SMILE: Smart Machine learning Imaging for Lesion Evaluation

Misbahus Abrar
Student# 1010329716
misbahus.abrar@mail.utoronto.ca

Ali Alyaseri
Student# 1006932492
ali.alyaseri@mail.utoronto.ca

Faraz Imran
Student# 1007996028
faraz.imran@mail.utoronto.ca

Haider Abbas
Student# 1006910988
haiderw.abbas@mail.utoronto.ca

Abstract

The SMILE project aims to address the challenges of detecting dental caries on panoramic radiographs by developing a machine learning-based dental assistant.
—-Total Pages: 6

1 Introduction

Cavities are one of the most common chronic diseases in the world, affecting an estimated 2.3 billion people with permanent teeth (GBD 2019 Dental Disorders Collaborators, 2020). If left untreated, they can lead to pain, tooth loss, and expensive dental procedures, stressing individuals and healthcare systems. Panoramic dental X-rays are often used to check for cavities during routine visits. However, reading these wide images can be time-consuming and subjective. Image quality varies from clinic to clinic, and detection accuracy may drop when dentists are busy or less experienced (Pasa et al., 2019). Early signs of decay are often faint and easy to miss, which can result in delayed or incorrect diagnoses. Our project, Smart Machine-learning Imaging for Lesion Evaluation (SMILE), is a dental assistant powered by machine learning. It is designed to automatically detect cavities in panoramic dental X-rays and localize the affected tooth by drawing a bounding box around it. The goal is to assist dentists by pointing out possible problem areas, making diagnoses more efficient. This system can support dentists during exams, act as a second opinion, help train dental students, and provide a more consistent level of care across different clinics. Deep learning, using convolutional neural networks (CNNs), is well-suited for this task. These networks can learn key features directly from image data, outperform traditional systems that use hand-crafted rules, and detect and locate issues in one pass. Unlike earlier systems that only detect whether a cavity is present (Moutselos et al., 2023), SMILE shows exactly where the cavity is located. By combining accuracy with practical use, SMILE could improve early diagnosis, reduce chair time, and make routine dental care more efficient. It also has potential applications in large-scale dental screening programs, especially in areas with limited access to dental specialists.

2 Visualization of SMILE

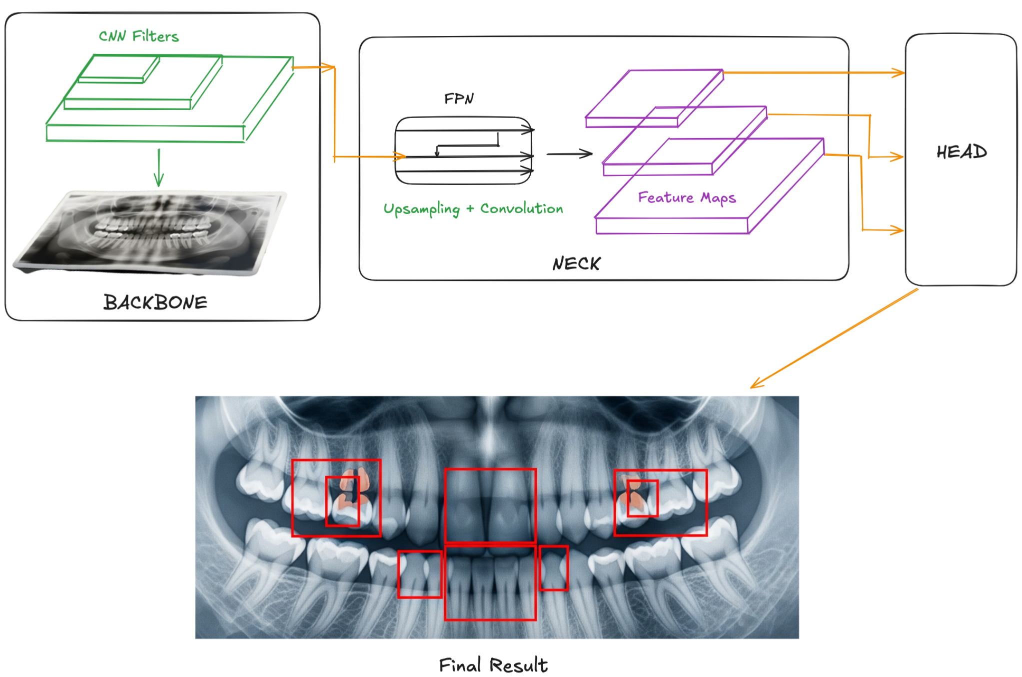

Figure 1: Architectural diagram of the SMILE Model

3 Background and Related Work

Deep learning has been widely used in medical image analysis tasks involving X-rays. A recent review of chest X-ray research showed that convolutional neural networks (CNNs) can perform at a similar level to radiologists for detecting conditions like pneumonia and COVID-19 (Meedeniya et al., 2022). These results highlight how machine learning can support medical professionals with accurate and fast image interpretation.

In another study, a deep learning model was used in an emergency department to help identify serious chest conditions on X-rays (Hwang et al., 2019). The model improved prioritization decisions and made it easier for clinicians to detect urgent cases. This shows that AI tools can be useful in real-time clinical settings, not just in controlled research environments.

In dentistry, CNNs have also been applied to panoramic X-rays. One study developed a model that automatically detects and labels each tooth, helping organize the image for further analysis, such as cavity detection (Tuzoff et al., 2019).

Another study trained an object-detection model on dental X-rays to identify different conditions, showing how deep learning can work well even with complex dental images (Almalki et al., 2022).

PaXNet applied deep learning to detect caries on panoramic X-rays using an ensemble of pretrained networks and capsule classifiers. In testing, it achieved 86% accuracy, but it only predicts whether a cavity exists, not where it is or on which tooth (Haghanifar et al., 2020). This reveals a gap: PaXNet lacks localization and tooth-level labelling, capabilities that would be valuable for clinical interpretation. These examples show that deep learning is a proven approach for analyzing medical and dental X-rays, and they support our decision to use it for detecting and labelling cavities in panoramic dental images.

4 Design Considerations

4.1 Architecture

Detecting and localizing dental cavities within panoramic OPG X-ray images is an object detection problem, for which we will employ a YOLO (You Only Look Once) model, based on the YOLOv8 architecture. We selected YOLO because it is known for its high speed and accuracy, making it well-suited for medical imaging analysis.

Our model architecture will consist of three main components as outlined in Figure 1. The backbone is a deep Convolutional Neural Network (CNN) responsible for extracting key visual features from the input X-ray image like edges, textures, and complex patterns relevant to dental structures. We plan to use a backbone pre-trained on a large-scale image dataset. This approach will allow our model to use generalized feature extraction capabilities which we will then fine-tune for the specific task of identifying dental anomalies. Secondly, the neck's purpose is to aggregate and refine the feature maps generated by the backbone. Lastly, the detection head consumes the feature maps output by the neck and performs the actual detection. This involves predicting the coordinates of the bounding boxes for potential cavities.

4.2 Data Processing

We will aggregate and standardize data from three distinct public sources. Our goal is to create a unified dataset with consistent image dimensions and a single, standardized annotation format (YOLO .txt), suitable for training our YOLOv8 model. The sources are:

- Dental Radiography Analysis and Diagnosis Dataset from Kaggle.
- Dental OPG XRAY Dataset from Mendeley Data.
- DC1000 Dataset compiled by a research team.

Our data processing pipeline will involve the following steps for each dataset:

KAGGLE DATASET PROCESSING:

This dataset contains 1,271 images (512x256 pixels) with a single annotations.csv file. We will first parse the CSV file and filter it to retain only the rows where the class is 'cavities'. For each relevant image, we will convert the bounding box coordinates from the provided Pascal VOC format (xmin, ymin, xmax, ymax) to the normalized YOLO format (class_id, x_center, y_center, width, height). Since we are only detecting one class, the class_id will be 0. The conversion will be done using the image dimensions. A separate .txt annotation file will be generated for each image containing cavity annotations.

MENDELEY DATASET PROCESSING:

This dataset contains 232 enhanced radiographs (640x340 pixels), each with an accompanying .txt annotation file in YOLO format. We will identify the class number corresponding to "caries" (cavities). We will then process each .txt file, keeping only the lines that match the caries class ID. We will standardize the class ID to 0 to maintain consistency with the other datasets. The bounding box coordinates are already in the required YOLO format and do not need conversion.

DC1000 DATASET PROCESSING:

This dataset includes 597 high-resolution images (2943x1435 pixels) where cavities are identified by segmentation masks (separate black and white images). For each mask, we will use a computer vision library to find the contours of the white regions (which represent cavities). For each contour detected, we will calculate the tightest possible bounding box (x, y, width, height). These bounding box coordinates will then be normalized and converted to the YOLO format, similar to the Kaggle dataset. A .txt annotation file will be generated for each corresponding X-ray image. All images from the three sources will be resized to a standard input size required by the YOLOv8 model (e.g., 640x640 pixels), and the labels will be scaled accordingly. We will use padding to maintain the original aspect ratio of each image to prevent distortion. This process will yield a large, clean, and standardized dataset which we will use to train our cavity detection model.

4.3 BASELINE MODEL

To measure how much our deep learning model improves performance, we will compare it to a traditional baseline that does not use deep learning. This baseline will use a combination of Histogram of Oriented Gradients (HOG) for feature extraction and a Support Vector Machine (SVM) for classification. This method provides a simple benchmark by decoupling feature detection from decision-making. We will start by preparing a set of image patches from our training data. Patches that contain cavities will be cropped from the areas marked in the annotations. To balance the dataset, we will also create a larger set of patches from areas without cavities. For each patch, we will calculate its HOG features, which capture edge and shape information by analyzing the direction and strength of gradients in the image. These features are then transformed into a fixed-length vector that the classifier can use. Next, we will train a linear SVM to separate cavity patches from non-cavity ones. Once trained, we will use a sliding window to move across new images. The model will decide at each step whether the patch under the window contains a cavity. Because multiple overlapping windows might detect the same cavity, we will use Non-Maximum Suppression (NMS) to combine these into a single final prediction box.

4.4 ETHICAL CONSIDERATIONS

While developing a tool to assist in dental diagnostics holds great promise, it is crucial to consider the ethical implications and limitations of our system. The most significant risk is a diagnostic error. A false negative (the model failing to detect a real cavity) could lead to delayed treatment and worsening of the patient's condition. A false positive (the model incorrectly identifying a cavity) could lead to unnecessary patient anxiety. It must be clear that SMILE is an assistive tool, and the final diagnostic responsibility remains with the qualified dental professional.

Secondly, our model is trained on publicly available datasets, which represent a very small sample size of the general population. This limited data may not capture the full diversity in dental morphology across different ages, ethnicities, and sexes, potentially leading to performance issues for underrepresented patient groups and contributing to healthcare disparities. Moreover, there is a risk that clinicians may become over-reliant on SMILE, trusting its output without applying their own expert judgment. This could lead to a decrease in vigilance and the potential to miss subtle signs of disease that the model is not trained to detect. Lastly, although we are using anonymized public data for this project, a real-world deployment of SMILE would handle sensitive patient health information. The model would need to be designed in strict compliance with privacy regulations like HIPAA and PIPEDA to protect patient confidentiality.

5 Project Management

5.1 Project Plan

Our team communicates through a dedicated Discord server, with tasks planned and assigned weekly on Saturdays using a group polling system. To ensure fairness, we vote simultaneously and rotate the order of first picks each week. Tasks are chosen based on individual interests while balancing the overall workload, and the chance to choose tasks first rotates among all members.

We provide status updates every three days and hold frequent check-ins, which also serve as a contingency for smooth handovers if a member is unavailable. All work is version-controlled with Git, and conflicts are resolved collaboratively to avoid overwriting contributions. The project is divided into clear stages with assigned members and deadlines, as shown in Table 1.

Task	Assignee(s)	Description	Projected Time (hrs)	Internal Deadline	Course Deadline
Stage 0 – Team Setup					
Team formation & role sheet	All	Agree on roles and weekly meeting slot	8	May 16	
Stage 1 – Idea & Proposal					
Research survey	Faraz, Haider	Summarise related work and candidate datasets	6	May 29	
Problem statement & branding	Ali	Craft SMILE problem statement and acronym	2	Jun 1	
Architecture design	Abrar	Sketch full pipeline (YOLO v8 + baseline)	3	Jun 5	
Formal proposal	Ali, Haider	Submit proposal	12	Jun 13	Project Proposal
Stage 2 – Data Processing & Baseline					
Annotation converter	Haider	Script to unify label formats	8	Jun 24	
HOG + SVM baseline model	Faraz	Baseline detector notebook with metrics	6	Jun 27	
Dataset split & resize pipeline	Abrar	Automated train/val/test organizer	4	Jun 30	
Stage 3 – Minimum Viable Product (MVP)					
YOLO v8 fine-tuning	Ali	Train the first working detector	10	Jul 6	
Hyper-parameter sweep	Faraz	Grid search for the best YOLO settings	6	Jul 9	
Baseline vs YOLO comparison	Haider	Report PR curves and timing	5	Jul 11	Progress Report
Stage 4 – Advanced Features					
Tooth-level overlay & viz	Abrar	Visualize detections per tooth	5	Jul 22	
Federated-learning extension	Ali	Prototype privacy-preserving training	6	Aug 1	
Cross-clinic generalisation test	Haider, Faraz	Evaluate the model on unseen clinic data	8	Aug 5	
Stage 5 – Presentation & Publication					
Final report	Ali, Haider	Write the final paper	16	Aug 12	
Demo video & slide deck	Faraz, Abrar	Prepare project showcase	12	Aug 13	
Dry-run rehearsal	All	Full practice run-through	6	Aug 14	
Final presentation & code release	All	Present and publish the GitHub Repository		Aug 15	Final Deliverable

Table 1: Architectural diagram of the SMILE Model

5.2 Risk Register

Risk	**Likelihood**	**Impact**	**Mitigation Plan**
High Rate of False Negatives (Missed Cavities)	Medium	High	A model that fails to detect existing cavities has significant clinical risk. To mitigate this, we will apply extensive data augmentation (rotation, brightness shifts) to increase data volume and variety, and consider adjusting classification thresholds.
Data Quality Issues and Labeling Errors	High	High	Inconsistent or incorrect annotations from our combined data sources can severely degrade model performance. Our primary mitigation is a thorough exploratory data analysis, including manual inspection of a random subset of annotations, to identify and correct errors before training.
Model Training Delays & Compute Limits	Medium	Medium	We will use Google Colab for GPU access, but training can still be slow. We will start with smaller model versions for rapid prototyping, implement model checkpointing to save progress, and optimize batch sizes to maximize the use of available compute resources.
Uneven Team Contribution	Low	High	A team member dropping the course or failing to contribute could jeopardize project timelines. Our modular task assignments and detailed documentation on GitHub are designed to allow for a quick handover. Regular weekly meetings will ensure accountability and allow for early intervention and task reallocation if needed.

Table 2: An overview of the risks anticipated for this project.

5.3 Github Link

https://github.com/Ali-Alyaseri/APS360-Dental-Divot-Detection

References

Yasser E. Almalki et al. Deep learning models for classification of dental diseases using orthopantomography x-ray images. *Sensors*, 22(19):7262, 2022.

GBD 2019 Dental Disorders Collaborators. Global, regional, and national burden of oral conditions in 1990–2019: a systematic analysis. *The Lancet*, 20(1):1–15, 2020.

Atefeh Haghanifar, Haikal Ko, Muhammad Moidu I. Yousaf, Seong eun Kwon, and U. Rajendra Acharya. Paxnet: Dental caries detection in panoramic x-ray using ensemble transfer learning and capsule classifier. arXiv preprint arXiv:2012.00383, 2020.

Eui Jin Hwang, Sang Hyun Park, and Ki Hwan Jin. Development and validation of deep-learning-based automated detection algorithm for major thoracic diseases on chest radiographs. *JAMA Network Open*, 2(3):e191095, 2019.

Duleep Meedeniya, Maheshi Bandara, and Indika Perera. Chest x-ray analysis empowered with deep learning: A systematic review. *Applied Soft Computing*, 126:109383, 2022.

Rafail Moutselos et al. Deep learning and artificial intelligence in dental diagnostic imaging: a review. *Artificial Intelligence in Dentistry*, 1:41–54, 2023.

Florin Pasa, Vladimir Golkov, Yoshua Pfeiffer, and Daniel Cremers. Efficient deep network architectures for fast chest x-ray tuberculosis screening and visualization. *Scientific Reports*, 9(1), 2019.

Sergey Tuzoff et al. Tooth detection and numbering in panoramic radiographs using convolutional neural networks. *Dentomaxillofacial Radiology*, 48(4):20180078, 2019.