diagnostic tools.

Keywords: Self-supervised Learning, Vision Transformers, Dental Caries, Masked Autoencoders, Bitewing Radiographs.

Table of Contents

List of Figures

List of Tables

List of Abbreviations

AUC-ROC – Area Under the Receiver Operating Characteristic curve

CBCT – Cone Beam Computed Tomography

CNN – Convolutional Neural Networks

CRUD – Create, Read, Update, Delete

FOTI – Fibre Optic Transillumination

JPEG – Joint Photographic Experts Group

MAE – Masked Autoencoder

NILT – Near-Infrared Light Transillumination

ROC – Receiver Operating Characteristic

SimCLR – Simple Framework for Contrastive Learning of Visual Representations

SimMIM – Sparse inpainting for Masked Image Modeling

SSL - Self-Supervised Learning

UI – User Interface

UM-MAE – Uniform Masking – Masked Autoencoder

VGG16 – Visual Geometry Group 16

ViTs – Vision Transformers

YOLO – You Only Look Once

Chapter 1: Introduction

1.1 Background Information

The World Health Organization estimates that around 2.5 billion individuals suffer from untreated dental caries in permanent teeth (*WHO*, 2022). Undetected caries can cause pain, infection, tooth loss, and higher healthcare costs. Early diagnosis allows for conservative intervention that can prevent extensive damage and maintain oral health (Mahesh Batra & Reche, 2023).

Dental radiographs, particularly bitewing images, are commonly used for caries detection because they can reveal hidden interproximal and occlusal lesions not visible during clinical examination. However, interpreting these images is highly subjective. Studies have shown significant inter- and intra-observer variability among dental professionals, especially in identifying early-stage interproximal and occlusal lesions (Bayrakdar et al., 2022). As such, diagnostic accuracy heavily relies on the experience of the clinician.

AI methods have emerged as a possible solution to this. Convolutional Neural Networks have dominated image-based medical analysis. In dentistry, CNNs such as VGG-16 and U-Net have demonstrated success in caries detection and segmentation tasks, achieving high sensitivity and specificity in bitewing radiographs (Bayrakdar et al., 2022). However, they depend on large quantities of annotated training data, which is difficult to obtain in clinical contexts due to privacy concerns and the expertise required for annotation.

To address these shortcomings, researchers are exploring vision transformers, which were originally developed for natural language processing and later adapted for computer vision (Dosovitskiy et al., 2021). Unlike CNNs, ViTs split images into sections and use a self-attention mechanism to model long-range dependencies throughout the image. This global perspective is particularly beneficial for analysing complex structures in medical images. Furthermore, self-supervised learning enables models to learn useful information from unlabelled data. This is especially important in medicine since labelled datasets are often scarce. SSL methods such as Masked Autoencoders have been shown to achieve strong performance with minimal supervision (He et al., 2021).

However, the specific application of self-supervised Vision Transformers to dental caries detection from bitewing radiographs remains underexplored. While ViTs and SSL have each shown potential individually, their combination in the dental diagnostic context has yet to be fully investigated. This gap is particularly significant considering the demand for generalizable, interpretable, and data-efficient AI tools in dentistry.

Addressing this research gap could lead to new diagnostic tools that are more robust to variability in image quality and more accessible in data-constrained environments. It may also help understand how modern AI architectures can support clinical decision-making in oral healthcare.

1.2 Problem Statement

Most existing dental caries detection AI models rely on Convolutional Neural Networks, which, despite their strong performance, require large, annotated datasets and often struggle to generalize across different imaging settings and patient populations (Bayrakdar et al., 2022). Furthermore, their limited receptive field prevents them from fully capturing the global context of radiographic structures, and their decision-making processes remain opaque, reducing clinical trust and adoption.

Vision transformers offer a compelling architectural alternative, as they are designed to model long-range dependencies across an image. When combined with self-supervised learning, they present an opportunity to address the issue of data scarcity in medical AI. However, their application in dental radiography, particularly for bitewing caries detection, has not been systematically investigated (Almalki & Latecki, 2022).

This research addresses this gap by developing and evaluating a self-supervised Vision Transformer model for dental caries detection from bitewing radiographs.

1.3 Objectives

This section outlines the objectives this research aims to achieve as well as the questions it aims to answer.

1.3.1 General Objective

To develop and evaluate a self-supervised Vision Transformer-based model for accurate and data-efficient detection of dental caries from bitewing radiographs.

1.3.2 Specific Objectives

i. To evaluate current methodologies for automated dental caries detection from radiographic images.
ii. To analyse the challenges facing existing dental caries detection systems.

iii. To review and analyse deep learning and self-supervised learning techniques applied to dental radiographs in related works.
iv. To evaluate the limitations of deep learning and self-supervised learning dental caries detection systems in related works.
v. To design and develop a self-supervised vision transformer model for dental caries detection in bitewing radiographic images.
vi. To test and evaluate the model's performance using standard evaluation metrics.

1.4 Research Questions

i. What are the existing methods for automated dental caries detection in radiographic images?
ii. What are the limitations of current automated dental caries detection systems?
iii. How have deep learning and self-supervised learning been applied to dental radiographs for caries detection?
iv. What are the limitations of current deep learning and self-supervised learning techniques for dental caries detection?
v. How was the self-supervised vision transformer model designed and implemented for dental caries detection from bitewing radiographs?
vi. How does the developed model perform based on standard evaluation metrics?

1.5 Justification

The use of AI, especially deep learning in medical imagery has grown over the years. In dentistry, CNNs such as VGG-16, and U-Net have been applied to detect and segment dental caries in bitewing radiographs (Bayrakdar et al., 2022). More recently, vision transformers have been adapted to medical domains due to their ability to capture long-range dependencies in images (Dosovitskiy et al., 2021). Self-supervised learning methods such as masked autoencoders (He et al., 2021) have been shown to produce robust feature representations in domains with data scarcity, such as dental radiology.

Despite these advances, current research on vision transformers in dentistry remains limited, especially for early-stage dental caries detection. CNNs still dominate the field, but they are constrained by their need for large, annotated datasets and limited contextual understanding. Moreover, few studies have explored the combined power of vision transformers and self-supervised learning in dental diagnostics. The integration of these techniques has not yet been systematically studied for caries detection in bitewing radiographs.

The lack of generalizable and data-efficient models leads to inconsistent diagnostic outcomes, especially in resource-limited settings where expert annotation is scarce. Clinicians may miss early lesions, resulting in preventable disease progression, higher treatment costs, and compromised oral health outcomes. Additionally, the dominance of black-box CNN models impedes trust and interpretability in clinical practice.

This study documents the development and evaluation of a self-supervised Vision Transformer architecture tailored to dental caries detection in bitewing radiographs. The use of Vision Transformers reduces reliance on large, annotated datasets while maintaining high diagnostic performance. This work assists dental clinicians and researchers by providing a more interpretable, generalizable, and data-efficient tool for dental caries detection. It also contributes to the broader research on self-supervised Vision Transformers in medical imaging.

1.6 Scope

This study focuses on developing a self-supervised Vision Transformer model for detecting dental caries from bitewing radiographs. The research explored the use of Masked Autoencoders to enable effective feature learning without relying on large amounts of labelled data. The model was trained and evaluated on two publicly available datasets, and was developed to perform binary classification, that is, identifying the presence or absence of caries based solely on radiographic images. Additionally, the system incorporates attention visualization to provide clinical interpretability by highlighting regions suspected to have carious lesions in the radiographs.

1.7 Limitations and Delimitations

The study is limited by the quality and diversity of the datasets used, which may not fully capture variations in patient demographics, imaging devices, or clinical conditions. Furthermore, the lack of access to real-time clinical settings limits the evaluation to standard performance metrics rather than actual diagnostic workflows. The model performs binary classification and does not provide lesion segmentation, precise localization, or dental radiographic severity grading such as (E1, E2, D1, D2, and D3 stages).

The research is delimited to bitewing radiographs and produces four primary outputs: binary classification (caries present/absent), confidence scores, attention heatmaps and clinical recommendations. Other radiographic types and dental conditions are outside the scope. The study also focuses on Vision Transformer architectures and self-supervised learning methods, excluding traditional CNN-based and fully supervised models from development.

Chapter 2: Literature Review

2.1 Introduction

This chapter analyses existing literature relevant to the development of self-supervised vision transformers for dental caries detection using radiographic images. Section 2.2 discusses the current diagnostic methods and their limitations. Section 2.3 explores related works, while section 2.4 identifies research gaps. Finally, section 2.5 illustrates and explains the conceptual framework used in this research.

2.2 The Current Methods in Dental Caries Detection

Dental caries is characterized by the localized destruction of tooth enamel, dentin, and cementum due to acid-producing bacterial activity. It arises from complex interactions among biofilm, fermentable carbohydrates, tooth surfaces, and time (Warreth, 2023). The detection of carious lesions at an early stage is necessary for initiating non-invasive treatment. Traditional diagnostic methods, while commonly employed, often suffer from subjectivity and low sensitivity for developing lesions, and depend heavily on practitioner experience (Ghodasra & Brizuela, 2025).

2.2.1 Visual and Tactile Examination

This is the most common method used to diagnose dental caries. It involves inspecting the tooth for surface changes such as white spot lesions, discoloration and surface texture (Al Saffan, 2023). Surfaces that are rough and have chalk-like edges are likely to have caries. On the other hand, surfaces that are smooth and have shiny edges are likely to be caries-free (Sadikoglu et al., 2020). This difference is illustrated in Figure 2.1, which differentiates active and inactive carious lesions.

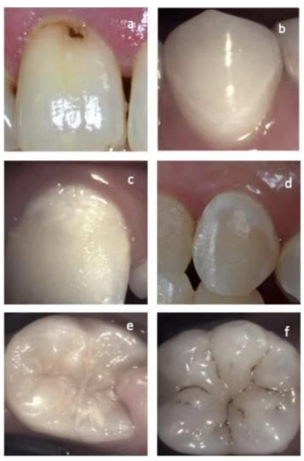


Figure 2.1: Difference between an active and an inactive lesion (Abdelaziz, 2023)

Despite its widespread use, visual and tactile examination has significant limitations. The method is inherently subjective, heavily reliant on practitioner expertise, and varies from dentist to dentist. Its diagnostic accuracy diminishes in early proximal lesions or cases where caries develop beneath intact enamel surfaces (Al Saffan, 2023). Moreover, the use of sharp probes has been criticized for potentially causing enamel microfractures or converting non-cavitated lesions into cavitated ones.

2.2.2 Radiographic Techniques

Radiographs, particularly bitewing radiographs, are widely used to complement clinical diagnosis (Al Saffan, 2023). It allows visualization of interproximal and occlusal surfaces that are not readily visible in a clinical examination. Carious lesions appear as radiolucent (dark) zones due to mineral loss. Clinicians categorize carious lesions into five radiographic stages based on their depth of their depth of penetration. E1 represents damage confined to the enamel's outer layer, while E2 indicates progression into the enamel's inner layer. D1 describes lesions that have reached only the outer third of the dentin, D2 shows advancement into the middle third, and D3 represents the deepest stage where the lesion has penetrated the inner third of the dentin (Al Saffan, 2023).

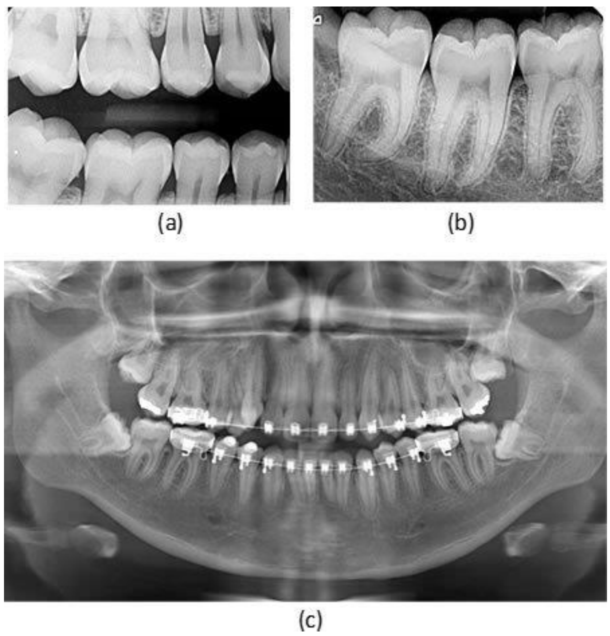


Figure 2.2: Types of Dental Radiographs (Jader et al., 2018)

However, radiographic caries detection has lower sensitivity but higher specificity for identifying initial proximal caries lesions (Al Saffan, 2023). This results in false-negative outcomes, leading to delayed diagnosis and increasing the risk of lesion progression (Walsh et al., 2021).

2.2.3 Transillumination Techniques

Transillumination is a non-invasive optical diagnostic technique that uses visible or near-infrared light to detect carious lesions. The method relies on differences in optical properties of healthy and demineralized tooth structures. Carious lesions scatter and absorb more light than healthy enamel and dentin, resulting in darker areas in the transmitted image (Al Saffan, 2023). There are two primary forms of transillumination used in dental practice: Fibre-Optic Transillumination (FOTI) and Near-Infrared Light Transillumination (NILT). FOTI uses a focused beam of white light directed at the tooth via a fibre-optic probe, where carious lesions appear as shadows due to increased light scattering. NILT uses near-infrared wavelengths (780-1500 nm) to penetrate deeper into tooth tissues, offering improved visualization of lesions without radiation exposure (Marchini et al., 2020).

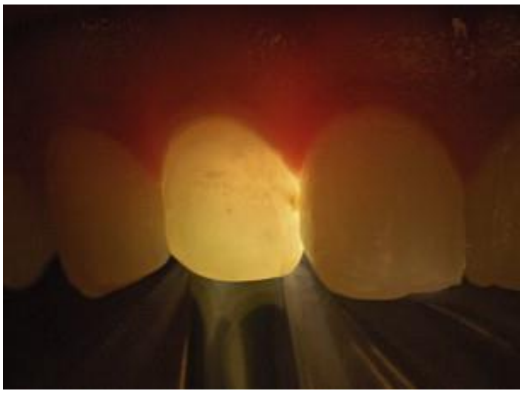


Figure 2.3: Fibreoptics Transillumination of a tooth (Rochlen & Wolff, 2011)

Despite these advances, transillumination remains sensitive to tooth anatomy since accurately interpreting deep or overlapping lesions can be challenging. It is therefore best used to complement radiographic or clinical methods rather than a standalone diagnostic tool (Elsawaf et al., 2024).

2.3 Related Works

2.3.1 Self-Supervised Learning in Dental Imaging

(Zanini et al., 2024) proposed a self-supervised approach to detect dental caries in cone beam computed tomography images. The research aimed to overcome the challenge of few annotated data in dental imaging. The model used a ResNet-18 encoder trained via SimCLR, a contrastive learning technique that encourages the model to group similar image representations together and separate dissimilar ones. It was later fine-tuned with 500 labelled images and achieved an F1-score of 88.42%, precision of 90.44%, and sensitivity of 86.67%, outperforming its supervised counterpart by a margin of over 5.5%. Figure 2.4 shows the comparison of dental caries classification.

Study		Metrics			SSL	Images	Classification type	Image type
		F1-Score	Precision	Sensitivity				
Taleb et al. [16]		–	–	57.90%	Yes	343	Binary	Bitewing
ForouzeshFar et al. [17]		93.00%	93.00%	95.00%	No	6032	Binary	Bitewing
Liu et al. [18]		88.60%	–	89.40%	No	12,524	Binary	Periapical
Oztekin et al. [19]		91.61%	87.33%	96.32%	No	13,870	Binary	Panoramic
Ezhov et al. [15]		–	–	72.85%	No	4398	Binary	CBCT
Esmaeilyfard et al. [14]		97.30%	–	96.50%	No	2355	Binary, 3-Class, 4-Class	CBCT
Zanini et al. [13]		86.20%	–	83.35%	No	493	ICDAS	CBCT
Our approach	Macro	88.42%	90.44%	86.67%	Yes	493	ICDAS	CBCT
	Binary	99.65%	99.30%	100.00%				

Figure 2.4: Comparison of Dental Caries Classification (Zanini et al., 2024)

2.3.2 Masked Image Modelling in Panoramic Radiographs

(Almalki & Latecki, 2022) applied Swin transformers with masked autoencoder pretraining to panoramic radiographs. The research aimed to improve restoration and classification of teeth and dental restorations. Two SSL frameworks, SimMIM and UM-MAE, were implemented using the Swin transformer backbone. The model was pretrained on unlabelled panoramic radiographs using masked image modelling, where random patches were masked and the model learned to reconstruct them. It was then fine-tuned on labelled data for tooth and restoration detection tasks. The final model reported an accuracy of 90.4% for tooth detection and 88.9% for restoration identification.

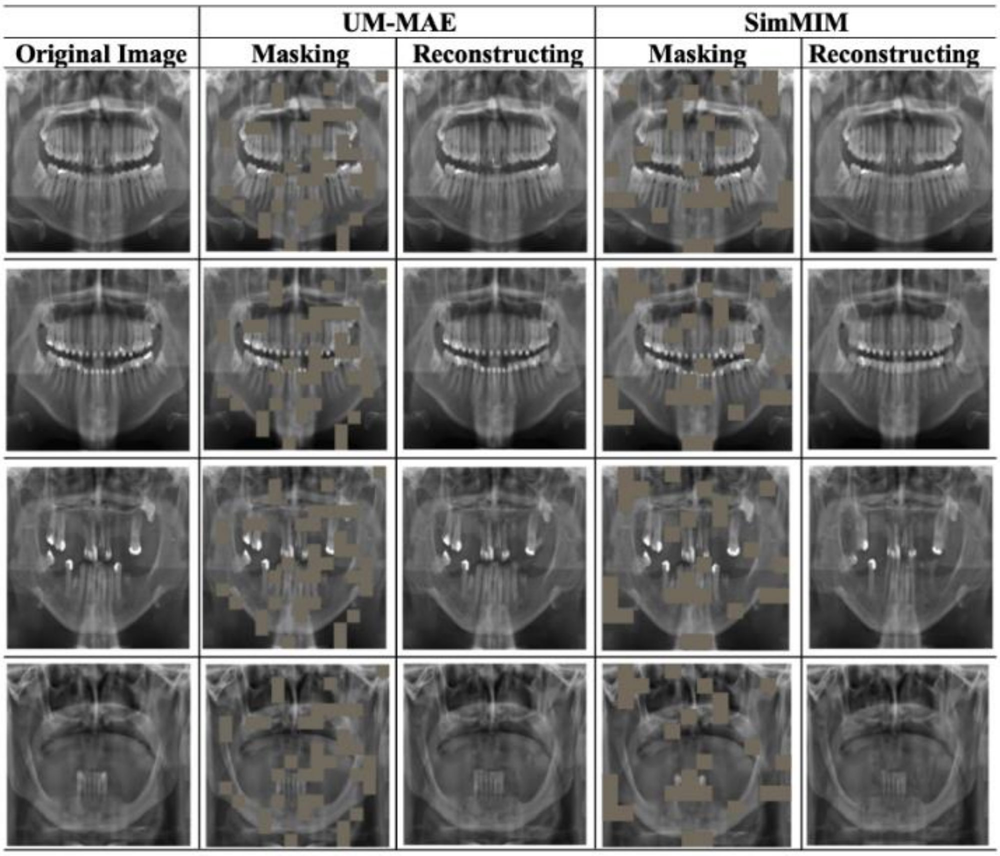


Figure 2.5: SimMIM and UM-MAE reconstruction results (Almalki & Latecki, 2022)

Initialization	Backbone	Pre-training Data	AP^{box}	AP^{mask}
Random	Swin-B	None	77.0	76.1
Supervised	Swin-B	IN-1K w/ Labels	80.3	79.2
UM-MAE	Swin-B	IN-1K	88.3	85.7
SimMIM	Swin-B	IN-1K	**90.4**	**88.9**

Figure 2.6: Results after augmenting dental restorations (Almalki & Latecki, 2022)

2.3.3 Tooth Type Enhanced Transformer for Children Caries Diagnosis

(Zhou et al., 2023) developed the Tooth-Type Enhanced Swin Transformer for diagnosing caries in children's panoramic radiographs. The model introduced domain-specific priors by integrating anatomical knowledge of different tooth types, that is; molars, canines, and incisors into the transformer pipeline. This enhancement allowed the model to better capture structural and morphological cues associated with paediatric dental diseases. The model achieved an area under the curve of 92.3%, outperforming CNNs and standard Swin transformer baselines.

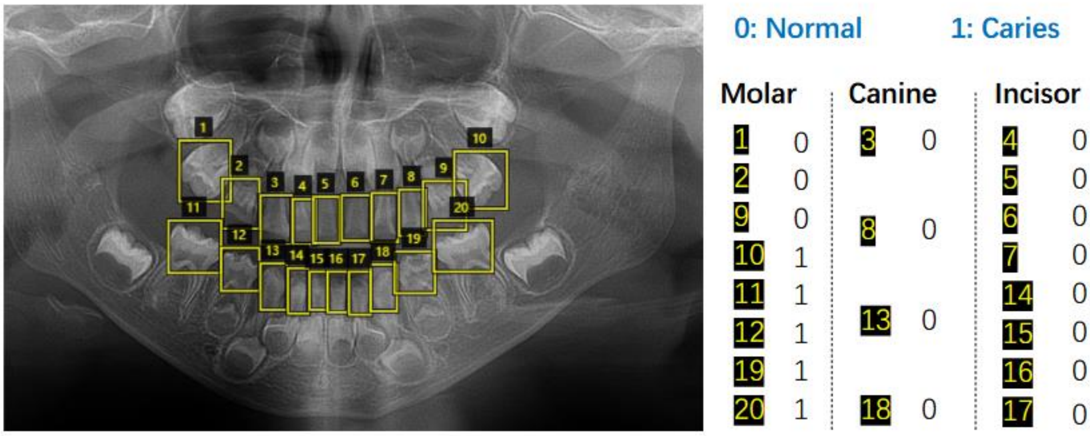


Figure 2.7: Extracting and labelling each tooth on a panoramic radiograph (Zhou et al., 2023)

Methods	Accuracy	Precision	Recall	F1	AUC
S-Transformer	0.8272	0.8576	0.7994	0.8275	0.8991
T2S-Transformer	0.8557	0.8832	0.8317	0.8567	0.9223

Figure 2.8: Performance comparison between the T2S-Transformer and the naïve S-Transformer (Zhou et al., 2023)

2.4 Gaps in Related Works

2.4.1 Self-Supervised Learning in Dental Imaging

The study in section 2.3.1 focuses on cone-beam computed tomography which is different from bitewing radiographs. CBCT captures 3D volumetric data, unlike the 2D bitewing radiographs commonly used in routine dental caries detection. Therefore, the representations learned using CBCT may not translate to 2D bitewing images due to differences in image characteristics such as contrast, resolution, and dental anatomy visibility (MacDonald & Telyakova, 2024).

2.4.2 Masked Image Modelling for Dental Panoramic Radiographs

The study in section 2.3.2 is restricted to panoramic radiographs, and model performance on bitewing data remains untested. Furthermore, the generalizability of masked image modeling to other radiographic modalities, such as bitewing images, has not been explored, reducing its clinical applicability.

2.4.3 Tooth Type Enhanced Transformer for Children Caries Diagnosis

The model in section 2.3.3 relied entirely on supervised learning, which required annotated datasets. These datasets are difficult to obtain because of ethical issues and are time-consuming to annotate.

2.5 Conceptual Framework

Based on the identified gaps in existing research, this project employs self-supervised deep learning with Vision Transformers for automated dental caries detection from bitewing radiographs. First, a Vision Transformer is pretrained on unlabelled dental x-rays using a Masked Autoencoder framework, where 75% of image patches are randomly masked and the model learns to reconstruct them. This pretraining phase enables the model to learn robust representations of dental anatomy without requiring manual annotations.

Second, the pretrained model is fine-tuned on a smaller labelled dataset for binary caries classification using cross-entropy optimization. During inference, the system generates four types of outputs: binary classification (caries present/absent), confidence score, attention heatmap showing regions suspected to have carious lesions, and clinical recommendations. The attention visualization provides interpretability by highlighting which regions influenced the model's decision, supporting clinical diagnosis and building trust in the AI system.

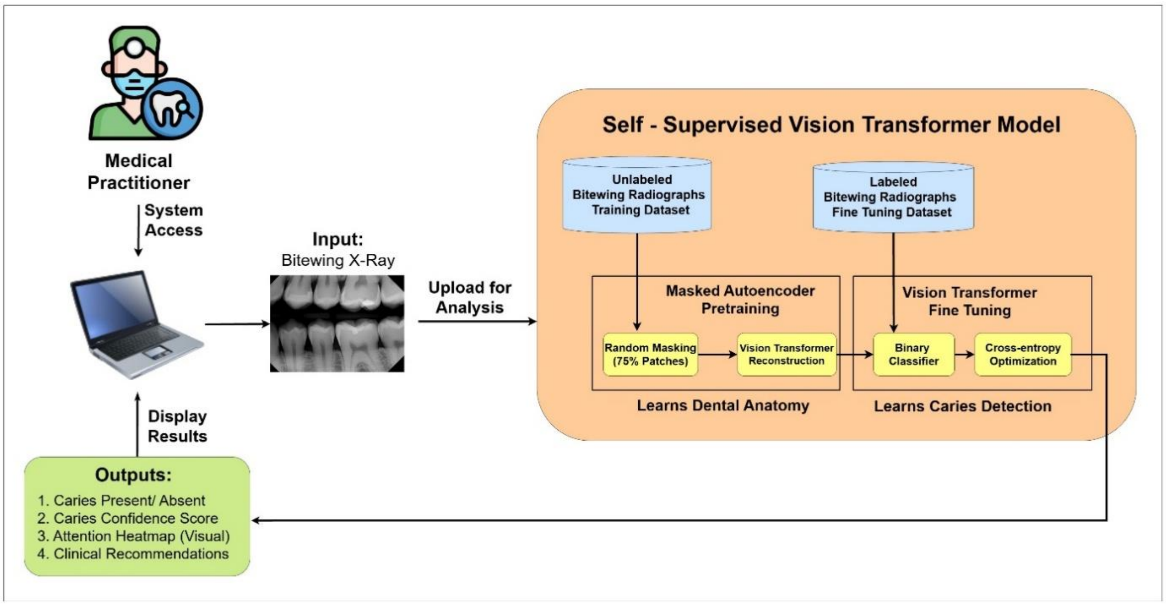


Figure 2.9: Conceptual Framework

Chapter 3: Methodology

3.1 Introduction

This chapter discusses the methodology adopted for the development of a self-supervised vision transformer model for dental caries detection from bitewing radiographs. It presents the experimental research paradigm used in this study.

The methodology covers developmental stages such as the research paradigm, collection of data and preparation, model development and validation, system analysis and design diagrams, tools and techniques used, and the deliverables produced by system.

3.2 Research Paradigm

This research adopts an experimental research paradigm as mentioned in section 3.1, focusing on iterative model training and validation to evaluate the effectiveness of using self-supervised learning and vision transformers for dental caries detection. The study documents the hypothesis that self-supervised pretraining using masked autoencoding enhances caries detection performance in scenarios with limited labelled data.

The experimental approach involves acquiring a radiographic dataset, applying preprocessing techniques, pretraining a vision transformer model using unlabelled data, fine-tuning the model with labelled data, and finally, evaluating the model. Each step is intended to contribute to the validation of hypothesis under controlled conditions.

3.2.1 Data Acquisition

This study used two datasets for different stages of model development. The tooth-number01 dataset from Kaggle (Sharma, 2025) was used for self-supervised pretraining. It contains 16,826 unlabelled bitewing dental x-ray images split into 13,460 training images, 1,682 validation images and 1,684 testing images. This dataset was used to train the Masked Autoencoder to learn general dental anatomy features such as tooth texture and structure, enamel patterns, spatial relationships between teeth and natural shadows and contours.

The bitewing-3my0p dataset from Roboflow Universe (Project-hjkow, 2024) was used for supervised fine-tuning and evaluation. It contains 624 bitewing x-ray images with YOLO-format bounding box annotations indicating caries regions. This dataset is split into 474 training images (61.4% with caries), 75 validation images (50.7% with caries), and 75 test images (53.3% with caries).

3.2.2 Data Preprocessing

All images were resized to a fixed resolution of 224 x 224 pixels to match the Vision Transformer input requirements. Standard ImageNet normalization statistics (mean = [0.485, 0.456, 0.406], std = [0.229, 0.224, 0.225]) were used.

For the pretraining phase, data augmentation techniques such as horizontal flipping (p = 0.5), random rotation (+-10degrees), and colour jittering (brightness and contrast +-20%) were applied to improve model generalization.

For the fine-tuning phase, YOLO-format annotation files were processed to create binary classification labels: images with bounding box annotations were labelled as “has caries” (Class 1), while those without annotations were labelled as “no caries” (Class 0). The same augmentation techniques were applied during training while the validation and test datasets only used to resize and normalization techniques.

3.2.3 Model Training

The model was trained in two phases. The first phase involved pretraining a Masked Autoencoder with a Vision Transformer backbone using the unlabelled tooth dataset. Following (He et al., 2021), 75% of image patches were randomly masked, and the model learned to reconstruct the original image content, encouraging the learning of contextual dental anatomy features. The pretraining used a batch size of 32, learning rate of 1.5 x 10^-4, and ran for 45 epochs with AdamW optimization.

The second phase involved fine-tuning the pretrained encoder for caries detection using the labelled bitewing x-ray images dataset. The encoder was connected to a classification head consisting of three fully connected layers (768 - 512 - 256 - 2) with ReLU activations and dropout (p = 0.3) for regularization. The model was trained for binary classification using weighted cross-entropy loss to handle class imbalance, with a learning rate of 1 x 10^-4, batch size of 16, and cosine annealing learning rate scheduling for 50 epochs. The model checkpoint achieving the highest F1-score on the validation dataset was saved for final testing.

3.2.4 Model Validation and Testing

The fine-tuned model was evaluated using 20% of the labelled dataset. Classification metrics such as accuracy, precision, recall, specificity, F1-score, and confusion matrix analysis were used to evaluate the model’s performance. Furthermore, the Area Under the ROC Curve and ROC curve visualizations were generated to assess the model’s discrimination ability across all classification thresholds.

For interpretability, the model uses attention visualization to extract the native self-attention weights from the final transformer block. By analysing the CLS token's attention distribution across image patches, spatial heatmaps are generated to visualize which regions of the dental x-ray the model focuses on during feature extraction, thus helping the users to identify the regions suspected to have dental caries.

3.3 Development Methodology and Justification

This research used the agile scrum framework to manage the iterative development of the system. Scrum is particularly suitable for experimental AI system due to its flexibility and emphasis on short feedback loops (Hema et al., 2020). The project implementation was done by a single person, and the supervisor assumed the product owner role, contributing to sprint reviews and validating outputs against research goals.

Scrum was selected due to its iterative, flexible, and user-focused approach, which is well suited for AI systems (Hema et al., 2020). Scrum also promotes close collaboration with stakeholders (Almalki & Latecki, 2022), which ensures timely feedback and iterative improvement to the system. As will be clarified from section 3.3.2 to section 3.3.5, the project development was broken down into sprints, allowing for continuous integration of model improvements, user interface refinement, and feedback from the supervisor. This helped build a reliable system.

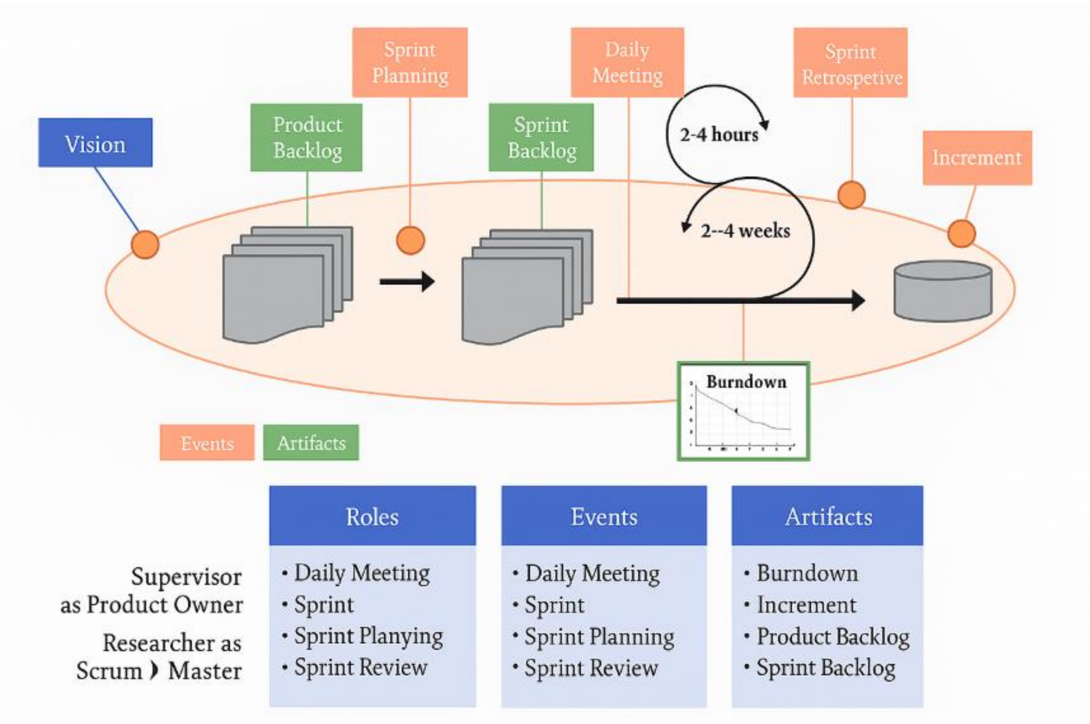


Figure 3.1: Scrum Framework Adapted for Solo AI System Development (Sassa et al., 2023)

3.3.1 Scrum Planning

This defined the scope and objectives of each sprint. At the beginning of each two-week sprint, the researcher outlines clear goals and tasks to be completed within that timeframe (Sassa et al., 2023). These goals are drawn from a dynamic product backlog consisting of tasks such as data preprocessing, model training and evaluation. The sprint backlog guides the tasks for each sprint and is also obtained from the product backlog. This is illustrated in Figure 3.1.

3.3.2 Daily Scrum Meeting

Since this was a single person project, the traditional team-based scrum meeting was adapted into a single-person development log. As illustrated in figure 3.1, each day, the researcher reflected on what was accomplished on the previous day, planned for that day, and addressed any obstacles encountered. These reflections were documented for follow-up reasons.

3.3.3 Sprint Review

Sprint reviews were done at the end of each sprint with the supervisor acting as a product owner. As illustrated in figure 3.1, the researcher demonstrated the progress made and received feedback. The feedback helped steer the next sprint by refining priorities and ensuring iterative improvement.

3.3.4 Sprint Retrospective

This was done following the sprint review to reflect on the progress made, as illustrated in figure 3.1. The researcher evaluated what went well during the sprint, the challenges encountered, and improvements that could be made in the following sprint.

3.4 System Analysis Diagrams

3.4.1 Use Case Diagram

This diagram showcased in section 4.3.1 outlines the relationship between the system and its primary users. It helped identify the core functionalities, such as logging in, uploading the image, and viewing the results. It provided a high-level overview of system functionalities and served as a reference for the developer to ensure all necessary features were implemented.

3.4.2 Sequence Diagram

This diagram showcased in section 4.3.2 shows the order of operations between the user and system components. It reflects how the system handles image input, passes it through the model, and returns predictions. This diagram helps identify the order of execution and potential timing or dependency issues of the system.

3.4.3 Context Diagram

The context diagram, showcased in section 4.3.3, gives an overview of external entities interacting with the system. It also shows the data flow to and from the model, thus sets the boundaries of the system and helps to identify necessary inputs and outputs.

3.4.4 Entity Relationship Diagram

This diagram, showcased in section 4.3.4, models how different data entities such as users, x-ray scans, model outputs, and reports relate to each other. This diagram helps design the database schema, ensuring that data storage is normalized, relationships are properly defined, and future queries can be executed efficiently.

3.5 System Design Diagrams

3.5.1 Logical Database Schema

This diagram, showcased in section 4.4.1, defines the structure of the database. It captures how data such as users, patients, and analysed records are stored and interrelated. Furthermore, it ensures data consistency, integrity, and scalability.

3.5.2 System Architecture

This diagram, showcased in section 4.4.2, outlines the technical structure of the system. It illustrates how the medical practitioner and system administrator interact with the web interface, which communicates with the classification model and the database. This ensures that the system is modular, scalable, and secure.

3.5.3 System Wireframes

The system wireframes, showcased in section 4.4.3, visually represent the user interface prior to implementation. They guide the frontend development and help ensure a user-friendly experience. The main interfaces will include login, upload, and results interface.

3.6 System Deliverables

3.6.1 Model

As explained in section 3.2, the project focuses on creating a model for identifying dental caries through the application of self-supervised learning techniques combined with vision transformer architecture.

3.6.2 User Interface

As explained in sections 3.4 and 3.5, the project aims to build an interactive web-based system that allows medical practitioners to login securely, upload bitewing x-ray images, view dental caries detection results, visualize model attention via attention visualization heatmaps, access analysis history and generate and export analysis reports.

3.6.3 Documentation

The project aims to compile comprehensive research documentation detailing the problem, the developed system, and the evaluation of the system.

3.7 System Development Tools and Techniques

This project proposes to integrate various development tools and techniques in model development, training, and evaluation.

3.7.1 Python

Python served as the core programming language, supporting useful libraries such as pandas and matplotlib for data processing, modeling and visualization.

3.7.2 PyTorch

PyTorch served as the primary deep learning framework, providing dynamic computation graphs ideal for rapid prototyping and research in self-supervised learning and computer vision.

3.7.3 Google Colab

Google Colab provided a cloud-based computational resources such as GPUs and TPUs, enabling model training without local hardware limitations.

3.7.4 OpenCV Library

The OpenCV library was used for image preprocessing tasks such as resizing, normalization, and augmentation which were essential for preparing the bitewing radiographs for model input.

3.7.5 Django Framework

Django served as the backend web framework for the system. It provided user authentication, database management, and RESTful API endpoints for communication between frontend and the database and model.

3.7.6 React.js

In this system, React was used to develop the front-end interface through which medical practitioners log in, upload dental x-rays, and visualize prediction results and attention heatmap explanations in an interactive and user-friendly way.

3.7.7 MySQL

Structured Query Language was used to manage and query the system's database. It stores user credentials, uploaded scan records, model prediction outputs, and system logs.

Chapter 4: System Analysis and Design

4.1 Introduction

This chapter documents the functional and non-functional requirements, system analysis diagrams (use case diagram, sequence diagram, entity relationship diagram, and context diagram), and system design diagrams (logical database schema, system architecture diagram, and system wireframes) that were considered during project development.

4.2 System Requirements

This section discusses the system requirements implemented in the system.

4.2.1 Functional Requirements

i. **Authentication Module** - This module manages user access to the system through a secure registration and login process. During registration, dentists provide their personal information including first and last name, as well as account credentials consisting of a username, email address, and password. To ensure account security, the following validation rules have been enforced: usernames must have at least three characters and must be unique within the system, while passwords must have at least eight characters, one uppercase, one lowercase, one number, and one special character. Once a dentist submits their registration information, the system sends a verification email to the provided email address. The email contains an activation link that the user must click to complete the registration process. After clicking the verification link, the user is redirected to the login page with their account fully activated.

   When a dentist logs in with a valid username and password for the first time, the system prompts them to set-up two factor authentication using a time-based one-time password authenticator application such as Microsoft Authenticator, Google Authenticator, or Authy. The system generates a QR code that the user scans with their chosen authenticator app to establish the connection.

   After configuring the authenticator app, the system generates and displays a set of backup authentication codes. These codes can be downloaded and are stored securely for use in situations where the authenticator app is unavailable. To complete the login process, users must enter the six-digit authentication code currently displayed in their authenticator app. If the code is correct, the system grants access and redirects the user to the system dashboard.

All passwords in the system are hashed using the PBKDF2 algorithm with a SHA-256 hash function. This protects the user accounts from unauthorized access in case of a database breach.

ii. **System Dashboard Module** - This module gives the dentist an overview of the total analyses done in the system, the number of cases with positive caries findings, and the total number of patients registered.

    The dashboard also provides navigation links to the upload single image, upload multiple images, and manage patients modules. As well as the settings module and logout button.

iii. **Patient Record Management Module** - This module manages patient information and allows for creation, viewing, updating and deleting patient records. It also displays the patient ID, name, gender, date of birth, and number of scans analysed.

iv. **X-ray Image Upload and Analysis Module** - This module allows the dentists to upload bitewing radiographs in standard formats (JPG/PNG) with a maximum file size of 10MB per image. The dentist can either upload a single x-ray image or multiple x-ray images. The dentist is required to select a patient, upload a dental x-ray image, select the image type, tooth region (optional), and clinical notes (optional). Afterwards, they can begin the image analysis process.

    The dental x-ray is processed through the self-supervised Vision Transformer model and yields four outputs: binary classification (caries present/absent), caries confidence score, attention heatmap visualization, and clinical recommendations for both doctor and patient.

v. **Dental Caries Results Module -** This module displays the outputs of the dental x-ray analysis mentioned in point (iii) above. It also includes the scan date, image type and tooth region. Finally, it includes an export button which can generate a pdf for the results of the analysis.

4.2.2 Non-Functional Requirements

i. **Performance -** The system has a robust database that can support many users and have a reliable response time. The system can process batch image

ii. **Security** – All passwords have been hashed using the PBKDF2 algorithm with a SHA-256 hash and access control has been enforced through two-factor authentication and role-based access control.

iii. **Usability** - The user interface is intuitive and easy to use.

iv. **Portability and compatibility** – The system requires a device that can support internet connection as well as the latest version of any browser.

4.3 System Analysis Diagrams

4.3.1 Use Case Diagram

The use case diagram illustrates the main actors (Dentist, Admin) and their interactions with the system. Core use cases include user registration, x-ray upload, image prediction, viewing results, and report generation. Admins manage user accounts.

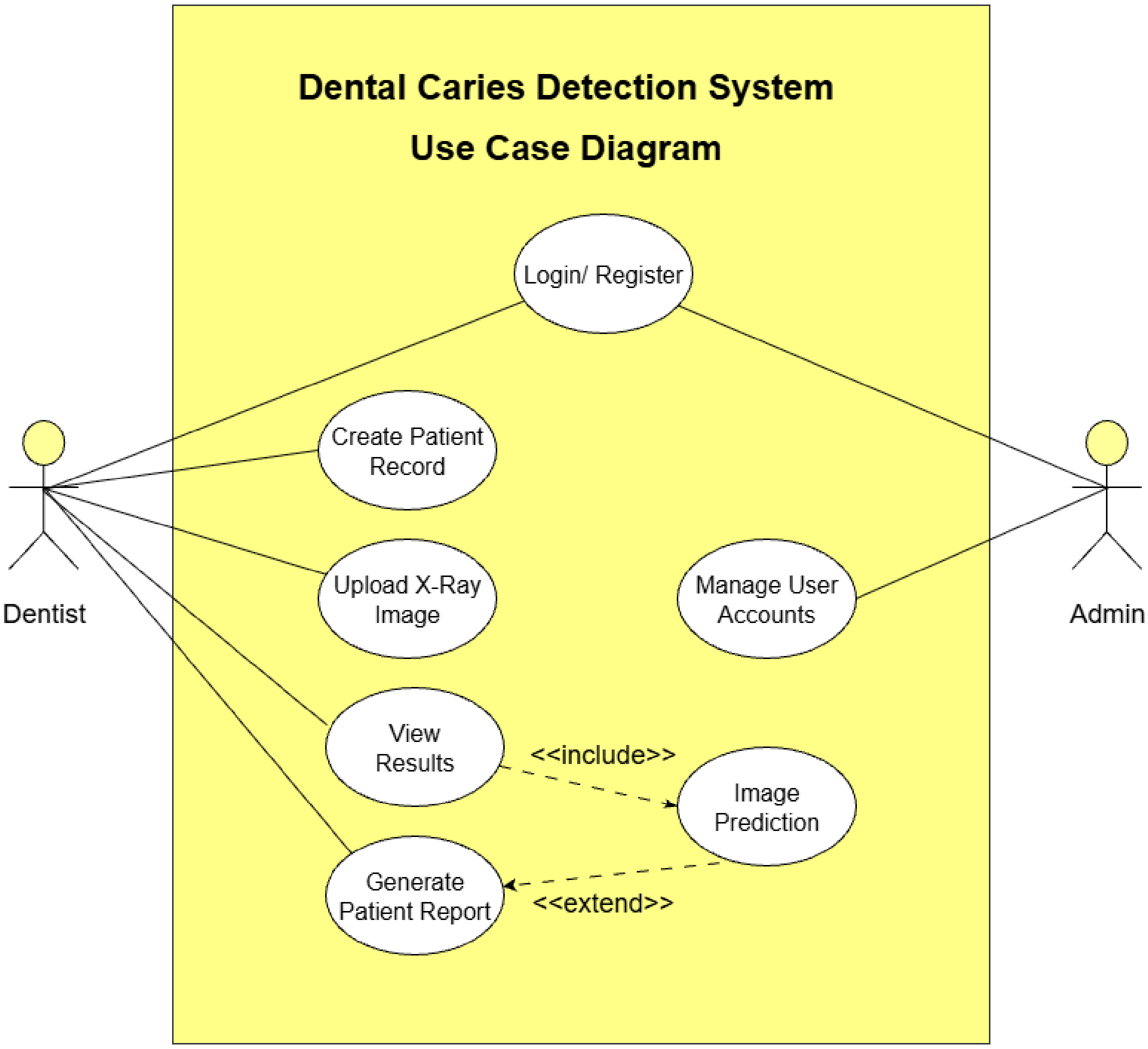


Figure 4.1: Use Case Diagram

4.3.2 Sequence Diagram

This diagram shows step-by-step interactions for registering, logging in, uploading an image, receiving classification results, and retrieving past records. It highlights communication between dentist, system, and the vision transformer model.

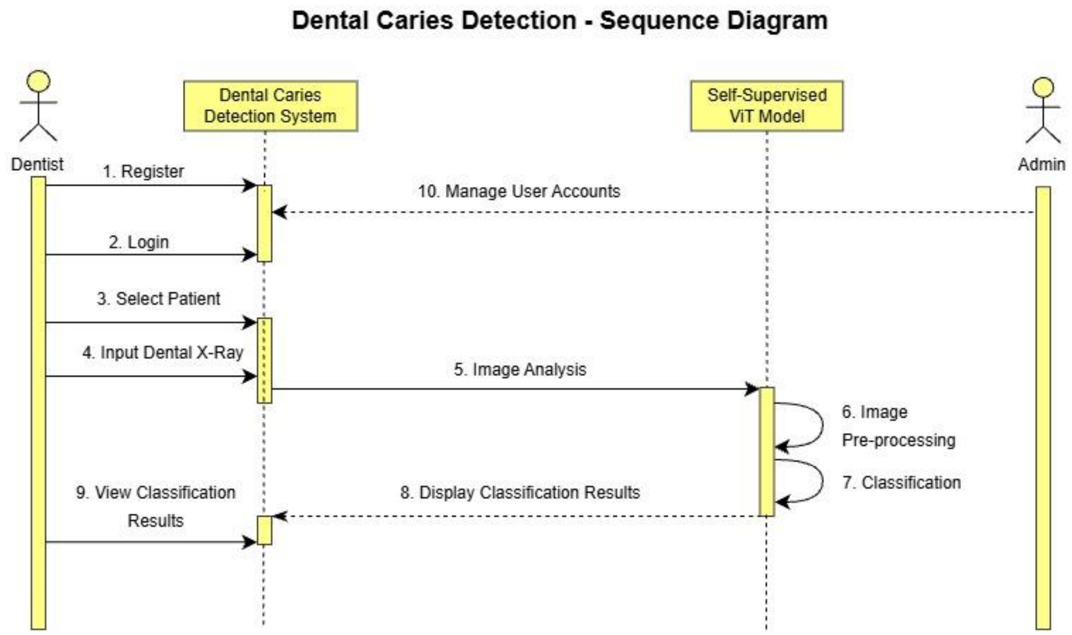


Figure 4.2: Sequence Diagram

4.3.3 Context Diagram

This diagram demonstrates how external entities (dentist and admin) interact with the Dental Caries Detection System.

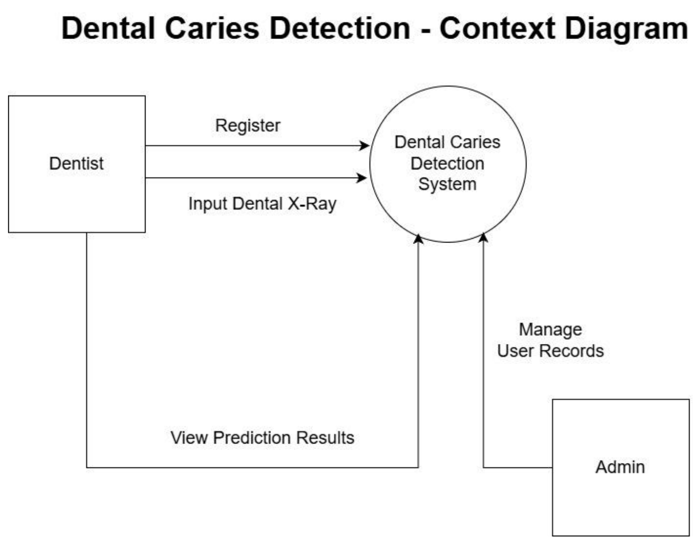


Figure 4.3: Context Diagram

4.3.4 Entity Relationship Diagram (ERD)

This diagram defines relationships between different entities of the system.

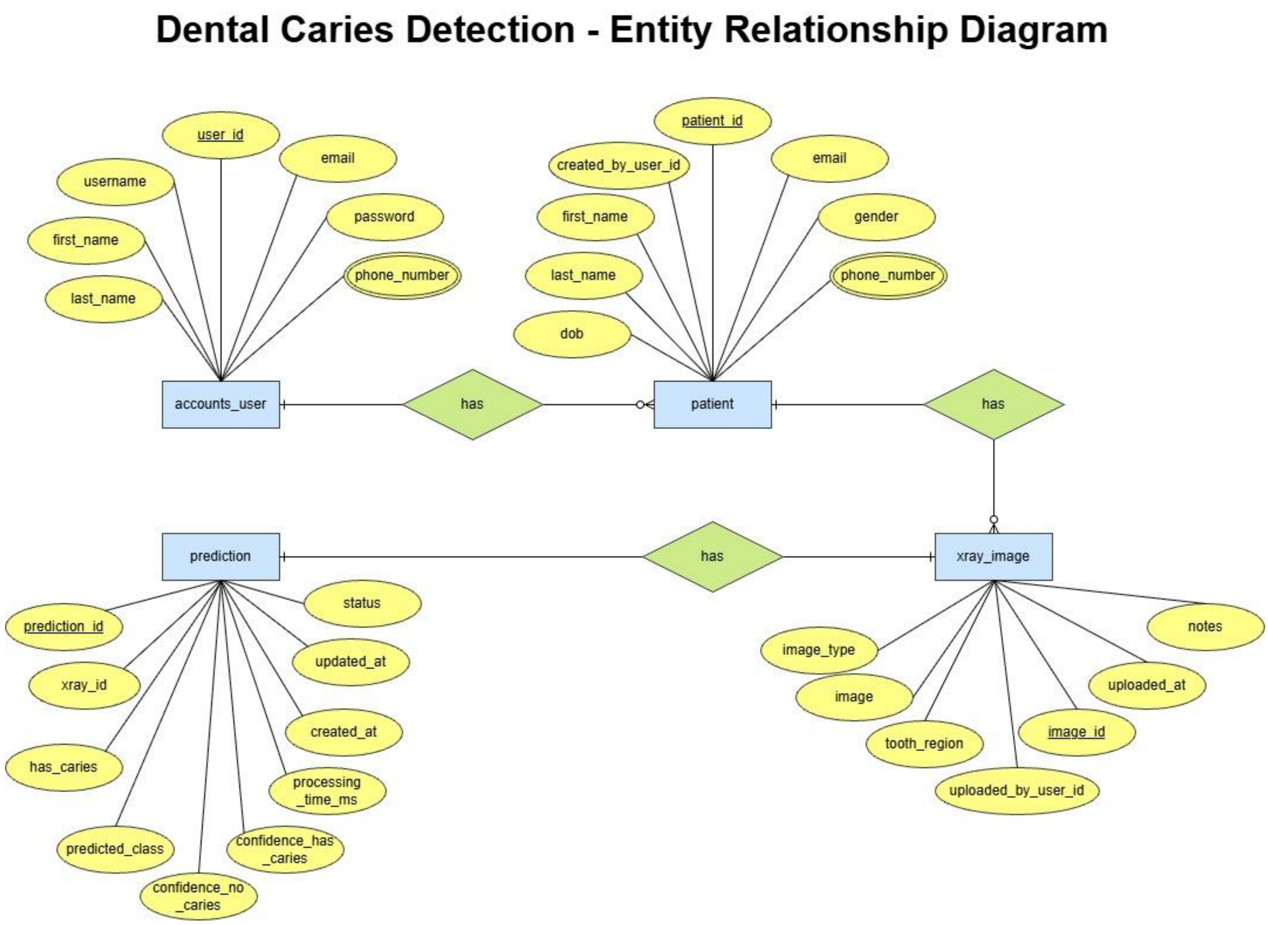


Figure 4.4: Entity Relationship Diagram

4.4 System Design Diagrams

4.4.1 Logical Database Schema

This diagram provides a schema-level representation with tables, attributes, keys, and relationships. Entities align with the ERD.

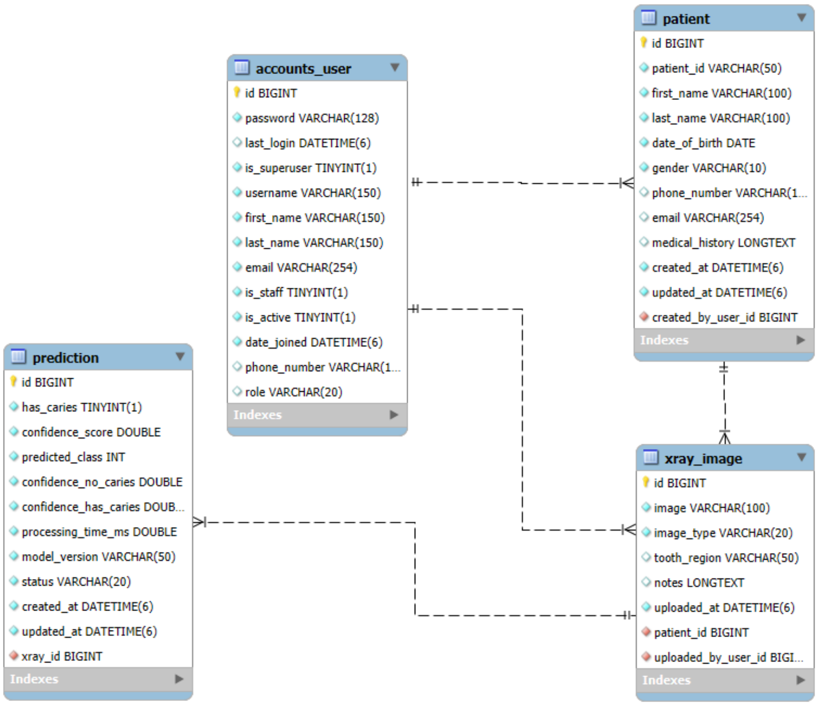


Figure 4.5: Logical Database Schema

4.4.2 System Architecture

This diagram illustrates how users access the web application, which communicates with the ViT model and the backend database over the internet. Key components include frontend UI, Django backend, ML Model, and storage.

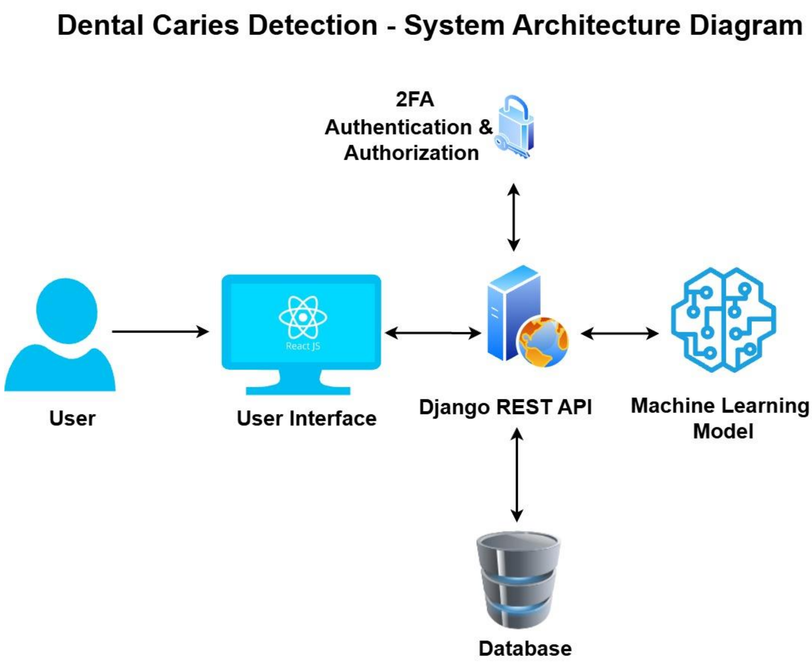


Figure 4.6: System Architecture Diagram

4.4.3 System Wireframes

i. Landing Page Wireframe

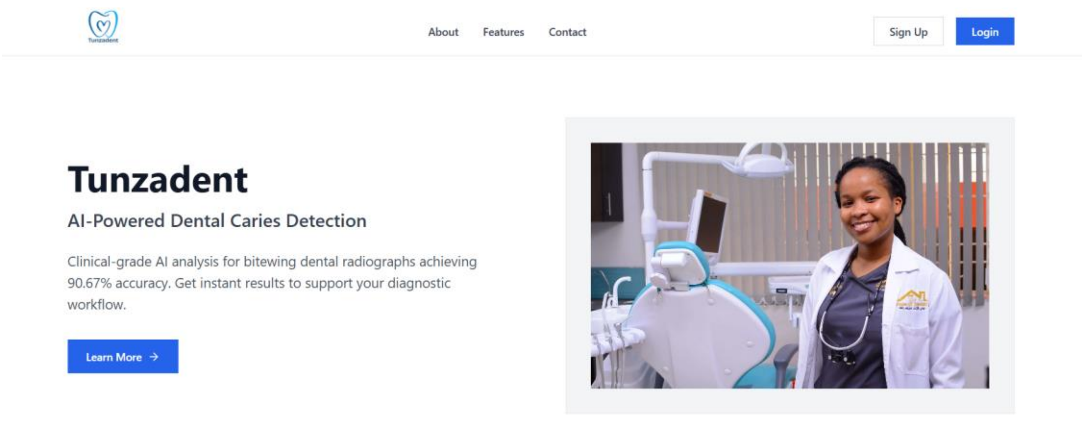


Figure 4.7: Landing Page Wireframe

ii. Registration Page Wireframe

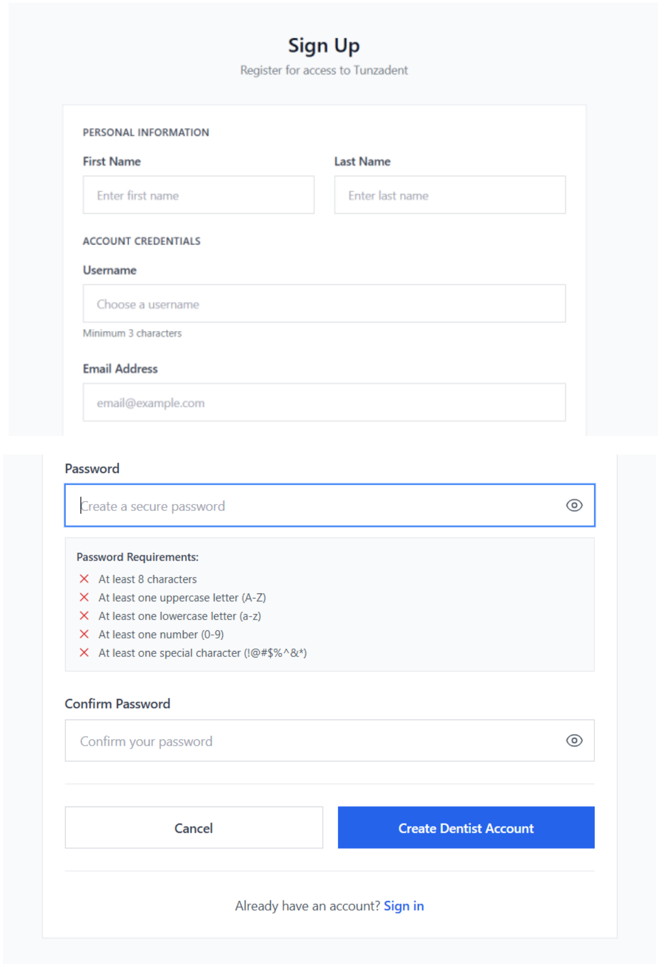


Figure 4.8: Registration Page Wireframe

iii. Login Page Wireframe

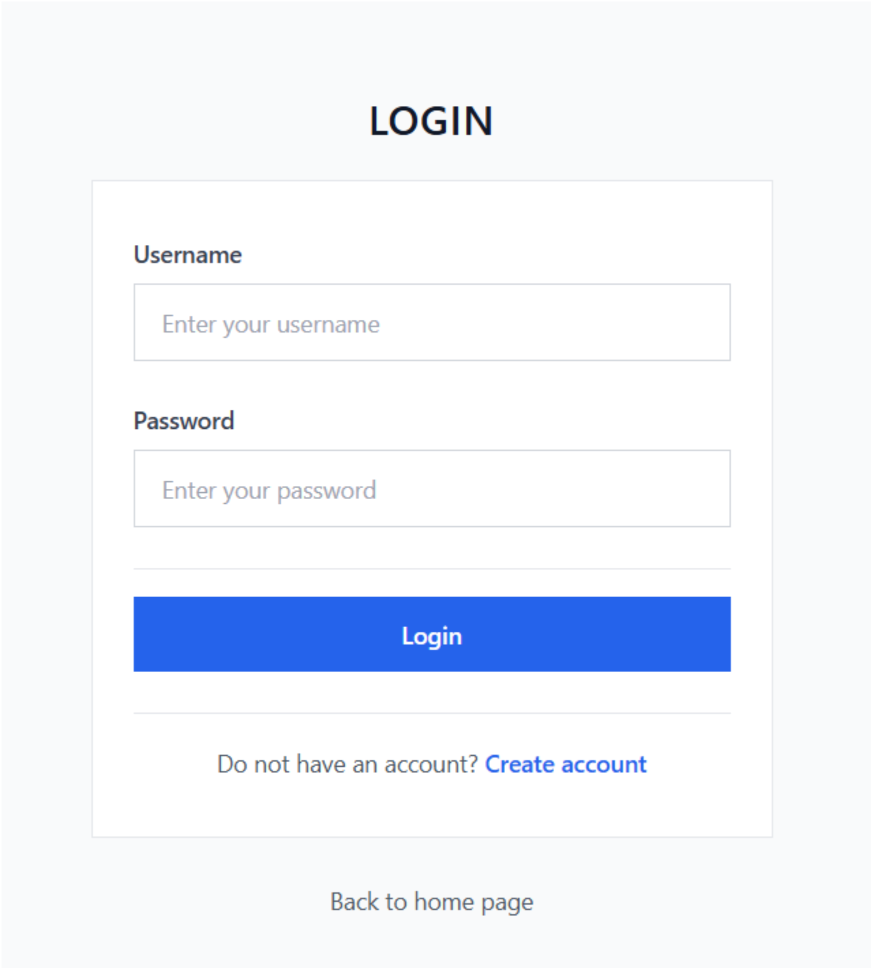


Figure 4.9: Login Page Wireframe

iv. Dashboard Wireframe

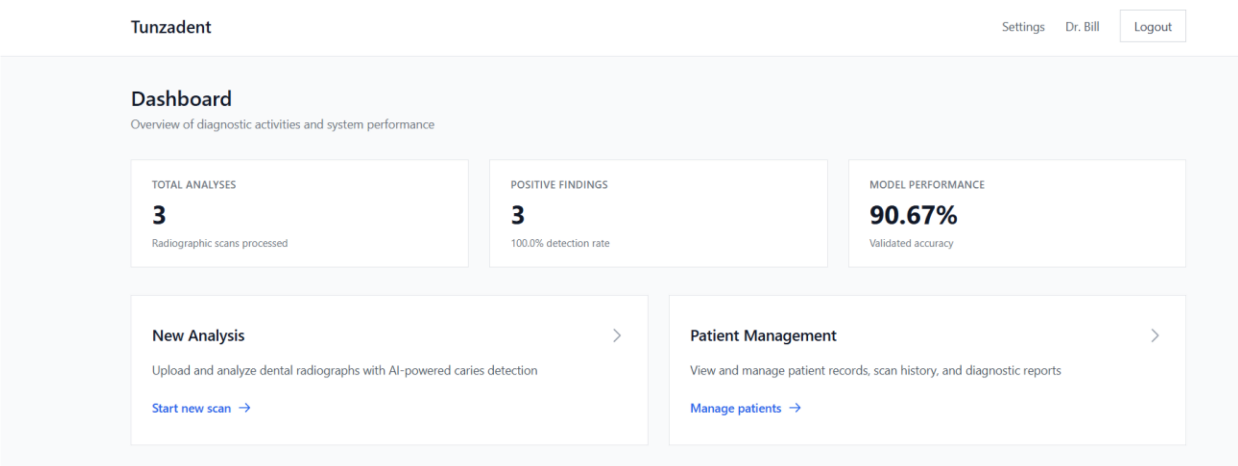


Figure 4.10: Dashboard Wireframe

v. Upload X-Ray Wireframe

This wireframe demonstrates the X-ray image upload functionality for dentists.

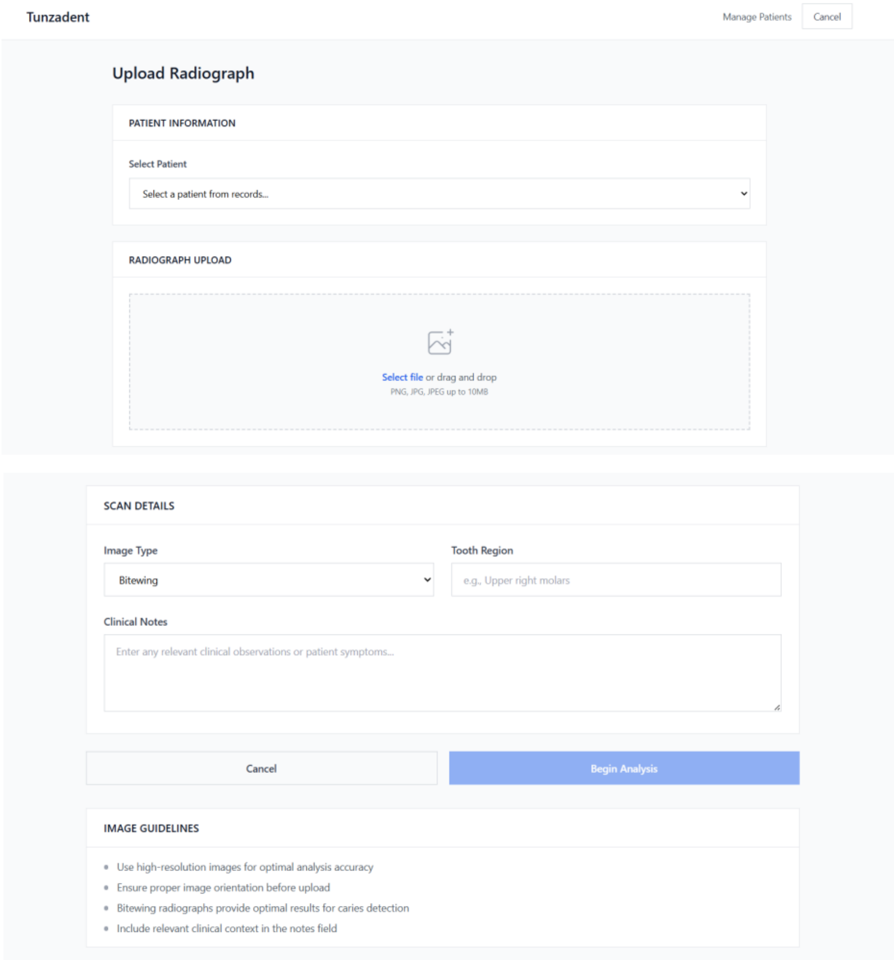


Figure 4.11: Upload X-Ray Wireframe

vi. Scan History Wireframe

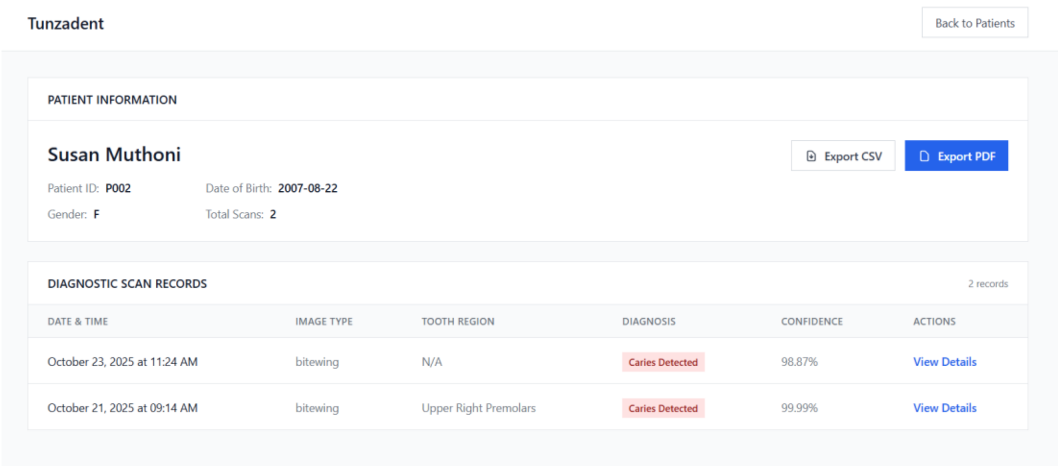


Figure 4.12: Scan History Wireframe

vii. Add Patient Wireframe

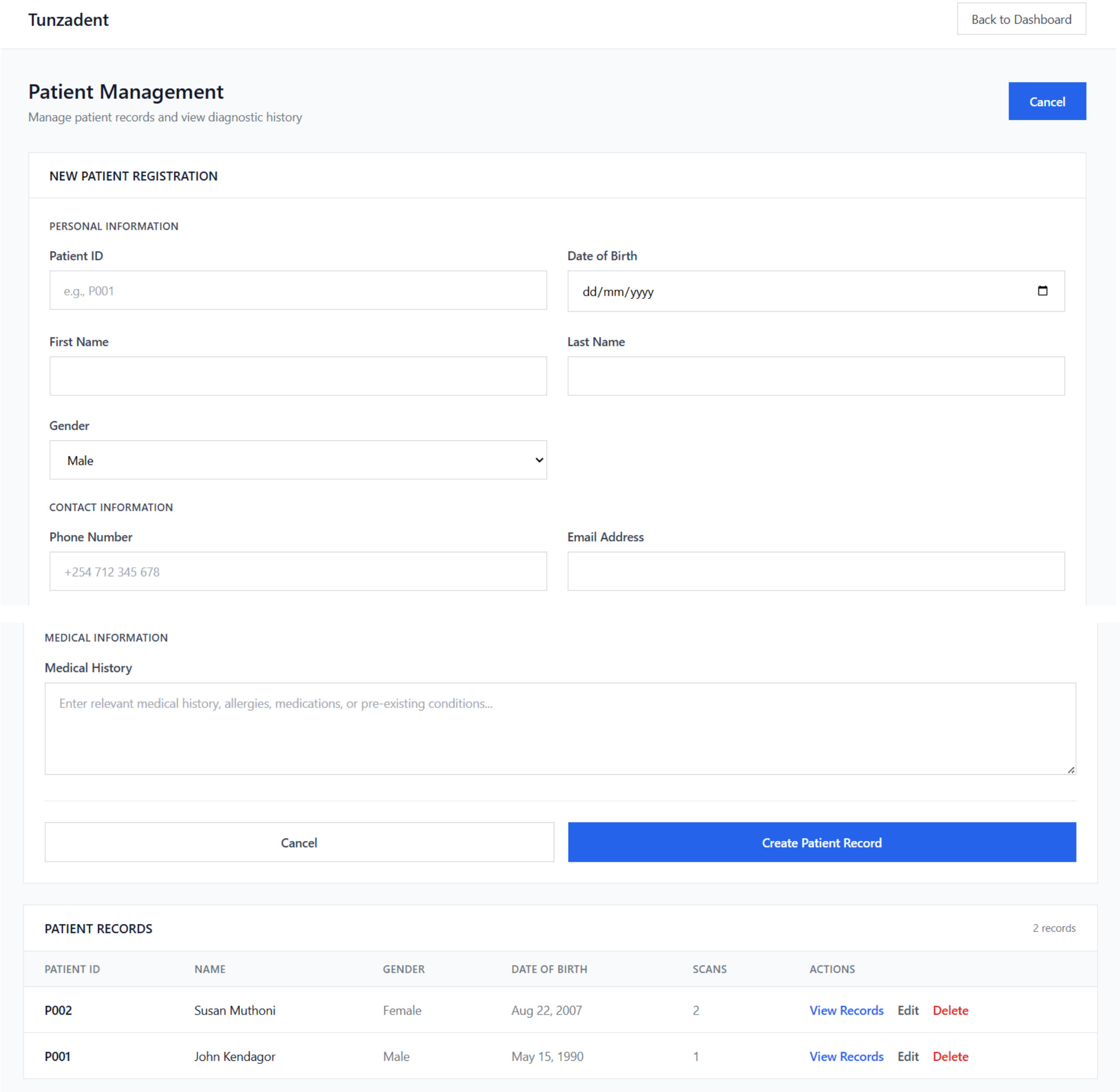


Figure 4.13: Add Patient Wireframe

Chapter 5: System Implementation and Testing

5.1 Introduction

This chapter explains the implementation and testing of the self-supervised vision transformer model for dental caries detection. It details the implementation environment, dataset description, model training and evaluation, testing methodologies and version control implemented throughout the project.

5.2 Description of the Implementation Environment

The implementation environment consists of hardware and software components used to develop and train the model.

5.2.1 Hardware Specifications

The table below gives a summary of the hardware requirements that were used to implement the system.

Table 5.1: Hardware Specifications

Hardware	Justification
GPU (NVIDIA Tesla T4)	Accelerated the training time of the model.
RAM (16GB)	Necessary to handle batch processing of x-ray radiographs during model training.
Storage (Google Drive)	Used to store checkpoints, models, datasets, and results across sessions.

5.2.2 Software Specifications

The table below gives a summary of the software requirements that were used to implement the system.

Table 5.2: Software Specifications

Software	Justification
Windows 10/11	Used to run Google Colab and the required specifications.

Google Colab	Provided free GPU access that was essential for training the model.
Python	Provided the necessary libraries for data manipulation and model training.
PyTorch	Deep learning framework used to implement the MAE architecture and Vision Transformer backbone.

5.3 Description of the Dataset

The model used two datasets: an unlabelled dataset for self-supervised pretraining and a labelled dataset for supervised fine-tuning.

5.3.1 Pre-Training Dataset

The pretraining phase employed the tooth-number01 dataset from Kaggle, which had 16,826 unlabelled dental radiographic images. The dataset was organized into three splits: training set having 13,460 images (80%), validation set having 1,682 images (10%), and testing set having 1,684 images (10%).

Although the original dataset contained annotation files for tooth detection, these labels were not utilized during the pretraining phase, as the MAE approach learns representations from the image data itself through reconstruction of masked patches.

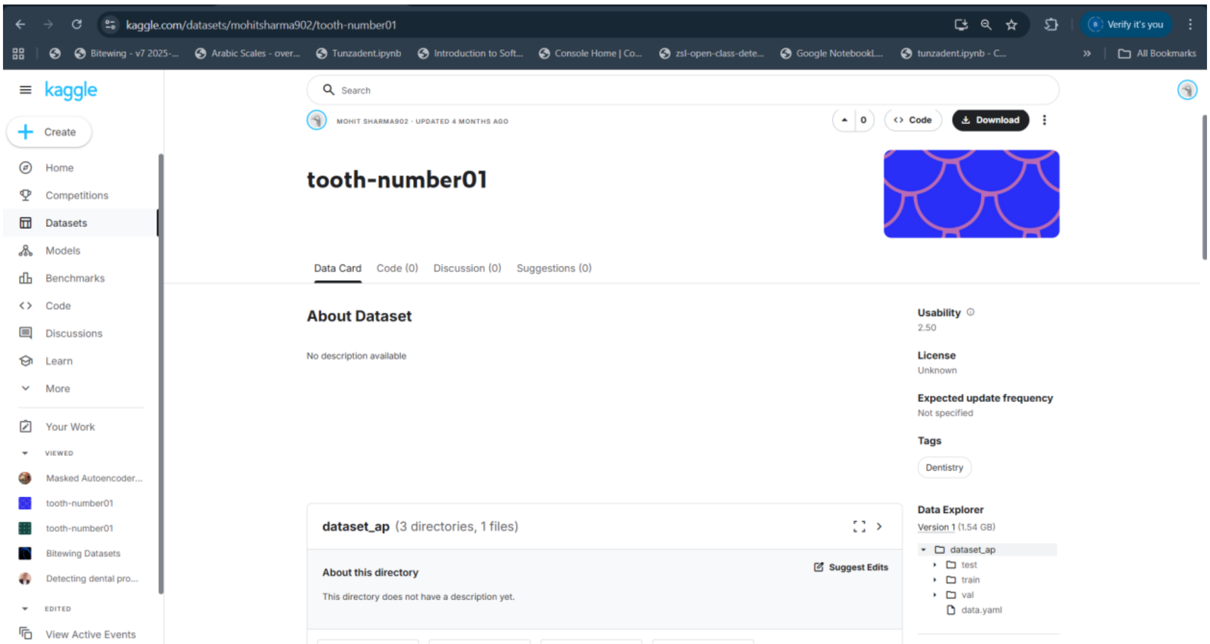


Figure 5.1: The Pre-Training Dataset Source (Sharma, 2025)

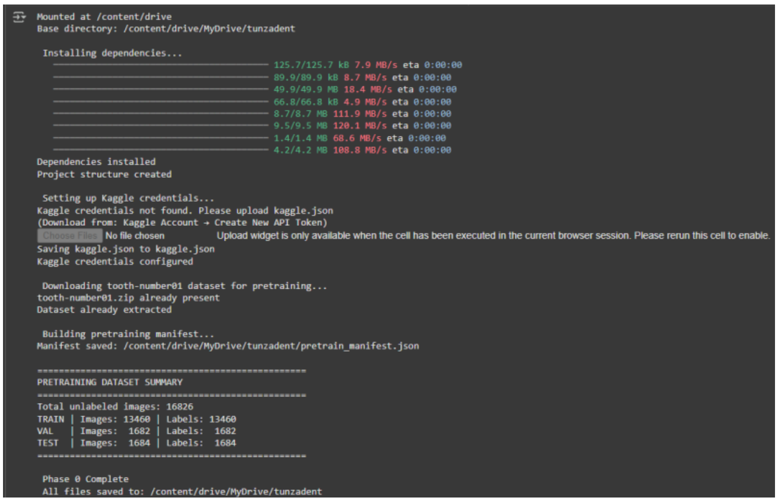


Figure 5.2: The Pre-Training Dataset Exploratory Analysis

5.3.2 Fine-tuning Dataset

The fine-tuning phase utilized the bitewing-3my0p dataset from Roboflow. This dataset contained 624 bitewing radiographs with YOLO format annotations indicating the presence and location of carious lesions.

The dataset was organized into three splits: training set having 474 images (291 with caries, 183 without caries), validation set having 75 images (38 with carries, 37 without caries), and testing set having 75 images (40 with caries, 35 without caries).

The dataset had 59.1% images with caries; this slight class imbalance was addressed through weighted loss functions during training.

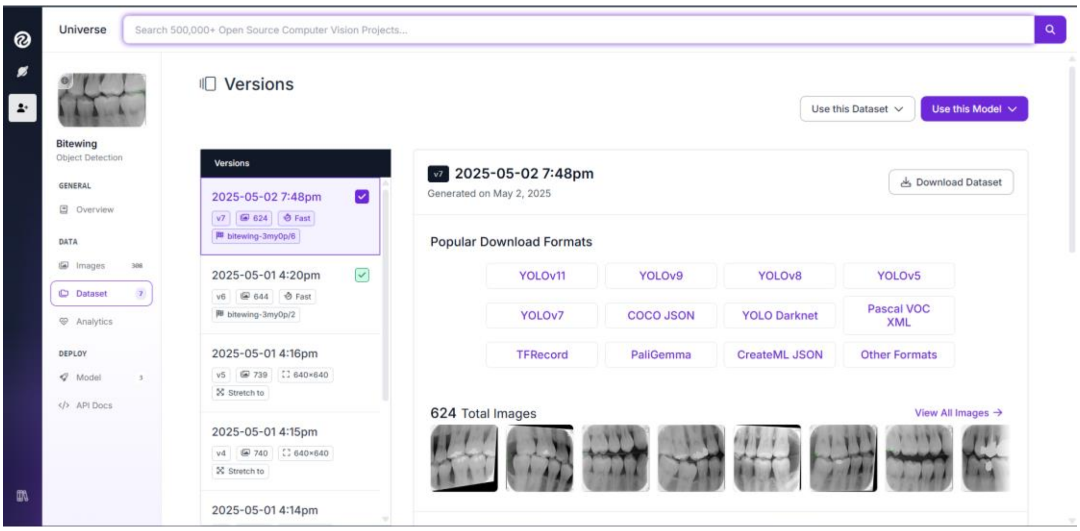


Figure 5.3: The Fine-Tuning Dataset Source (Project-hjkow, 2024)

```
Roboflow Authentication
Get your API key from: https://app.roboflow.com/settings/api

Enter your Roboflow API key: 2tC6jzSUFSZg0fo51siF
Authenticated with Roboflow

Downloading bitewing-3my0p dataset...
This may take a few minutes...

loading Roboflow workspace...
loading Roboflow project...
Downloading Dataset Version Zip in /content/drive/MyDrive/tunzadent/bitewing_caries to yolov8:: 100%|██████████| 24493/24493 [00:02<00:00, 9721.11it/s]

Extracting Dataset Version Zip to /content/drive/MyDrive/tunzadent/bitewing_caries in yolov8:: 100%|██████████| 1260/1260 [00:08<00:00, 143.15it/s]
Dataset downloaded to: /content/drive/MyDrive/tunzadent/bitewing_caries

Verifying dataset structure...

TRAIN Split:
  Total images: 474
  With caries:  291
  Without:      183

VALID Split:
  Total images: 75
  With caries:  38
  Without:      37

TEST  Split:
  Total images: 75
  With caries:  40
  Without:      35

Found data.yaml configuration file
```

Figure 5.4: The Fine-Tuning Dataset Exploratory Analysis

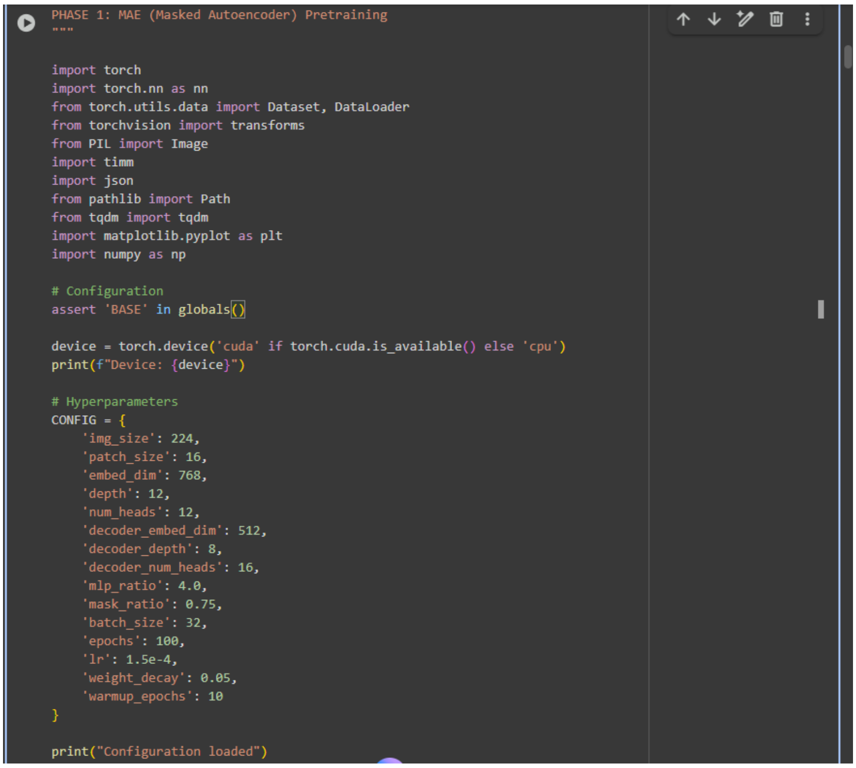
5.4 Description of Training and Evaluation

The model development followed a two-stage approach: self-supervised pretraining using Masked Autoencoders, followed by supervised fine-tuning for binary classification.

5.4.1 Data Preprocessing

All images underwent standardized preprocessing to ensure consistency:

1. Resizing: All images were resized to 224 x 224 pixels to match the vision transformer input requirements.
2. Normalization: Pixel values were normalized using ImageNet statistics (mean = [0.485, 0.456, 0.406], std = [0.229, 0.224, 0.225]).
3. Data Augmentation: The training set was augmented with random horizontal flips (probability = 0.5), random rotations (+- 10 degrees), and colour jitter (brightness = 0.2, contrast = 0.2).
4. Label Conversion: YOLO bounding box annotations were converted to binary classification labels (0 = no caries, 1 = has caries) based on the presence of annotation files.



```
PHASE 1: MAE (Masked Autoencoder) Pretraining
"""

import torch
import torch.nn as nn
from torch.utils.data import Dataset, DataLoader
from torchvision import transforms
from PIL import Image
import timm
import json
from pathlib import Path
from tqdm import tqdm
import matplotlib.pyplot as plt
import numpy as np

# Configuration
assert 'BASE' in globals()

device = torch.device('cuda' if torch.cuda.is_available() else 'cpu')
print(f"Device: {device}")

# Hyperparameters
CONFIG = {
    'img_size': 224,
    'patch_size': 16,
    'embed_dim': 768,
    'depth': 12,
    'num_heads': 12,
    'decoder_embed_dim': 512,
    'decoder_depth': 8,
    'decoder_num_heads': 16,
    'mlp_ratio': 4.0,
    'mask_ratio': 0.75,
    'batch_size': 32,
    'epochs': 100,
    'lr': 1.5e-4,
    'weight_decay': 0.05,
    'warmup_epochs': 10
}

print("Configuration loaded")
```

Figure 5.5: Data Preprocessing Configurations

5.4.2 Masked Autoencoder Pretraining

The masked autoencoder pretraining phase enabled the model to learn meaningful representations of the dental formula without requiring labelled data. The architecture was configured to have an image size of 224 x 224 pixels, patch size of 16 x16 pixels (196 patches per image), 768 embedding dimension, 12 transformer blocks of encoder depth, 12 number of attention heads, 512 decoder embedding dimension, 8 transformer blocks of decoder depth, 16 decoder attention heads and 75% masking ratio.

The model was trained using the following hyperparameters: batch size: 32, epochs: 100, learning rate: 1.5 x 10^-4, weight decay: 0.05, warmup epochs: 10, optimizer: Adam W(B1 = 0.9, B2 = 0.95), loss function: mean squared error on masked patches only.

The pretraining process was monitored continuously to detect signs of overfitting. After 45 complete epochs, early stopping was implemented. The final pretrained model with a loss of 0.2317 demonstrated effective learning of dental radiograph features.

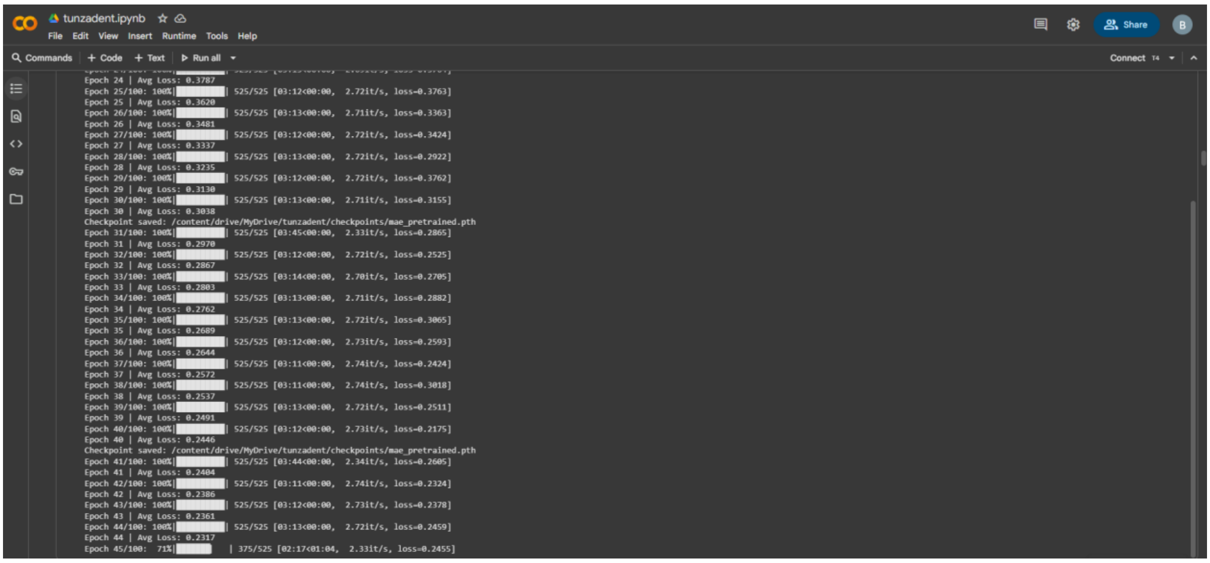


Figure 5.6: Masked Autoencoder Pretraining Completion

5.4.3 Fine-tuning for Caries Classification

The pretrained MAE encoder was adapted for binary classification through the addition of a classification head and fine-tuning on the labelled caries dataset. The encoder from the pretrained MAE was retained with its learned weights.

The classification head consisted of linear layer of 768 - 512 dimensions, ReLU activation, dropout (p=0.3), linear layer: 512 – 256, output layer: 256 – 2 classes. The following hyperparameters were used: batch size: 16, epochs: 50, learning rate: 1 x 10^-4, weight decay: 0.01, optimizer: AdamW, loss function: weighted cross-entropy loss, learning rate scheduler: cosine annealing (T_max = 50, n_min = 1 x 10^-6).

To address the class imbalance (59.1% positive cases), weighted cross-entropy loss was used with class weights inversely proportional to class frequencies. The encoder weights were not frozen, allowing for task-specific adaptation of the learned representations. The fine-tuning process achieved the following: best validation F1-score of 0.8933, final training accuracy of 96.20%, final validation accuracy of 85.33%.

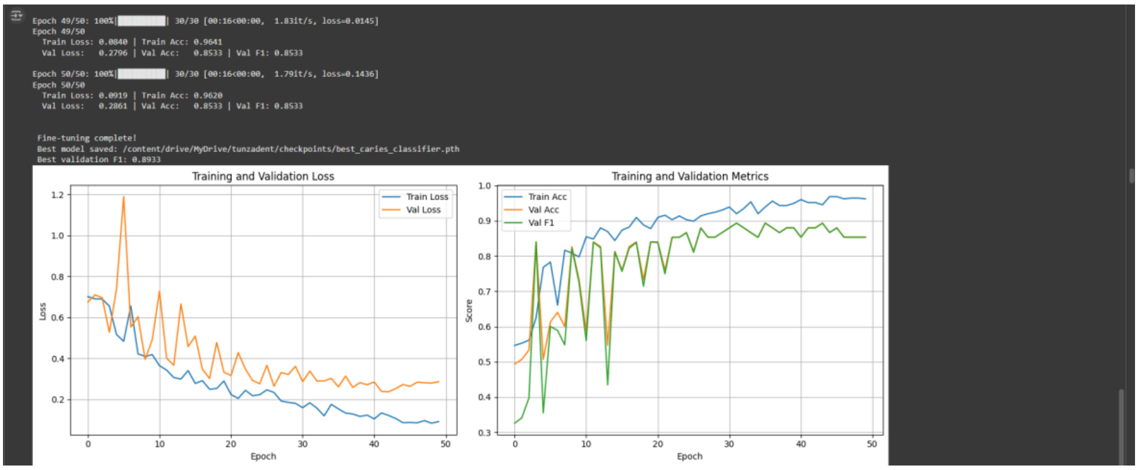


Figure 5.7: Fine-tuning for Caries Classification Completion

5.4.4 Model Evaluation Metrics

The model was evaluated using standard classification metrics. Table 5.3 below outlines the performance metrics achieved.

Table 5.3: Evaluation Metrics

Metric	Percentage
Accuracy	90.67%
Precision	90.24%
Recall	92.50%
F1-Score	91.36%
AUC-ROC	96.57%

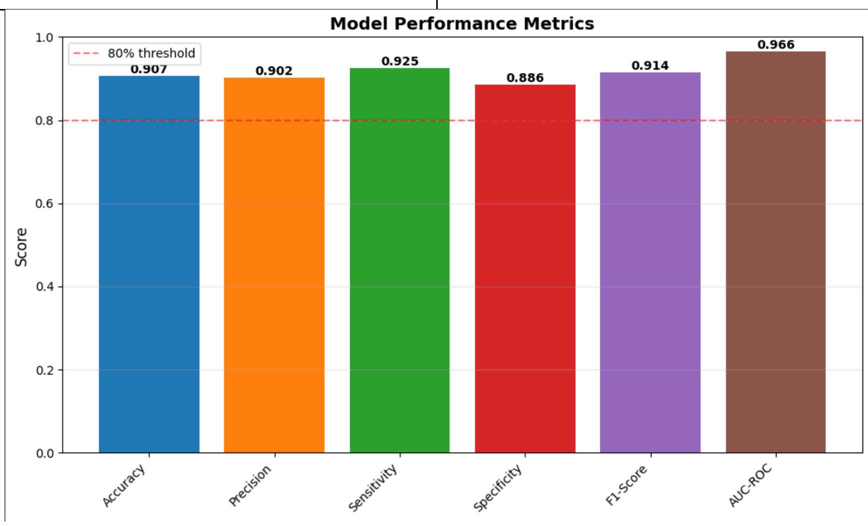


Figure 5.8: Model Performance Metrics

The confusion matrix provides a detailed insight into the model's classification patterns. Figure 5.9 shows the model's confusion matrix.

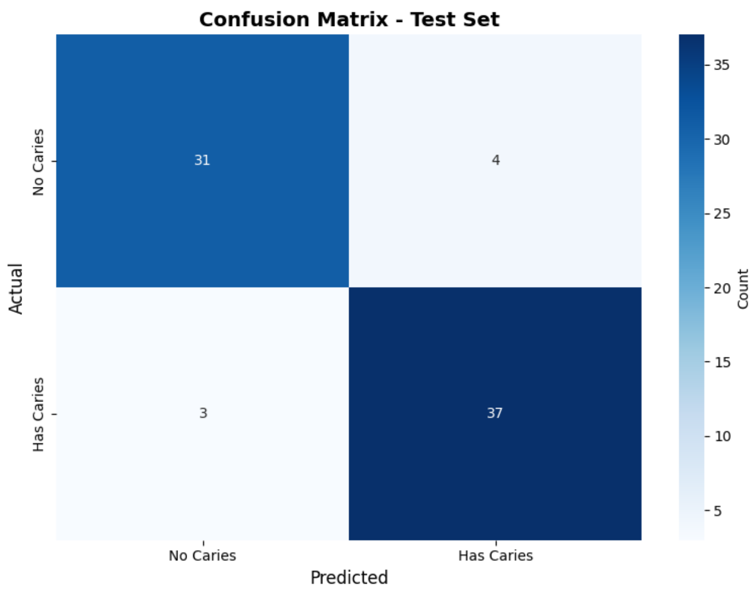


Figure 5.9: Model Confusion matrix

5.5 System implementation

5.5.1 Landing Page

This is the first interface the user encounters when accessing the system. It includes several sections: the navigation bar, hero section, about section, clinical performance metrics, diagnostic workflow, system features, frequently asked questions, contact information, and the footer. This page provides an overview of the system and guides users to subsequent modules.

5.5.2 Registration

This module allows new users to create an account by providing their first and last name, username, email address, and password. After submitting the registration form, the system sends a verification email containing an activation link. Once the user clicks the link, their account is activated, and they are redirected to set up two-factor authentication using an authenticator app such as Google Authenticator. After completing the authenticator setup, the user can proceed to the login page.

5.5.3 Login Page

This module provides fields for entering a username and password. Upon submitting valid credentials, the system requests a 6-digit code from the user's registered authenticator app. After entering a valid code, the user is logged in successfully.

5.5.4 Dashboard

The dashboard provides a summary of key system metrics, including total analyses conducted, number of cases with detected findings, and total number of registered patients. It also contains navigation buttons that lead to the image upload module and the patient management module. From this page, the user can access account settings or log out.

5.5.5 Manage Patients

This module enables users to create, view, update, and delete patient records. Each record includes patient ID, name, gender, date of birth, and the number of scans associated with the patient.

5.5.6 Upload Single Radiograph

This module allows the dentist to upload a single bitewing radiograph for analysis. The user selects the patient, uploads the radiograph, specifies the image type, tooth region, and optionally adds clinical notes. The image can then be submitted for analysis.

5.5.7 Upload Multiple Radiographs

This module enables the dentist to upload multiple radiographs for batch analysis. After selecting the patient and uploading multiple images, the system processes each image sequentially.

5.5.8 View Patient Report

This module displays the results of the image analysis. It includes whether caries were detected, the confidence score of the prediction, attention heatmaps highlighting regions of interest, analysis metadata such as scan date, image type, and tooth region.

The system also generates clinical recommendations based on confidence thresholds: **> 90% has_caries_confidence** - Severe recommendations, **70%–90% has_caries_confidence** - Moderate recommendations, **< 70% has_caries_confidence** - Mild recommendations, **> 85% no_caries_confidence** - Positive reinforcement recommendations. If severity is uncertain - Suggestion to recheck the image are provided. These outputs assist the dentist to chart a way forward for the patient.

5.5.9 Export Patient Report

This module allows users to export analysis reports in CSV or PDF formats for offline storage, sharing, or integration with patient records.

5.6 Description of Testing

This section gives a summary of how the system was tested and the outcomes of the testing process.

5.6.1 Testing Paradigm

The system was tested using back box testing which focuses on providing input and observing the outputs to verify that the system behaves correctly.

5.6.2 Testing Results

This section will discuss the testing outcomes of the web application that was used to demonstrate the model developed in this research.

Table 5.4: Registration Test Case

Test Case ID	Test Description	Test Data	Expected Results	Actual Results	Pass/ Fail
REG-001	Valid user registration	**Username**: bchakairu **Email:** bchakairu@gmail.com **Password:** Hippocratus123 **First Name:** Bill **Last Name:** Chakairu **Role:** Dentist	The user account should be created successfully. Verification email sent to provided email address. Success message displayed.	The user account should be created successfully. Verification email sent to provided email address. Success message displayed.	Pass
REG-002	Registration with existing username	**Username**: bchakairu **Email:** essie.wanja@gmail.com **Password:** EssieW@nj@123 **First Name:** Esther **Last Name:** Wanja **Role:** Dentist	Error message: A user with that username already exists. Registration fails.	Error message: A user with that username already exists. Registration fails.	Pass

REG-003	Registration with existing email	**Username**: cnjoki **Email:** bchakairu@gmail.com **Password:** C@therine123 **First Name:** Catherine **Last Name:** Njoki **Role:** Dentist	Error message: User with this email already exists. Registration fails.	Error message: User with this email already exists. Registration fails.	Pass
REG-004	Registration with weak password	**Username**: pmunyao **Email:** pmunyao@gmail.com **Password:** 123 **First Name:** Paul **Last Name:** Munyao **Role:** Dentist	Error message: Password must be at least 8 characters. Registration fails.	Error message: Password must be at least 8 characters. Registration fails.	Pass
REG-005	Registration with mismatched passwords	**Username**: kirungu **Email:** kirungu@gmail.com **Password:** Kevin123 **Confirm Password:** Kevin456 **First Name:** Kevin **Last Name:** Irungu **Role:** Dentist	Error: Passwords do not match. Registration fails.	Error: Passwords do not match. Registration fails.	Pass
REG-006	Registration with invalid email format	**Username**: jkimeu **Email:** jkimeu.com **Password:** Jason123 **Confirm Password:** Jason123 **First Name:** Jason **Last Name:** Kimeu **Role:** Dentist	Error: Please enter a valid email address. Registration fails.	Error: Please enter a valid email address. Registration fails.	Pass

REG-007	Registration with empty required fields.	**Username**: lwangui **Email:** lucy.wangui@gmail.com **First Name:** Lucy **Last Name:** Wangui	Error: Please fill out this field. Registration fails.	Error: Please fill out this field. Registration fails.	Pass

Table 5.5: Login Test Cases

Test Case ID	Test Description	Test Data	Expected Results	Actual Results	Pass/ Fail
LOG-001	Valid login credentials (verified email, 2FA not set up)	**Username**: bchakairu **Password:** Hippocratus123	Redirect to 2FA setup page. QR code displayed for authenticator app.	Redirect to 2FA setup page. QR code displayed for authenticator app.	Pass
LOG-002	Valid login credentials (verified email, 2FA already set up)	**Username**: bchakairu **Password:** Hippocratus123	2FA code input field displayed. Prompt to enter 6-digit code.	2FA code input field displayed. Prompt to enter 6-digit code.	Pass
LOG-003	Invalid username	**Username**: b_chakairu **Password:** Hippocratus123	Error message: Invalid credentials. Login fails.	Error message: Invalid credentials. Login fails.	Pass
LOG-004	Invalid password	**Username**: bchakairu **Password:** Hippocratus001	Error message: Invalid credentials. Login fails.	Error message: Invalid credentials. Login fails.	Pass

LOG-005	Login with unverified email	**Username**: smutua **Password:** Sandra123	Error message: Please verify your email before logging in. Login fails.	Error message: Please verify your email before logging in. Login fails.	Pass
LOG-006	Empty username field	**Username**: " " **Password:** Hippocratus001	Error message: Please fill out this field. Login fails.	Error message: Please fill out this field. Login fails.	Pass
LOG-007	Empty password field	**Username**: bchakairu **Password:** " "	Error message: Please fill out this field. Login fails.	Error message: Please fill out this field. Login fails.	Pass

Table 5.6: Patient Management Test Cases

Test Case ID	Test Description	Test Data	Expected Results	Actual Results	Pass/ Fail
PAT-001	Add new patient with valid data	**Patient ID:** P001 **First Name:** John **Last Name:** Kendagor **DOB:** 1990-05-15 **Gender:** Male **Phone:** +254712345678 **Email:** jk@gmail.com	Patient added successfully. Success message displayed. Patient appears in patient list.	Patient added successfully. Success message displayed. Patient appears in patient list.	Pass
PAT-002	Add patient with duplicate ID	**Patient ID:** P001 **First Name:** Susan **Last Name:** Muthoni	Error message: "Patient ID with this patient	Patient ID with this patient id already exists. Patient not added.	Pass

		DOB: 2007-08-22 **Gender:** Female **Phone:** +254772354678 **Email:** sw@gmail.com	id already exists.” Patient not added.		
PAT-003	Add patient with missing required fields	**Patient ID:** “ ” **First Name:** Jack **Last Name:** Mutwiri **DOB:** 2010-03-12 **Gender:** Male **Phone:** +254734564678 **Email:** jm@email.com	Error messages for each required field. Patient not added.	Error: Please fill out this field. Patient not added.	Pass
PAT-004	View patient list	Navigate to patients page	List of all registered patients displayed with ID, name, gender, DOB, and X-ray count.	List of all registered patients displayed with ID, name, gender, DOB, and X-ray count.	Pass

Table 5.7: X-Ray Upload Test Cases

Test Case ID	Test Description	Test Data	Expected Results	Actual Results	Pass/Fail
XRY-001	Upload valid bitewing X-ray (JPG)	**Patient:** P001 **File:** bitewing1.jpg (valid bitewing, 2MB) **Image Type:** Bitewing **Tooth Region:** Upper right molars	File uploads successfully. Preview displayed. Success message shown.	File uploads successfully. Preview displayed. Success	Pass

				message shown.	
XRY-002	Upload without selecting patient	**Patient:** (none selected) **File:** bitewing2.jpg (valid bitewing, 2MB) **Image Type:** Bitewing **Tooth Region:** Upper right molars	Error message: "Please select an item in the list." Upload fails.	Error message: "Please select an item in the list." Upload fails.	Pass
XRY-003	Upload without selecting file	**Patient:** P002 **File:** (none selected) **Image Type:** Bitewing **Tooth Region:** Upper right molars	Upload button disabled.	Upload button disabled.	Pass

Table 5.8: Caries Prediction Test Cases

Test Case ID	Test Description	Test Data	Expected Results	Actual Results	Pass/Fail
PRD-001	Predict on X-ray with caries	Upload bitewing X-ray containing visible caries	Model detects caries. Results page shows: "Caries Detected", confidence score given.	Model detects caries. Results page shows: "Caries Detected", confidence score given.	Pass
PRD-002	Predict on X-ray without caries	Upload bitewing X-ray with healthy teeth	Model detects no caries. Results page shows: "No Caries Detected", confidence score given.	Model detects no caries. Results page shows: "No Caries Detected", confidence score given.	Pass
PRD-003	View prediction details	After prediction, view results page	Display shows: presence/absence, confidence score, processing time, model version,	Display shows: presence/absence, confidence score, processing time, model version,	Pass

			patient info, X-ray metadata.	patient info, X-ray metadata.	
PRD-004	Multiple predictions for same patient	Upload 3 different X-rays for same patient	Each prediction saved separately. Patient's X-ray count updates correctly. History accessible.	Each prediction saved separately. Patient's X-ray count updates correctly. History accessible.	Pass

5.7 Github Documentation

GitHub was used in this project to manage version control, track progress, and maintain a clear record of feature development.

5.7.1 Branches Used and Their Purpose

Table 5.9: Git branches used and their Purpose

Branch Name	Purpose
feat/sprint-1-email-verification	Added email verification feature.
feat/sprint-1-totp-2fa	Implemented time-based one-time password two-factor-authentication.
feat/sprint-1-auth-ui	Developed the user interface for authentication (login, registration).
feat/sprint-1-landing-page	Created the landing page for the system.
feat/sprint-2-patient-crud-api	Built backend APIs for patient CRUD operations.
feat/sprint-2-patient-crud-ui	Developed frontend UI for patient CRUD.
feat/sprint-2-upload-radiographs-ui	Created UI for uploading radiographs.
feat/sprint-3-scan-history	Implemented history of patient scans UI.
feat/sprint-3-ml-integration	Integrated machine learning model for caries prediction.
feat/sprint-4-dashboard-analytics	Developed analytics dashboard.
feat/sprint-5-export-pdf	Added feature to export reports as PDF.
lizztt-patch-1, lizztt-patch-2	Update readme file.

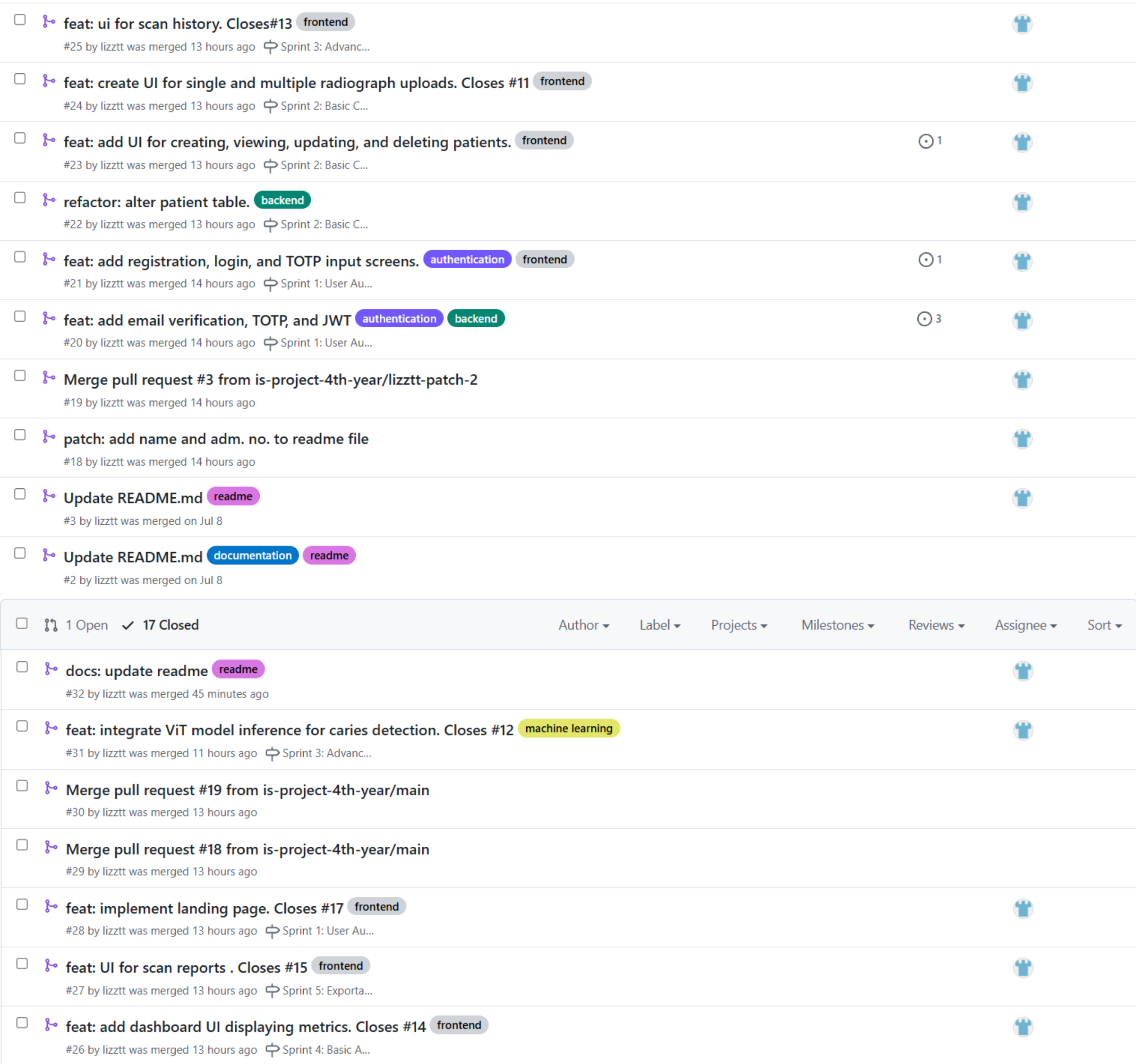


Figure 5.10: Git Merge Pull Requests Flow

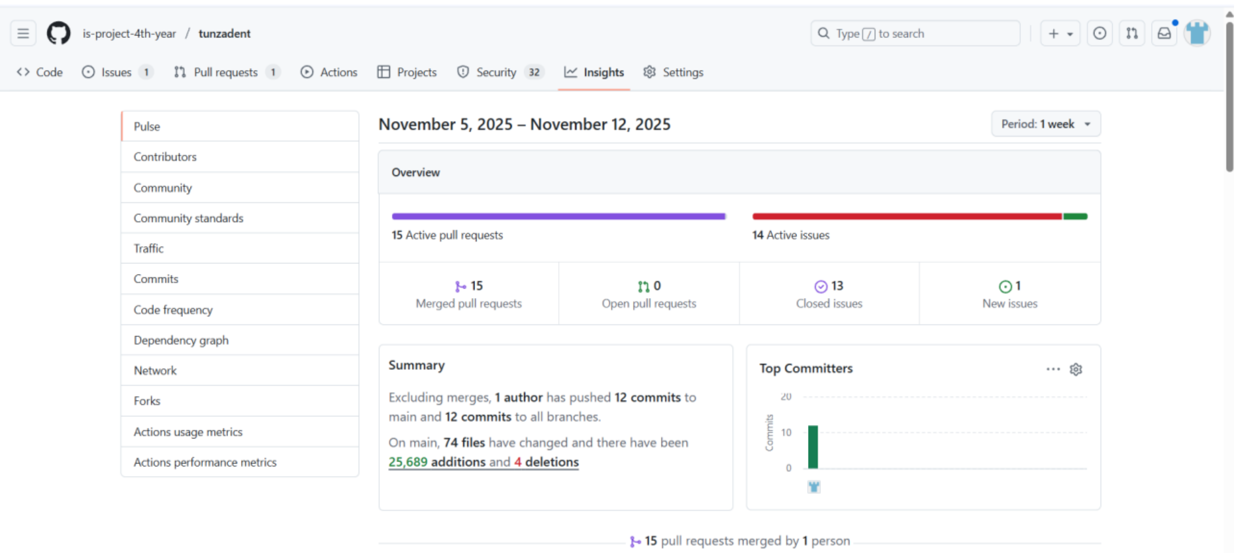


Figure 5.11: Git Insights

Chapter 6: Conclusions, Recommendations and Future Works

6.1 Conclusions

This research developed and tested a self-supervised Vision Transformer model to detect dental caries from bitewing radiographs. The model performed well, achieving **90.67% accuracy, 91.36% F1-score, and 96.57% AUC-ROC**. These results show that the model can reliably identify carious regions in dental images, demonstrating the potential of self-supervised learning for tasks with limited labelled data.

6.2 Recommendations

To make the system more useful in a clinical setting, it is recommended that it be connected to existing hospital or dental clinic systems. This would allow patient data to flow smoothly between the diagnostic tool and electronic health records. The current rule-based recommendation system could also be improved by integrating large language models (LLMs) to suggest a wider range of treatment options, while keeping the final decision with the clinician.

6.3 Future Works

Future research could focus on extending the system in several ways. First, moving from simple binary detection to caries severity grading would allow the model to distinguish between different stages of caries, supporting better treatment planning and monitoring over time. Second, x-ray image segmentation could be applied to detect exact lesion boundaries, rather than just highlighting areas of concern, which would help in measuring lesion size and tracking changes in follow-up visits.

References

Abdelaziz, M. (2023). Detection, Diagnosis, and Monitoring of Early Caries: The Future of Individualized Dental Care. *Diagnostics*, *13*(24), 3649. https://doi.org/10.3390/diagnostics13243649

Al Saffan, A. D. (2023). Current Approaches to Diagnosis of Early Proximal Carious Lesion: A Literature Review. *Cureus*. https://doi.org/10.7759/cureus.43489

Almalki, A., & Latecki, L. J. (2022). *Self-Supervised Learning with Masked Image Modeling for Teeth Numbering, Detection of Dental Restorations, and Instance Segmentation in Dental Panoramic Radiographs* (No. arXiv:2210.11404). arXiv. https://doi.org/10.48550/arXiv.2210.11404

Bayrakdar, I. S., Orhan, K., Akarsu, S., Çelik, Ö., Atasoy, S., Pekince, A., Yasa, Y., Bilgir, E., Sağlam, H., Aslan, A. F., & Odabaş, A. (2022). Deep-learning approach for caries detection and segmentation on dental bitewing radiographs. *Oral Radiology*, *38*(4), 468–479. https://doi.org/10.1007/s11282-021-00577-9

Dosovitskiy, A., Beyer, L., Kolesnikov, A., Weissenborn, D., Zhai, X., Unterthiner, T., Dehghani, M., Minderer, M., Heigold, G., Gelly, S., Uszkoreit, J., & Houlsby, N. (2021). *An Image is Worth 16x16 Words: Transformers for Image Recognition at Scale* (No. arXiv:2010.11929). arXiv. https://doi.org/10.48550/arXiv.2010.11929

Elsawaf, A. E., Deri, A. Y. A., Armanious, P. S., Khasawneh, A. M., AlKhaja, A. M., Yasin, A. R., Al-Rawi, N. H., Kawas, S. A., & Shetty, S. R. (2024). Efficiency of Near-Infrared Technology in the Clinical Detection of Carious Lesions: A Systematic Review. *European Journal of Dentistry*, *18*(01), 014–025. https://doi.org/10.1055/s-0043-1761187

Ghodasra, R., & Brizuela, M. (2025). Dental Caries Diagnostic Testing. In *StatPearls*. StatPearls Publishing. http://www.ncbi.nlm.nih.gov/books/NBK574510/

He, K., Chen, X., Xie, S., Li, Y., Dollár, P., & Girshick, R. (2021). *Masked Autoencoders Are Scalable Vision Learners* (Version 3). arXiv. https://doi.org/10.48550/ARXIV.2111.06377

Hema, V., Thota, S., Naresh Kumar, S., Padmaja, C., Rama Krishna, C. B., & Mahender, K. (2020). Scrum: An Effective Software Development Agile Tool. *IOP Conference Series: Materials Science and Engineering*, *981*(2), 022060. https://doi.org/10.1088/1757-899X/981/2/022060

Jader, G., Fontineli, J., Ruiz, M., Abdalla, K., Pithon, M., & Oliveira, L. (2018). Deep Instance Segmentation of Teeth in Panoramic X-Ray Images. *2018 31st SIBGRAPI Conference on Graphics, Patterns and Images (SIBGRAPI)*, 400–407. https://doi.org/10.1109/SIBGRAPI.2018.00058

MacDonald, D., & Telyakova, V. (2024). An Overview of Cone-Beam Computed Tomography and Dental Panoramic Radiography in Dentistry in the Community. *Tomography*, *10*(8), 1222–1237. https://doi.org/10.3390/tomography10080092

Mahesh Batra, A., & Reche, A. (2023). *A New Era of Dental Care: Harnessing Artificial Intelligence for Better Diagnosis and Treatment—PubMed*. https://pubmed.ncbi.nlm.nih.gov/38143639/

Marchini, L., Reynolds, J. C., Caplan, D. J., Sasser, S., & Russell, C. (2020). Predictors of having a dentist among older adults in Iowa. *Community Dentistry and Oral Epidemiology*, *48*(3), 240–247. https://doi.org/10.1111/cdoe.12521

Project-hjkow. (2024). *Bitewing-3my0p* [Dataset]. https://universe.roboflow.com/project-hjkow/bitewing-3my0p/dataset/7

Rochlen, G. K., & Wolff, M. S. (2011). Technological Advances in Caries Diagnosis. *Dental Clinics of North America*, *55*(3), 441–452. https://doi.org/10.1016/j.cden.2011.02.018

Sadikoglu, Cyprus, & Nicosia. (2020). White Spot Lesions: Recent Detection and Treatment Methods. *Cyprus Journal of Medical Sciences*, *5*(3), 260–266. https://doi.org/10.5152/cjms.2020.1902

Sassa, A. C., Almeida, I. A. D., Pereira, T. N. F., & Oliveira, M. S. D. (2023). Scrum: A Systematic Literature Review. *International Journal of Advanced Computer Science and Applications*, *14*(4). https://doi.org/10.14569/IJACSA.2023.0140420

Sharma, M. (2025). *Tooth-number01* [Dataset]. Kaggle.Com. https://www.kaggle.com/datasets/mohitsharma902/tooth-number01

Walsh, T., Macey, R., Riley, P., Glenny, A.-M., Schwendicke, F., Worthington, H. V., Clarkson, J. E., Ricketts, D., Su, T.-L., & Sengupta, A. (2021). Imaging modalities to inform the detection and diagnosis of early caries. *Cochrane Database of Systematic Reviews*, *2021*(12). https://doi.org/10.1002/14651858.CD014545

Warreth, A. (2023). Dental Caries and Its Management. *International Journal of Dentistry*, *2023*, 1–15. https://doi.org/10.1155/2023/9365845

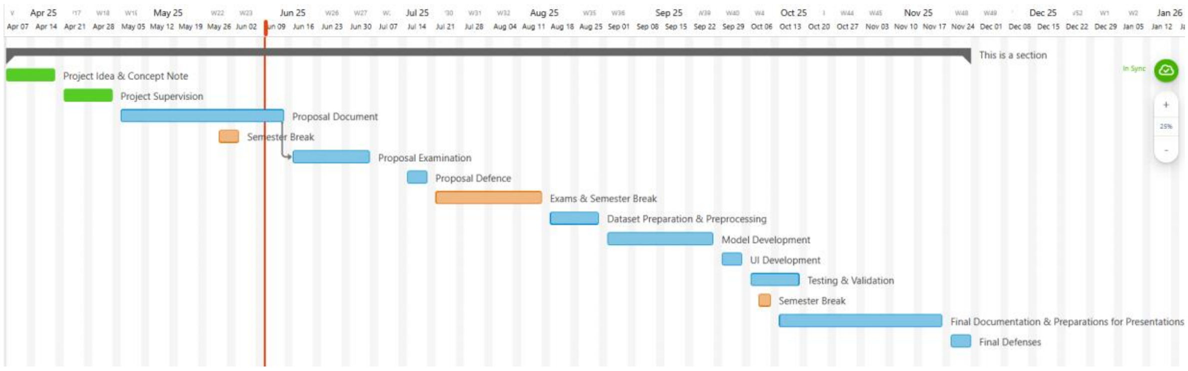
WHO. (2022, November 18). https://www.who.int/news/item/18-11-2022-who-highlights-oral-health-neglect-affecting-nearly-half-of-the-world-s-population

Zanini, L. G. K., Rubira-Bullen, I. R. F., & Nunes, F. de L. D. S. (2024). Enhancing dental caries classification in CBCT images by using image processing and self-supervised learning. *Computers in Biology and Medicine*, *183*, 109221. https://doi.org/10.1016/j.compbiomed.2024.109221

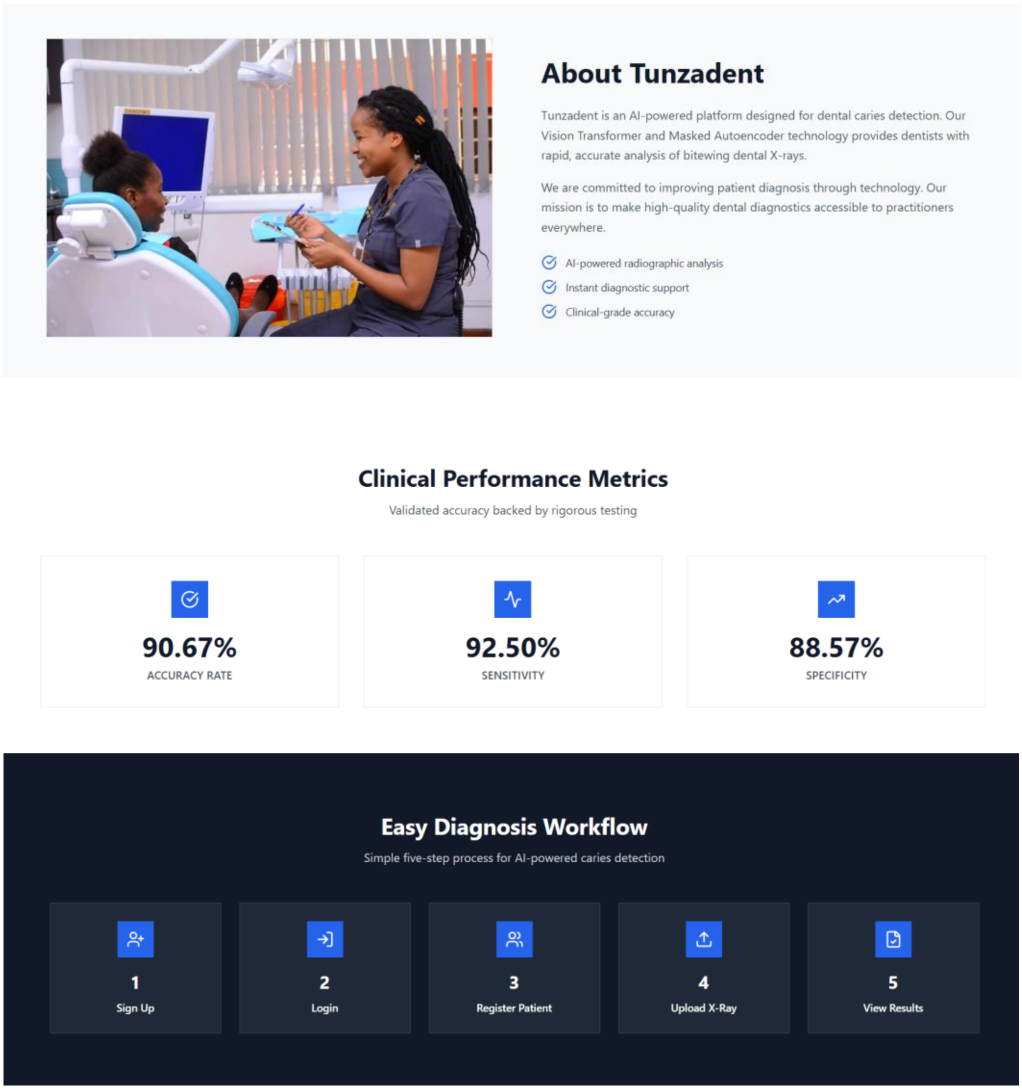
Zhou, X., Yu, G., Yin, Q., Yang, J., Sun, J., Lv, S., & Shi, Q. (2023). Tooth Type Enhanced Transformer for Children Caries Diagnosis on Dental Panoramic Radiographs. *Diagnostics*, *13*(4), 689. https://doi.org/10.3390/diagnostics13040689

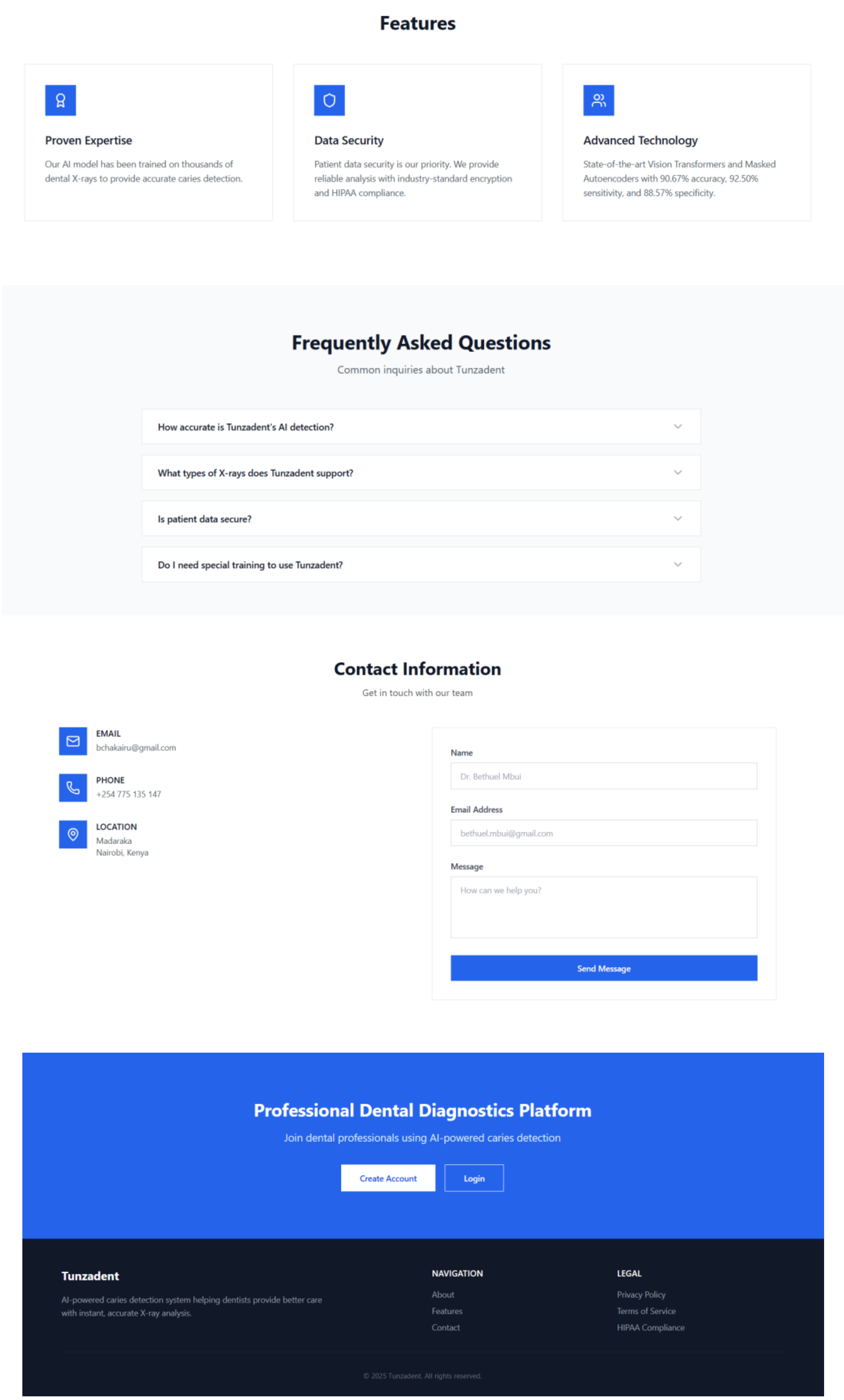
Appendix

Appendix 1: Gantt Chart

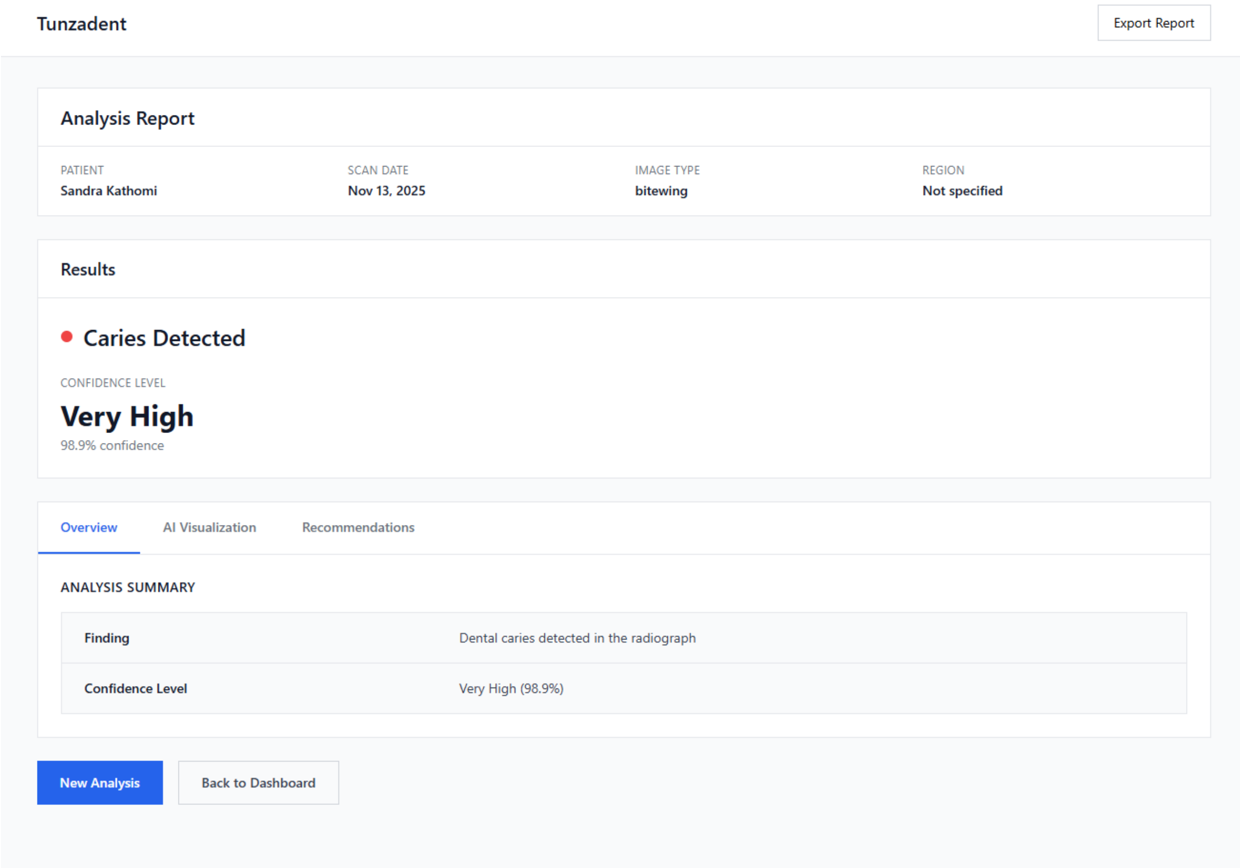


Appendix 2: Landing Page Wireframes

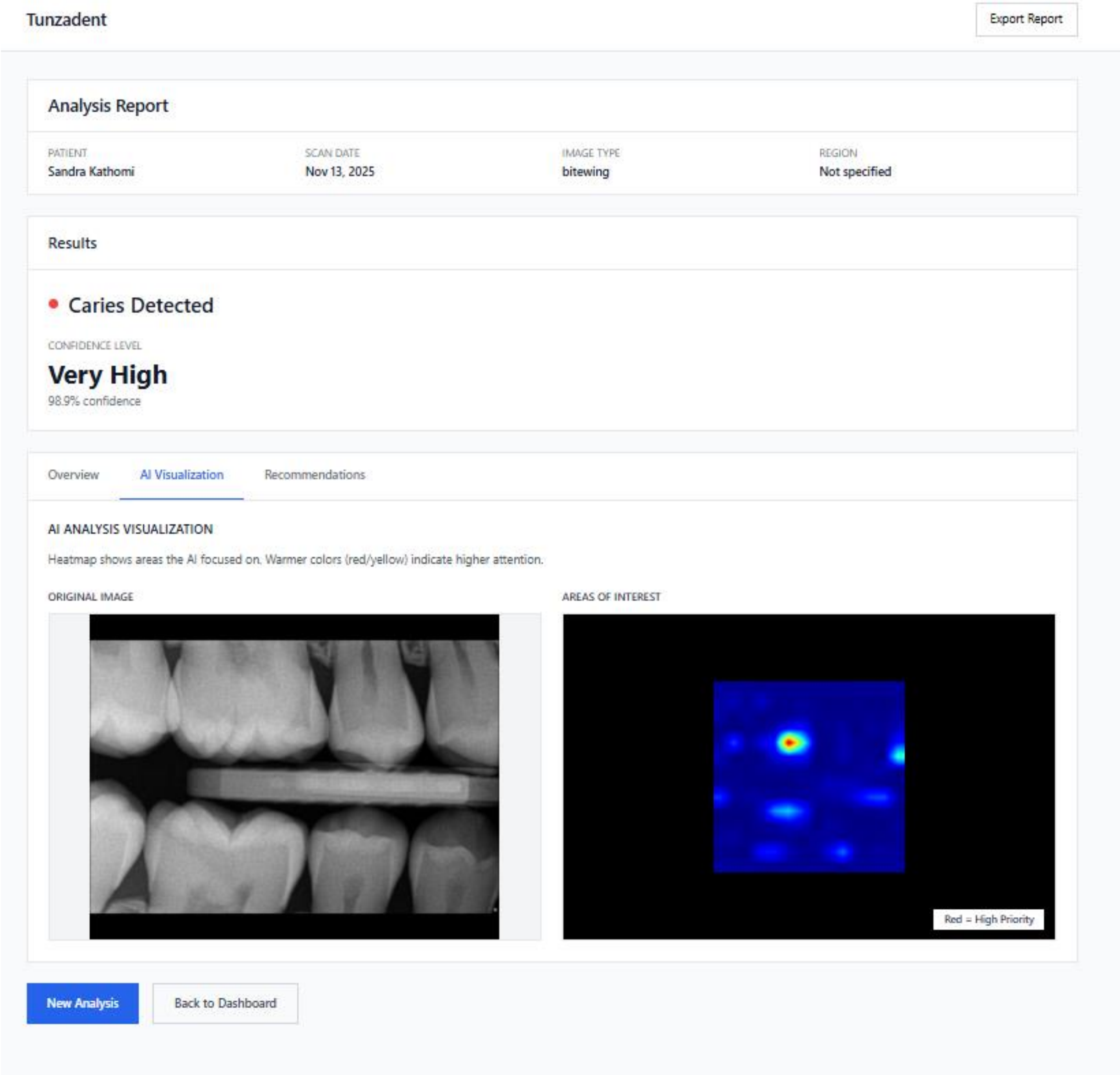



Features
Proven Expertise
Our AI model has been trained on thousands of dental X-rays to provide accurate caries detection.
Data Security
Patient data security is our priority. We provide reliable analysis with industry-standard encryption and HIPAA compliance.
Advanced Technology
State-of-the-art Vision Transformers and Masked Autoencoders with 90.67% accuracy, 92.50% sensitivity, and 88.57% specificity.
Frequently Asked Questions
Common inquiries about Tunzadent
How accurate is Tunzadent's AI detection?
What types of X-rays does Tunzadent support?
Is patient data secure?
Do I need special training to use Tunzadent?
Contact Information
Get in touch with our team
EMAIL
bchakairu@gmail.com
PHONE
+254 775 135 147
LOCATION
Madaraka
Nairobi, Kenya
Name
Dr. Bethuel Mbui
Email Address
bethuel.mbui@gmail.com
Message
How can we help you?
Send Message
Professional Dental Diagnostics Platform
Join dental professionals using AI-powered caries detection
Create Account
Login
Tunzadent
AI-powered caries detection system helping dentists provide better care with instant, accurate X-ray analysis.
NAVIGATION
About
Features
Contact
LEGAL
Privacy Policy
Terms of Service
HIPAA Compliance
© 2025 Tunzadent. All rights reserved.

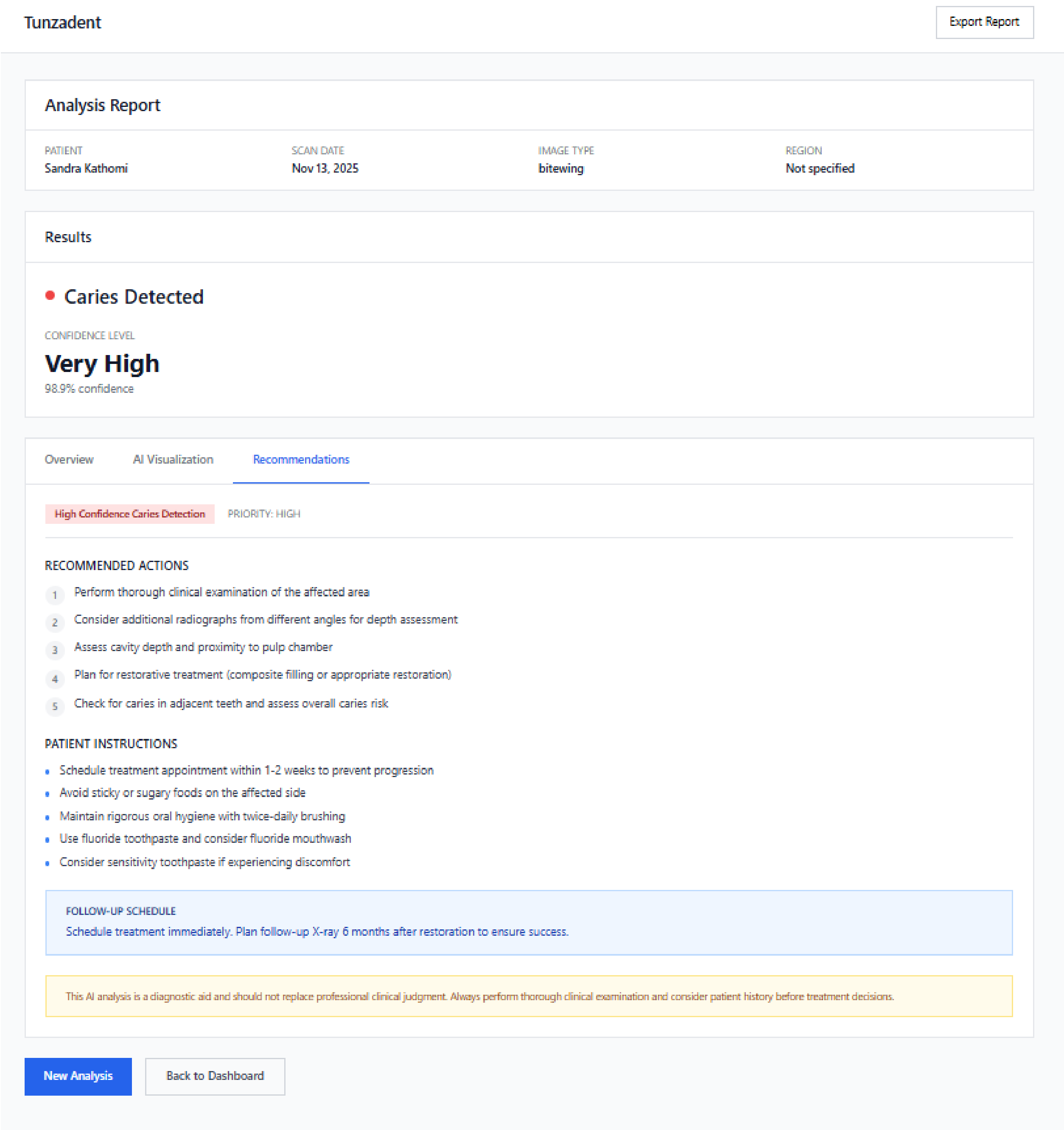
Appendix 3: Scan Report Wireframe - Overview



Appendix 4: Scan Report Wireframe - AI Visualization



Appendix 5: Scan Report Wireframe - Recommendations



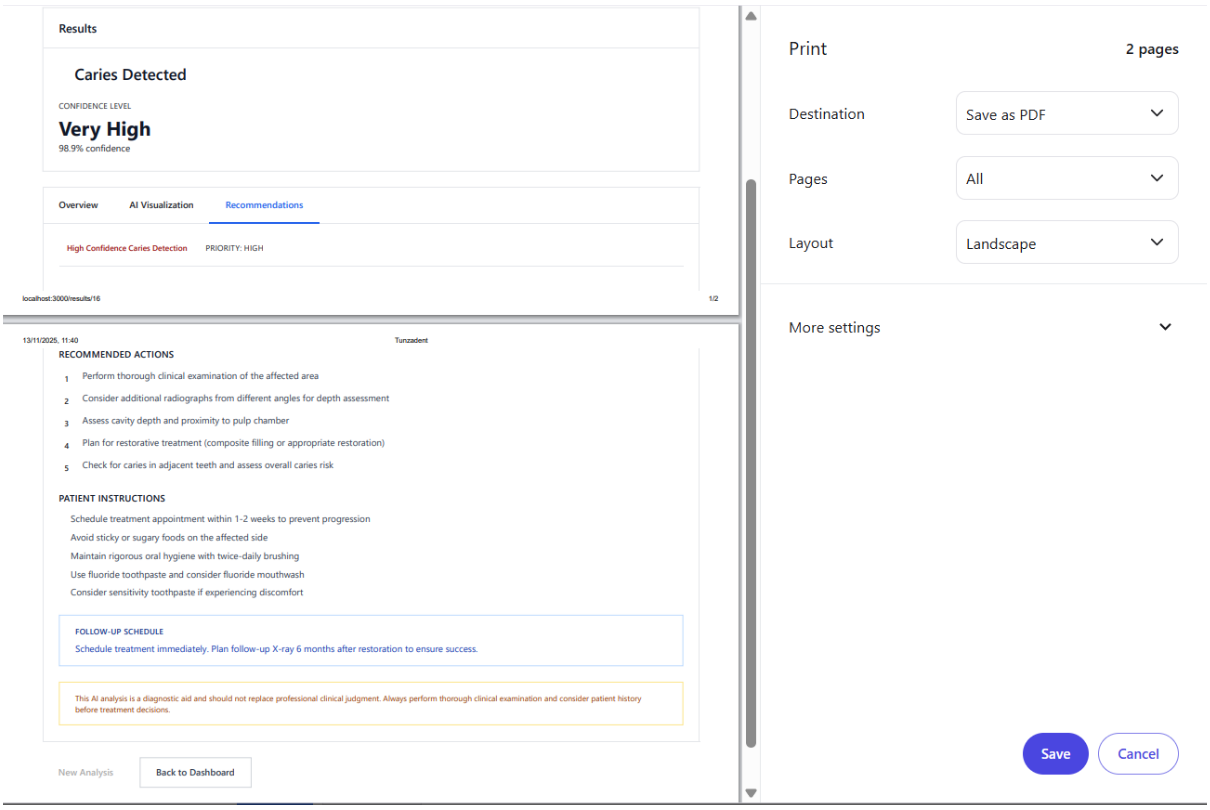
Appendix 6: Scan Report Recommendations Generation Logic

```python
def generate_recommendations(prediction_data):
    """
    Generate clinical recommendations based on prediction results

    Args:
        prediction_data: Dict with 'has_caries' and 'confidence_score'

    Returns:
        dict: Structured recommendations for clinician and patient
    """
    has_caries = prediction_data['has_caries']
    confidence = prediction_data['confidence_score']

    recommendations = {
        'severity': None,
        'clinical_actions': [],
        'patient_advice': [],
        'follow_up': None,
        'urgency_level': 'low',
        'disclaimer': (
            'This AI analysis is a diagnostic aid and should not replace professional clinical judgment. '
            'Always perform thorough clinical examination and consider patient history before treatment decisions.'
        )
    }

    if has_caries:
        if confidence >= 0.90:
            recommendations.update({
                'severity': 'High Confidence Caries Detection',
                'urgency_level': 'high',
                'clinical_actions': [
                    'Perform thorough clinical examination of the affected area',
                    'Consider additional radiographs from different angles for depth assessment',
                    'Assess cavity depth and proximity to pulp chamber',
                    'Plan for restorative treatment (composite filling or appropriate restoration)',
                    'Check for caries in adjacent teeth and assess overall caries risk'
                ],
                'patient_advice': [
                    'Schedule treatment appointment within 1-2 weeks to prevent progression',
                    'Avoid sticky or sugary foods on the affected side',
                    'Maintain rigorous oral hygiene with twice-daily brushing',
                    'Use fluoride toothpaste and consider fluoride mouthwash',
                    'Consider sensitivity toothpaste if experiencing discomfort'
                ],
                'follow_up': 'Schedule treatment immediately. Plan follow-up X-ray 6 months after restoration to ensure success.'
            })
```

```python
        elif confidence >= 0.70:
            recommendations.update({
                'severity': 'Moderate Confidence Caries Detection',
                'urgency_level': 'medium',
                'clinical_actions': [
                    'Perform detailed visual and tactile examination',
                    'Consider additional diagnostic tests (transillumination, laser fluorescence)',
                    'Monitor closely if early-stage caries without cavitation',
                    'Assess patient caries risk factors (diet, oral hygiene, fluoride exposure)',
                    'Consider preventive measures versus immediate intervention'
                ],
                'patient_advice': [
                    'Schedule appointment within 2-4 weeks for thorough examination',
                    'Increase brushing frequency to twice daily with proper technique',
                    'Use fluoride mouthwash daily',
                    'Reduce frequency of sugar and acidic food/drink consumption',
                    'Consider dental sealants for at-risk teeth'
                ],
                'follow_up': 'Re-evaluate in 3-6 months with follow-up X-ray if monitoring approach is chosen.'
            })
        else:
            recommendations.update({
                'severity': 'Possible Early-Stage Caries',
                'urgency_level': 'low',
                'clinical_actions': [
                    'Perform careful clinical examination for early signs',
                    'Look for white spot lesions, surface roughness, or staining',
                    'Consider remineralization therapy with high-fluoride products',
                    'Assess patient oral hygiene practices and dietary habits',
                    'May monitor before intervention if very early stage'
                ],
                'patient_advice': [
                    'Enhance oral hygiene routine with proper brushing technique',
                    'Use high-fluoride toothpaste (1450ppm or prescription strength)',
                    'Increase flossing frequency to daily',
                    'Reduce acidic and sugary food/drink consumption between meals',
                    'Consider calcium and phosphate supplements for remineralization'
                ],
                'follow_up': 'Monitor with follow-up X-ray in 6-12 months. Focus on prevention and remineralization.'
            })
```

```python
    else:
        if confidence >= 0.85:
            recommendations.update({
                'severity': 'Healthy - No Caries Detected',
                'urgency_level': 'low',
                'clinical_actions': [
                    'Confirm with visual examination during routine check-up',
                    'Continue routine preventive care and monitoring',
                    'Reinforce good oral hygiene practices',
                    'Schedule regular check-ups as per standard protocol'
                ],
                'patient_advice': [
                    'Maintain current oral hygiene routine',
                    'Continue brushing twice daily for 2 minutes',
                    'Floss daily to prevent interproximal caries',
                    'Attend regular dental check-ups every 6 months',
                    'Continue balanced diet with limited sugar intake'
                ],
                'follow_up': 'Routine check-up and X-ray as per standard recall interval (typically 6-12 months).'
            })
        else:
            recommendations.update({
                'severity': 'Uncertain - Further Examination Recommended',
                'urgency_level': 'medium',
                'clinical_actions': [
                    'Perform thorough clinical examination',
                    'Consider retaking X-ray if image quality is suboptimal',
                    'Check for borderline lesions or incipient caries',
                    'Assess overall caries risk and preventive needs'
                ],
                'patient_advice': [
                    'Schedule follow-up examination within 3-4 months',
                    'Maintain preventive care routine',
                    'Monitor for any tooth sensitivity or discomfort',
                    'Report any changes in symptoms promptly'
                ],
                'follow_up': 'Clinical follow-up in 3-4 months with repeat radiograph if indicated.'
            })

    return recommendations
```

Appendix 7: Scan Report Export - PDF



Appendix 8: Scan Report Export - CSV