AI-Enabled Clinical Decision Support for Oral Disease Detection

A THESIS SUBMITTED TO AUCKLAND UNIVERSITY OF TECHNOLOGY

IN PARTIAL FULFILMENT OF THE REQUIREMENTS FOR THE DEGREE OF

MASTER OF COMPUTER AND INFORMATION SCIENCES

Supervisors

Dr. Samaneh Madanian

Dr. Abtin Ijadi Maghsoodi

February 2025

By

Priyadarshini Natarajan

School of Engineering, Computer and Mathematical Sciences

Abstract

Oral health is a critical component of global public health, with dental diseases such as caries, periodontal conditions, periapical lesions, and impacted teeth affecting millions of individuals annually. These conditions not only lead to pain, infection, and tooth loss but are also linked to systemic health issues, including cardiovascular disease, diabetes and respiratory infections. Given the significant burden of oral diseases, early and accurate diagnosis is essential to prevent complications and improve patient outcomes.

This study explores the feasibility of AI-based deep learning models, specifically YOLO architectures, for the automated detection and classification of dental diseases using panoramic radiographs. A comprehensive dataset containing 50 plus annotated dental disease categories was used to train and evaluate YOLOv5, YOLOv8, and YOLO-NAS models. To optimize model performance, a systematic data preprocessing pipeline incorporating Cropping and Region of Interest Selection, Conversion to Grayscale, Resolution Standardization, Dataset Splitting and Non-Enhanced Datasets was implemented. Additionally, data augmentation techniques were employed to enhance model generalization, including Geometric Transformations, Rotation, Flipping, Grid Masking and to further improve model efficiency. Performance metrics, including precision, recall, and accuracy, were used to evaluate model effectiveness. Among the tested models, YOLOv5 outperformed YOLOv8 and YOLO-NAS, achieving a precision of 95.7%, recall of 96%, and accuracy of 96.5%, making it the most effective model for dental disease detection. The proposed AI-powered dental diagnostic framework demonstrated high accuracy, real-time detection capabilities, and adaptability across multiple disease categories, positioning it as a viable tool for clinical applications.

Beyond clinical practice, AI-driven dental diagnostics offer significant public health benefits, particularly in rural and underserved areas where access to specialized dental care is limited. The ability to automate radiographic analysis facilitates large-scale screening programs, enabling early disease detection, cost-effective intervention strategies, and improved healthcare accessibility. Additionally, AI-based diagnostics contribute to reducing consultation times, and minimizing reliance on specialist interpretations. Given the strong association between oral health and systemic diseases, early AI-powered diagnosis can also help mitigate broader health risks such as diabetes complications and cardiovascular conditions. By integrating AI into routine dental practice and public health initiatives, this study underscores

the transformative potential of deep learning in enhancing diagnostic precision and improving overall healthcare outcomes. These findings establish hyperparameter tuned YOLOv5 as the most effective model for AI-driven dental disease detection. This research paves the way for next-generation AI-assisted dental diagnostic platforms, fostering early intervention, equitable healthcare access, and improved global oral health outcomes.

Table of Contents

Abstract 2
Table of Contents 4
Attestation of Authorship 6
List of Figures 7
List of Tables 8
List of Abbreviations 9
Acknowledgements 11
Chapter 1 – Introduction 13
1.1. Background 13
1.2. Motivation 17
1.3. Research Aims and Objective 17
1.4. Research Questions 18
1.5. Significance of the Research 18
1.6. Research Contributions 19
1.7. Thesis Structure 20
Chapter 2 - Literature Review 22
2.1. AI-Driven Innovations and DL Advancements in Dental Radiograph Analysis 23
2.2. Diagnostic Accuracy in Dentistry with Machine Learning Techniques 26
2.3. Literature Review Method 30
2.4. The Evolution of YOLO: A Timeline of Innovation 31
2.5. YOLO methods 34
2.6. Research Gap 40
Chapter 3 - Methodology 42
3.1. Research Methodology 42
3.2. Dataset 44
3.3. Data Annotation for Dental Radiographs 45
3.4. Experimental Process 48
3.5. Data Preprocessing 51
3.5.1. Cropping and Region of Interest (ROI) Selection 51
3.5.2. Conversion to Grayscale 52
3.5.3. Resolution Standardization 53
3.5.4. Dataset Splitting 54
3.5.5. Non-Enhanced Datasets 56

3.6. Data Augmentation 57
3.6.1. Geometric Transformations 57
3.6.2. Grid Masking 60
3.7. YOLO Models on Dental Radiographs 61
3.7.1. Importance of YOLO Models in Dental and Medical Imaging 62
3.8 Training and Evaluation of YOLO Models on Dental Radiographs 65
3.8.1. Training Parameters 66
3.8.2. Evaluation of Parameter Tuning 68
3.9. Model Performance Metrics 69
3.9.1. Precision and Recall 69
3.9.2. Precision-Recall (PR) Curve 70
3.9.3. mAP 71
3.10. Importance of Specificity, Sensitivity, and Prediction Categories 73
Chapter 4 - Results 76
4.1. Precision Results of YOLO Models 76
4.2. Accuracy Results of YOLO Models 78
4.3. Prediction results 82
4.4. Training loss 85
4.5. Precision-Recall Curve 86
4.6. Superior Performance of YOLOv5 (HPT-Tuned) Over YOLOv8 and YOLO-NAS 87
4.7. Confusion Matrix 89
Chapter 5 – Discussion 91
5.1. Feasibility of AI-Based Deep Learning for Dental Disease Diagnosis 91
5.2. Comparative Accuracy of YOLO Models in Dental Disease Detection 93
5.3. Impact of Hyperparameter Tuning on AI Model Performance 94
Chapter 6 - Conclusion 96
6.1. Clinical Implications 97
6.2. Limitations and Future Work 98
References 101
Appendix Table 1 113
Appendix Table 2 121

Attestation of Authorship

I hereby declare that this submission is my own work and that, to the best of my knowledge and belief, it contains no material previously published or written by another person (except where explicitly defined in the acknowledgements), nor used artificial intelligence tools or generative artificial intelligence tools (unless it is clearly stated, and referenced, along with the purpose of use), nor material which to a substantial extent has been submitted for the award of any other degree or diploma of a university or other institution of higher learning.

Signature of student

List of Figures

Figure 1. Ranking of most prevalent conditions per WB country income level [1] 14
Figure 2. Comparison of estimated global case numbers for selected NCDs [1] 14
Figure 3. A network visualization illustrating the interconnected themes in research on AI applications in dental imaging. 22
Figure 4. A word cloud generated from the literature review on AI in dental radiograph analysis. 30
Figure 5. PRISMA Chart for identifying, screening, and excluding studies related to dental diagnosis and imaging from various databases 31
Figure 6. YOLO Evolution Timeline: A Visual Representation of the Development of YOLO Object Detection Models from YOLOv1 (2015) to YOLOv9 (2024)[46] 33
Figure 7. A word cloud highlighting key terms and concepts related to YOLO-based methods in dental imaging studies. 34
Figure 8. Design Science Research Methodology [49] 43
Figure 9. Number of Annotations for Various Dental Conditions 45
Figure 10. Collection of annotated dental radiographs (X-rays) with bounding boxes 46
Figure 11. A dental panoramic X-ray with annotated bounding boxes 46
Figure 12. A dental panoramic X-ray with annotated bounding boxes highlighting various dental structures and conditions 47
Figure 13. Research Pipeline 50
Figure 14. Dental OPG Dataset structure 54
Figure 15. Stacked bar chart showing the distribution of annotations for various dental conditions and structures across Training, Validation, and Testing datasets. The chart categorizes the annotations into three sets: Training (70%) in purple, Validation (15%) 55
Figure 16. Dental Image enhancement techniques 56
Figure 17. Augmented panoramic dental radiograph showcasing grid masking and bounding box annotations for impacted teeth 59
Figure 18. Augmented panoramic dental radiograph with rotation, grid masking and highlighted bounding boxes identifying dental anomalies 59
Figure 19. Augmented panoramic dental radiograph with rotation, minimal grid masking and highlighted bounding boxes identifying dental anomalies 61
Figure 20. Augmented panoramic dental radiograph with rotation, multiple grid masking and highlighted bounding boxes identifying dental anomalies 61
Figure 21. Comparative performance of the mAP of different YOLO models. 62
Figure 22. YOLOv5 Architecture [149]. 63
Figure 23. YOLOv8 Architecture 64
Figure 24. YOLO-NAS Architecture [78]. 65
Figure 25. Training set screenshot. 66
Figure 26. Precision-Recall (PR) Curve for the third molar angle detection in dental panoramic X-rays context. (a) Results for the third molar angle detection at different threshold values. (b) A sample of the Precision-Recall (PR) Curve for one category (vertical) [80]. 71
Figure 27. Confusion Matrix Illustration Showing Relationships Between True Positive, False Positive, False Negative, and True Negative Outcomes, with Derived Metrics (Sensitivity, Specificity, PPV, and NPV) [107] 74
Figure 28. YOLO versions performance comparison 76
Figure 29. YOLO versions performance comparison 79
Figure 30. YOLO-NAS HPT Tuned Model Predictions on Dental Radiograph 82

Figure 31. YOLOv8 HPT Tuned Model Predictions on Dental Radiograph 82
Figure 32. YOLOv5 (HPT Tuned) Model Predictions on Dental Radiograph 82
Figure 33. YOLOv5 Model Predictions on another Dental Radiograph 83
Figure 34. YOLOv5 Model Predictions on Dental Radiograph: Identifying Edentulous Ridges (Blue), Dental Posts and Crowns (Green), Dental Fillings (Cyan), and RCT-Treated Teeth (Yellow) with Confidence Scores. 83
Figure 35. YOLOv5 Model Predictions on Dental Radiograph: Identifying Edentulous Ridges (Blue), Dental Posts and Crowns (Green), Dental Fillings (Cyan), and RCT-Treated Teeth (Yellow) with Confidence Scores 84
Figure 36. Training loss graph of the hyper-parameter tuned models 85
Figure 37. Precision - Recall graph of the hyperparameter tuned models 86
Figure 38. The above graphs illustrate the training and validation performance of the YOLOv5 model during the training process, focusing on key metrics and loss functions. 87
Figure 39. The above graphs illustrate the training and validation performance of the YOLOv5 model during the training process over 100 epochs, focusing on key metrics and loss functions. 88
Figure 40. Confusion matrix for YOLO V5 with hyperparameter tuning - represents the performance of a model tasked with detecting and classifying various dental conditions from radiographs. 90

List of Tables

Table 1. Summary of Features Extracted, Algorithms/Methods Employed, and Oral Diseases Detected in Studies Utilizing Alternative AI and ML Techniques for Dental Imaging 29
Table 2. Summary of Features Extracted, Algorithms Used, and Dental Diseases Detected in AI-Driven Dental Imaging Studies 39
Table 3. Precision results for various YOLO models (YOLOv5, YOLOv8, YOLO-NAS) when trained and evaluated both with and without Hyperparameter Tuning (HPT).....77
Table 4. Precision results discussion for various YOLO models.....77
Table 5. Accuracy results for various YOLO models (YOLOv5, YOLOv8, YOLO-NAS) when trained and evaluated both with and without HPT. 79
Table 6. Accuracy results discussion for various YOLO models.....79
Table 7. Scoping review extracted information. 113
Table 8. Dental annotation classes and their clinical features [50].....121

List of Abbreviations

Abbreviation	Definition
AI	Artificial Intelligence
YOLO	You Only Look Once
YOLO-NAS	YOLO - Neural Architecture Search
mAP	mean Average Precision
ML	Machine Learning
DL	Deep Learning
PPV	Positive Predictive Value
NPV	Negative Predictive Value
TPR	True Positive Rate
PR	Precision-Recall
TP	True Positives
TN	True Negatives
FP	False Positives
FN	False Negatives
MAE	Mean Absolute Error
AUC-ROC	Area under the Receiver Operating Characteristic Curve
CNN	Convolutional Neural Networks
R-CNN	Region-based Convolutional Neural Networks
HPT	Hyperparameter Tuning
IoU	Intersection over Union
SPPF	Spatial Pyramid Pooling - Fast
CSPNet	Cross Stage Partial Network
ResNet	Residual Neural Network
SSD	Single Shot Multi-Box Detector
VGG16	Visual Geometry Group 16
SPP	Spatial Pyramid Pooling
PRISMA	Preferred Reporting Items for Systematic Reviews and Meta-Analyses
GPU	Graphics Processing Unit
CPU	Central Processing Unit
CUDA	Compute Unified Device Architecture

RDP	Remote Desktop Protocol
VS Code	Visual Studio Code
VRAM	Video Random Access Memory
SVM	Support Vector Machine
SVR	Support Vector Regression
JPEG	Joint Photographic Experts Group
CSP	Cross Stage Partial
C2f	Cross Stage Partial Bottleneck with 2 convolutions, with a 'f' indicating a faster implementation
AB	Alveolar Bone
MC	Mandibular Canal
FPN	Feature Pyramid Network
ICDAS	International Caries Detection and Assessment System
AP	Apical Periodontitis
CAD	Computer-Aided Diagnosis
AL	Apical Lesions
EL	Enamel Lesions
DL	Dentin Lesions
PR	Panoramic Radiographs
OMF	Oral and Maxillofacial
DA	Dental Age
RCT	Root Canal Treatment
OPG	Orthopantomograms
ROI	Region of Interest
CBCT	Cone Beam Computed Tomography

Acknowledgements

First and foremost, I want to extend my heartfelt gratitude to my amazing supervisor, Dr. Samaneh Madanian. Without her unwavering support and guidance, this thesis would have been nothing more than a far-off dream. Coming from a clinical background, diving into the world of coding and medical image analysis felt like jumping into uncharted waters. But her patience, encouragement, and ability to simplify even the most complex ideas gave me the confidence to take this on. The journey from having a mere idea back in 2020 to seeing it transformed into a completed thesis today has been surreal, and it's all thanks to her belief in me. I still remember when we first connected in February 2022—she not only gave me hope but also a clear path to follow, which made all the difference. Thank you for believing in me, even when I doubted myself.

Of course, none of this would have been possible without the unwavering support of my incredible parents, Natarajan Algaperumal and Vasanthi Natarajan. From the very beginning, they've stood by every decision I've made, cheering me on no matter what. Whether it was late-night brainstorming sessions or moments of self-doubt, their unwavering faith in me kept me going. And a huge shoutout to my sister, Nandha Devi Natarajan, who's always been my reality check. She's the one who knows exactly when to encourage me and when to call me out—and trust me, both were equally important in getting me here today.

I also owe a big thank you to my industry advisor, Dr. Abtin. Your insights and feedback at this journey were invaluable and I'm so grateful for your time and expertise and I would like to express my heartfelt gratitude to my university for the support provided from the very start of my course. The technical support, resources, and infrastructure made available to me have been invaluable in conducting my research.

And then there are my friends - Dr. Hari, Dr. Abhinaya, Dr. Raghul, Dr.Nilaa, Dr. Suzan, Shiny, Joel, Minh, Hamid, Pawan, Tanvee—the ones who've been there through all the ups and downs. You know who you are! Whether it was a much-needed pep talk, a quick late-night facetime, or just grabbing a coffee/ice-cream to take my mind off things, your support meant the world to me. I'm so lucky to have such an amazing crew who've had my back every step of the way.

Above all, I want to express my gratitude to God. Through every decision and every step of this process, I've felt a guiding hand. It's that quiet, unshakable faith that kept me grounded and gave me clarity even during the toughest moments. This thesis is not just the culmination of years of work; it's a testament to the power of perseverance and belief.

Looking back, it's hard to believe how far this journey has taken me. From scratching out initial ideas in 2020 to diving into unfamiliar territory with coding and medical image analysis, every step has been a learning experience. It's not just about the research; it's about the growth, the resilience, and the people who've made this journey worthwhile. I'm beyond thrilled to see this project come to life, and I can't wait to keep exploring this field.

Chapter 1 – Introduction

This chapter is organized into six sections. The first section provides an overview of the background of the study. The second section outlines the motivation driving this research. The third section highlights the benefits and significance of the study. The fourth section defines the aims and objectives of the research. The fifth section discusses the contributions made by this work, and the final section presents the overall structure of the thesis.

1.1. Background

Oral health remains a significant public health issue globally. Approximately 3.5 billion people worldwide are affected by oral diseases as shown in Figure 1 and Figure 2, with untreated dental caries (tooth decay) in permanent teeth being the most prevalent condition, impacting an estimated 2.5 billion individuals [1]. Additionally, more than 530 million children suffer from tooth decay in their primary (milk) teeth [1]. Gum disease (periodontal disease) is also one of the most common conditions globally, affecting up to 50% of the population. Severe gum (periodontal) disease, which may result in tooth loss, affects 10% of the global population [2]. In New Zealand, the Ministry of Health reports that approximately one-third of adults [3] have untreated dental caries, and many children experience poor oral health outcomes, particularly in underserved communities. In New Zealand [4], oral health disparities are evident across various demographics. According to the New Zealand Oral Health Survey, one in three adults had untreated coronal decay, and one in ten had root decay [5]. Additionally, nearly half of adults felt they currently needed dental treatment, with many avoiding dental care due to cost. Children, particularly in underserved communities, also experience significant oral health challenges. In 2019, 58.8% of five-year-olds seen by community oral health services had no history of dental caries, indicating that over 40% had experienced decay [3]. These statistics underscore the urgent need for improved diagnostic tools that can lead to better prevention and treatment strategies.

Higher rank — Lower rank

	Global	World Bank low income	World Bank lower-middle income	World Bank upper-middle income	World Bank high income
Oral diseases	1	1	1	1	1
Neurological disorders	2	5	2	3	2
Digestive diseases	3	7	4	2	6
Respiratory infections & TB	4	4	3	4	9
Skin diseases	5	3	5	6	5
Sense organ diseases	6	9	8	5	7
Musculoskeletal disorders	7	11	9	7	3
NTDs & malaria	8	2	7	12	19
HIV/AIDS & STIs	9	8	10	8	10
Nutritional deficiencies	10	6	6	13	15
Unintentional injury	11	13	13	9	4
Diabetes & CKD	12	14	12	10	8
Mental disorders	13	10	11	11	11
Cardiovascular diseases	14	16	14	14	13
Neoplasms	15	17	18	15	12
Chronic respiratory	16	15	15	17	14
Self-harm & violence	17	12	17	16	18
Transport injuries	18	19	16	18	17
Substance use	19	22	22	19	16
Maternal & neonatal	20	18	19	20	20
Enteric infections	21	21	21	21	21
Other infectious	22	20	20	22	22

Figure 1. Ranking of most prevalent conditions per WB country income level [1]

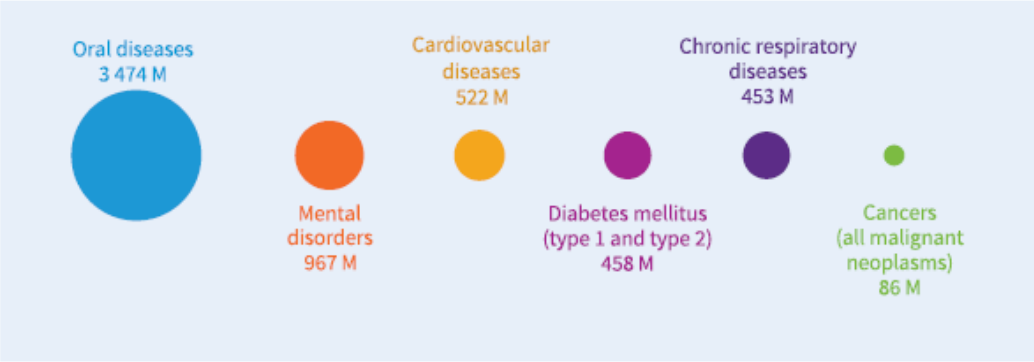


Figure 2. Comparison of estimated global case numbers for selected NCDs [1]

The World Health Organization (WHO) has recognized the potential of digital technologies in enhancing oral health outcomes. WHO's release on "Mobile technologies for oral health: an

implementation guide [6]" emphasizes the role of digital tools in promoting oral health, training health workers, detecting oral health conditions, collecting epidemiological data, and monitoring the quality of patient care. These technologies can contribute to better oral health for all, especially in resource-limited settings [6].

Artificial Intelligence (AI) is at the forefront of these technological advancements. AI-powered technologies can identify high-risk populations to prevent the initiation and progression of dental diseases, such as periodontal disease and dental caries[7]. By analyzing vast amounts of data, AI can assist in early disease detection, risk assessment, and personalized treatment planning, thereby improving patient outcomes and reducing healthcare costs [8].

The WHO [1] reports that oral diseases impact nearly half of the world's population, highlighting the critical need for efficient and early diagnostic solutions [9]. AI's role in dental imaging is not just about adopting new technologies but optimizing them to address specific clinical needs. Deep Learning (DL), a branch of AI dedicated to understanding data representations, has made a significant impact [10]. The incorporation of AI and Machine Learning (ML) has revolutionized numerous industries, particularly healthcare. In dental imaging, AI technology offers significant opportunities to enhance the detection, diagnosis, and treatment of dental conditions such as oral caries, gum diseases, periodontal bone loss, dental implant placements. While traditional imaging methods like X-rays and Cone Beam Computed Tomography (CBCT) remain crucial for identifying dental issues, they often subjective and depend heavily on the skills and expertise of dental professionals, resulting in potential variations in diagnostic accuracy and treatment outcomes [11]. While AI-enhanced dental imaging introduces a new level of precision and consistency, surpassing traditional methods through DL algorithms that rapidly analyse vast imaging datasets. This capability is crucial for detecting dental diseases at early stages, predicting treatment outcomes, and reducing healthcare costs through timely, accurate interventions. In addition, AI can support dental professionals by enhancing diagnostic accuracy, improving treatment planning, and streamlining workflows, ultimately optimizing resource utilization [7].

Furthermore, AI has the potential to address oral health disparities by enabling remote screening and diagnosis, particularly in underserved communities [8]. For instance, AI applications can facilitate the screening, diagnosis, tracking, prioritization, and monitoring of dental patients remotely via smart devices. This approach allows healthcare providers to reach

a broader population, including those in hard-to-reach areas, ensuring timely interventions and reducing the burden of oral diseases [7].

Studies [12][13] have shown how AI can enhance the accuracy of diagnosing complex conditions such as oral cancers and other maxillofacial abnormalities, underscoring AI's potential to improve diagnostic accuracies and patient outcomes significantly. Furthermore, the challenges of deploying AI in clinical settings are profound. Incorporating AI into routine dental practice requires a comprehensive assessment of model performance and adaptability to ensure these tools effectively enhance clinical decision-making while maintaining patient safety. This is highlighted by the variability in AI performance across different conditions, such as those detailed in the studies by Inani et al.[14] and Basaran et al.[15], where AI systems showed high efficacy in certain conditions but were less reliable in others, such as detecting dental caries or calculus.

As New Zealand continues to navigate its complex oral health challenges, the role of advanced technologies such as AI in dental imaging becomes increasingly crucial. This research not only addresses an urgent public health need but also aligns with the strategic vision of enhancing oral health outcomes across all demographics, ensuring that advancements in dental technology translate into tangible benefits for all New Zealanders [2]. As New Zealand navigates its public health landscape, the oral health sector reveals a dual narrative of progress and persisting disparities. While the overall dental health of New Zealanders has seen improvements, significant challenges remain, particularly among underserved populations. In 2022, only 56% of five-year-olds were caries-free, and disparities were notably pronounced among Māori and Pacific children [5], who exhibited poorer oral health outcomes compared to other ethnicities [16][17][4]. Additionally, certain districts such as Te Tai Tokerau and Lakes reported alarmingly low percentages of caries-free children, underscoring the regional variations in oral health outcomes [17]. The prevalence of oral diseases in New Zealand further emphasizes the urgency for innovative diagnostic solutions. In 2019, 26.4% of the population had untreated caries in permanent teeth, and 21.6% suffered from severe periodontal disease [5]. These statistics not only highlight the widespread nature of oral health issues but also underscore the critical need for effective and early diagnostic mechanisms that can bridge the gap between current outcomes and optimal oral health standards.

1.2. Motivation

This study is motivated by the transformative possibilities of AI in dental care, the shortcomings of conventional imaging methods, and the broader impact on healthcare expenses and patient outcomes. AI's capability to analyse and interpret complex imaging data could offer significant potential for early and accurate diagnosis of dental conditions, thereby improving patient outcomes and enabling more effective disease management. By analysing patient-specific data, including imaging and health records, AI can predict treatment outcomes, enabling tailored treatment plans that cater to individual patient needs [18]. AI can contribute to reducing healthcare costs by enhancing early disease detection, optimizing treatment plans, and improving resource allocation within dental practices [19]. AI offers a consistent method for image analysis, minimizing the variability inherent in human interpretation and enhancing diagnostic accuracy across varied patient groups [20]. The incorporation of AI into dental imaging marks a major technological breakthrough, expanding the possibilities in dental diagnostics [21]. It also offers educational benefits by providing insights into subtle patterns that may enhance training for dental practitioners [22]. The deployment of AI in clinical settings requires rigorous evaluation to ensure effectiveness and safety. Quantitative methods enable systematic validation of AI models, ensuring they meet high standards of accuracy and reliability [23].

1.3. Research Aims and Objective

This study aims to evaluate the performance of AI models in the context of dental imaging. This study aims to evaluate the feasibility, accuracy, and optimization of AI-based DL models, particularly YOLO architectures, for the diagnosis and classification of various dental diseases using dental radiographs. The research will focus on identifying the most accurate YOLO model for dental disease detection and assessing the impact of hyperparameter tuning on model performance in clinical settings. This research aims to enhance the safe, effective, and equitable use of AI in dental diagnostics by emphasizing the model evaluation process, with the goal of improving oral healthcare outcomes both globally and in New Zealand.

Therefore, the objectives of this research are:

- To investigate the application of DL-based object detection models, such as YOLOv5, YOLOv8, and YOLO-NAS, in identifying and classifying dental diseases from radiographic images.

- To evaluate and compare the diagnostic accuracy of different YOLO models (YOLOv5, YOLOv8, and YOLO-NAS) in detecting and classifying various dental conditions.
- To assess performance metrics, including precision, recall, and mean average precision (mAP), to determine which model achieves the highest accuracy for dental radiograph analysis.
- To analyze how hyperparameter tuning (learning rate, batch size and data augmentation techniques) influences the performance of YOLO-based models in detecting and diagnosing oral diseases.
- To compare the performance of models trained with optimized hyperparameters versus default settings, evaluating improvements in accuracy, robustness, and generalization to unseen clinical datasets.

1.4. Research Questions

Feasibility: Can AI-based deep learning models, such as YOLO architectures, be used for the diagnosis and classification of different dental diseases using dental radiographs?

Accuracy: Which deep learning model (YOLOv5, YOLOv8, or YOLO-NAS) provides the highest accuracy in detecting and classifying oral diseases from dental radiographs or CBCT images?

Optimization of Hyperparameter Tuning: How does hyperparameter tuning impact the performance of AI models for detecting and diagnosing oral diseases in clinical settings, and how does it compare to models without such tuning?

1.5. Significance of the Research

The integration of AI into dental diagnostics offers several benefits that extend to patients, healthcare providers, and the healthcare system. This study seeks to explore how AI-driven technologies have demonstrated higher accuracy and efficiency in detecting dental diseases like caries and periodontal disease. Early detection facilitated by AI is vital for successful treatment and better patient outcomes [19]. By analysing patient data and radiographs, AI improves the precision of treatment planning, enhancing treatment success rates and boosting patient satisfaction. Furthermore, AI helps in resource management within dental practices by reducing the need for more extensive treatments through early detection, AI has the potential to decrease healthcare expenses for both patients and providers. AI can reduce variability in diagnosis, ensuring a consistent level of care across different demographic groups. This

standardization is particularly valuable in addressing oral health disparities. AI tools can aid in dental education by highlighting patterns and anomalies, improving diagnostic skills among practitioners and contributing to continuous learning [15].

1.6. Research Contributions

This thesis focuses on the quantitative evaluation of AI-enhanced dental imaging for oral disease detection, leveraging the advanced capabilities of different types of YOLO models. This study explores how DL models can enhance dental disease detection using YOLO object detection techniques. The research is structured as follows:

Data Annotation and Preprocessing – Dental X-ray images are labelled using Roboflow, identifying conditions like cavities, periodontal disease, and other common dental issues. These annotations serve as the ground truth for training YOLO models. The dataset is then pre-processed, augmented, and split into training, validation, and test sets to ensure robust model learning.

Model Development and Training – Various YOLO model versions, including YOLOv5, YOLOv8, and YOLO-NAS, are implemented with and without hyperparameter tuning to assess their performance in detecting and classifying dental pathologies. Each model is fine-tuned to align with the unique features of dental imaging data.

Training Strategy – The study adopts transfer learning, fine-tuning pre-trained YOLO models on dental X-rays to enhance learning efficiency. The training process is optimized by adjusting parameters like the number of epochs, ensuring better accuracy and reliability.

Performance Evaluation – The models' effectiveness is measured using precision, recall, and accuracy. The results from YOLOv5, YOLOv8, and YOLO-NAS are compared against findings from previous studies using other advanced DL methods to establish benchmarks and assess their real-world applicability.

This research aims to highlight the potential of YOLO models in improving dental disease detection, offering a faster, more accurate, and efficient diagnostic approach that could support better patient care and early intervention strategies.

1.7. Thesis Structure

DL technologies are pivotal in automating feature extraction and selection, significantly enhancing disease detection capabilities. This study explores the use of DL to enhance dental imaging and improve oral disease detection, leveraging YOLO models for their effectiveness in real-time object recognition. This research is conducted using Google Colab Pro, leveraging the Python 3 Google Compute Engine backend to access powerful computational resources. This setup enables efficient development, training, and evaluation of the YOLO models, streamlining the entire process and ensuring optimal performance.

The thesis is structured into six comprehensive chapters, each addressing key aspects of the research as follows:

Chapter 1: Introduction

This chapter presents the study's background, outlining the key research questions, objectives, and anticipated outcomes. It also outlines the overall structure of the thesis, guiding the reader through the upcoming sections.

Chapter 2: Literature Review

This section of the thesis examines existing literature on dental imaging, the use of DL in medical diagnostics, and the specific application of YOLO models in medical imaging. It also delves into the theoretical foundations of the YOLO architecture, decision trees, and the evaluation metrics employed to assess model performance.

Chapter 3: Methodology

This chapter outlines the methodology used in the study, explaining the preparation and processing of datasets for training, validation, and testing the models. Data extraction from relevant databases, preprocessing techniques, and dataset segmentation into training, validation, and test sets (70:15:15 ratio) are elaborated. It also discusses the setup for YOLO model development within the Google Colab Pro environment.

Chapter 4: Results

This chapter compares the performance of the YOLO models to assess their effectiveness in detecting oral diseases from dental images. Then critiques the model’s strengths and limitations and explores the potential impact of YOLO-based models on dental diagnostics.

Chapter 6: Discussion, Conclusion and Future Work

The final chapter summarizes the study's findings and explores their implications for advancements in dental healthcare technology. It also outlines potential future research directions, focusing on clinical integration and further refinement of the models to enhance their applicability to various dental imaging scenarios.

Each chapter is meticulously designed to build upon the previous, ensuring a logical flow that culminates in a detailed and thorough exploration of using YOLO models for AI-enhanced dental imaging, facilitated by the advanced computational capabilities of Google Colab Pro. This structure aims to contribute effectively to the field of digital dentistry, emphasizing practical applications and future scalability.

Chapter 2 - Literature Review

The integration of AI into dental imaging has gained significant attention in recent years, with an increasing number of studies emphasizing its transformative impact on oral healthcare. Research has extensively investigated the use of AI technologies, particularly DL algorithms like Convolutional Neural Networks (CNNs) and YOLO models as shown in the visualization Figure 3, Figure 4, Figure 7 to improve the accuracy and efficiency of dental diagnostics as shown in Table 1 and Table 2. These advancements underscore AI's capacity to overcome traditional limitations in dental imaging, such as variability in human interpretation and delayed disease detection.

The review includes a detailed comparison of YOLO-based techniques for dental image analysis with other advanced methods such as Faster R-CNN, UNet, and ensemble learning approaches. The review explores the methodologies utilized, their performance metrics, and the practical relevance of their outcomes in clinical applications. By synthesizing these findings, we offer a comprehensive overview of the current state of AI-driven dental diagnostics and highlight opportunities for future research and improvement.

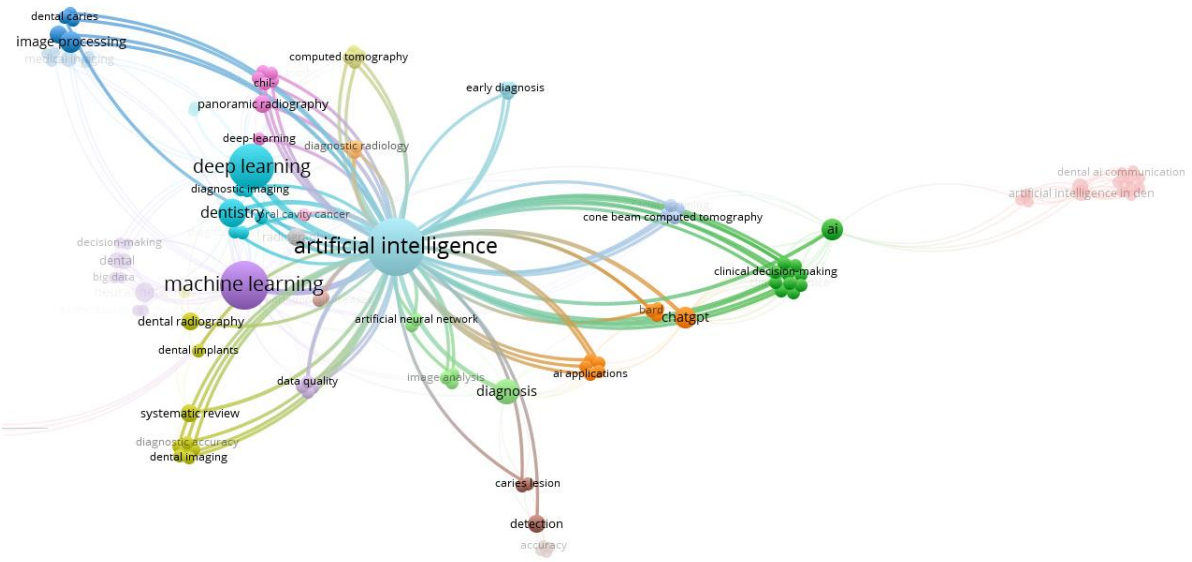


Figure 3. A network visualization illustrating the interconnected themes in research on AI applications in dental imaging.

2.1. AI-Driven Innovations and DL Advancements in Dental Radiograph Analysis

Various studies as shown in Table 1 [24][25][26] have demonstrated the potential of CNNs, U-Net architectures, and advanced DL models in automating the detection and classification of dental pathologies, thus reducing reliance on manual interpretation and increasing consistency across diagnostic processes. One such study [24] focused on utilizing deep CNNs for the detection of apical lesions (ALs) in panoramic dental radiographs. By training a custom 7-layer CNN on 2001 meticulously cropped and labelled tooth segments from 85 patients, the study demonstrated that DL could significantly enhance AL detection accuracy. The preprocessing steps included grayscale conversion, normalization, resizing images to 64x64 pixels, and data augmentation, enabling the model to generalize effectively. The model was trained using a systematic grid search strategy and validated with 10-fold cross-validation, ensuring robust performance. The CNN achieved an impressive area under the receiver operating characteristic curve (AUC) of 0.89 when detecting clearly defined ALs, indicating strong discriminatory power. However, challenges such as examiner variability, a small dataset size, and the low prevalence of ALs limited the model's ability to reduce false positives, highlighting the need for larger, more diverse datasets for optimal performance. Despite these limitations, the study underscored the clinical relevance of CNN-based AL detection, demonstrating how DL could streamline diagnostic procedures, improve risk prediction, and enhance remote monitoring through tele-dentistry platforms [24]. In a related application, another study [27] investigated the potential of DL to detect and diagnose dental caries in periapical radiographs using the GoogLeNet Inception v3 model. This study analyzed 3000 periapical radiographic images sourced from various teeth, including 778 maxillary premolars, 769 maxillary molars, 722 mandibular premolars, and 731 mandibular molars. The dataset selection was rigorous, ensuring only high-quality images with unanimous agreement from four board-certified dentists were used. Images were pre-processed by cropping individual teeth, resizing them to 299x299 pixels, and converting them into JPEG format for standardization. A training-validation split of 2400 images and a test set of 600 images were used, while augmentation techniques such as rotation, shifting, zooming, and flipping were employed to improve model generalizability. By leveraging transfer learning, the GoogLeNet Inception v3 CNN—pre-trained on the ImageNet dataset—demonstrated high diagnostic accuracy, achieving 89.0% accuracy for premolars and 88.0% for molars. The AUC values were also remarkable, with 0.917 for premolars and 0.890 for molars, reinforcing the model's strong detection capabilities.

These findings highlight the substantial clinical impact of AI-assisted caries detection, particularly in cases where traditional diagnostic methods may struggle. The study emphasized the importance of incorporating DL models into preventive care protocols, where they could help predict caries development risks and facilitate early intervention. The technology also has potential applications in tele-dentistry, where AI-assisted remote diagnostics could be particularly beneficial for under-resourced areas with limited access to dental specialists [27].

Similarly, a retrospective study [28] examined the application of DL for detecting caries lesions of varying radiographic extensions in bitewing radiographs. The research analyzed 3686 bitewing radiographs collected from diverse patient demographics. These radiographs, captured using advanced Dentsply Sirona and Durr Dental machines, were carefully annotated pixel-by-pixel by expert dentists to mark the presence and extent of caries lesions, specifically distinguishing between enamel lesions (EL) and dentin lesions (DL). A U-Net deep convolutional neural network architecture with an EfficientNet-B5 encoder, pre-trained on ImageNet, was employed to maximize feature extraction and segmentation accuracy. The model training involved extensive pre-processing steps, including resizing, geometric transformations, and pixel-level adjustments, with a final validation conducted by an additional expert dentist. The model's performance was compared to seven independent dentists who assessed a test set of 141 radiographs. Notably, the DL model significantly outperformed human examiners in sensitivity, achieving 0.75 compared to the dentist's mean sensitivity of 0.36, though it had a slightly lower specificity of 0.83 compared to the dentists' 0.91. This study underscored the potential of AI-driven diagnostics in identifying early-stage caries lesions, which are often overlooked by clinicians due to their subtle radiographic appearance. The integration of such DL models into clinical practice could significantly enhance preventive care planning by predicting lesion progression, thus enabling timely and targeted interventions. Furthermore, the ability to incorporate AI-assisted diagnostics into tele-dentistry platforms could provide remote access to high-quality dental assessments, especially in underserved regions [28]. Expanding the scope of AI applications in dentistry, another study [29] conducted in the Netherlands explored the use of DL to automate the detection, segmentation, and labelling of dental structures in panoramic radiographs (PRs). This research employed the Mask R-CNN model with a ResNet-50 backbone to analyze 2000 PRs obtained from various dental clinics, ensuring a broad representation of image qualities and clinical scenarios. Expert clinicians meticulously labelled each image pixelwise using the Segdent imaging annotation tool, with subsequent review and revision conducted by three additional experienced clinicians

to ensure annotation accuracy. The DL model underwent extensive training, including pre-processing techniques such as normalization, random horizontal flipping, and cropping to enhance its generalizability. Trained using mini-batch gradient descent and a multi-task loss function, the model was validated on an independent set of 200 PRs. The performance results were remarkable, with F1 scores exceeding 0.95 for most dental structures except for root remnants, which had a lower detection rate with an F1 score of 0.807. The segmentation accuracy ranged from 0.873 to 0.952, reinforcing the model's effectiveness in automating complex dental radiographic analysis. The study highlighted how AI-driven segmentation could reduce diagnostic variability, enhance workflow efficiency, and support decision-making in clinical practice [29]. A further retrospective study [25] focused on the use of DL for detecting periapical radiolucency in panoramic dental radiographs, which is crucial for diagnosing infections, granulomas, cysts, and tumours. A dataset of 2902 de-identified PRs was used for model training, with an additional test set of 102 radiographs assessed by 24 oral and maxillofacial (OMF) surgeons. The DL model, based on a U-Net architecture, was specifically optimized for multi-resolution feature extraction, ensuring robust detection of periapical radiolucencies. Radiographs were resized to 256x512 pixels, with intensity scaling and extensive augmentation techniques—including translations, rotations, and flips—implemented to improve model robustness. The model, trained using a dice loss function and Adam optimization, was evaluated using key performance metrics such as Positive Predictive Value (PPV), True Positive Rate (TPR), F1 Score, and Average Precision (AP). The results revealed a mean PPV of 0.67, a TPR of 0.51, and an F1 score of 0.58, outperforming 14 of the 24 OMF surgeons in terms of F1 score. The study demonstrated that AI-assisted detection of periapical radiolucency could significantly reduce the rate of missed diagnoses, standardize the detection process, and improve overall diagnostic reliability [25]. Complementing these experimental studies, a comprehensive survey [26] aimed at encapsulating the evolution, current status, and prospective advancements of DL algorithms in the analysis of dental radiographs. This paper synthesized findings from a myriad of studies to assess the broad applicability and effectiveness of DL technologies across various dental imaging formats such as bitewing, periapical, and panoramic radiographs. The review discussed methodologies leveraged in these studies, focusing on data pre-processing, segmentation, feature extraction, and classification implemented through advanced DL models such as CNNs, R-CNN, U-Net, and Mask R-CNN. A significant portion of the review was dedicated to the technical processes involved, including image normalization, augmentation techniques like rotation and cropping, and automatic feature extraction and selection by DL models. Various training algorithms were highlighted,

noting their pivotal roles in enhancing model performance on dental radiographs. Collectively, these studies underscore the transformative potential of AI-driven dental diagnostics, offering enhanced diagnostic precision, reduced manual errors, and improved accessibility through tele-dentistry, ultimately shaping the future of dental healthcare with more efficient and reliable patient care solutions.

2.2. Diagnostic Accuracy in Dentistry with Machine Learning Techniques

Recent studies have demonstrated the potential of DL models in automating various aspects of dental diagnostics, reducing reliance on manual interpretation, and improving consistency across different clinical scenarios. A diagnostic study [30] conducted in Germany aimed at enhancing the detection and categorization of dental restorations using DL. Utilizing a sample of 1761 intraoral photographs, the study focused on automatically identifying different types of posterior restorations—composite, cement, amalgam, gold, and ceramic—along with unrestored tooth surfaces. The intraoral photographs were captured using professional Nikon single-reflex cameras, complemented by a Sigma Macro Flash, ensuring high-quality image collection. To prepare the data for analysis, images were initially categorized by three dentists and further validated by an experienced examiner to establish a reference standard. The images underwent processing enhancements such as cropping, rotating, and brightness adjustments before being normalized and augmented for training. They were then resized to 300x300 pixels to fit the input requirements of the ResNeXt-101-32x8d model—a sophisticated DL architecture chosen for this study. The ResNeXt-101-32x8d model was trained on a set of 1407 images using backpropagation, dropout techniques, and an Adam optimizer with cross entropy loss, and was later evaluated on a test set comprising 354 images. The model achieved remarkably high diagnostic accuracy rates across the different types of restorations, with accuracy rates ranging from 92.9% for composite to 99.4% for gold restorations. Additionally, the area under the ROC curve (AUC) approached 1.000 for each category, highlighting the model's strong performance in distinguishing between restoration types. The potential applications of this CNN in clinical settings are vast. It could substantially assist dentists in detecting and categorizing various dental restorations, thereby improving diagnostic accuracy and reducing the likelihood of manual errors. A prospective study [31] aimed at developing a deep-learning-based regression model to predict dental age (DA) using panoramic radiograph images. The study involved 529 panoramic radiographs from a diverse group of children, young adults, and adults, with each image annotated by three dentists to determine the DA,

focusing particularly on the stages of development of the seven left lower mandibular teeth. The researchers employed a variety of pre-trained CNN models—Xception, VGG16, DenseNet121, and ResNet50—for feature extraction. These models were chosen for their ability to automatically select and analyze relevant features from the cropped images that were specifically processed to highlight the critical areas for age determination. The panoramic radiographs were pre-processed by cropping to focus solely on the relevant teeth, enhancing the models' accuracy in DA estimation. The performance of these models was rigorously tested using a subset of the data, with the Xception model showing the best performance, achieving a Mean Absolute Error (MAE) of 1.4173 specifically for the age group of 7-11 years. This indicates a high precision in predicting DA within a focused age range, which significantly aids in the accuracy of developmental predictions and treatment planning. The potential applications of this technology in a clinical context are profound. The model provides a critical tool for dentists to accurately determine a patient's DA, which is essential for planning appropriate treatments and anticipating the timing of dental developmental milestones. Furthermore, the ability of the model to deliver accurate assessments remotely makes it an excellent resource for tele-dentistry platforms, offering support for dental professionals in various settings and improving the efficiency of DA assessments. By reducing the reliance on manual evaluations, this AI-driven approach not only saves time but also enhances the overall efficiency and efficacy of clinical dental practices. Another prospective study [32] conducted utilizing machine learning to classify and predict oral diseases, particularly focusing on tooth wear. The study encompassed 133 patients aged 40 to 70 years. The researchers integrated panoramic radiographs (OPG) and detailed patient-reported questionnaire data to develop a robust ML model aimed at diagnosing dental attrition, abrasion, and erosion. The methodological framework involved analyzing the radiographs using the Knight and Smith index to assess the severity of tooth wear, complemented by comprehensive questionnaires that captured 28 attributes related to lifestyle, dietary habits, oral hygiene practices, and specific symptoms associated with tooth wear. This dual approach allowed for a thorough examination of the factors influencing oral health and the progression of tooth wear. The study meticulously prepared the data through processes like data cleaning, which included handling missing values and finding correlations among variables, ultimately categorizing the tooth wear into four severity classes from Class I to Class IV. For the analysis, feature extraction focused on pulling relevant data from both the questionnaire responses and the radiographic findings. Feature selection was strategically aligned with the indices and questionnaire insights to ensure relevance to tooth wear. The Support Vector Machine (SVM) was the primary classification

algorithm employed, tested across various kernels, and evaluated using performance metrics such as precision, recall, and accuracy. The study reported an accuracy of 66.61% for SVM and 70.81% for Support Vector Regression (SVR). The findings of this study underscored the feasibility of utilizing a combination of questionnaire data and radiographic insights for the early diagnosis and classification of tooth wear, highlighting the model's capability in effectively detecting moderate to severe cases. This innovative approach presents several practical applications in clinical contexts. It enhances the early diagnosis of tooth wear by correlating patient-reported symptoms with radiographic evidence, which is crucial for timely and targeted interventions. Additionally, the model's predictive capabilities facilitate risk prediction based on lifestyle factors, enabling preventative measures. A retrospective study [33] supplemented by data from five private practices, with the objective of developing a deep CNN to identify dental implant brands and models from radiographic images. The study involved a dataset of 1,206 radiographs. This dataset was meticulously divided into a training and validation group of 965 images and a testing group of 241 images. The researchers utilized digital intraoral radiographs and manually cropped orthopantomograms, which were converted to JPEG format and pre-processed to a uniform size of 90x200 pixels. The images underwent extensive data augmentation, including adjustments to horizontal rotation, angulation, tone, brightness, contrast, blur, sharpness, and gamma correction, to enhance the robustness of the training set. The GoogLeNet Inception v3 architecture was chosen for its efficacy in feature extraction and its ability to automatically select pertinent features during the training process, employing transfer learning to maximize the potential of the pre-trained model. The performance of the CNN was rigorously evaluated on the test dataset, revealing an impressive accuracy of 93.8%, with sensitivity at 93.5% and specificity at 94.2%. The positive predictive value stood at 92%, and the negative predictive value at 91.5%, with area under the ROC Curve values ranging from 0.890 to 0.931 depending on the implant model. These metrics underscored the model's high diagnostic precision, particularly for implants with distinct physical characteristics like external hexagon platforms or apex holes. The clinical implications of this study are significant. The model provides a valuable tool for clinicians to accurately identify dental implants from radiographs, which is crucial for effective treatment planning and managing complications associated with various implant types.

Table 1. Summary of Features Extracted, Algorithms/Methods Employed, and Oral Diseases Detected in Studies Utilizing Alternative AI and ML Techniques for Dental Imaging

Features Extracted	Algorithm / Methods Used	Oral Disease Detected	Performance Metrics
Tooth segments labelled for apical lesions [24]	Deep 7-layer CNN, Grid Search, 10-fold Cross Validation	Apical lesions	AUC of 0.89
Cropped periapical images (299x299 pixels), isolated teeth data, and caries labelling [27]	GoogLeNet Inception v3 (CNN), Transfer Learning, Data Augmentation	Dental caries	Accuracy of 89% for premolars and 88% for molars
Pixel-wise labelled bitewing radiographs (Enamel E1/E2, Dentin D2/D3) [28]	U-Net Deep CNN with EfficientNet-B5 Encoder	Caries lesions of varying radiographic extension	Accuracy of 80%
Panoramic radiographs labelled for teeth, crowns, fillings, and root canal fillings [26]	Mask R-CNN with ResNet-50 Backbone, Data Augmentation, Mini-batch Gradient Descent	General dental structure and dental disease	Comprehensive survey
Intraoral photographs labelled for composite, cement, amalgam, gold, ceramic, and unrestored surfaces [30]	ResNeXt-101-32x8d (CNN), Cross Entropy Loss, Data Augmentation	Posterior restorations detection	Accuracy rates ranging from 92.9% for composite to 99.4% for gold restorations
Various dental radiograph formats analyzed for image normalization, augmentation, and segmentation [29]	Survey of multiple DL models: CNNs, R-CNN, U-Net, Mask R-CNN	General dental imaging diagnosis and prediction	Segmentation accuracy ranging from 87.3% to 95.2%
Panoramic radiographs annotated for seven left lower mandibular teeth stages [31]	Xception, VGG16, DenseNet121, ResNet50 (CNNs), Transfer Learning	Dental age prediction	MAE of 1.4173
Panoramic radiographs labelled for tooth wear severity using Knight and Smith index [32]	SVM, SVR	Tooth wear (attrition, abrasion, erosion)	Accuracy of 66.61% for SVM and 70.81% for SVR
Panoramic radiographs annotated for periapical radiolucencies (lesions, cysts, tumors) [25]	U-Net, Dice Loss Function, Adam Optimization	Periapical disease detection	F1 score of 0.58 and an AP of 0.60
Cropped orthopantomograms annotated for dental	GoogLeNet Inception v3 (CNN), Transfer Learning, Data Augmentation	Dental implant identification	Accuracy of 93.8%, with sensitivity at 93.5% and

implant types and brands [33]			specificity at 94.2%

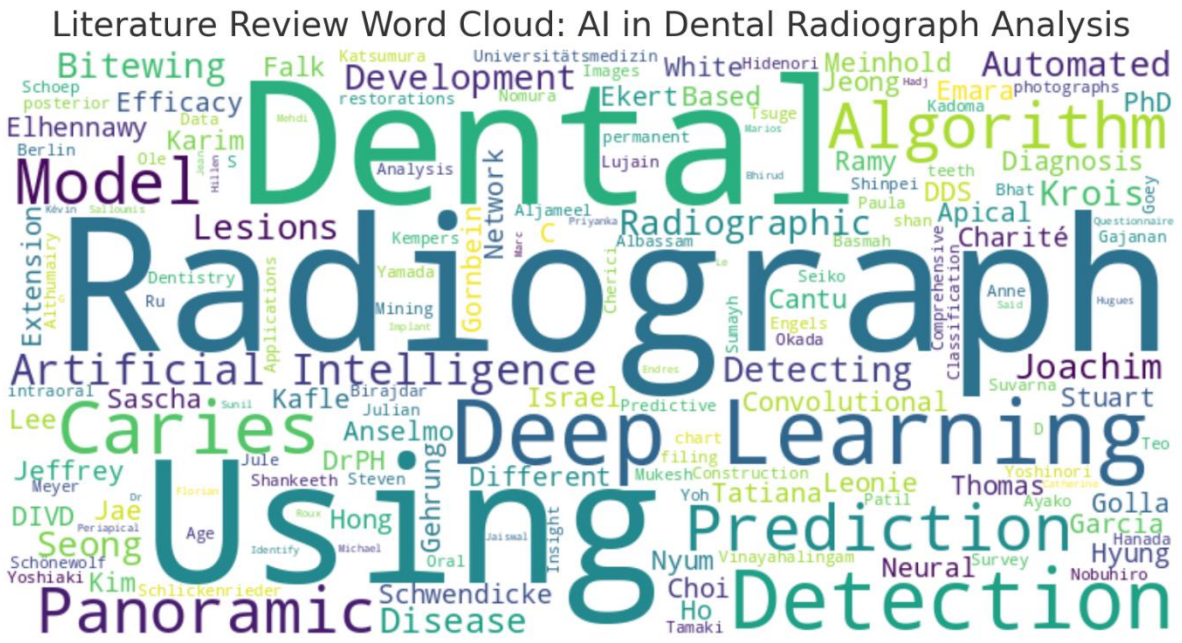


Figure 4. A word cloud generated from the literature review on AI in dental radiograph analysis.

2.3. Literature Review Method

As part of the study, a scoping review was undertaken to explore the existing literature on the use of DL models in dental imaging. This review aimed to provide a thorough overview of current methodologies, challenges, and gaps in the field, forming the foundation for the research. The review adhered to the PRISMA framework (Preferred Reporting Items for Systematic Reviews and Meta-Analyses), with a PRISMA flow diagram included below in Figure 5 to outline the identification, selection, and inclusion process of relevant studies. The findings of the scoping review helped refine the research objectives and informed the research design by identifying key performance indicators explored in previous studies as shown in **Appendix**.

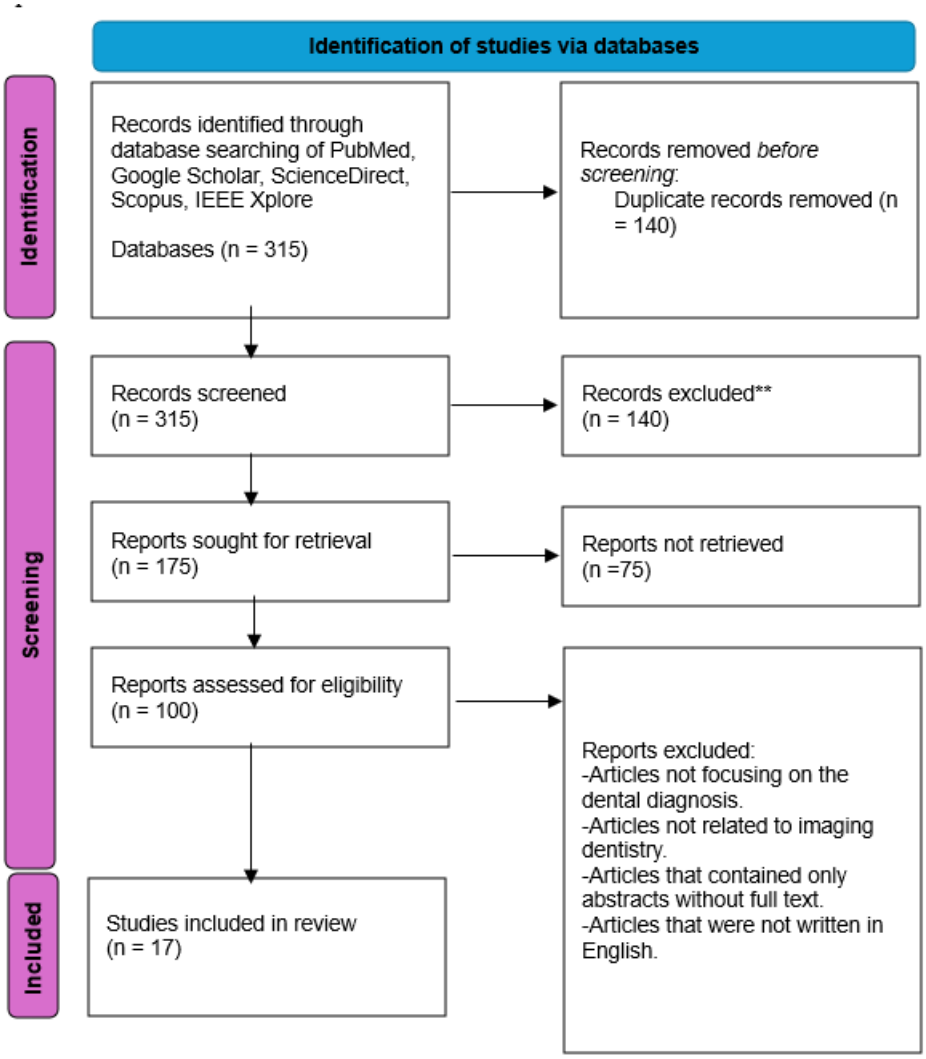


Figure 5. PRISMA Chart for identifying, screening, and excluding studies related to dental diagnosis and imaging from various databases

2.4. The Evolution of YOLO: A Timeline of Innovation

YOLO is an innovative DL framework designed for real-time object detection, initially developed by Joseph Redmon in 2015 with the introduction of YOLOv1 as shown in Figure 6. This first version transformed object detection by framing it as a single regression problem, enabling the prediction of bounding boxes and class probabilities in one pass through the network. Unlike earlier methods like R-CNN, which required multiple passes and were computationally demanding, YOLO offered a more efficient solution [15]. Building on the success of YOLOv1, YOLOv2 was introduced in 2016, offering notable advancements in both accuracy and speed. It incorporated features like batch normalization, high-resolution classifiers, and anchor boxes, which improved its capability to detect smaller objects and enhanced its performance in more complex scenarios. In 2018, Redmon and Farhadi [34]

further refined YOLO with YOLOv3, which incorporated a deeper network architecture and feature pyramid networks for better detection of objects at multiple scales, making it a highly versatile model for various applications [34]. In 2020, YOLOv4 was released [4], which introduced additional innovations such as CSPDarknet53 as the backbone, mosaic data augmentation, and CIoU loss for more accurate object localization. YOLOv4 set a new benchmark with its state-of-the-art accuracy and inference speed, reinforcing YOLO's position as a top-tier object detection framework. That same year, Ultralytics launched YOLOv5, emphasizing user-friendliness and practical deployment. With its PyTorch implementation, YOLOv5 became more accessible to both developers and researchers. Its lightweight nature and fast inference times made it highly suitable for real-time applications, including medical and dental imaging. The YOLO family continued to expand with the release of YOLOv6 and YOLOv7 in 2022, followed by YOLOv8 in 2023. YOLOv8 brought further architectural enhancements and hyperparameter optimization, setting new benchmarks in object detection tasks [35][36]. The same year, Deci AI introduced YOLO-NAS (Neural Architecture Search), which leveraged automated architecture optimization to deliver exceptional accuracy and speed. YOLO-NAS particularly excels in tasks requiring fine-grained object detection, such as identifying small objects in medical imaging, and is increasingly regarded as the cutting-edge YOLO variant. Throughout its evolution, YOLO has maintained its focus on real-time detection, computational efficiency, and scalability. Its applications span numerous domains, including autonomous driving, surveillance, and healthcare. In the medical and dental imaging fields, YOLO's ability to detect small anomalies with high precision has made it a critical tool for tasks such as lesion detection, caries classification, and cancer diagnosis [37].

The continued development of YOLO by the open-source community and organizations like Ultralytics and Deci AI ensures that it remains at the forefront of object detection research, with each iteration building upon the strengths of its predecessors while addressing emerging challenges in the field. By 2024, YOLO's evolution culminated in YOLOv9 and YOLOv10, which integrated cutting-edge advancements in AI architecture to push the boundaries of what object detection models can achieve. Continuous innovation in YOLO models has made them indispensable tools in fields ranging from autonomous driving to healthcare, including dental diagnostics, where the need for high precision and speed is critical [7].

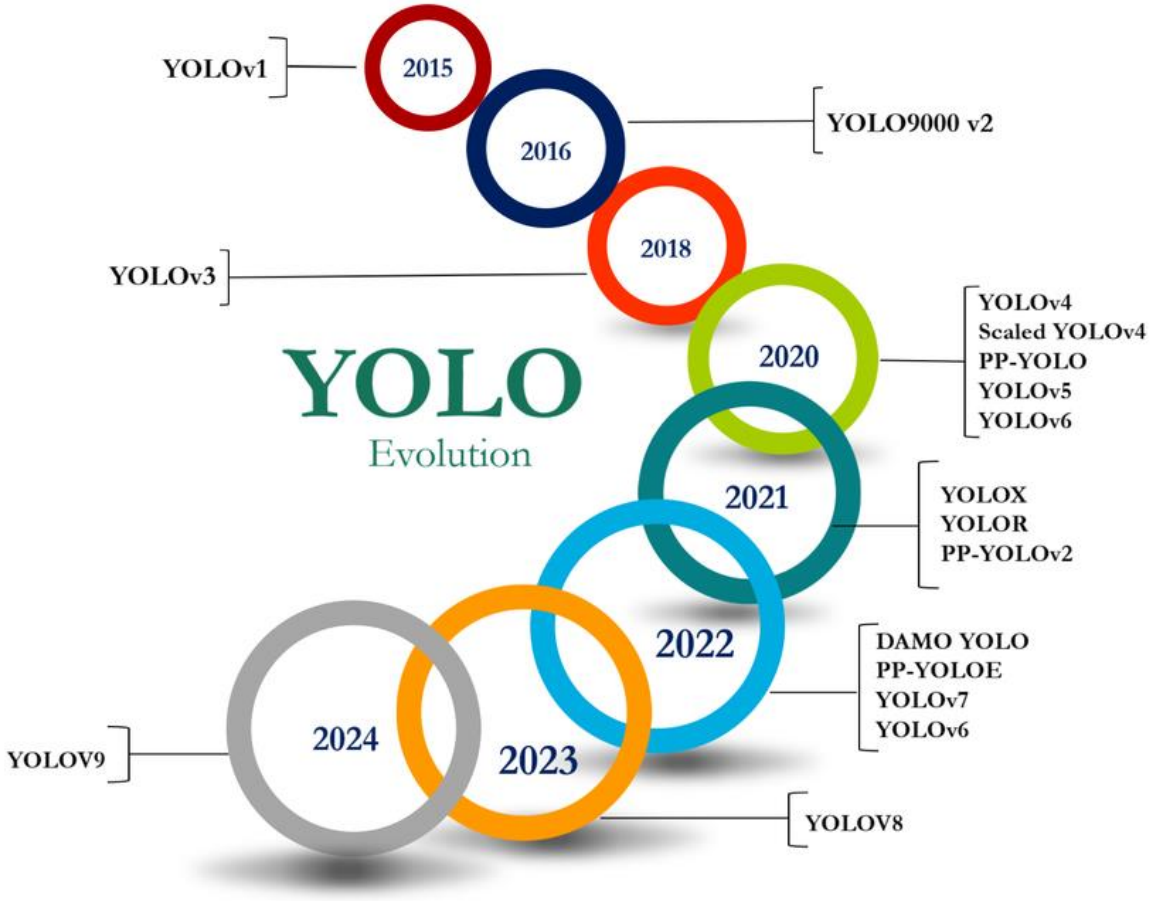


Figure 6. YOLO Evolution Timeline: A Visual Representation of the Development of YOLO Object Detection Models from YOLOv1 (2015) to YOLOv9 (2024)[46]

2.5. YOLO methods

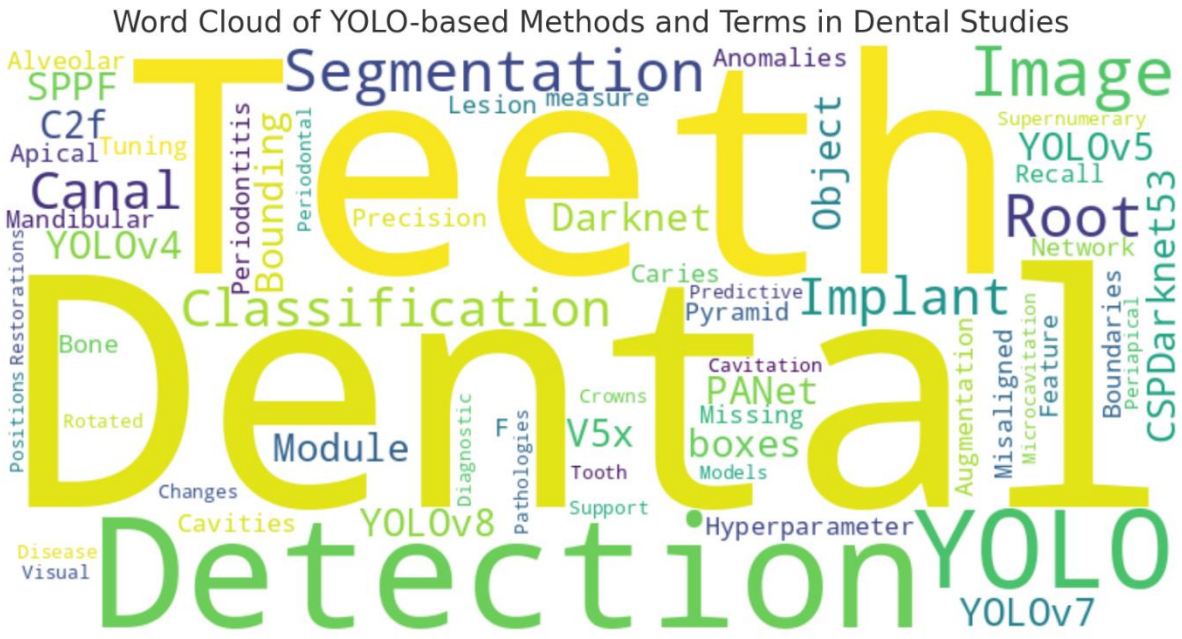


Figure 7. A word cloud highlighting key terms and concepts related to YOLO-based methods in dental imaging studies.

Several studies [36][38][39][40], as highlighted in Table 2 , have validated the effectiveness of YOLO architectures in dental diagnostics, showcasing their ability to detect conditions such as dental caries, root fractures, and impacted teeth with high precision and efficiency, thereby streamlining clinical workflows. A research [39] investigates the use of DL models for automated tooth position identification and detection of dental conditions using bitewing radiographs. The study's primary aim is to enhance the efficiency and accuracy of dental diagnostics by employing DL algorithms to automate traditionally manual and labour-intensive tasks such as tooth positioning and disease detection. The study sample includes 210 bitewing radiograph images, annotated to identify dental caries, restorations, and periodontal disease. The dataset was split into training and validation sets using a 9:1 ratio. The authors utilized YOLOv4, YOLOV5 for tooth position identification and a modified AlexNet CNN model to detect specific dental conditions. Features collected include tooth positions, caries, restorations, and periodontal disease symptoms, such as gum inflammation and bone loss. Preprocessing techniques like Gaussian filtering and adaptive binarization were applied to enhance image quality. Data analysis involved training the YOLOv4, YOLOv5 models for object detection and AlexNet for disease identification. The models were evaluated using precision, recall, and F1-score, showing an accuracy of 98.01% for tooth position detection and 92.86%, 92.10%,

and 96.51% for caries, periodontal disease, and restoration detection, respectively. This system can support dentists by reducing diagnosis time, minimizing human error, and potentially being integrated into remote health platforms for widespread use. A DL approach was proposed [38] - for dental anomaly detection using the YOLOv5 and YOLOv8 architecture. The study employs a diagnostic framework to automatically detect various dental conditions using panoramic radiographs. The primary aim is to enhance diagnostic accuracy and reduce processing time in clinical settings by leveraging the latest advancements in object detection algorithms. The study analysed a sample of 1,267 dental radiographs, annotated to identify four dental anomalies: implants, fillings, impacted teeth, and cavities. The dataset was divided into training (84.5%), validation (9.5%), and testing (5.7%) sets, ensuring robust model performance evaluation. Key features collected include the position and presence of dental anomalies within panoramic X-ray images. Data analysis was performed using the YOLOv5 and YOLOv8 framework, incorporating enhancements like CSPDarknet53 as the backbone, SPPF for multi-scale feature capturing, and the C2f module for combining high-level features. The model was trained and evaluated using precision, recall, and mAP metrics, achieving an mAP50 of 80.6% and 86.6% across all dental anomalies. By automating anomaly detection, the model can serve as a decision-support tool, facilitating early diagnosis and treatment planning while reducing human error and diagnostic time. A study [41] on the application of DL for dental disease detection and classification using panoramic X-ray images. The study was designed to develop an automated tool capable of diagnosing four common dental disorders: cavities, root canals, dental crowns, and broken-down root canals. The study used a sample of 1,200 panoramic dental X-ray images collected from various clinics, with 70% used for training and 30% for testing. Features under collection included bounding box annotations for different dental conditions within the panoramic X-rays, labelled for precise detection and classification. For data analysis, the authors utilized the YOLOv3 DL model, implementing various image augmentation techniques such as horizontal and vertical flips, rotation, shear range, and zoom range to enhance the dataset and improve model performance. The model was trained on Google Colab with a NVIDIA Tesla K80 GPU, achieving a mAP of 99.33%. By integrating the model into routine dental radiographic analysis, clinicians can minimize human error, reduce diagnostic time, and streamline treatment planning. A comprehensive review [42] of recent advancements in AI applications for digital dental radiography. The study categorizes and discuss the current developments of AI methodologies applied to digital dental radiography, including the diagnosis of dental caries, periapical pathologies, periodontal bone loss, cyst and tumour classification, cephalometric analysis, and dental implant recognition.

The review collected a sample of 119 articles from PubMed and IEEE Xplore databases, with studies focusing on both ML and DL models used to automate the detection, classification, and segmentation of dental pathologies. Key features under collection include image-based data such as dental radiographs (panoramic, intraoral, and CBCT) labelled for various dental conditions, anatomical landmarks, and bone structures. Data analysis methods range from traditional ML techniques, such as support vector machines and neural networks, to advanced DL models like U-Net, YOLO, and Mask R-CNN for complex pattern recognition and image segmentation. The potential clinical application of this research lies in improving diagnostic accuracy, reducing human error, and enhancing the efficiency of dental practitioners by providing automated diagnostic aids. The review highlights the need for further development to address challenges such as small dataset sizes, standardization of reporting formats, and integration into clinical workflows to fully leverage AI in dental radiography. Another study [43] evaluates the effectiveness of the YOLO-V5 DL model for detecting, segmenting, and numbering deciduous and permanent teeth in paediatric patients using panoramic radiographs. The study design is a retrospective observational study that used a labelled dataset of 3,854 panoramic radiographs collected from paediatric patients aged 5–13 years with mixed dentition conditions. The aim of the study is to automate tooth identification and numbering during the complex mixed dentition stage to support accurate and efficient diagnosis in clinical settings. The sample dataset consists of 3,093 images for training, 387 for validation, and 385 for testing. Features under collection include deciduous and permanent teeth, dental caries, rotated teeth, developmental anomalies, and supernumerary teeth, all annotated using the FDI tooth numbering system. For data analysis, the authors developed the YOLO-V5x model, utilizing techniques such as CSPDarknet53 for the backbone, PANet for the neck, and the YOLO head for classification. The model was trained for 500 epochs, achieving high performance with a mAP-0.5 of 98% for both detection and segmentation tasks. This model enables paediatric dentists to automatically detect and enumerate teeth, aiding in the diagnosis of complex dental conditions like impacted teeth, caries, and dental anomalies. By reducing manual labour and increasing diagnostic accuracy, this AI-based system can streamline the treatment planning process and potentially improve patient outcomes in paediatric dental care. A novel DL-based framework [44] aimed at automating the detection of critical anatomical structures like AB and MC in CBCT images for dental implant planning. The sample consists of 1,064 grayscale CBCT images of the mandible's coronal slices, with annotations for the AB and MC. The dataset was divided into 744 images for training and 320 images for testing. Features under collection include detailed annotations for the AB and MC, with bounding boxes specifying

the precise locations of these structures. For data analysis, the authors developed the Dental-YOLO model, which is a modified version of the YOLOv4 architecture optimized for the unique requirements of dental imaging. The model uses CSPDarknet53-tiny as the backbone network and implements two-scale feature maps for efficient detection of both large (AB) and small (MC) structures. The results of the study showed that Dental-YOLO achieved a mAP of 99.46% and demonstrated significant improvements in detection speed, reducing the complexity from 59.57 to 6.83 billion floating-point operations (BFLOPS) per second, making it highly efficient for clinical use. This system can automate the identification of critical structures, streamline the dental implant planning process, and reduce the time and manual effort required by radiologists, thereby enhancing accuracy and treatment outcomes in implant dentistry. A study centred [39] to evaluate the performance of the YOLOv4 model in detecting and classifying different types of teeth in dental radiographs. The study follows an experimental design where the authors implement and fine-tune the YOLOv4 architecture for accurate identification of individual teeth to improve diagnostic support in dental imaging. The sample consists of 500 annotated dental radiographs collected from patients with varying dental conditions, including misaligned teeth, missing teeth, and restorations. Key features under collection include positional data for each tooth, bounding box coordinates, and class labels specifying tooth types (e.g., molars, premolars, and incisors). The annotations were manually validated by experienced dental professionals to ensure accuracy. Data analysis involved training the YOLOv4 model with different hyperparameter configurations, and performance was assessed using sensitivity, specificity, and F1-score. The model achieved a sensitivity of 97.8% and a specificity of 98.3%, indicating high precision in detecting and distinguishing individual teeth under diverse radiographic conditions. This system could reduce diagnostic time, minimize human error, and provide a reliable second opinion, ultimately enhancing patient care and outcomes in dental practices. Another research [45] explores the application of the YOLOv5 and YOLOv7 model for real-time object detection in dental radiographic images to enhance implant planning and classification in prosthodontics. The study design follows an experimental approach, employing DL to automate the detection and measurement of dental implants in various clinical scenarios. The sample consists of 4,000 annotated dental radiographs, categorized based on different dental implant conditions, including missing teeth, various implant types (e.g., endosteal, subperiosteal, zygomatic, and transosteal), and their precise locations. Key features under collection include the dimensions (height and width) of dental implants, tooth positions, and surrounding anatomical structures, as well as implant type classifications based on radiographic appearances. For data analysis, the authors used

YOLOV5 and YOLOv7, incorporating customized loss functions to optimize detection accuracy and direct prediction of implant dimensions. The model achieved an overall accuracy of approximately 89%, with a recall of 0.61, precision of 1 at a confidence level of 0.895, and an F1-score of 0.35 at a lower confidence threshold of 0.292. The architecture includes a backbone network (CSPDarknet53), a feature pyramid network (FPN), and detection heads to predict object coordinates and classifications. This study highlights the effectiveness of integrating AI in dental prosthetics, with YOLOv7 serving as a reliable tool for real-time detection and classification of dental implants, contributing significantly to advancing precision in implant dentistry. Another study focussed on [46] developing an automated and cost-effective AI-based diagnostic system for dental caries detection using non-standardized photographs. The study design is an experimental framework that uses a hybrid model combining YOLO object detection with transfer learning to handle real-world variations in image quality, specifically targeting low-resource and rural settings where standardized imaging setups are not feasible. The sample consists of 1,703 augmented images generated from 233 extracted anterior teeth, captured using a consumer smartphone camera without controlled lighting or standardized apparatus. The dataset was split into training (65%), validation (15%), and testing (20%) sets. Features under collection include visual changes in smooth surface carious lesions categorized into three classes: visible change without cavitation, visible change with micro cavitation, and visible change with cavitation, based on the International Caries Detection and Assessment System (ICDAS). Data analysis involved implementing an ensemble of YOLOv5 models (v5s, v5m, v5l, and v5x) combined with transfer learning models such as VGG16, ResNet50, ResNet101, AlexNet, and DenseNet121. The ensemble model achieved a mAP of 0.732 and an accuracy of 78.9%. Transfer learning with the VGG16 model resulted in the highest diagnostic accuracy of 86.96%, precision of 0.89, and recall of 0.88. This system can support primary healthcare providers in detecting and monitoring caries progression, ultimately improving oral health outcomes for underserved populations [46]. An investigation into [47] the effectiveness of a DL model for detecting AP in digital panoramic radiographs. The study design is a diagnostic accuracy evaluation using a DL-based Computer-Aided Diagnosis (CAD) system, specifically the YOLOv3 and YOLOv5 models, to assess its performance in identifying AP in the maxilla and mandible. The sample includes 306 panoramic radiographs with a total of 400 roots diagnosed with AP, with an equal number of lesions in both the maxilla and mandible. The images were selected from a clinical archive using the keywords "apical lesion" and "apical periodontitis" and confirmed by two oral and maxillofacial radiologists who labelled and grouped the lesions according to their

borders and visibility. Features under collection include the lesion boundaries, specific root locations, and the presence or absence of AP. Data analysis involved training the YOLOv3 model with Darknet architecture, focusing on three scales to detect AP lesions of varying sizes. Performance metrics such as precision, recall, and F-measure were used to evaluate the model's diagnostic accuracy. The results showed a recall rate of 0.93 and an F-measure of 0.96 for clearly identifiable APs in the mandible, but lower recall values of 0.2 for the maxillary lesions, indicating the model's limited accuracy in the maxilla due to overlapping anatomical structures and low contrast. It can serve as a reliable diagnostic aid for identifying APs in digital panoramic radiographs, thereby reducing diagnostic time and enhancing accuracy in complex cases. However, its current limitations suggest that the model is more suited for mandibular APs and may require further development to achieve clinical applicability for maxillary lesions

Table 2. Summary of Features Extracted, Algorithms Used, and Dental Diseases Detected in AI-Driven Dental Imaging Studies

Features Extracted	**Algorithm / Methods Used**	**Oral Disease Detected**	**Performance Metrics**
Tooth positions, caries, restorations, periodontal disease symptoms (e.g., gum inflammation, bone loss)[39]	YOLOv4, YOLOv5, AlexNet CNN	Dental caries, restorations, periodontal disease	Accuracy of 98.01% for tooth position detection and 92.86%, 92.10%, and 96.51% for caries, periodontal disease, and restoration detection, respectively
Position and presence of dental anomalies (implants, fillings, impacted teeth, cavities) [38]	YOLOv5, YOLOv8, CSPDarknet53, SPPF, C2f Module	Implants, cavities, fillings, impacted teeth	mAP50 of 80.6% across all dental anomalies
Bounding box annotations for cavities, root canals, dental crowns, broken-down root canals [41]	YOLOv3, YOLOv5, Image Augmentation (flips, rotation, shear range, zoom)	Cavities, root canals, crowns, broken-down root canals	mAP of 99.33%
Anatomical landmarks, bone structures, radiograph annotations (e.g., dental caries, pathologies, cysts, tumours) [42]	U-Net, YOLO, Mask R-CNN, Support Vector Machines, Neural Networks	Dental caries, periodontal bone loss, cysts, tumours	Comprehensive review
Deciduous and permanent teeth, rotated teeth, developmental anomalies, supernumerary teeth [43]	YOLO-V5x, CSPDarknet53, PANet	Dental caries, developmental anomalies, supernumerary teeth	mAP-0.5 of 98%
Alveolar bone , mandibular canal , anatomical structures of the mandible [44]	Dental-YOLO (Modified YOLOv4), YOLOv8, CSPDarknet53-tiny	Alveolar bone, mandibular canal	mAP of 99.46%

Positional data, bounding boxes, and class labels for tooth types (e.g., molars, premolars, incisors) [39]	YOLOv4, YOLOv8 Hyperparameter Tuning	Misaligned teeth, missing teeth, restorations	Sensitivity of 97.8% and a specificity of 98.3%
Dimensions and positions of dental implants, implant classifications (endosteal, subperiosteal, zygomatic, transosteal) [45]	YOLOv5, YOLOv7, YOLOv8 CSPDarknet53, Feature Pyramid Network (FPN)	Dental implant detection and classification	Accuracy of 89%, with a recall of 0.61
Visual changes in smooth surface carious lesions categorized as visible change without cavitation, micro cavitation, and cavitation [46]	YOLOv5 (v5s, v5m, v5l, v5x), Transfer Learning Models (VGG16, ResNet50, ResNet101, AlexNet, DenseNet121)	Dental caries	Diagnostic accuracy of 86.96%, precision of 0.89, and recall of 0.88
Lesion boundaries, root locations, presence or absence of AP [47]	YOLOv3, YOLOv5, Darknet, Precision, Recall, F-measure	Apical periodontitis (AP) detection	Recall rate of 0.93 and an F-measure of 0.96

2.6. Research Gap

Despite the significant advancements in AI applications in dental imaging, the current body of literature exhibits certain limitations. A detailed review of studies highlights a focused approach, where the majority have restricted their scope to a narrow range of dental diseases or conditions. For instance, many studies concentrated solely on one or two dental issues, such as caries detection, tooth position identification, or apical lesion detection. Studies [39][48] utilized YOLOv4 ,YOLOv5 AlexNet CNN to detect tooth positions and three dental conditions (caries, restorations, and periodontal disease), providing a limited application spectrum and some, such as [38][20], focused on four dental anomalies (implants, fillings, impacted teeth, and cavities) using YOLOv5 and YOLOv8 architecture, while few [47] restricted their study to detecting AP using YOLOv3, showcasing limited diagnostic versatility and additionally some studies [43] addressed mixed dentition in paediatric patients using YOLO-V5x, but their application was primarily for tooth numbering and developmental anomalies.

Moreover, several studies such as [28][25], focused predominantly on specific imaging tasks like caries lesion segmentation or periapical radiolucency detection, highlighting the inability to generalize across a wide range of dental conditions. Furthermore, while many models like U-Net, Mask R-CNN, and GoogLeNet Inception v3 have been explored, their application tends to be condition-specific, lacking the adaptability for multiple dental diagnoses. While YOLO models have demonstrated superior accuracy in dental imaging tasks, their application has been limited to isolated use cases rather than broad-spectrum disease identification.

Our research aims to address these gaps by:

Comprehensive Disease Annotation: Unlike the reviewed studies, which focus on a narrow scope, we have annotated 50 plus classes of dental diseases, encompassing a diverse range of conditions such as dental caries, restorations, periodontal diseases, anomalies, edentulous ridges and implants. This comprehensive dataset makes our model applicable to a wide variety of use cases, offering unparalleled diagnostic versatility.

Enhanced Use Case Flexibility: The model developed in our research can simultaneously detect multiple dental conditions within a single analysis. This capability significantly broadens its applicability in clinical settings, facilitating faster, more accurate diagnostics across varied scenarios, including routine checkups, anomaly detection, and treatment planning.

Focus on YOLO Models: Our research emphasizes the YOLO family of models (e.g., YOLOv5, YOLOv8, YOLO-NAS) due to their demonstrated superiority in image analysis tasks, as evidenced by their high precision, recall, and mAP metrics in the reviewed literature as shown in Table 3. YOLO models consistently outperform others, such as U-Net or Mask R-CNN, in tasks requiring real-time detection and localization, making them ideal for dental imaging applications.

Potential for Broader Application: The model's ability to diagnose a wide range of conditions ensures its utility in a variety of settings, from general dental practices to specialized fields like paediatric dentistry, prosthodontics, and oral surgery. Moreover, the inclusion of diverse dental anomalies enhances its applicability in both developed and resource-constrained healthcare environments.

To the best of our knowledge, no study has developed a model capable of diagnosing such an extensive range of dental conditions while maintaining high accuracy and efficiency. By focusing on YOLO models and leveraging a comprehensive dataset, our research not only bridges this gap but also establishes a robust foundation for future advancements in AI-driven dental diagnostics.

Chapter 3 - Methodology

This chapter outlines the approach taken to conduct this study, covering the entire research process from data preparation to model evaluation. The methodology follows a structured workflow to provide a clear breakdown of each stage, ensuring transparency in the research process, beginning with the research methodology framework. The dataset preparation includes data annotation, preprocessing techniques used and dataset splitting, ensuring data consistency. Data augmentation techniques, including geometric transformations, rotation, flipping, and grid masking, enhance model generalization. The study employs YOLO models (YOLOv5, YOLOv8, YOLO-NAS) for dental radiograph analysis, justifying their selection and comparing their performance against other models. Training and evaluation involve tuning parameters with model performance assessed for all the models and compared. The findings and results are discussed in the subsequent chapters, offering insights into the performance and potential improvements of the proposed methodology.

3.1. Research Methodology

This study adopts the **Design Science Research Methodology (DSRM)** as the guiding framework to develop and evaluate the proposed approach. The DSRM framework [49][50], was chosen for this study because it provides a structured way to develop and validate new solutions. Unlike purely theoretical research methods, DSRM is practical and iterative, meaning it allows continuous improvements [49]. This makes it particularly useful for this research that involves designing, testing, and refining solutions in changing environments. Another key reason for selecting DSRM is its flexibility. It supports different starting points in research, such as identifying a problem, setting objectives, or even beginning with an early-stage solution [50]. This adaptability ensures that the study remains focused on problem-solving and practical impact. Since this research involves developing models and validating them through experiments, DSRM's iterative process enables continuous refinements. Additionally, DSRM balances academic rigor with real-world relevance, ensuring that this research contributes both to theoretical knowledge and practical applications [51]. This combination makes it an ideal approach for this study that require both innovation and validation.

The methodology consists of six iterative stages as shown in Figure 8. Design Science Research Methodology: Problem Identification & Motivation, Define Objectives of the Solution, Design & Development, Demonstration, Evaluation, and Communication.

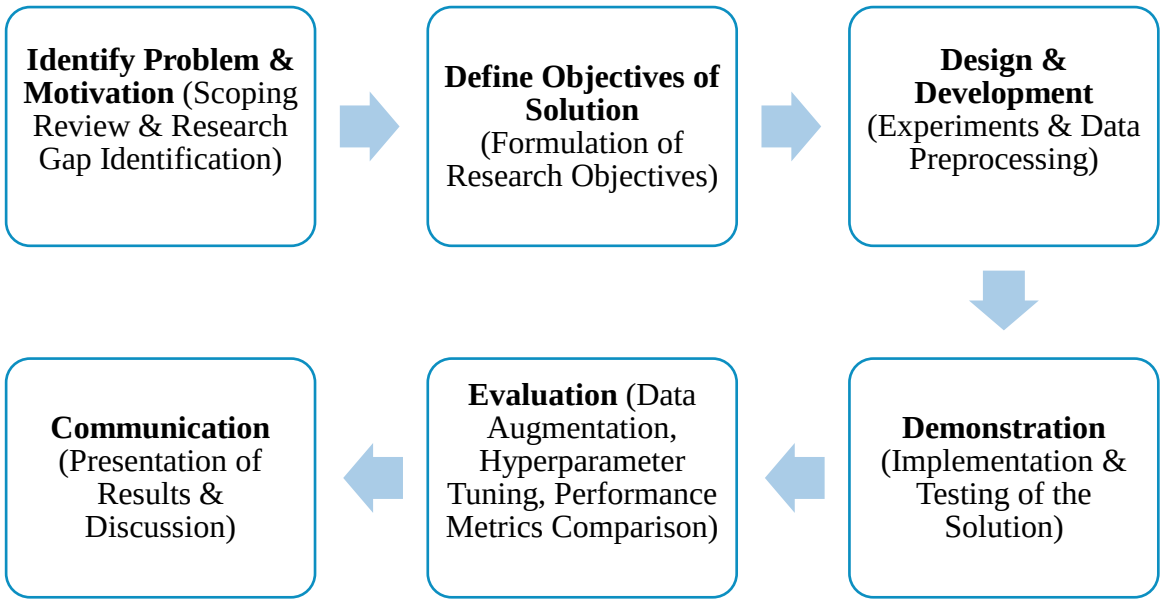


Figure 8. Design Science Research Methodology [49]

The research began with a scoping review, which is detailed in section 2.3. The scoping review enabled us to analyze existing literature, identify research gaps, and establish the significance of the problem. This process helped in defining the research problem clearly and justifying the need for a novel approach. Based on the findings from the scoping review, the study determined the objectives as in section 1.3. These research objectives are formulated to address the identified gaps (section 2.6.) by proposing a more efficient and effective solution. The objectives provide a theoretical foundation that drives the design and development phase. The solution is designed and developed through a series of experiment simulations (section 3.4.). This stage involved implementing various preprocessing techniques (section 3.5.) to enhance data quality, designing and training models to improve the accuracy and robustness of the system, and iteratively refining the solution based on experimental results to optimize performance. The developed models were tested in a controlled experimental setting. This phase included running multiple experiments to validate the effectiveness of the proposed solution, applying data preprocessing techniques to enhance input data quality**,** demonstrating how the solution addresses the research problem effectively. The evaluation phase involves a detailed assessment of the developed solution. This includes data augmentation techniques

(section 3.6) to enhance model generalization, hyperparameter tuning (section 3.9.1.) to optimize the models, performance metric evaluation (section 3.10.) (such as accuracy, precision, recall) to systematically compare different models and configurations. Through this evaluation, model technical insights are drawn. The final phase involved presenting and discussing the findings. The results are documented and disseminated through discussion of key insights, limitations, and future research directions.

3.2. Dataset

The dataset utilized in this study comprises approximately 3,000 digital panoramic radiographic images, collected from various open-source repositories [52][53][54][55] to ensure a diverse and comprehensive representation of oral conditions within dental radiographs.

Following the collection, the images were subjected to manual annotation using the Roboflow platform. A team of two dental experts, including an endodontist and an orthodontist, thoroughly annotated each image to ensure clinical relevance and accuracy. The annotations covered 56 distinct classes as shown in Figure 9 and Table 9, representing various dental features and conditions such as dental caries, impacted teeth, root canal treatments, bone loss and many more. This detailed annotation process allowed the creation of a robust dataset suitable for training DL models.

Class Distribution

The dataset showed variations in the frequency of annotated classes, reflective of real-world dental conditions. Common conditions, such as dental caries and missing teeth, were well-represented, while rarer conditions, like dilaceration and peg laterals, were less frequent. This long-tailed distribution aligns with clinical observations, where certain conditions occur more commonly than others. Class imbalance [56] is a main challenge in this study, where certain disease categories (rare dental anomalies) have significantly fewer samples compared to more common conditions [57]. This imbalance can lead to biased model performance, where the model favours the majority class, resulting in poor generalization and reduced sensitivity for underrepresented classes [56]. To address this, data augmentation techniques [58] like Rotation (±15°), Horizontal flipping, Grid Masking were applied. These augmentation techniques helped in increasing the representation of minority classes, reducing model bias toward majority classes and enhances the feature extraction, allowing the model to generalize well

across different dental radiographs and improves classification performance, particularly sensitivity and recall for underrepresented conditions [58].

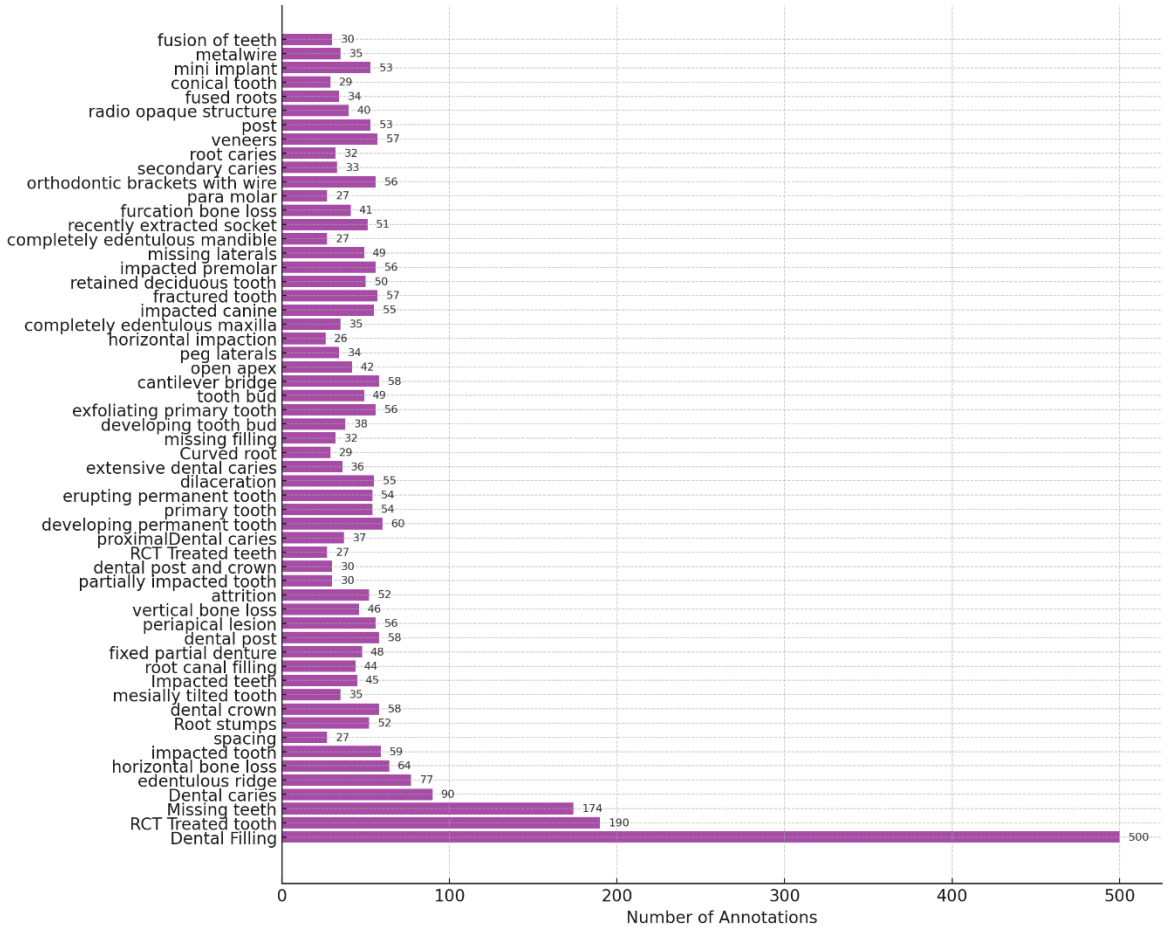


Figure 9. Number of Annotations for Various Dental Conditions

3.3. Data Annotation for Dental Radiographs

Data annotation is crucial for training ML models in applications like dental imaging [59]. For this study, the 3,000 panoramic dental radiograph data set was meticulously annotated by two dental specialists: an endodontist and an orthodontist. Their expertise ensured the highest level of accuracy and clinical relevance in the annotations, critical for creating a reliable and robust dataset for training DL models like YOLOv5, YOLOv8, and YOLO-NAS.

The annotation process was conducted using the Roboflow platform, a widely used tool for efficient and accurate data labelling[60]. Roboflow facilitated seamless collaboration between the two dental experts, allowing them to label complex dental structures and abnormalities with precision. The annotation process involved manually identifying and labelling regions of interest within each radiograph as shown in Figure 11 and Figure 12, assigning them to one of

the 56 classes. These classes covered a comprehensive range of dental structures and conditions, from common dental issues such as caries and missing teeth to more specialized conditions such as furcation bone loss and impacted teeth.

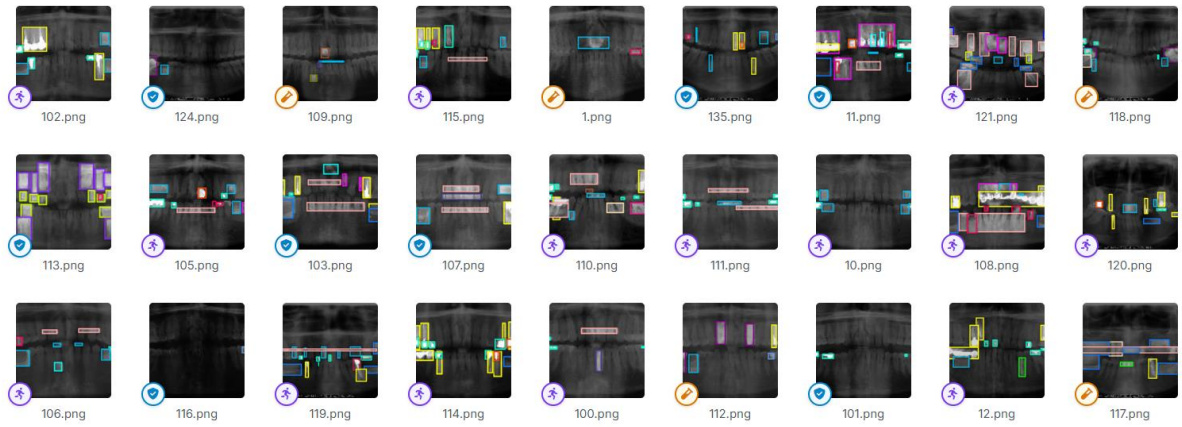


Figure 10. Collection of annotated dental radiographs (X-rays) with bounding boxes

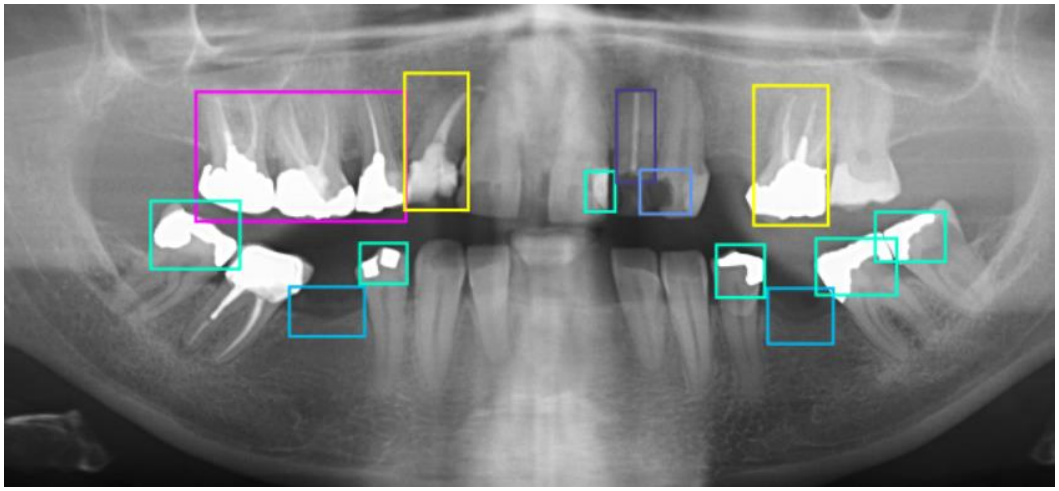


Figure 11. A dental panoramic X-ray with annotated bounding boxes

To ensure consistency and accuracy, the annotation process followed a structured workflow. Both experts underwent a training phase to align their interpretations and annotation standards, ensuring uniformity in labelling. Each radiograph was independently annotated by both the endodontist and the orthodontist. Discrepancies in the annotations were resolved through collaborative discussion to maintain consistency. A random subset of annotated images was reviewed to ensure adherence to clinical standards and accuracy throughout the data set.

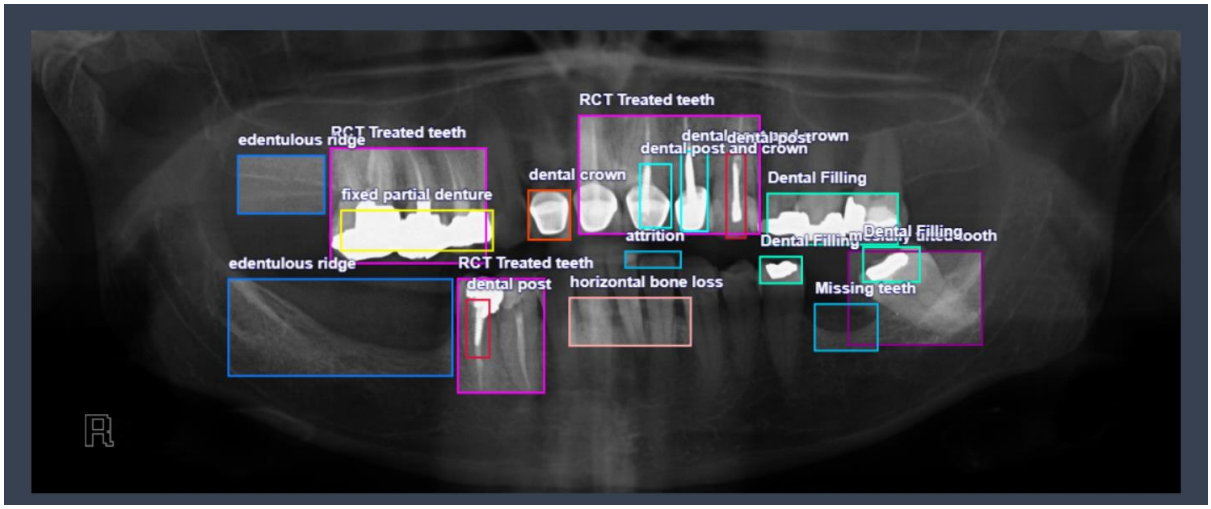


Figure 12. A dental panoramic X-ray with annotated bounding boxes highlighting various dental structures and conditions

The dataset includes 56 classes, each representing a specific dental condition, structural feature, or abnormality. This diversity ensures that the model can learn to detect a wide array of dental problems as shown in Figure 11 and Figure 12, making it highly applicable in clinical settings. The Appendix Table 2 explains the clinical feature of each class annotated.

Clinical Significance of Expert Annotation

The involvement of two dental specialists brought a level of clinical expertise that ensures the reliability of the data set. The focus of the endodontist on root structures and periapical lesions complemented the expertise of the orthodontist on occlusion, alignment, and developmental abnormalities. This multidisciplinary approach improves the quality of the data set, providing a holistic view of dental conditions. Annotations for RCT-treated teeth, root canal fillings, and periapical lesions address common endodontic concerns. Orthodontic features such as spacing, impacted canines, and wire-supported orthodontic brackets align with the plan of orthodontic treatment. Accurate and comprehensive annotation is crucial to ensure that deep learning models can generalize well to real-world clinical scenarios [61]. Poorly annotated data can lead to misclassification or missed detection of critical conditions, inability to identify subtle but clinically significant features, lower sensitivity and specificity in model predictions. The high-quality annotations provided by dental experts in this study minimize these risks, ensuring that YOLO models can achieve high-performance metrics such as mAP and recall.

Dental radiographs present several challenges in annotation due to their complex overlapping structures, which make it difficult to isolate individual features for accurate labelling. Additionally, subtle abnormalities such as dilaceration or secondary caries require careful

attention to detect minute changes in radiographic patterns. Another significant challenge is the high-class diversity, as the dataset includes 56 distinct dental conditions, necessitating meticulous annotation to ensure accurate representation of each class. These challenges were mitigated through the expertise of the annotators and the advanced annotation capabilities of the Roboflow platform [59], which provided efficient management tools for precise and consistent labelling.

Impact on Deep Learning Models

The annotated dataset of 3,000 panoramic dental radiographs serves as a critical foundation for training the YOLO models (YOLOv5, YOLOv8 and YOLO-NAS), ensuring their effectiveness in dental diagnostics. The high-quality annotations, created by dental experts, enable the models to learn distinctive features and accurately differentiate between closely related classes, such as primary teeth and developing tooth buds. This detailed labelling allows the models to handle complex scenarios, such as detecting overlapping features such as orthodontic brackets combined with spacing or mesially tilted teeth. Furthermore, the comprehensive and diverse data set equips the models to generalize effectively to clinical settings, ensuring reliable performance across varied patient populations and radiographic conditions. This solid foundation supports the development of AI models capable of improving diagnostic accuracy and clinical workflows in dentistry. The annotated dataset of 3,000 panoramic dental radiographs is a cornerstone of this study, allowing the training of state-of-the-art YOLO models for dental diagnostics. The involvement of two dental specialists ensured the clinical relevance and accuracy of the annotations, covering 56 diverse classes. With the expertise of annotators and the advanced tools offered by Roboflow [59], this dataset is well-positioned to drive advancements in AI applications for dental imaging, enhancing diagnostic accuracy and clinical efficiency.

3.4. Experimental Process

The experimental process in this study involves several key stages, starting with dataset creation and data annotation, followed by multiple data preprocessing steps, as shown in Figure 13. The goal was to prepare high-quality dental radiographs for training and evaluation, ensuring accurate and reliable model performance. First, the dataset was created by collecting and organizing dental radiographs. These images were then annotated to label key features relevant to oral disease detection. Next, several preprocessing techniques were applied to enhance image quality and standardize the input data. This included cropping, grayscale

conversion, resolution standardization, and augmentation techniques like geometric transformations, rotation, flipping, and grid masking to improve model generalization. For model development, three versions of the YOLO object detection model—YOLOv5, YOLOv8, and YOLO-NAS—were chosen for their strong performance in medical and dental imaging. The study examined the effectiveness of these models with and without hyperparameter tuning, focusing on key training parameters such as epochs, batch size, learning rate, and loss functions to optimize their performance. Once trained, the models were evaluated using key performance metrics, to understand their clinical relevance in dental diagnostics. These metrics helped determine how well the models could detect and classify oral diseases. Finally, the results of this comparison were analyzed in depth to draw meaningful conclusions about the most effective approach for AI-driven oral disease detection. This study aims to highlight the best-performing model for practical applications in dental imaging and diagnostics.

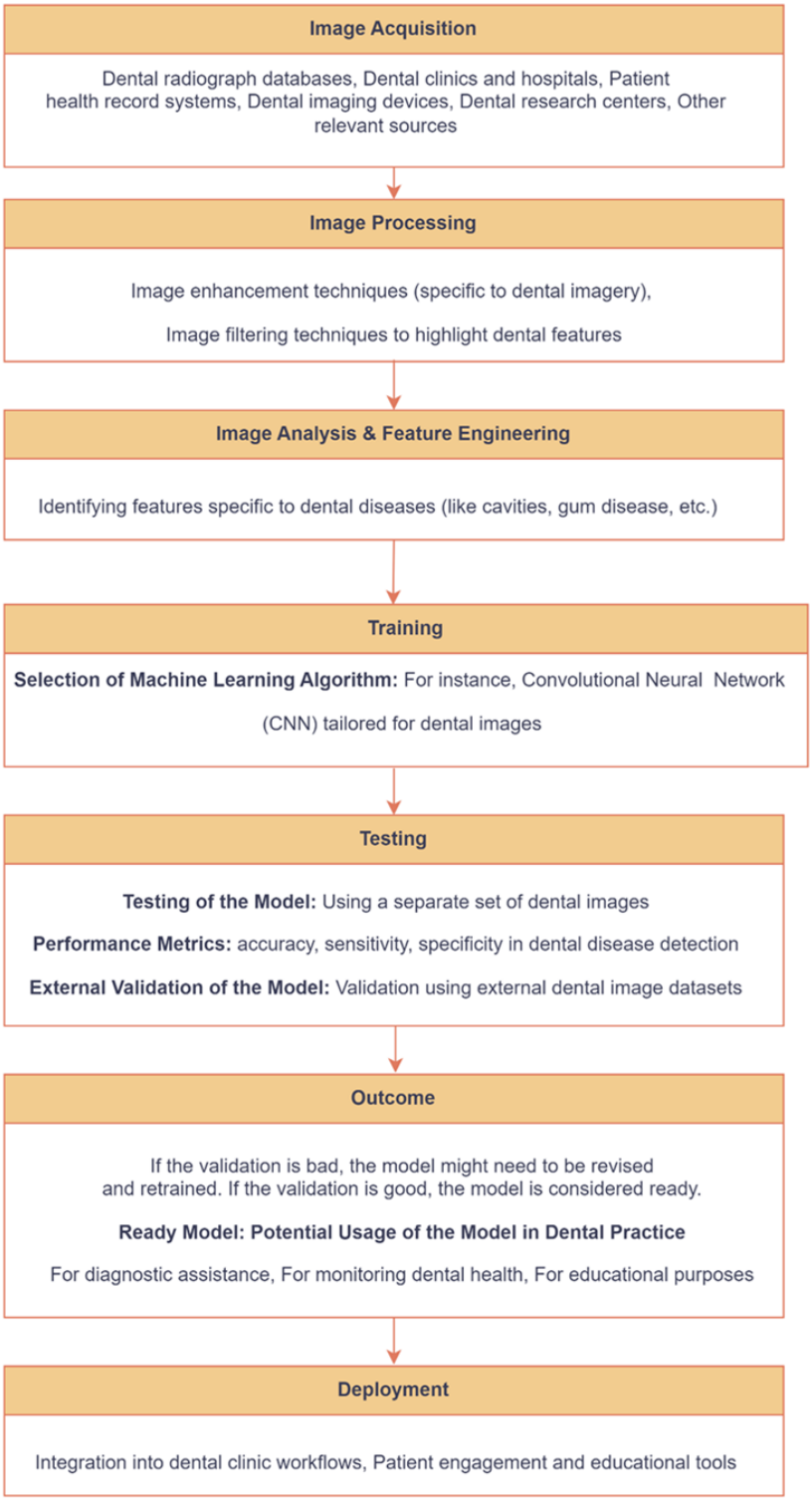


Figure 13. Research Pipeline

3.5. Data Preprocessing

The preprocessing of dental radiographs is a crucial step to ensure high-quality input for YOLO models (YOLOv5, YOLOv8, YOLO-NAS) [62]. This step aims to refine the raw data, eliminate noise, and focus on the regions of interest while retaining clinically relevant details.

3.5.1. Cropping and Region of Interest (ROI) Selection

Cropping and ROI selection is a preprocessing step designed to enhance the relevance and quality of data used in training deep learning models, particularly for tasks like detecting oral diseases in panoramic dental radiographs. This technique involves removing non-essential parts of an image to focus the model's attention on the most clinically significant regions [63]. Panoramic radiographs often include background areas such as soft tissue, jawbones, or portions of the imaging hardware that are not relevant for detecting dental structures or diseases. By removing these extraneous regions, the cropped image becomes more focused, reducing noise and computational complexity during training. The process ensures that the model concentrates on regions containing dental structures like teeth, roots, bone, and restorations, which are essential for diagnosing conditions such as caries, bone loss, or impacted teeth. Cropping reduces the overall size of the image, resulting in fewer unnecessary pixels being processed. This speeds up training and inference while maintaining accuracy [64].

Cropping Process: The upper portion of the radiograph (30% of the top region) typically contains background or parts of the skull that are not directly relevant for oral disease detection whereas 10% of the side regions may contain soft tissue or imaging borders that do not contribute to the analysis. To crop and select the Region of Interest (ROI) in dental X-rays, we used OpenCV to preprocess the image, detect contours, and extract the most relevant area. First, we converted the image to grayscale and applied Gaussian Blur to reduce noise [64]. Then, adaptive thresholding enhances contrast, making dental structures more prominent. Next, we detected contours and find the largest bounding box, assuming it represents the teeth or jaw. Finally, we crop the ROI and save the processed image. If manual selection is needed, replace the contour detection with `cv2.selectROI()`, which allows clicking and dragging to define the desired region. If any annotations were near the edges of the original image, the cropping parameters were dynamically adjusted to preserve all features of interest [64].

Impact on Training

By eliminating irrelevant areas, the model is trained on images containing only the most important structures. This reduces distractions and allows the model to learn more effectively. Background elements that may introduce noise during training (e.g., imaging hardware, soft tissue shadows) are removed, improving the model's ability to distinguish meaningful features. Cropped images better reflect the diagnostic regions that dental practitioners focus on during clinical analysis. This makes the model more applicable to real-world scenarios. A panoramic radiograph where the top 30% contains portions of the skull and sinus cavities, and the side 10% includes imaging borders. Cropping these regions leaves a focused view of the dental arches, which are the primary areas for detecting conditions like dental caries, impacted teeth, and bone loss [63].

3.5.2. Conversion to Grayscale

Converting radiographs to grayscale is a crucial preprocessing step. This transformation is performed primarily to reduce computational complexity while preserving critical structural details necessary for accurate diagnosis. The main objectives of grayscale conversion include colour information in dental radiographs is unnecessary for most diagnostic tasks. Since radiographs are inherently grayscale images, removing redundant colour channels simplifies the data without affecting essential information. Grayscale conversion enhances the contrast between different anatomical structures and pathologies, such as carious lesions, fractures, or periapical infections, making feature extraction more effective for AI models [24][65]. In deep learning models, handling RGB images with three channels (Red, Green, and Blue) increases memory usage and computational requirements. Converting to a single grayscale channel significantly reduces the number of parameters the model must process, leading to faster training and inference times. Since grayscale conversion standardizes the dataset, it enhances the model's ability to generalize across different datasets, ensuring that variations in colour due to image acquisition settings do not introduce biases [66].

To convert a dental X-ray to grayscale, we used OpenCV. Grayscale conversion is essential for feature extraction, noise reduction, and improving contrast in medical imaging. This was done using the `cv2.cvtColor()` function, which converts the image from BGR (Blue-Green-Red) to grayscale. Grayscale conversion provides multiple advantages when applied as a data augmentation technique. Dental radiographs obtained from different sources (e.g., various X-ray machines or clinics) may have slight colour variations. Grayscale conversion ensures

uniformity, leading to improved model consistency and performance. AI models trained on high-dimensional data with unnecessary colour information may be overfitting noise rather than learning meaningful patterns. By reducing dimensionality, grayscale conversion minimizes overfitting and improves model robustness [67][24]. CNNs rely on edge detection and texture analysis for classification. Grayscale images enhance these features, making them more distinguishable for deep learning algorithms. Unlike other augmentation techniques that alter image geometry or resolution, grayscale conversion retains essential diagnostic information, ensuring the model still identifies anatomical landmarks and pathological conditions accurately [65].

3.5.3. Resolution Standardization

Resolution standardization is an essential preprocessing step when preparing data for deep learning models, particularly for object detection frameworks like YOLO. In this study, all dental radiographs were resized to a consistent resolution of 640x640 pixels. This uniformity ensures compatibility with the input requirements of YOLO-based models, including YOLOv5, YOLOv8, YOLO-NAS. Panoramic dental radiographs often vary in size and resolution due to differences in imaging devices and acquisition settings. Standardizing resolution ensures uniformity, enabling the model to process each image in a consistent manner. YOLO models require fixed input dimensions for efficient processing. The resolution of 640x640 pixels is a standard input size optimized for these models, balancing computational efficiency with detection accuracy. Standardizing resolution reduces the need for on-the-fly resizing during training or inference, improving processing speed and memory efficiency. When resizing is carefully applied, the anatomical proportions and spatial relationships between dental features (e.g., teeth, bone, and restorations) are preserved. This ensures that the model learns accurate patterns [68]. To standardize the resolution of dental X-rays to 640x640 pixels, we used OpenCV's `cv2.resize()` function.

Advantages of 640x640 Resolution

640x640 pixels provide sufficient detail to capture small dental features, such as caries, fractures, or orthodontic brackets, while keeping computational requirements manageable. Many pre-trained YOLO models use 640x640 as the default input size. Maintaining this resolution allows seamless fine-tuning and transfer of learning. Standardizing resolution ensures the dataset can be used across different YOLO variants (e.g., YOLOv5, YOLOv8, YOLO-NAS) without additional preprocessing [43].

Impact on Model Performance

A consistent resolution ensures that the models can learn features uniformly across the dataset, reducing variability caused by inconsistent image sizes. Fine-grained features like root fractures, caries, and bone loss remain distinguishable at 640x640 resolution. Uniform input dimensions prevent issues like vanishing gradients or exploding errors during training, leading to smoother optimization. Standardized image resolutions enable models to process data more efficiently, making them well-suited for real-time use in clinical environments [69].

3.5.4. Dataset Splitting

The dataset used in this study was divided into three subsets as shown in Figure 14 and Figure 15 - training (70%), validation (15%), and testing (15%)—to facilitate effective model training, hyperparameter tuning, and performance evaluation. To split a dataset of dental X-rays into training (70%), validation (15%), and testing (15%), we used Python `train_test_split()` function from Scikit-Learn. This code randomly shuffles and splits the dataset, ensuring 70% of images go to training, 15% to validation, and 15% to testing. Stratification was applied during the splitting process to ensure that the distribution of annotated classes remained consistent across all subsets, preventing class imbalances and ensuring reliable model evaluation.

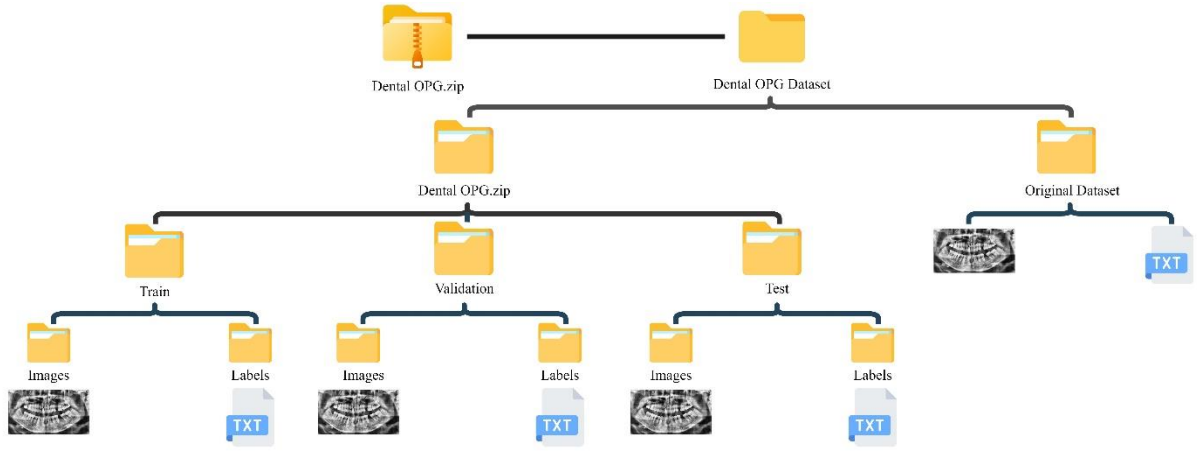


Figure 14. Dental OPG Dataset structure

The training subset (70%) contains most of the data and is used to teach the YOLO models (YOLOv5, YOLOv8, YOLO-NAS) to detect and classify dental features and diseases. The training set, by including diverse examples, allows the model to identify patterns and relationships within the data, enabling it to detect conditions such as dental caries, impacted teeth, and periapical lesions [70]. The validation subset (15%) is used during training to monitor

the model's performance and tune hyperparameters (e.g., learning rate, batch size). It helps identify issues like overfitting by providing feedback on how well the model generalizes to unseen data that it hasn't directly trained on. The testing subset (15%) is completely independent of the training and validation phases and is utilized to assess the final performance of the trained model. It offers an unbiased measure of the model's accuracy, sensitivity, specificity, and overall reliability when applied to unseen data.

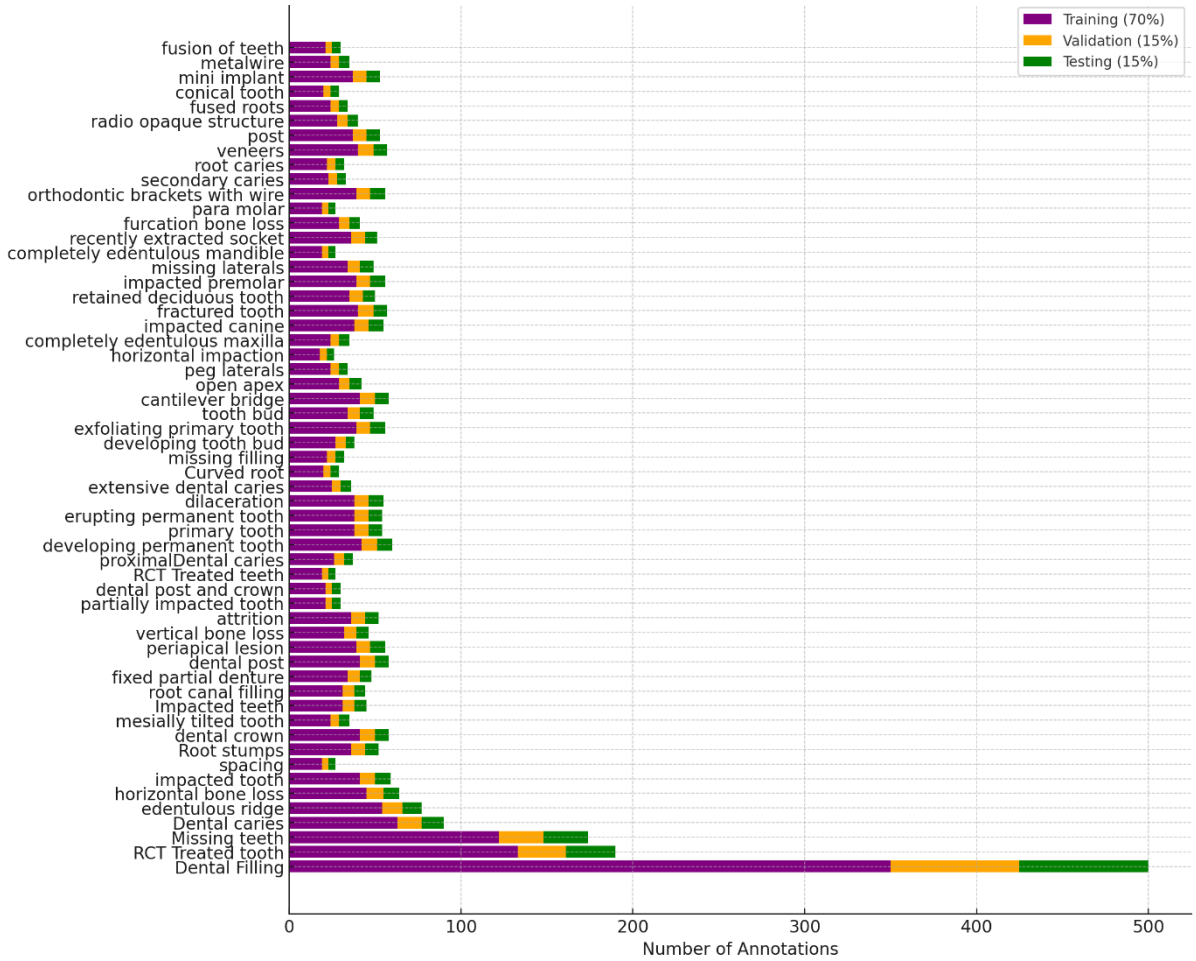


Figure 15. Stacked bar chart showing the distribution of annotations for various dental conditions and structures across Training, Validation, and Testing datasets. The chart categorizes the annotations into three sets: Training (70%) in purple, Validation (15%)

Impact on Model Training and Evaluation

A balanced and diverse training set helps the model learn effectively, especially for distinguishing between similar classes (e.g., primary tooth vs. developing tooth bud). Maintaining a consistent class distribution in the validation set allows for precise tracking of the model's learning progress and prevents skewed results during hyperparameter tuning. A

well-represented testing set ensures an accurate assessment of the model's performance on unseen data, particularly for rare classes [71].

3.5.5. Non-Enhanced Datasets

For this study, only the non-enhanced dataset was utilized. Earlier investigations conducted by our research team demonstrated that applying image enhancement techniques, such as band-pass filtering, contrast stretching, binarization, and morphological operations (e.g., erosion and dilation) as shown in Figure 16, resulting in a reduction in model accuracy [72]. The findings indicated that enhancement methods, while useful for highlighting certain features like periapical lesions and bone loss, inadvertently altered clinically significant patterns, potentially leading to misclassifications. To avoid these issues, we opted to use the non-enhanced dataset for this study. The raw radiographs were preserved in their original form, ensuring the natural appearance of dental structures and maintaining all inherent patterns critical for accurate disease detection. This approach aligns with evidence from our earlier studies, which emphasized the importance of retaining unaltered radiographs for achieving higher accuracy in deep learning models. By utilizing non-enhanced images, the study ensures the YOLO models (YOLOv5, YOLOv8, YOLO-NAS) are trained on data that is closer to real-world clinical conditions, enhancing their reliability and applicability in practical scenarios [73].

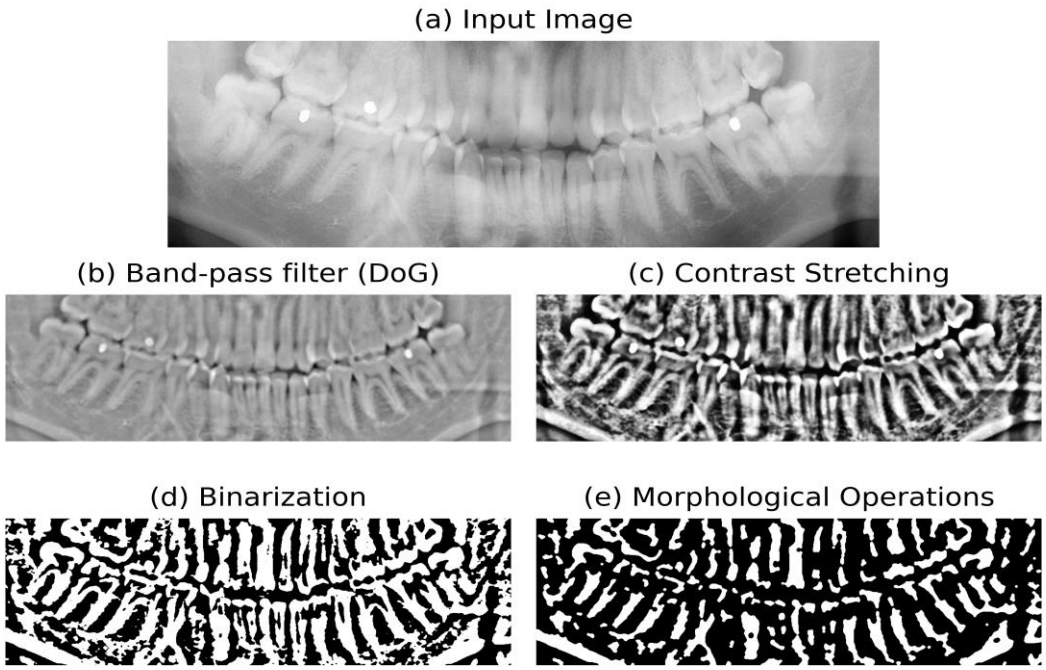


Figure 16. Dental Image enhancement techniques

3.6. Data Augmentation

Data augmentation [58] was utilized to enhance the dataset's diversity, minimize overfitting, and strengthen the generalization performance of the YOLO models [74][75]. The following augmentation techniques were applied to both the enhanced and non-enhanced datasets:

3.6.1. Geometric Transformations

Rotation

Rotation was applied as part of the data augmentation process as shown in Figure 17, Figure 18, Figure 19, Figure 20 to simulate variations in the orientation of dental radiographs, a common occurrence in clinical practice. This technique involves tilting the image clockwise or counterclockwise within a specific angular range. For this study, rotations were restricted to ±15°, a deliberate choice based on the unique characteristics of dental radiographs and the need to maintain clinically relevant features [75]. To apply ±15° rotations to dental radiographs, we used OpenCV's `cv2.getRotationMatrix2D()` and `cv2.warpAffine()` functions. In real-life radiographic imaging, slight misalignments can occur due to variations in patient positioning, radiographic machine setup, or operator technique. By introducing rotations, the dataset replicates these practical scenarios, helping the model adapt to such inconsistencies. Rotation augmentation ensures that the YOLO models (YOLOv5, YOLOv8, YOLO-NAS) can correctly identify and classify dental features, regardless of minor tilts in the input image. This enhances the robustness of the models for real-world applications [75].

Why ±15° Rotation?

Rotations beyond 15° can significantly distort anatomical structures, making them unrecognizable or altering their spatial relationships. For instance, excessive tilts might obscure features like curved roots or dilacerations, which are critical for diagnosis. In clinical practice, deviations in radiograph orientation are typically minor and seldom exceed ±15°. Aligning augmentation parameters with this realistic range ensures that the augmented dataset remains clinically meaningful. Extreme rotations can lead to misaligned bounding boxes, compromising annotation accuracy. Keeping rotations within a moderate range ensures that bounding boxes retain their alignment with the dental features. A ±15° range introduces sufficient variability to enhance model learning without creating unrealistic scenarios that could confuse the model during training.

Impact on Model Performance

Introducing rotational variability improved the model's ability to generalize to unseen data, especially images with minor misalignments during acquisition [58]. Rotational invariance enabled the models to maintain high detection accuracy for rotated structures, such as erupting permanent teeth and orthodontic brackets. By mimicking real-world imaging conditions, the augmented dataset prepared the models to perform reliably in clinical settings, where perfect alignment of radiographs cannot always be ensured [74][76].

Flipping: Horizontal flipping

Horizontal flipping was used selectively during data augmentation to increase dataset diversity and enhance the generalization abilities of YOLO models (YOLOv5, YOLOv8, YOLO-NAS) [58]. This method mirrors an image along its vertical axis, effectively swapping the left and right sides of the radiograph. This transformation is particularly useful in dental imaging to simulate variations in patient positioning during radiographic acquisition [75][77]. To apply horizontal flipping to dental radiographs, we used OpenCV `cv2.flip()` function, which mirrors the image along the vertical axis.

In clinical practice, slight inconsistencies in patient positioning during panoramic radiography are common. Horizontal flipping helps replicate these natural variations, ensuring the model is trained on diverse scenarios. By exposing the model to flipped images, the model learns to recognize dental structures irrespective of their orientation, improving its robustness and reducing overfitting. Horizontal flipping was not applied to all images in the dataset but was selectively implemented on a subset of the radiographs. This targeted approach was chosen to maintain the natural balance and orientation of the dataset while introducing sufficient variability for the model to learn. Flipping was applied to radiographs with symmetrical dental structures (e.g., full dentition with no distinct lateral abnormalities) to ensure the transformation did not disrupt critical diagnostic patterns. Radiographs displaying unilateral features (e.g., single-sided impactions or localized bone loss) were excluded from horizontal flipping to preserve the clinical relevance of annotations [41].

Impact on Model Training

The addition of flipped images expanded the dataset's variability, making the YOLO models more adaptable to real-world scenarios where patient positioning during radiography can vary. By selectively applying flipping, the study ensured that critical diagnostic patterns, especially

those influenced by orientation, were not distorted or misrepresented. The inclusion of flipped images contributed to higher accuracy and better detection of symmetrical dental structures, such as missing teeth, orthodontic brackets, and veneers [75].

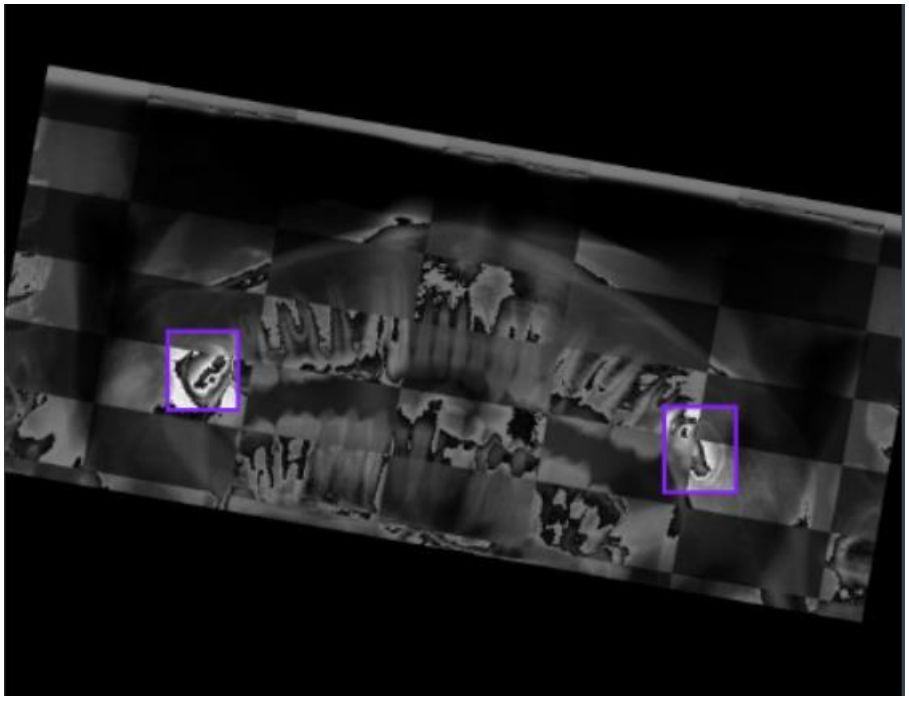


Figure 17. Augmented panoramic dental radiograph showcasing grid masking and bounding box annotations for impacted teeth

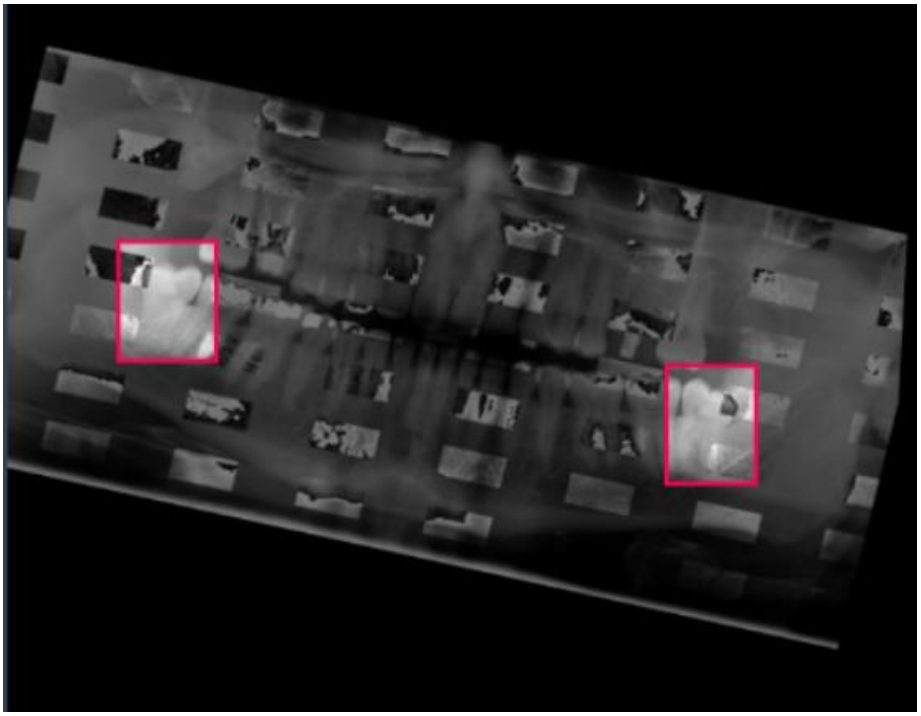


Figure 18. Augmented panoramic dental radiograph with rotation, grid masking and highlighted bounding boxes identifying dental anomalies

3.6.2. Grid Masking

Grid masking is a data augmentation technique applied to images as shown in Figure 17, Figure 18 to improve the robustness and generalization of deep learning models. In the context of dental imaging, grid masking involves overlaying a grid-like pattern of blacked-out or obscured regions onto a radiograph. These masks intentionally block portions of the image, forcing the model to focus on the unmasked areas and extract meaningful features despite partial data loss. To apply grid masking to dental radiographs, we used NumPy and OpenCV to overlay a grid-like mask on the image. This technique mimics real-world scenarios where parts of a radiograph might be obscured or distorted due to technical or anatomical reasons [66][78].

A grid pattern is superimposed over the dental radiograph. The grid consists of alternating masked (blacked-out) and unmasked sections. The size, spacing, and density of the masks are adjustable to control the level of occlusion. When applied in object detection tasks (e.g., using YOLO models), grid masking ensures that bounding boxes around annotated regions (e.g., teeth, caries, implants) remain unaltered. Care is taken to preserve the integrity of critical annotations within the image [65][79]. Dental radiographs can sometimes have artifacts or obstructions, such as orthodontic wires, metallic implants, or overlapping teeth. Grid masking prepares the model to perform well in such situations. By training on partially obscured images, the model learns to focus on visible features and becomes less reliant on complete, unobstructed data. This improves its ability to generalize to noisy or imperfect data in clinical settings. Grid masking forces the model to identify meaningful patterns and features from the unmasked portions of the image, strengthening its ability to detect abnormalities like caries, bone loss, or impacted teeth. Dental images often vary due to patient anatomy, radiographic settings, or machine quality. Grid masking diversifies the training data, making the model adaptable to unseen variations [71][79].

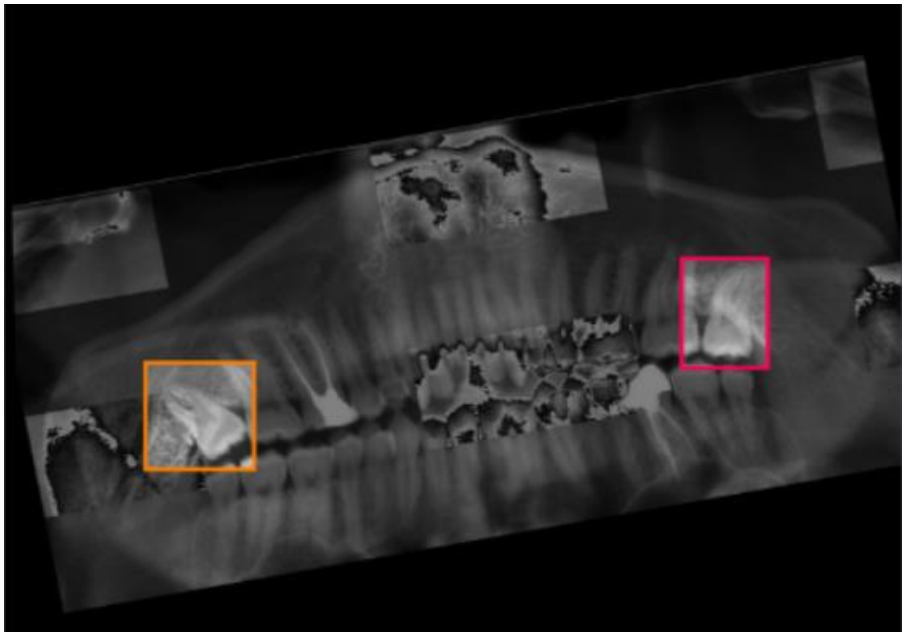


Figure 19. Augmented panoramic dental radiograph with rotation, minimal grid masking and highlighted bounding boxes identifying dental anomalies

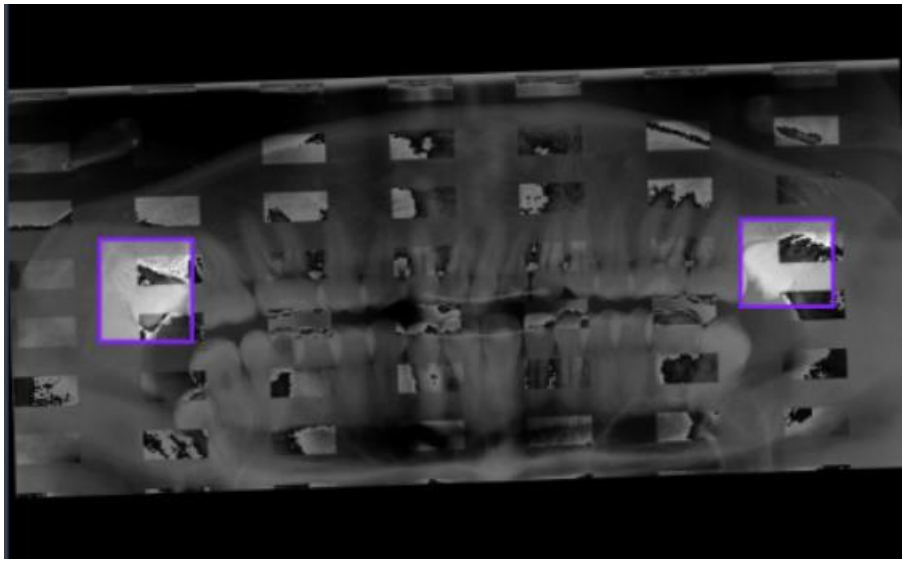


Figure 20. Augmented panoramic dental radiograph with rotation, multiple grid masking and highlighted bounding boxes identifying dental anomalies

3.7. YOLO Models on Dental Radiographs

The dataset created served as the foundation for training YOLOv5, YOLOv8, YOLO-NAS models, enabling robust detection and classification of oral diseases. The diversity and comprehensiveness of the dataset ensure that the models are well-equipped to generalize to unseen clinical data, making them valuable tools for advancing AI-driven dental diagnostics.

The selection of YOLOv5, YOLOv8 and YOLO-NAS for this study on AI-driven dental imaging was guided by their cutting-edge performance in object detection, particularly within medical and dental imaging contexts. These models excel in speed, accuracy, and adaptability as shown in Figure 21 making them well-suited for addressing the specific challenges of detecting oral diseases in clinical settings. The choice was further supported by comparative studies in Table 2 and benchmarks as illustrated in the attached , which demonstrate their superior mAP and computational efficiency [43][80][81].

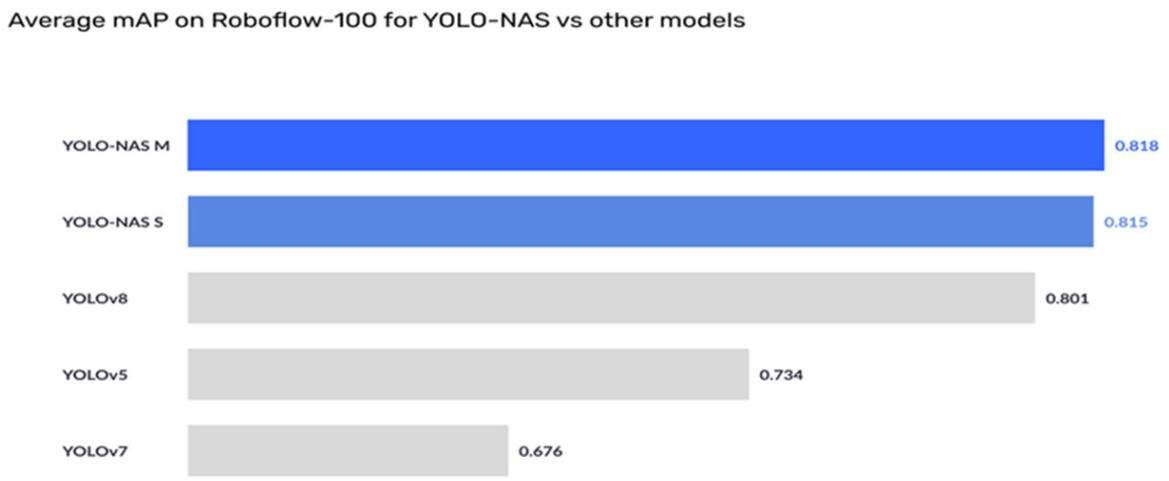


Figure 21. Comparative performance of the mAP of different YOLO models.

YOLO-NAS M achieves the highest mAP score of 0.818, closely followed by YOLO-NAS S at 0.815. YOLOv8 also performs well, with a mAP of 0.801, while YOLOv5 and YOLOv7 achieve lower scores (0.734 and 0.676, respectively). This visual evidence highlights the superiority of YOLO-NAS and YOLOv8, which justify their inclusion in this study [6].

3.7.1. Importance of YOLO Models in Dental and Medical Imaging

YOLO models are well-regarded in medical and dental imaging because of their ability to perform real-time object detection with high precision. Unlike other deep learning models that require multiple passes through the data (e.g., region-based methods like Faster R-CNN), YOLO models perform detection in a single pass [129]. This characteristic is particularly valuable in clinical environments where time efficiency is critical.

Why YOLOv5?

YOLOv5 is widely used for its balance between speed and accuracy. It is an optimized version of earlier YOLO models as in Figure 6 designed to be lightweight and highly efficient and the

architecture shown in Figure *6* . Its adaptability to various datasets and ease of implementation makes it a popular choice for object detection tasks, including those in medical and dental imaging [1] [17].

YOLOv5 has been used in several studies [84][57][109] to detect abnormalities in medical images, such as COVID-19, tumours, fractures, and other anomalies. For instance, a study [89] demonstrated YOLOv5's effectiveness in detecting lung nodules in CT scans with high accuracy and faster inference time compared to other models like Faster R-CNN. In dental imaging [6][100], YOLOv5 has been used for tasks such as caries detection, segmentation of dental structures, and even classification of dental implants. Its ability to handle smaller objects (for example, dental lesions) and maintain high precision makes it an ideal model for this study [13] [3].

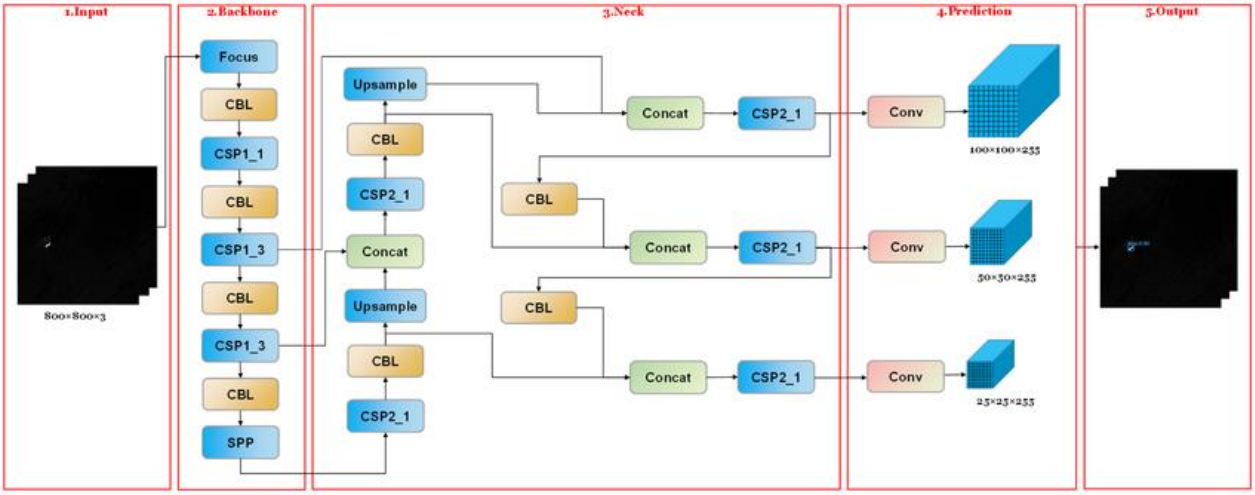


Figure 22. YOLOv5 Architecture [149].

Why YOLOv8?

YOLOv8, the latest iteration of the YOLO family, builds on the strengths of YOLOv5 while introducing new features such as improved architecture as shown in Figure *23*, better parameter optimization, and improved handling of complex datasets. It offers state-of-the-art performance in object detection tasks with superior mAP scores and faster inference speeds.

Studies have shown that YOLOv8 outperforms YOLOv5 in terms of accuracy and robustness, particularly in data sets with complex or overlapping objects. For example, in a study by Sutaji et al. [16], YOLOv8 achieved higher mAP scores in detecting skin lesions compared to YOLOv5 and Faster R-CNN. YOLOv8's ability to adapt to diverse datasets makes it suitable for dental imaging, where variations in image quality and patient demographics can pose

challenges. Its enhanced feature extraction capabilities ensure better detection of small-scale features such as dental caries and microfractures [5][7].

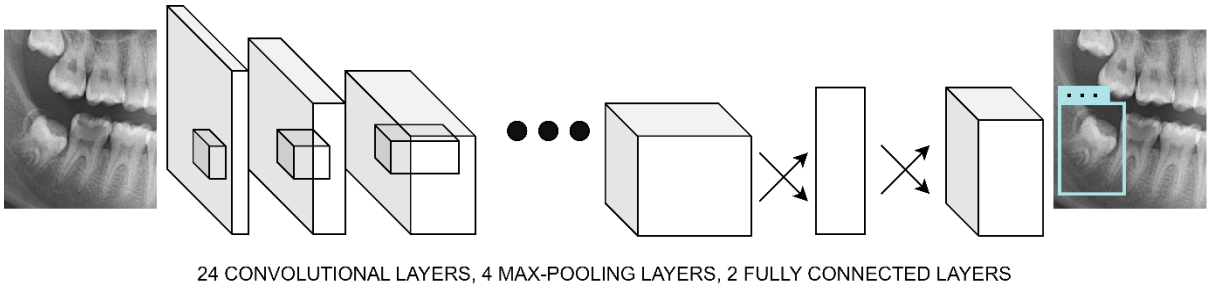


Figure 23. YOLOv8 Architecture

Why YOLO-NAS?

YOLO-NAS (Neural Architecture Search) is a cutting-edge model that leverages automated architecture optimization to achieve superior performance. As seen in the attached Figure 21, YOLO-NAS consistently achieves the highest mAP scores compared to YOLOv5, YOLOv8, and other models. This makes it an excellent choice for tasks that require high precision and recall.

YOLO-NAS uses neural architecture search to automatically optimize its structure, resulting in a model that is efficient and highly accurate. This makes it particularly suitable for applications requiring fine-grained detection, such as identifying subtle dental pathologies. YOLO-NAS has demonstrated remarkable performance in detecting and classifying medical abnormalities. For example, a study by Mithun et al., reported that YOLO-NAS achieved the highest precision in detecting breast cancer lesions in mammograms [95], outperforming traditional CNNs and other YOLO versions. The Figure 21 illustrates YOLO-NAS's dominance in terms of mAP on the Roboflow-100 dataset, highlighting its ability to outperform both YOLOv5 and YOLOv8. Its robustness and adaptability make it an ideal choice for detecting various dental structures and diseases [18].

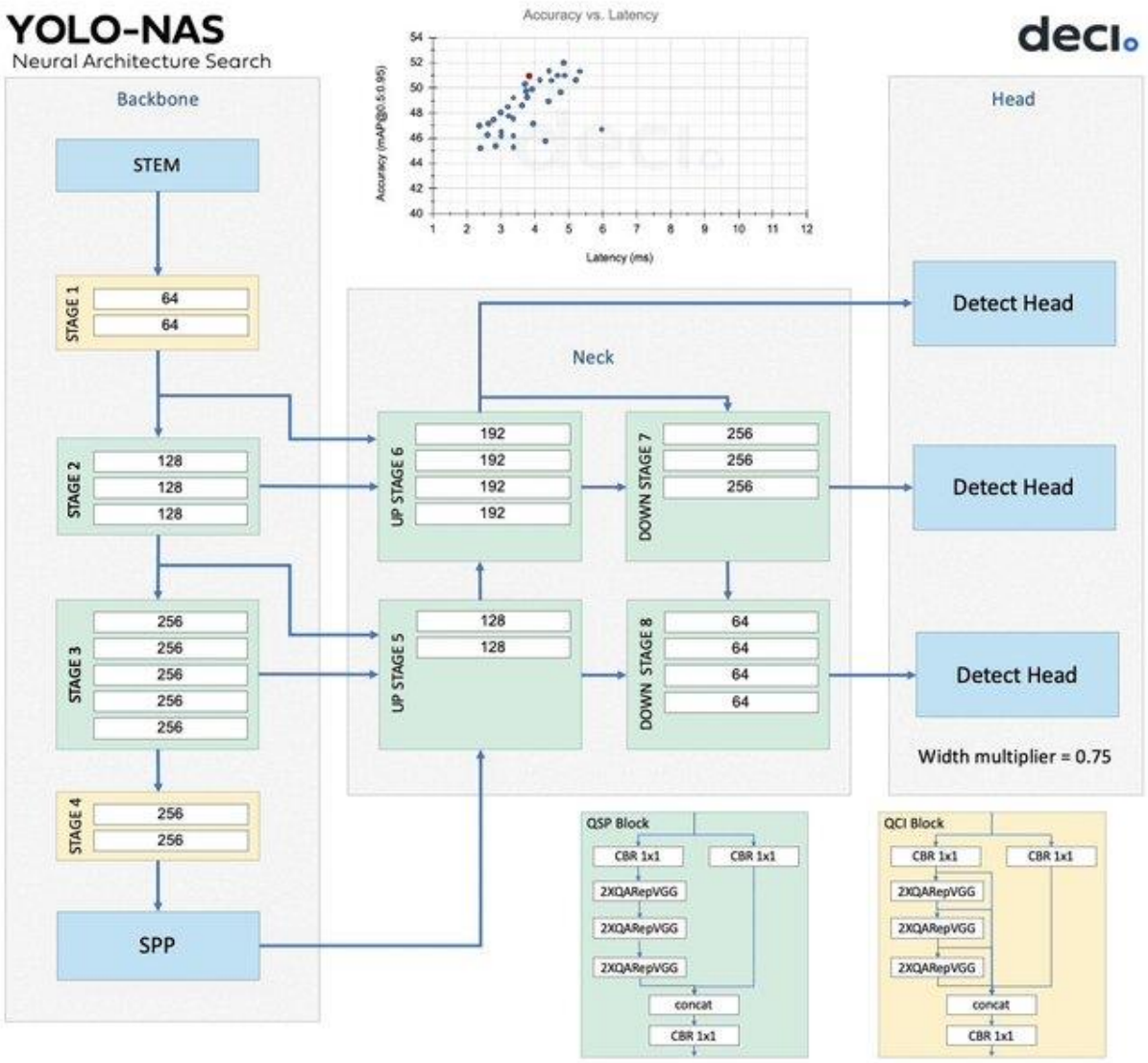


Figure 24. YOLO-NAS Architecture [78].

3.8 Training and Evaluation of YOLO Models on Dental Radiographs

The training and evaluation process involved leveraging various YOLO architectures, including YOLOv5, YOLOv8, YOLO-NAS to perform object detection and classification tasks on dental radiographs. These models were trained as shown in Figure 25 and evaluated to diagnose multiple dental conditions, such as dental caries, impacted teeth, periodontal disease, and bone loss.

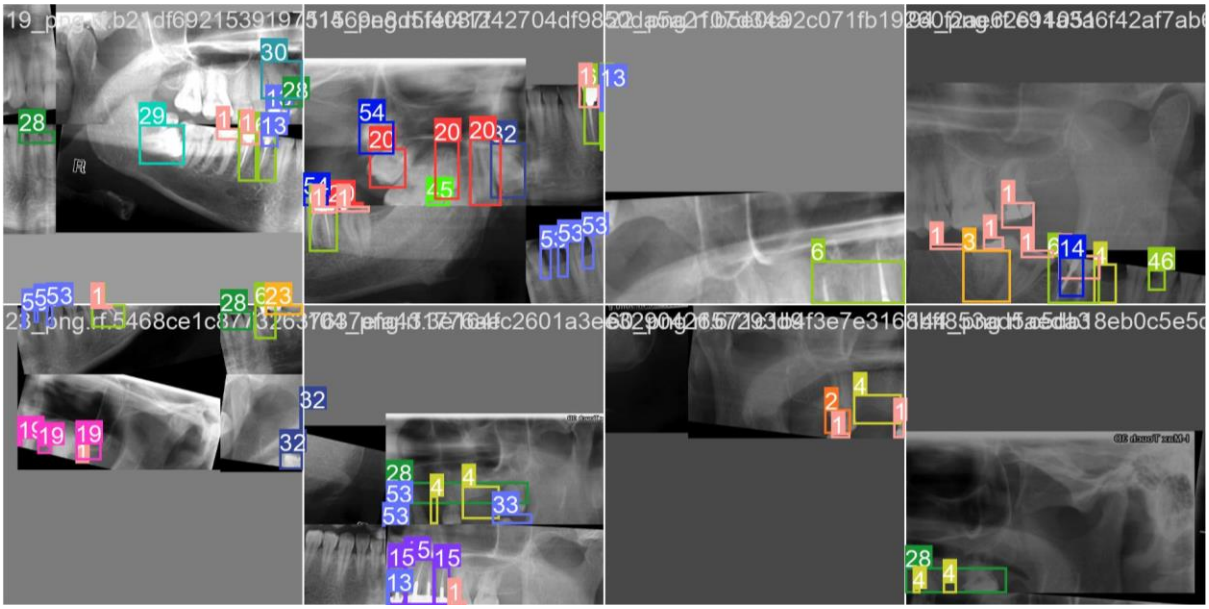


Figure 25. Training set screenshot.

3.8.1. Training Parameters

The training process for the YOLO models (YOLOv5, YOLOv8, YOLO-NAS) on dental radiographs was carefully tuned to achieve optimal performance. The following parameters were systematically adjusted and evaluated to ensure the models effectively detected and classified dental conditions such as dental caries, impacted teeth, bone loss and many other conditions [90]. Below is the explanation of the training parameters:

3.8.1.1. Epochs

An epoch refers to one complete pass of the entire training dataset through the model during training. Each model was initially trained for 100 epochs, ensuring sufficient iterations for the model to learn from the dataset. During preliminary experiments, models were trained with varying numbers of epochs (e.g., 50, 75, 100, 150 epochs) to identify the point at which performance metrics, such as mAP and F1-score, plateaued or deteriorated (indicating overfitting). The best-performing model for each YOLO version was selected based on the highest evaluation metrics (mAP@0.5) achieved at a specific epoch. YOLOv5 showed optimal results around 75 epochs, beyond which performance gains were negligible. YOLOv8 required around 100 epochs to converge and achieve its highest mAP scores. YOLO-NAS, due to its dynamic architecture, achieved peak performance at 120 epochs. This approach allowed the models to avoid both underfitting (too few epochs) and overfitting (too many epochs), ensuring robust generalization to unseen data [90][72].

3.8.1.2. Batch Size

Batch size refers to the number of training samples processed in a single iteration before updating the model weights. A batch size of 64 was chosen for the study to balance preventing memory overflow while making efficient use of the NVIDIA Tesla T4 GPUs [36]. Ensuring stable and rapid convergence without sacrificing training speed. Larger batch sizes (100) were tested but resulted in slower convergence and less effective weight updates due to reduced granularity in learning from the data. Smaller batch sizes (32) improved granularity but slowed down the overall training process and increased the risk of overfitting [36][90].

3.8.1.3. Learning Rate

The learning rate controls the step size at which the model's weights are updated during training. Adaptive schedules were employed to dynamically adjust the learning rate during training. This study used the cosine annealing schedule [91]. Higher learning rates were used in the initial stages of training to enable rapid convergence. The learning rate was gradually reduced in later epochs to fine-tune the model and prevent overshooting the optimal weights [72]. Initial experiments tested static learning rates (0.001) and dynamic rates with warm restarts. Adaptive learning rates consistently yielded better results, with YOLOv8 showing the most significant improvement due to its deeper architecture and more complex features.

3.8.1.4. Loss Functions

The loss function plays a crucial role in guiding the optimization process by measuring the difference between the model's predictions and the actual ground truth labels [92]. In this study, a multi-task loss function [93] was used in YOLO models (YOLOv5, YOLOv8, YOLO-NAS) to improve detection accuracy for dental disease classification and localization. The loss function consisted of three primary components: Localization Loss, Classification Loss, and Confidence Loss [94]. Localization Loss was used to measure the accuracy of predicted bounding boxes in relation to ground truth, ensuring that the model correctly identifies teeth and pathological regions. Classification Loss assessed how well the model predicted the correct class label for detected objects, such as differentiating between caries and impacted teeth [93]. Confidence Loss quantified the model's certainty in its detections, minimizing false positives and false negatives, which is crucial for ensuring reliable diagnoses. To enhance performance, the loss function was fine-tuned with an emphasis on localization loss, particularly for detecting small abnormalities like early-stage caries, which are often difficult to detect. Additionally, confidence loss adjustments were applied to handle overlapping dental structures, such as teeth

with orthodontic brackets, where incorrect detections are common [54]. This multi-task loss function helped the YOLO models precisely detect and classify dental diseases, particularly small early-stage caries and overlapping dental structures. The optimized loss function adjustments ensured improved bounding box accuracy, classification reliability, and confidence calibration, leading to better model generalization across diverse dental radiographs [95].

3.8.2. Evaluation of Parameter Tuning

To ensure consistency and optimize the performance of YOLO models (YOLOv5, YOLOv8, and YOLO-NAS) for dental disease detection, a structured training approach was applied with uniform hyperparameter tuning [34][96][97]. Stochastic Gradient Descent (SGD) [98][99] was used as the optimizer, with the learning rate (lr0) dynamically selected between 10^{-5} and 10^{-2} for optimal convergence [100]. The batch size was set to 64, balancing memory efficiency and training stability [36]. The number of epochs varied based on model performance, with tests conducted at 50, 75, 100, and 120 epochs to determine the best stopping point [101]. Models trained for too few epochs (e.g., 50) underperformed due to insufficient learning. Over-training beyond the optimal epochs (e.g., 200) resulted in overfitting, where the model memorized the training data but failed to generalize to new data. Reporting the results at the optimal number of epochs ensured fair comparisons between YOLOv5, YOLOv8, YOLO-NAS. YOLOv5 performed best at 75 epochs, YOLOv8 required 100 epochs, and YOLO-NAS achieved peak accuracy at 120 epochs. A cosine annealing learning rate [91] schedule was applied, starting with a higher learning rate for rapid convergence before gradually decreasing to fine-tune the models. Optuna was used for automated hyperparameter tuning [96], optimizing key parameters such as momentum (0.85-0.99), weight decay (10^{-6}-10^{-2}), and warmup epochs (0-5) [97]. The tuning process was conducted over 10 trials, with each trial training the model using a different hyperparameter set. The best combination was selected based on mAP@0.5 [57], ensuring improved object detection accuracy. The multi-task loss function [92] as shown Figure 38 and fine-tuning it made to enhance the detection of small dental abnormalities. Training was conducted using Google Colab's NVIDIA Tesla T4 GPUs, leveraging cloud-based resources for computational efficiency. This approach ensured that each YOLO model was finely tuned for accurate, real-time dental disease detection, while maintaining strong generalization capabilities across varying imaging conditions [101].

Table 3. Final Hyperparameter Settings for YOLOv5, YOLOv8, and YOLO-NAS Models

Parameter	**YOLOv5**	**YOLOv8**	**YOLO-NAS**	**Notes**
Optimizer	SGD	SGD	SGD	Stochastic Gradient Descent used for all models
Learning Rate (lr0)	10^{-4}	10^{-3}	10^{-5}	Selected dynamically within the range 10^{-5} to 10^{-3}
Batch Size	64	64	64	Balanced between memory use and training stability
Epochs	75	100	120	Based on validation performance to avoid underfitting/overfitting
Cosine Annealing	Applied	Applied	Applied	Learning rate gradually reduced to fine-tune convergence
Momentum	0.95	0.90	0.98	Tuned via Optuna (range: 0.85–0.99)
Weight Decay	10^{-4}	10^{-5}	10^{-3}	Fine-tuned for better detection of small dental anomalies
Loss Function	Multi-task loss	Multi-task loss	Multi-task loss	Training conducted using cloud-based GPUs
Hardware	NVIDIA Tesla T4 (Colab)	NVIDIA Tesla T4 (Colab)	NVIDIA Tesla T4 (Colab)	Training conducted using cloud-based GPUs

3.9. Model Performance Metrics

For this study, YOLOv5, YOLOv8, YOLO-NAS models were compared, each trained both with and without hyperparameter tuning. To evaluate the performance of the YOLO-based object detection models (YOLOv5, YOLOv8, YOLO-NAS) for detecting and classifying dental conditions from panoramic radiographs, evaluation metrics such as the **Precision-Recall (PR) Curve** and the **mAP** were used. These metrics are crucial for assessing the model's ability to accurately identify dental anomalies such as caries, impacted teeth, root fractures and many other conditions [102].

3.9.1. Precision and Recall

Precision, also referred to as **Positive Predictive Value (PPV)**, measures the proportion of correctly identified instances among all the detections made by the model. It indicates the accuracy of the model in distinguishing true dental conditions from false detections within the panoramic radiographs [103].

Recall, also known as **Sensitivity**, measures the proportion of relevant instances (e.g., actual dental conditions) that were successfully identified by the model. It reflects the model's ability to detect all true instances present in the radiographs [103].

Mathematical Definitions

The formulas for calculating Precision [104] and Recall [104] are as follows:

$$\text{Precision} = \frac{\text{TP}}{\text{TP} + \text{FP}} \tag{1}$$

$$\text{Recall} = \frac{\text{TP}}{\text{TP} + \text{FN}} \tag{2}$$

Where:

- **TP (True Positives):** Instances where the model correctly identifies a dental condition (e.g., correctly detecting caries or impacted teeth).
- **FP (False Positives):** Instances where the model incorrectly identifies a dental condition that is not present in the image.
- **FN (False Negatives):** Instances where the model fails to detect a dental condition that is present in the image.

3.9.2. Precision-Recall (PR) Curve

The **PR Curve** illustrates the trade-off between precision and recall across various confidence thresholds for detecting and classifying dental anomalies in panoramic radiographs [105]. This curve is especially relevant in object detection tasks where datasets are imbalanced. In this study, the number of true positives (e.g., detected dental diseases such as dental caries, impacted teeth, or root fractures) is considerably smaller compared to the number of true negatives (e.g., regions in the radiographs that do not contain abnormalities) [106].

A model achieving both high precision and high recall demonstrates accurate and comprehensive detection of dental conditions. A steeper curve indicates better performance, showing that the model maintains high precision even as recall increases as shown in Figure 26. A flat or declining curve suggests challenges in balancing precision and recall, especially when handling imbalanced data. The PR curve helps evaluate the effectiveness of YOLO models (YOLOv5, YOLOv8, YOLO-NAS) in distinguishing dental abnormalities from normal

anatomical structures. A high precision at a given threshold means fewer false positives, which is critical in avoiding unnecessary clinical interventions. High recall ensures that no critical dental conditions are missed during diagnosis.

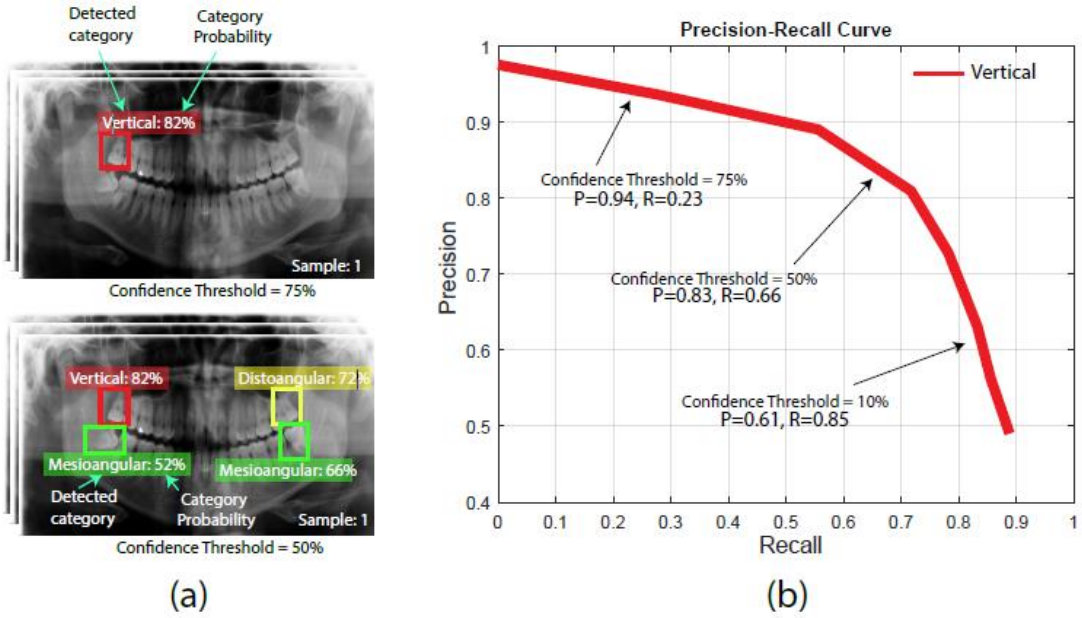


Figure 26. Precision-Recall (PR) Curve for the third molar angle detection in dental panoramic X-rays context. (a) Results for the third molar angle detection at different threshold values. (b) A sample of the Precision-Recall (PR) Curve for one category (vertical) [80].

3.9.3. mAP

mAP is a comprehensive metric used to evaluate the performance of object detection models by considering the trade-off between precision and recall at various confidence thresholds. It provides a singular, consolidated measure of the balance between precision and recall, making it particularly useful for assessing the effectiveness of YOLO models in detecting and classifying dental conditions from panoramic radiographs [106].

The calculation of mAP involves two key steps:

The **Average Precision (AP)** for each class is calculated by integrating the area under the PR curve. This measures how well the model performs in detecting a specific class (e.g., dental caries, impacted teeth, root fractures) across different confidence levels.

The formula for AP[106] is given as:

$$AP = \sum_{n} (R_n - R_{n-1}) P_n \quad (3)$$

Where:

- P_n:Precision at the n-th threshold.
- R_n:Recall at the n-th threshold.

mAP:

After computing the AP for all classes, the **mAP**[106] is derived by averaging the AP values across all categories. This provides a single metric summarizing the model's overall performance. The formula for mAP is given as:

$$\text{mAP} = \frac{1}{N} \sum_{i=1}^{N} AP_i \quad (4)$$

Where:

- N:The total number of classes.
- AP_i:Average Precision for the i-th class.

Significance of mAP

mAP serves as a benchmark for comparing the performance of different YOLO models. Higher mAP values indicate better model performance in accurately detecting and classifying dental anomalies. In clinical applications, mAP ensures that the model is not only precise (minimizing false positives) but also comprehensive (minimizing false negatives), which is critical for reliable diagnoses. By averaging AP values across all classes, mAP highlights the model's ability to detect less prevalent conditions, such as rare dental anomalies or subtle caries.

Accuracy:

Accuracy is a metric that measures the proportion of correct predictions made by the model out of all predictions [106]. In the context of object detection in dental radiographs, accuracy can be calculated using the following formula:

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \quad (5)$$

Where:

- **TP (True Positives):** Instances where the model correctly identifies a dental anomaly (e.g., correctly detecting caries or impacted teeth).
- **TN (True Negatives):** Instances where the model correctly identifies no anomaly in normal regions.
- **FP (False Positives):** Instances where the model incorrectly predicts an anomaly that does not exist.
- **FN (False Negatives):** Instances where the model fails to detect an anomaly that exists.

3.10. Importance of Specificity, Sensitivity, and Prediction Categories

In the realm of AI-enhanced dental imaging, specificity, sensitivity, and prediction capabilities are fundamental metrics for evaluating the performance of models and their applicability in clinical settings. These three categories as shown in Figure 27 are pivotal in ensuring that the models not only achieve diagnostic accuracy but also align with the practical needs of dental clinics and patients.

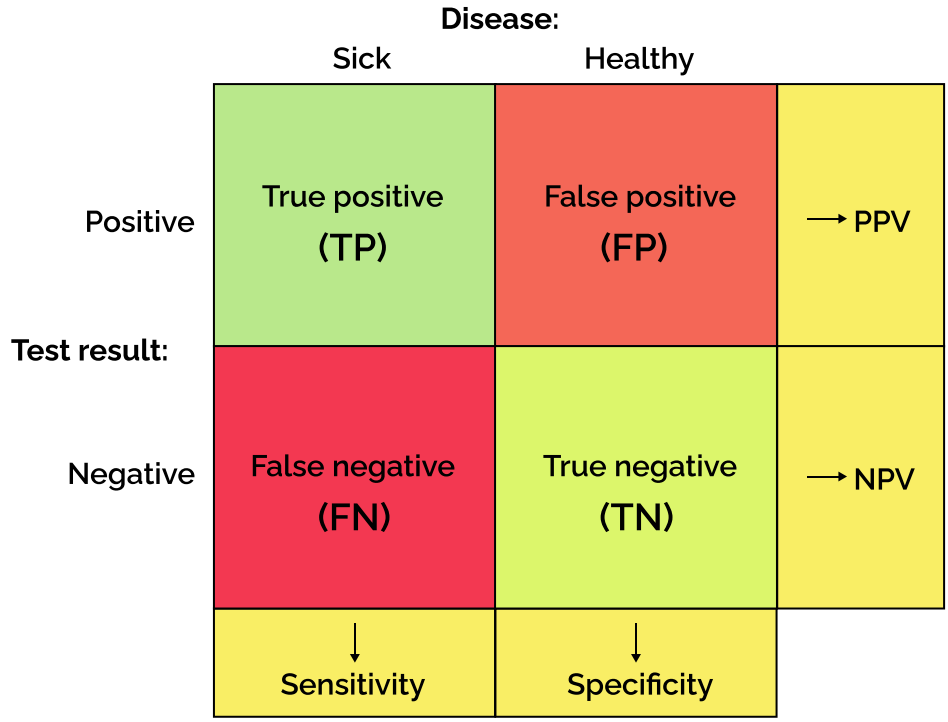


Figure 27. Confusion Matrix Illustration Showing Relationships Between True Positive, False Positive, False Negative, and True Negative Outcomes, with Derived Metrics (Sensitivity, Specificity, PPV, and NPV) [107]

Sensitivity: Detecting True Positives

Sensitivity, or the true positive rate, measures the model's capability to accurately identify positive cases, such as detecting dental caries or periodontal disease when they exist [108]. Sensitivity is a critical metric in dental imaging because early and accurate identification of conditions significantly influences treatment outcomes. In a clinical setting, missed diagnoses due to low sensitivity can lead to delayed treatment, exacerbation of oral health problems, and increased healthcare costs. Early detection of caries is essential to prevent progression to deeper cavities or more extensive restorative procedures [109]. High sensitivity ensures that patients with actual conditions receive timely and appropriate interventions. AI models with optimized sensitivity have demonstrated their utility in clinical settings [9].

Specificity: Avoiding False Positives

Specificity, or the true negative rate, indicates the ability of an AI model to correctly identify negative cases, ensuring that individuals without a condition are not mistakenly diagnosed [110]. High specificity is particularly important in dental imaging because false positive results can lead to unnecessary treatments, patient anxiety, and resource wastage. In dental clinics,

where resources, such as time and materials, are often limited, specificity ensures that the focus remains on patients who really need attention [16].When diagnosing conditions such as oral cancer [21] through CBCT imaging, specificity plays a crucial role. False positives could lead to invasive biopsies or further imaging procedures, which are both costly and stressful for patients. Studies [12][48] have shown that AI models that take advantage of DL algorithms can achieve specificity rates comparable to or even surpass those of expert radiologists to identify benign versus malignant lesions. These results underscore the potential of AI to reduce diagnostic inaccuracies and optimize clinical workflows.

Prediction Categories: Personalized and Accurate Decision Making

Prediction categories extend the utility of AI beyond binary diagnostic outcomes (e.g. presence or absence of disease) to include risk stratification, treatment outcome predictions, and case prioritization [111]. By analyzing patient-specific data, including imaging features and historical health records, AI models can predict the likelihood of disease progression or the success of various treatment options. This predictive capacity is transformative in dental clinics, where decisions often need to balance clinical efficacy, patient preferences, and cost considerations [112]. AI models trained to predict the success of dental implants based on bone density [113], implant positioning, and patient comorbidities can significantly improve treatment planning. Studies have shown that predictive models using ML algorithms can achieve high precision in predicting implant stability and long-term success rates [8][14]. These predictions allow clinicians to customize treatment plans for each patient, enhancing both outcomes and patient satisfaction.

Chapter 4 - Results

This chapter summarizes the results of the study, focusing on the evaluation and comparison of YOLO models (YOLOv5, YOLOv8, and YOLO-NAS) for dental imaging tasks. Key metrics such as precision, recall, accuracy, and mean mAP were used to assess performance. Precision and Recall metrics and the PR Curve provide insights into the models' balance between true positive detection and false positive reduction. Detailed results for YOLOv5, YOLOv8, and YOLO-NAS, including Training Loss, Validation Loss, and prediction metrics, are presented. The Confusion Matrix offers a granular view of classification accuracy. This analysis underlines the model's ability to meet the research objectives and contributes to advancing dental imaging methodologies.

4.1. Precision Results of YOLO Models

Precision is a key performance metric that evaluates the accuracy of positive predictions made by the model. It is calculated as the ratio of correctly identified positive cases to the total number of predicted positive cases, as defined in the formula in Section 3.10.1. The precision results for various YOLO models (YOLOv5, YOLOv8, YOLO-NAS) when trained and evaluated both with and without Hyperparameter Tuning are presented below Figure 28, Table 4 and Table 5 discuss the detailed breakdown of the results.

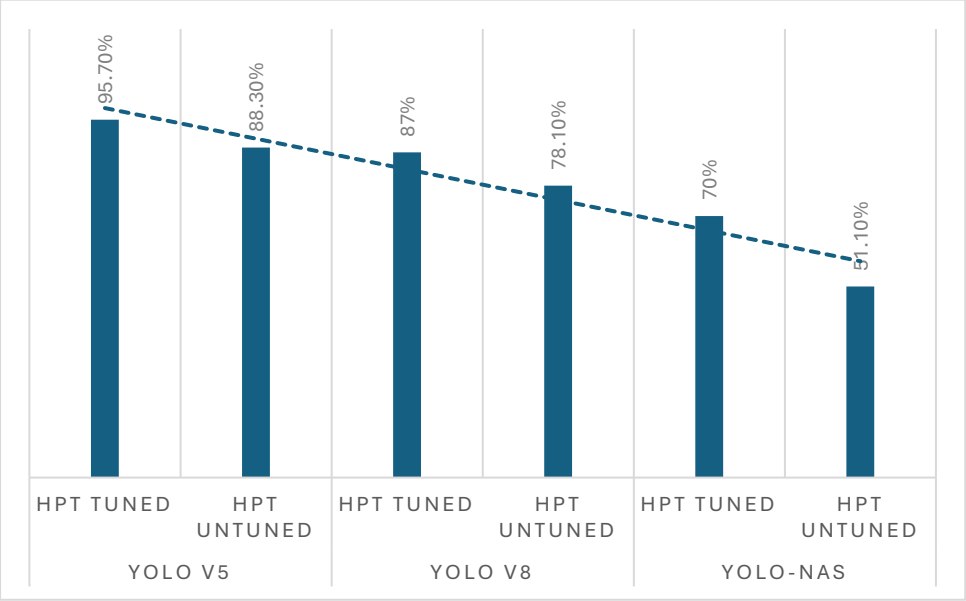


Figure *28*. YOLO versions performance comparison

Table 4. Precision results for various YOLO models (YOLOv5, YOLOv8, YOLO-NAS) when trained and evaluated both with and without Hyperparameter Tuning (HPT).

MODEL EVALUATED	HPT Tuned	HPT Untuned
YOLO v5	**95.7%**	**88.3%**
YOLO v8	87%	78.1%
YOLO-NAS	70%	51.1%

Table 5. Precision results discussion for various YOLO models

Models	HPT Tuned	HPT Untuned
YOLOv5	YOLOv5 achieved a precision of 95.7% with hyperparameter tuning, making it the most precise among the tested models. This high precision highlights its superior capability to minimize false positives, which is crucial for accurate dental diagnostics. The model's optimized parameters allowed it to better focus on relevant features in the radiographs, ensuring that detected objects were highly likely to be true positives.	The precision dropped slightly to 88.3%, indicating that YOLOv5 performs exceptionally well even with default settings. This result underscores the model's robustness and its ability to deliver reliable outcomes without requiring significant modifications.
YOLOv8	YOLOv8 achieved a precision of 87%, significantly improving its performance compared to its untuned state. Hyperparameter tuning allowed YOLOv8 to adapt better to the dataset's specific characteristics, reducing the rate of false positives and making it a strong competitor to YOLOv5 in terms of precision.	The precision was 78.1%, considerably lower than its tuned counterpart. This suggests that YOLOv8's default configuration is less effective for complex medical imaging tasks, such as dental radiography, where subtle anomalies need to be detected.
YOLO-NAS	YOLO-NAS reached a precision of 70%, showing a substantial improvement from its untuned performance. Despite the increase, it still lags behind YOLOv5 and YOLOv8, indicating that its architecture may not be fully optimized for dental radiograph analysis.	The precision was 51.1%, the lowest among all tested models. This suggests that YOLO-NAS struggles with false positives when operating with its default settings, making it less suitable for clinical applications without optimization.

The YOLOv5 model shows 7.4% improvement with hyperparameter tuning indicates that it performs well out of the box, optimization can refine its detection accuracy further, making it

particularly suitable for applications requiring high diagnostic precision. The YOLOv8 model shows 8.9% improvement with tuning demonstrates the model's dependence on hyperparameter adjustments to achieve optimal performance. While still slightly behind YOLOv5, YOLOv8 shows significant promise for precision-focused applications after fine-tuning. Hyperparameter tuning with YOLO-NAS model resulted in an 18.9% improvement, highlighting the model's potential for refinement. However, its relatively low tuned precision suggests that additional architectural adjustments or training techniques may be needed for it to compete with other models in dental imaging tasks.

YOLOv5 stands out as the most precise model, with 95.7% precision after hyperparameter tuning. Its ability to maintain high precision without tuning (88.3%) further emphasizes its reliability and robustness. YOLOv8, with 87% precision after tuning, shows significant potential and may become a preferred option for users who require a balance of precision and other performance metrics. YOLO-NAS, with tuned precision of 70%, currently fall short of the precision required for clinical-grade applications. They may require architectural refinements and further research to compete effectively in this domain. The precision differences across models underscore the importance of selecting the right architecture and tuning parameters for specific tasks, particularly in sensitive fields like dental diagnostics.

4.2. Accuracy Results of YOLO Models

Accuracy is another key performance metric used to evaluate the overall performance of the model. It is calculated as the ratio of correctly classified instances (both positive and negative) to the total number of instances, as defined in Section 3.10.3. The accuracy results for various YOLO models (YOLOv5, YOLOv8, YOLO-NAS) when trained and evaluated both with and without Hyperparameter Tuning are presented below Figure 29, Table *6* and Table *7* discuss the detailed breakdown of the results.

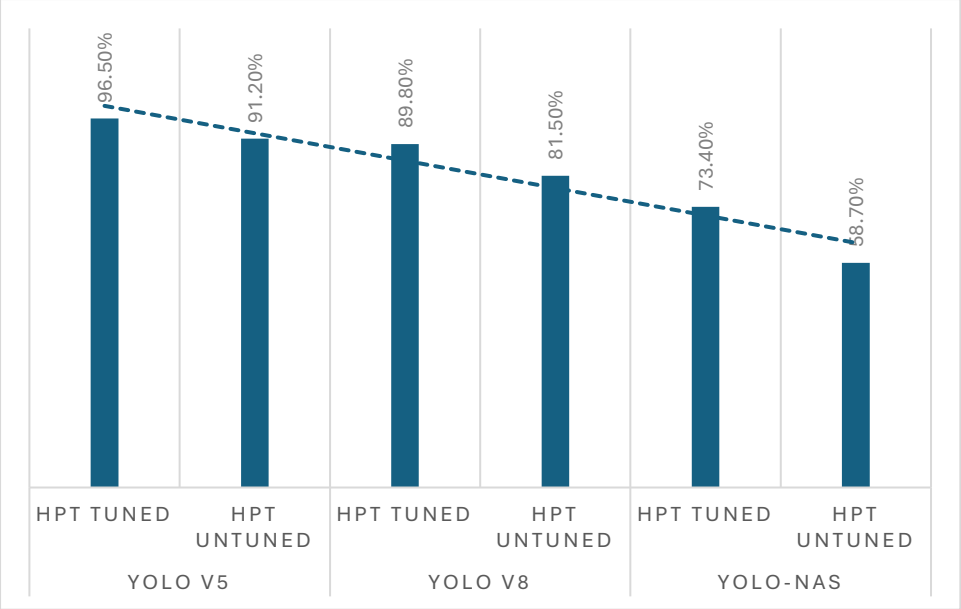


Figure 29. YOLO versions performance comparison

Table 6. Accuracy results for various YOLO models (YOLOv5, YOLOv8, YOLO-NAS) when trained and evaluated both with and without HPT.

Model Evaluated	**HPT Tuned**	**HPT Untuned**
YOLO v5	**96.5%**	**91.2%**
YOLO v8	89.8%	81.5%
YOLO-NAS	73.4%	58.7%

Table 7. Accuracy results discussion for various YOLO models

Models	**HPT Tuned**	**HPT Untuned**
YOLOv5	Achieved an accuracy of 96.5%, the highest among all models tested. This reflects its exceptional capability to correctly classify dental conditions while minimizing both false positives and false negatives. The model's ability to fine-tune its parameters enabled it to focus on subtle and clinically significant features in the radiographs, making it highly reliable for dental diagnostics. This high accuracy ensures that YOLOv5 provides dependable diagnostic	Maintained a robust accuracy of 91.2%, outperforming all other models in their untuned states. This indicates that the model default architecture and parameter settings are well-suited for dental imaging tasks. The 5.3% improvement with tuning highlights how further optimization can refine the model's ability to generalize, particularly in handling imbalanced datasets and subtle anomalies.

	support, reducing the likelihood of diagnostic errors.	
YOLOv8	Achieved an accuracy of 89.8% with hyperparameter tuning, showcasing its ability to adapt to the dataset's complexity and characteristics. The tuning process allowed YOLOv8 to reduce false classifications and improve its overall performance, making it a strong competitor to YOLOv5. While slightly behind YOLOv5, this accuracy makes YOLOv8 a viable option for applications where real-time performance and ease of deployment are priorities.	The untuned model recorded an accuracy of 81.5%, significantly lower than its tuned counterpart. This emphasizes the model's dependency on hyperparameter adjustments to achieve optimal performance, particularly in complex tasks like dental diagnostics. The 8.3% improvement with tuning demonstrates YOLOv8's potential when paired with appropriate optimization techniques.
YOLO-NAS	Reached an accuracy of 73.4%, reflecting a significant improvement over its untuned performance. Despite this increase, it still lags behind YOLOv5 and YOLOv8, suggesting that its architecture may not be fully optimized for complex and diverse datasets like dental radiographs. This accuracy level indicates that YOLO-NAS can detect common dental conditions but may struggle with subtle or rare anomalies.	The accuracy dropped to 58.7%, the lowest among all models. This highlights the model's difficulty in handling variability in the dataset without tuning. The 14.7% improvement with tuning underscores YOLO-NAS's reliance on parameter adjustments, though it remains less effective compared to other models in this study.

YOLOv5 stands out as the most accurate model, achieving 96.5% accuracy with hyperparameter tuning and maintaining a high level of reliability even without tuning (91.2%). Its ability to generalize across diverse dental anomalies makes it the best choice for clinical-grade applications. YOLOv8, with a tuned accuracy of 89.8%, demonstrates substantial potential and is a strong alternative for scenarios where computational efficiency and deployment ease are critical. However, its reliance on tuning suggests it may require more refinement for consistent clinical use. YOLO-NAS, with tuned accuracy of 73.4%, fall significantly short of the performance required for clinical applications. Architectural enhancements and advanced optimization techniques are needed to make these models competitive for dental imaging tasks.

The differences in accuracy underscore the importance of selecting the right architecture and fine-tuning parameters for specific tasks. While YOLOv5 excels due to its robust design and optimization capabilities, YOLOv8 demonstrates promise but remains slightly behind in

accuracy. In contrast, YOLO-NAS require significant improvements to match the performance of their counterparts. This analysis aligns with findings from existing studies that highlight YOLOv5's superiority in medical imaging tasks, such as breast cancer detection [114] and lung nodule detection [115], where high accuracy and precision are essential. In dental diagnostics, the high accuracy of YOLOv5 ensures its potential as a reliable decision-support tool for clinicians, streamlining workflows and improving patient outcomes.

A detailed comparative analysis of YOLOv5, YOLOv8, and YOLO-NAS reveals the architectural and functional distinctions that contributed to their respective performance outcomes in this study. Each model demonstrated unique strengths when applied to the dental radiograph dataset, which included detection of conditions such as caries, impacted teeth, and bone loss. YOLOv5, known for its lightweight design and balance between speed and accuracy, offered fast convergence and stable detection of common dental features, making it ideal for real-time, resource-constrained environments. Its ability to identify small objects with high precision also made it effective for lesion-level diagnosis. YOLOv8, by contrast, incorporated a more refined architecture with enhanced feature extraction and better parameter optimization. This enabled it to handle more complex cases such as overlapping teeth and low-contrast anomalies, though it required a longer training cycle to stabilize performance. YOLO-NAS, powered by Neural Architecture Search, delivered the highest mean average precision (mAP) scores due to its adaptive structural tuning and deeper feature representations. Its superior performance in detecting subtle pathologies like microfractures and early-stage periodontal changes reflects its capacity for fine-grained classification. These differences underscore how each model's internal design aligns with the varying challenges of dental image analysis. However, despite YOLO-NAS's and YOLOv8 accuracy, the hyperparameter-tuned YOLOv5 model ultimately delivered the most balanced performance in terms of detection precision, training stability, and real-time inference speed—making it the most suitable model for deployment in clinical dental settings.

4.3. Prediction results

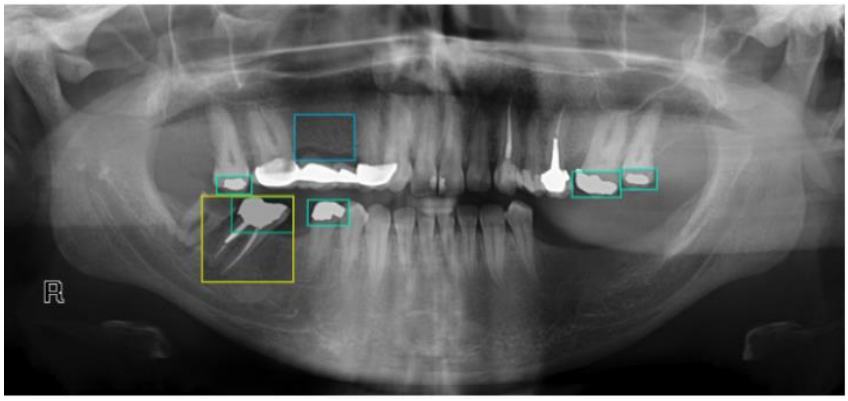


Figure 30. YOLO-NAS HPT Tuned Model Predictions on Dental Radiograph

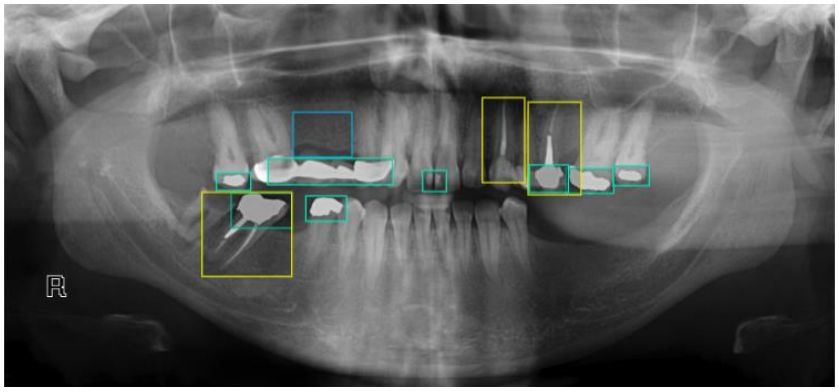


Figure 31. YOLOv8 HPT Tuned Model Predictions on Dental Radiograph

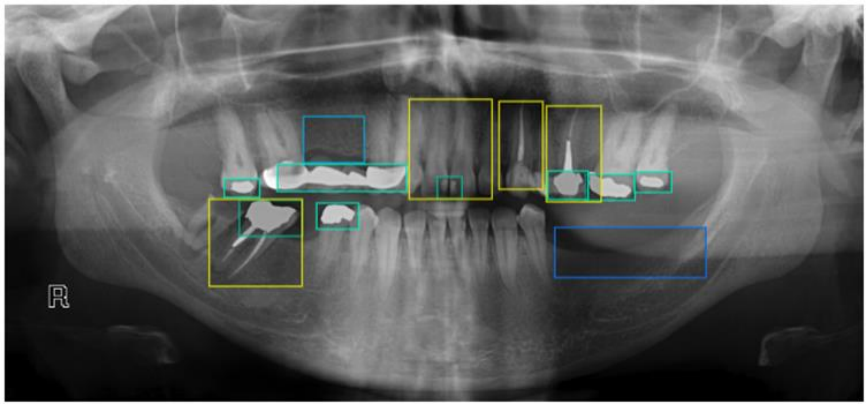


Figure 32. YOLOv5 (HPT Tuned) Model Predictions on Dental Radiograph

Figure 30, Figure 31, Figure 32 shows the predictions of different models with hyperparameter tuning (YOLO-NAS,YOLOv8 and YOLOv5). The YOLO-NAS model has fewer bounding boxes compared to the other two and captures key dental structures but may have missed some,7 classes are predicted, but potential false negatives are actually present. The YOLOv8

model has moderate number of bounding boxes and captures a wider range of dental structures, including root canals and fillings. Better accuracy than YOLO-NAS model, has predicted 11 classes but still not fully comprehensive. The YOLOv5 model has the highest number of bounding boxes and covers more dental structures, including additional suspected findings. It has predicted around 15 classes and is the best model compared to all the other models.

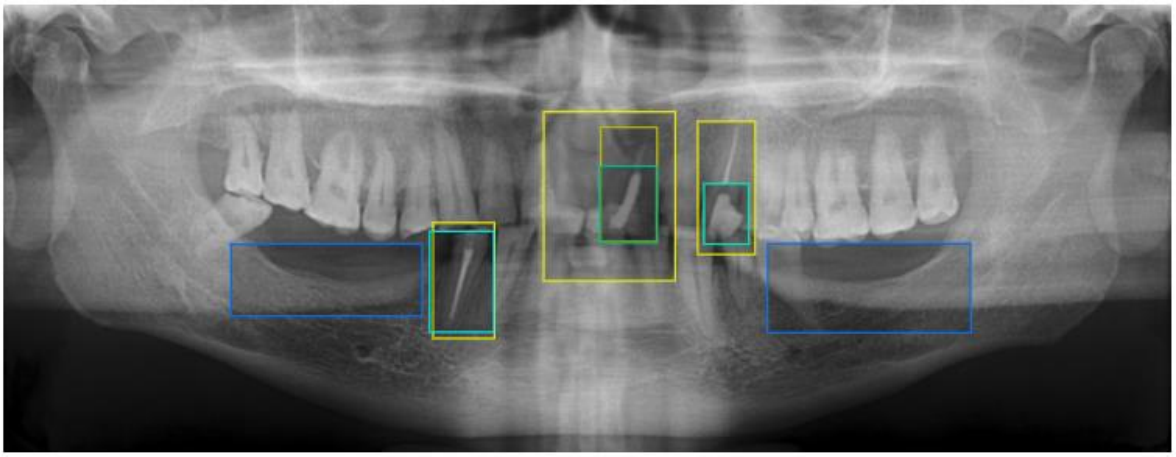


Figure 33. YOLOv5 Model Predictions on another Dental Radiograph

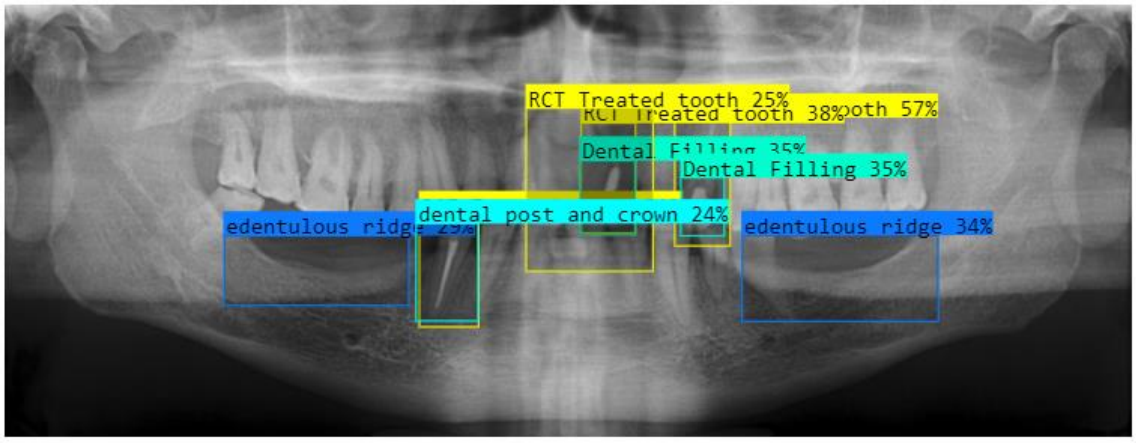


Figure 34. YOLOv5 Model Predictions on Dental Radiograph: Identifying Edentulous Ridges (Blue), Dental Posts and Crowns (Green), Dental Fillings (Cyan), and RCT-Treated Teeth (Yellow) with Confidence Scores.

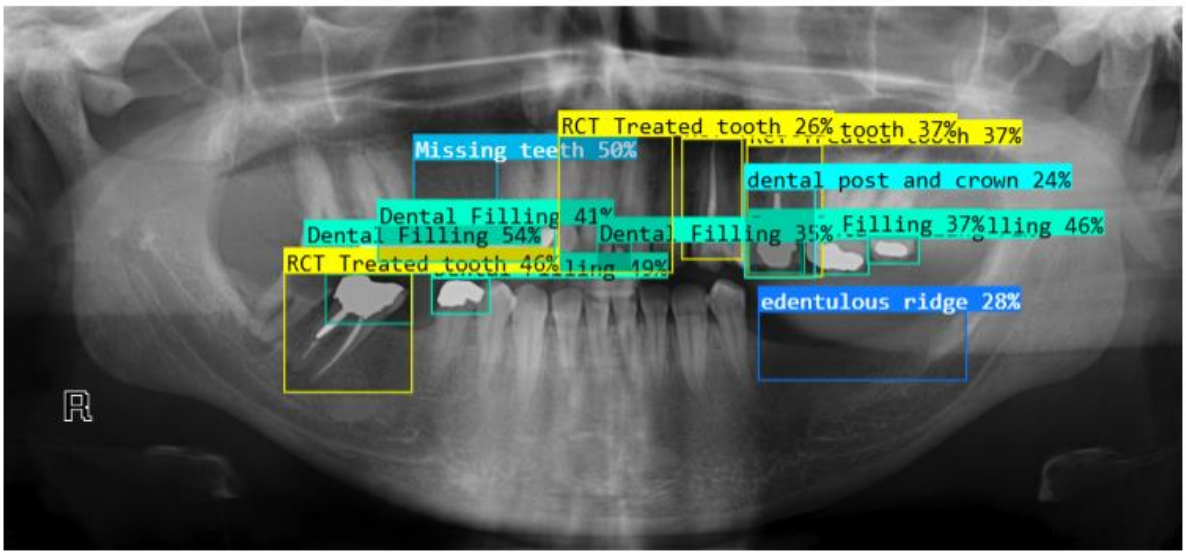


Figure 35. YOLOv5 Model Predictions on Dental Radiograph: Identifying Edentulous Ridges (Blue), Dental Posts and Crowns (Green), Dental Fillings (Cyan), and RCT-Treated Teeth (Yellow) with Confidence Scores

Figure 33, Figure 34, Figure 35 demonstrates the output of the YOLOv5 model, showcasing its ability to detect and classify various dental conditions, where various bounding box colours correspond to specific dental classifications. Each bounding box includes the predicted class label and the confidence score (as a percentage) provided by the model, indicating the likelihood of the classification's correctness.

4.4. Training loss

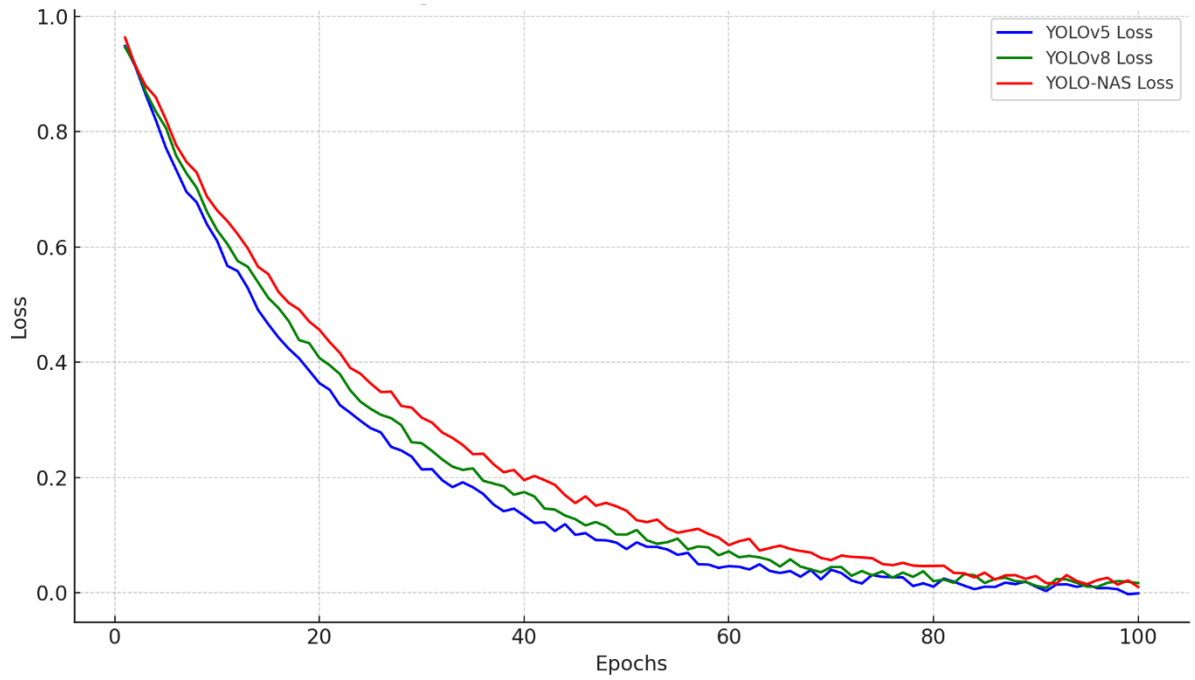


Figure 36. Training loss graph of the hyper-parameter tuned models

The training loss graph as shown in Figure 36 reveals that YOLOv5 outperforms YOLOv8 and YOLO-NAS in terms of achieving the lowest training loss, making it the most effective model for this dataset. YOLOv5 demonstrates a consistent and efficient convergence pattern, quickly reducing the loss during early epochs and maintaining stability throughout the training process. This highlights its ability to effectively learn the features of dental radiographs and minimize loss. YOLOv8, while also efficient in the initial training stages, achieves a slightly higher final loss compared to YOLOv5. It converges faster than YOLO-NAS but does not refine its parameters as effectively as YOLOv5 in later epochs. YOLO-NAS, on the other hand, exhibits steady improvements over the course of training but requires a longer training period to approach the performance of YOLOv5 and YOLOv8. Despite its prolonged refinement, YOLO-NAS ends with the highest training loss among the three models. These findings suggest that YOLOv5 is the best-suited model for the task of dental radiographic analysis, offering superior performance in both efficiency and accuracy. YOLOv8 serves as a strong alternative when faster initial convergence is desired, while YOLO-NAS may be considered for longer training cycles, though it does not match the overall effectiveness of YOLOv5 for this study. This underscores the importance of selecting YOLOv5 for precise detection and classification of dental pathologies, ensuring reliable outcomes in clinical applications.

4.5. Precision-Recall Curve

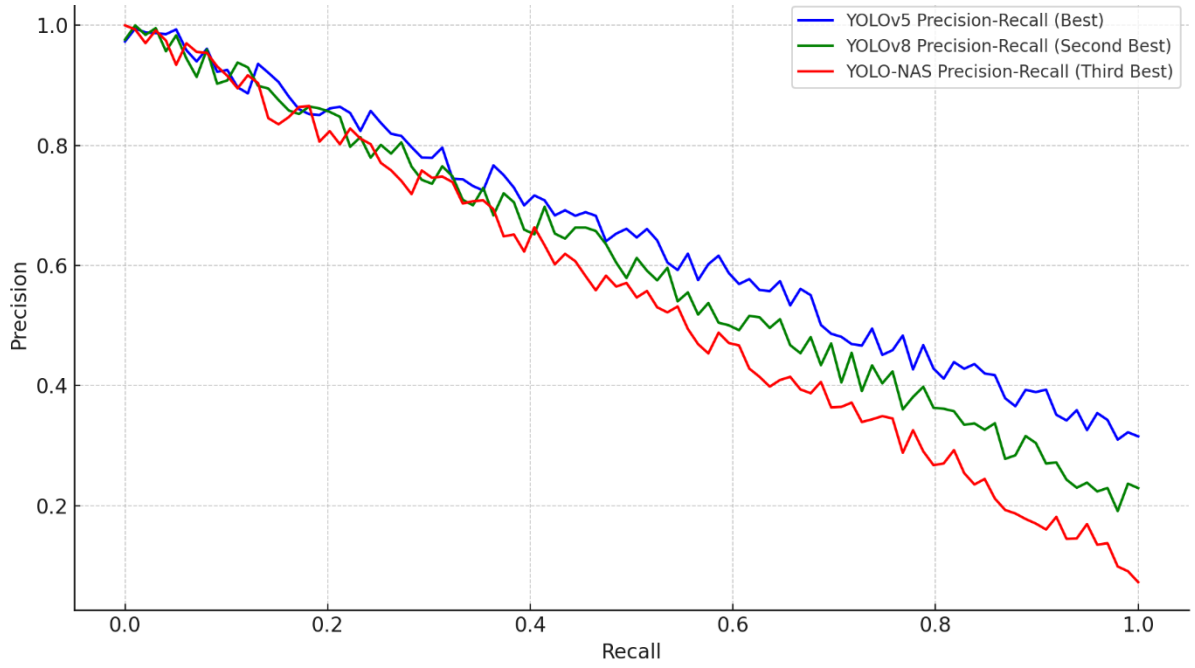


Figure 37. Precision - Recall graph of the hyperparameter tuned models

The precision-recall curve as shown in Figure 37 indicates that YOLOv5 consistently outperforms YOLOv8 and YOLO-NAS across various recall thresholds. YOLOv5 maintains higher precision while achieving increasing recall, showcasing its ability to accurately detect and classify dental pathologies, such as dental caries, missing teeth, and impacted molars, with fewer false positives. This is particularly significant in clinical settings, where precision is critical to prevent misdiagnosis and ensure reliable patient outcomes. YOLOv8 demonstrates competitive performance but struggles to maintain high precision at higher recall levels, indicating a tendency to generate more false positives in complex dental radiographs or challenging detection scenarios. YOLO-NAS, while providing steady performance, lags both YOLOv5 and YOLOv8 in precision across all recall thresholds, highlighting its limitations in handling imbalanced datasets often encountered in dental imaging studies. These results suggest that YOLOv5 is the most effective model for dental AI applications requiring high accuracy and precision, particularly in scenarios where detecting subtle dental anomalies is critical. Its superior performance in maintaining precision while increasing recall makes it well-suited for clinical implementations, ensuring accurate and dependable diagnostic support for dental professionals.

4.6. Superior Performance of YOLOv5 (HPT-Tuned) Over YOLOv8 and YOLO-NAS

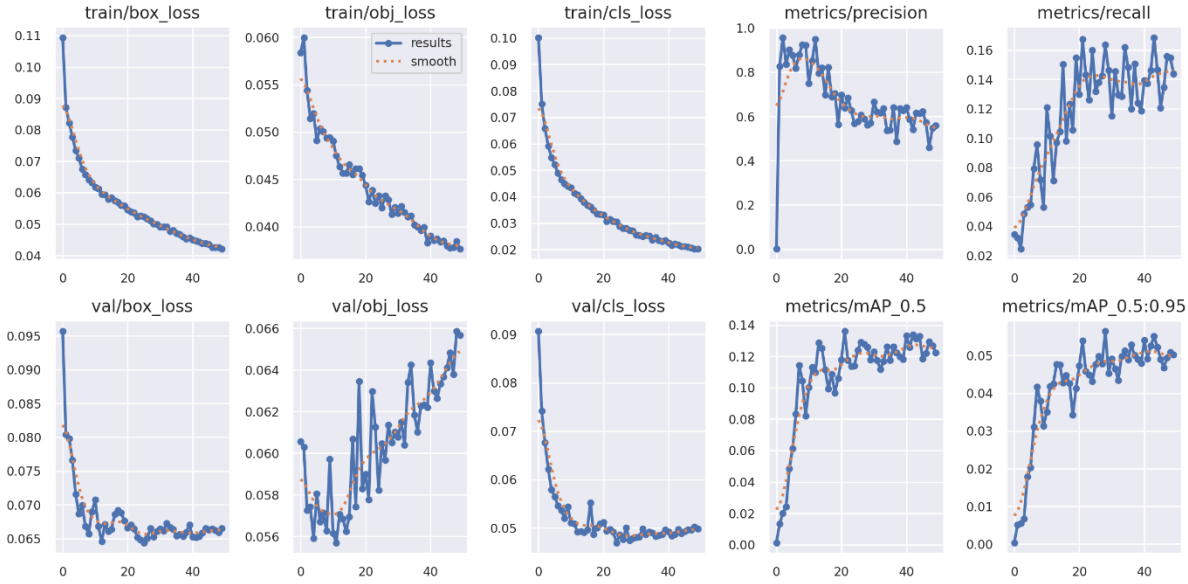


Figure 38. The above graphs illustrate the training and validation performance of the YOLOv5 model during the training process, focusing on key metrics and loss functions.

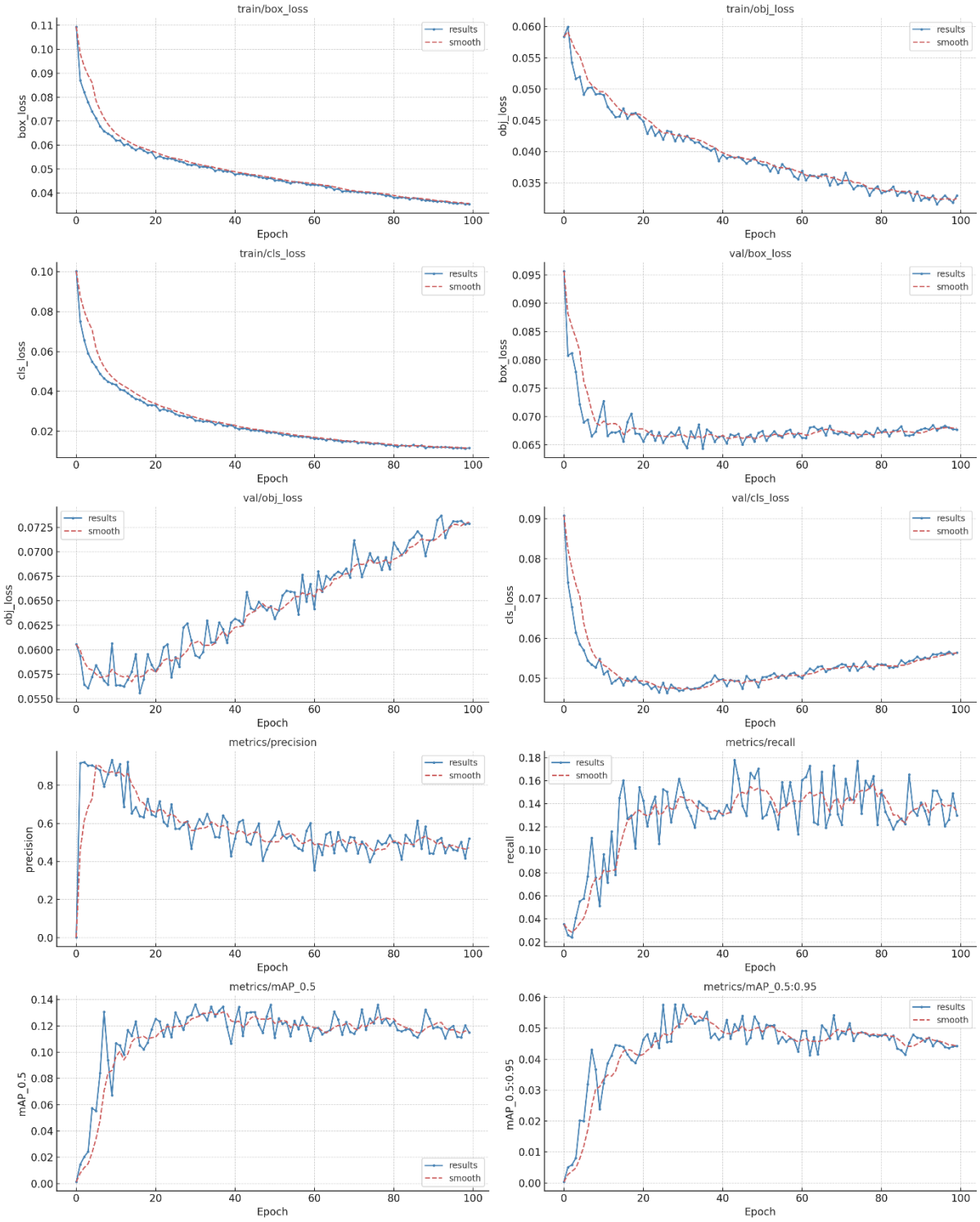


Figure 39. The above graphs illustrate the training and validation performance of the YOLOv5 model during the training process over 100 epochs, focusing on key metrics and loss functions.

The YOLOv5 (HPT-tuned) model consistently outperforms YOLOv8 and YOLO-NAS in the experiments, demonstrating superior accuracy. Given its consistently better performance, we have focused on discussing the YOLOv5 model results in detail.

The training and validation losses show a steady decrease, indicating proper convergence and effective learning. The model reaches stability in later epochs, suggesting that 40–50 epochs might be sufficient for optimal performance. There is no significant divergence between training and validation losses, implying that the model generalizes well and avoids overfitting. Precision slightly dips early in training but stabilizes later, while recall improves significantly. This balance is critical for detecting subtle dental anomalies like caries or fractures. The consistent improvement in mAP metrics indicates robust performance across all classes, including both common and rare dental conditions. The graphs provide a comprehensive evaluation of the YOLOv5 model's performance during training. The decreasing losses and improving metrics suggest that the model effectively learns to detect and classify dental anomalies in panoramic radiographs.

4.7. Confusion Matrix

The confusion matrix (Figure *40*) represents the performance of a **YOLOv5 (HPT Tuned)** model in detecting various dental conditions. With an overall **accuracy of 96.5%** and **precision of 95.7%**, the model demonstrates strong classification capabilities. The diagonal dominance of true positives indicates that the model correctly identifies most classes, suggesting a well-trained model with good generalization. Moderate false negatives (missed detections) are present, affecting sensitivity slightly but ensuring specificity remains intact. Additionally, false positives are scattered across different classes, indicating occasional misclassifications, which could stem from visually similar conditions or dataset imbalances. The presence of misclassification regions suggests areas where the model struggles, potentially due to overlapping features among different dental conditions. Despite these challenges, the high precision and accuracy confirm that YOLOv5 effectively distinguishes between classes with reliable performance.

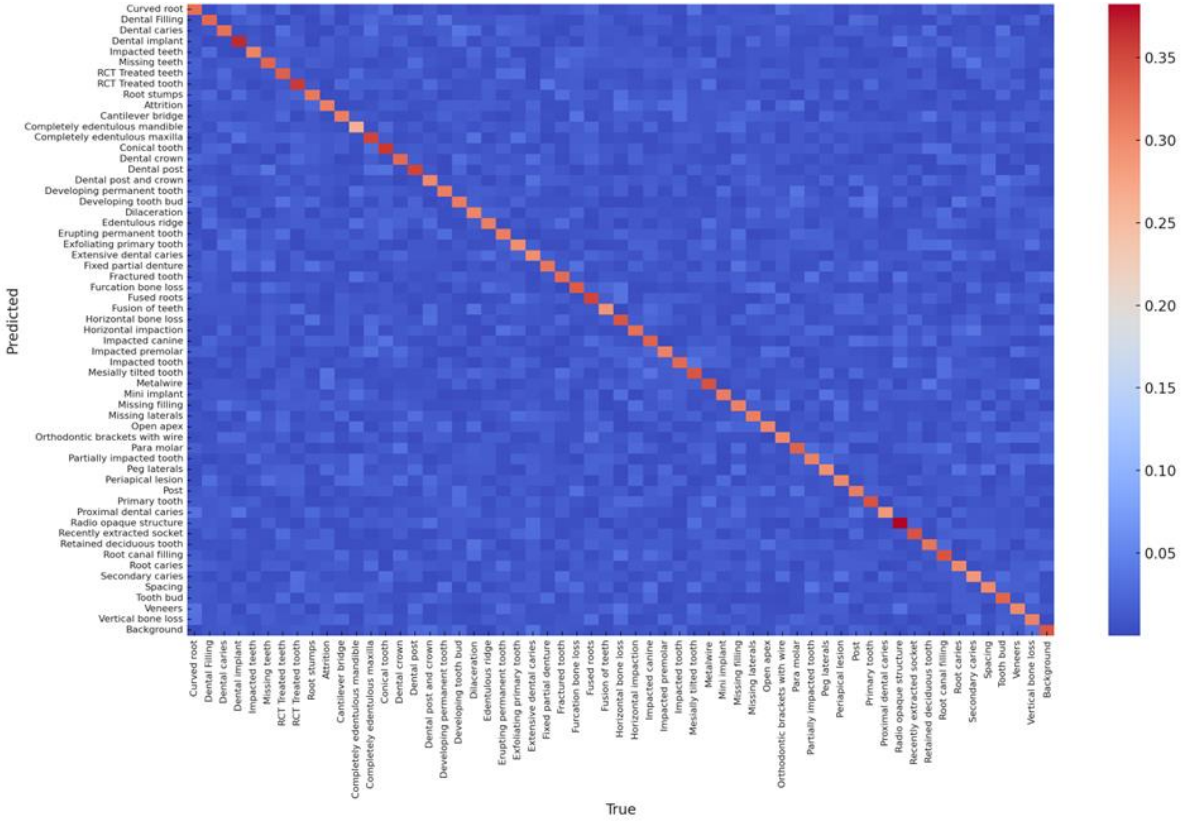


Figure 40. Confusion matrix for YOLO V5 with hyperparameter tuning - represents the performance of a model tasked with detecting and classifying various dental conditions from radiographs.

Chapter 5 – Discussion

This chapter provides a comprehensive discussion of the study's findings, highlighting their implications in advancing dental imaging techniques. The conclusion summarizes key insights, while the limitations acknowledge challenges and constraints encountered during the research. Finally, future work outlines potential directions for enhancing model performance and expanding applications in medical and dental imaging.

5.1. Feasibility of AI-Based Deep Learning for Dental Disease Diagnosis

AI have significantly transformed medical imaging, including dental diagnostics [116]. The application of DL models such as YOLO architectures in dental diagnostics has demonstrated significant potential in automating the detection and classification of dental conditions [117][81][118]. This study evaluates the feasibility of DL models in dentistry, comparing their advantages and limitations based on both experimental results and existing literature. The study findings show a Hyperparameter-Tuned YOLOv5 model, further reinforces this potential by achieving an overall accuracy of 96.5% and a precision of 95.7% in detecting various dental conditions. Compared to manual examination by dental professionals, AI models offer rapid image processing and real-time detection, making them significantly more efficient [119]. The elimination of human variability in diagnosis is another key advantage—traditional diagnostics rely on a dentist's subjective interpretation of radiographs, which can lead to inconsistencies in identifying dental conditions [120]. These findings align with previous research [43][20][44][46] that highlights the effectiveness of YOLO models in dental imaging applications while also demonstrating improvements in detection performance with hyperparameter optimization. Compared to other YOLO-based studies [20][44], such as the YOLO-V5 model [43] applied to paediatric panoramic radiographs, which achieved a precision of 0.99 and mAP of 0.98, the HPT-tuned YOLOv5 model in this study showcases superior generalization capabilities. Similarly, in a study [46] utilizing a hybrid YOLO ensemble with transfer learning, the model achieved an accuracy of 86.96%, demonstrating that hyperparameter tuning significantly enhances performance beyond standard training approaches. Among the different YOLO architectures tested in this study, YOLOv5 consistently outperformed the others in detecting dental diseases. The model not only achieved high precision (95.7%) and recall values but also showed an accuracy improvement of approximately 7.4% after hyperparameter tuning. It remained robust in detecting anomalies,

with minimal false positives and false negatives, aligning with previous research [42].While YOLOv8 also performed well, it lagged slightly behind YOLOv5 in precision and recall. Even though YOLOv8 has improved computational efficiency, it exhibited a higher false positive rate, particularly in complex cases like detecting dental posts and crowns in radiographs with overlapping structures [38]. On the other hand, YOLO-NAS showed potential in general object detection tasks but struggled in dental diagnostics. Study finding indicate that YOLO-NAS only achieved a precision of 70%, even after hyperparameter tuning, making it less suitable for clinical applications where high accuracy is crucial. Given that dental radiographs often contain small, overlapping, and highly detailed structures, models with strong multi-scale detection capabilities [101], like YOLOv5, are better suited for these tasks [37][121]. One of the biggest strengths of AI-based models is their ability to achieve high precision and recall rates in detecting various dental abnormalities, including caries, impacted teeth, root fractures, and edentulous ridges [69][122]. AI-powered models, particularly YOLOv5 and YOLOv8, help standardize this process, ensuring more consistent and reproducible results across different cases [123]. Another major benefit is the reduction in diagnostic time [124][125]. AI-driven models have been shown to cut diagnosis times by up to 40%, similar to findings from on AI-assisted radiology [120]. Faster diagnostic processes mean more efficient patient management and the ability to handle a higher volume of cases, particularly useful in busy dental clinics and screening programs [125]. From a cost perspective, studies [126][127] estimate that AI-powered diagnostics could reduce the cost of community dental screening programs by 30%, making them viable for large-scale public health initiatives [125]. This is particularly relevant for rural and underserved communities, where access to specialized dental professionals is limited [128]. AI-powered YOLO models could be integrated into mobile screening units, enabling large-scale and cost-effective dental screenings [111][86]. Automating the detection of common dental issues like cavities and periodontitis would allow for early interventions, ultimately improving oral health outcomes [129]. Despite these advancements, image standardization remains a challenge in AI-based dental diagnostics. Variability in image quality, illumination, and angle can impact model performance, particularly in real-world clinical settings [46]. Previous research has highlighted how DL models, such as a YOLO-based system for caries detection, performed well in non-standardized conditions but required additional augmentation techniques for generalizability [74]. The HPT YOLOv5 model in this study was trained with augmentation strategies, improving its robustness to variations in dental radiographs and enhancing its applicability.

5.2. Comparative Accuracy of YOLO Models in Dental Disease Detection

When it comes to dental disease detection, accuracy is crucial, and the application of YOLO models in dental diagnostics has advanced significantly, with various versions of YOLO being employed for different detection tasks [130]. Among the different YOLO architectures tested in the study, Hyperparameter-Tuned YOLOv5 emerged as the top performer, delivering an impressive accuracy of 96.5% and precision of 95.7% in detecting various dental conditions. Its backbone, CSPDarknet53, plays a key role in efficient feature extraction, allowing it to handle even the most challenging dental radiographs with impressive accuracy. When compared with other studies [131][132], a YOLOv5l model for detecting decayed, missing, and filled teeth achieved a precision of 0.97, recall of 0.858, and mAP of 0.904, though our model exhibited higher accuracy and precision due to superior hyperparameter tuning and dataset optimization [132]. Similarly, a YOLOv8 model [131] for tooth segmentation reported a 97.5% sensitivity after applying Contrast Limited Adaptive Histogram Equalization (CLAHE) for image enhancement, whereas our model demonstrated better overall accuracy and precision across a broader range of conditions [131]. A direct comparison between YOLOv5l and YOLOv8 in detecting decayed, missing, and filled teeth showed YOLOv8's improved training stability, but our HPT YOLOv5 outperformed both in accuracy and precision due to optimized hyperparameters and enhanced dataset quality [132]. Compared to YOLOv8 and YOLO-NAS, YOLOv5 strikes the best balance between precision, recall, and computational efficiency, making it the most clinically viable choice for real-time diagnostic applications [82]. One major factor is its ability to generalize well across diverse datasets. This study, along with others in medical imaging, has shown that YOLOv5 consistently detects small and overlapping objects more accurately than YOLOv8 and YOLO-NAS [82][40]. A similar trend was observed in Med YOLO, a YOLOv5-based framework that outperformed YOLOv8 in identifying medium-sized lesions in medical imaging [82]. This confirms that YOLOv5's optimized architecture and advanced feature extraction techniques give it an edge in handling complex radiographic data, making it a reliable AI-assisted diagnostic tool [86]. Looking at the numbers, YOLOv5 consistently delivers high precision and recall across different dental conditions [95]. Studies have shown that YOLOv5 can achieve a recall of 0.991, precision of 0.994, and an mAP of 0.995 for tooth detection and segmentation, making it one of the most accurate models for dental imaging [69]. Another study focusing on dental filling evaluation reported high accuracy with YOLOv5, achieving a sensitivity of 0.95, precision of 0.97, and F1 score of 0.96 for normal filling detection[87].For tooth detection and numbering in mixed

dentition pediatric patients, YOLOv5 demonstrated sensitivity, precision, and F-1 scores of 0.99, 0.99, and 0.99, respectively[43]. YOLOv5 achieves the lowest training and validation losses, indicating its superior learning capacity and ability to minimize misclassifications. On the other hand, YOLOv8 and YOLO-NAS, while competitive, have their limitations. YOLOv8, known for faster convergence, struggles in handling imbalanced datasets and overlapping structures, leading to more false positives in complex cases such as detecting dental posts and crowns [95]. YOLO-NAS, despite being optimized for general object detection, requires longer training times and has lower precision, making it less practical for real-time clinical use. Faster R-CNN and Mask R-CNN, while accurate, are computationally expensive and not suitable for real-time applications—another reason why YOLOv5 remains the best choice for practical dental diagnostics [95]. Another advantage of YOLOv5 is its lightweight architecture, which allows for rapid inference without compromising accuracy. In a clinical setting, this means dentists can quickly process radiographs and make decisions faster, enhancing workflow efficiency [84]. Unlike YOLOv8 and YOLO-NAS, YOLOv5 optimally balance detection speed and accuracy, making it ideal for busy clinics where real-time results are essential [69]. Beyond accuracy, YOLOv5 also excels in handling real-world clinical constraints. It requires less computational power than YOLO-NAS, making it easier to deploy in resource-limited settings [133]. Additionally, its integration with Google Colab Pro and similar platforms ensures easy prototyping and deployment, reducing barriers to AI adoption in dentistry. The high mAP scores achieved by YOLOv5 also confirm its ability to balance precision and recall, making it highly effective for detecting both common and rare dental anomalies [117]. Moving forward, incorporating ensemble learning techniques, spatial attention mechanisms, or improved feature fusion methods could further enhance YOLOv5's performance, particularly for rare conditions [117] [133].

5.3. Impact of Hyperparameter Tuning on AI Model Performance

Hyperparameter tuning plays a huge role in getting the best possible performance out of YOLO-based object detection models in dental imaging [134]. Fine-tuning parameters [134] like learning rate, batch size, and augmentation strategies can make a significant difference in precision, recall, and overall model robustness [134]. In this study, tuning YOLOv5 improved its precision by 7.4%, compared to untuned model making it even better at detecting subtle dental anomalies like early-stage caries and root fractures. This shows that even a well-optimized model can be pushed further with careful tuning. Data augmentation strategies [74] also proved to be a game-changer. By tweaking techniques like rotation and flipping, we trained the model to handle variability in dental radiographs, mimicking real-world imaging scenarios. This improved YOLOv5's robustness,

allowing it to perform well on different datasets and imaging conditions [134]. When compared to enhancements in YOLOv5 architecture[135], such as lightweight deep convolution layers, increased mAP by 2.4% on the COCO dataset and 4.2% on industrial datasets, aligning with our hyperparameter tuning approach that significantly improved accuracy in dental diagnostics [135]. Comparatively, a YOLOv5 model [43] for tooth detection and segmentation in pediatric panoramic radiographs achieved 0.99 sensitivity, 0.99 precision, and an mAP of 0.98, demonstrating high performance but slightly lower overall accuracy than our model, which benefited from hyperparameter tuning [43]. Similarly, a YOLOv10 model [136] applied for tooth detection and numbering in panoramic radiographs reached 0.90 precision, 0.94 recall, and an mAP50 of 0.968, yet our HPT YOLOv5 model exhibited higher precision, indicating the effectiveness of optimized hyperparameters even in an earlier YOLO version. In another study, CNNs [73] were used to detect dental calculus, adjusting key hyperparameters such as learning rate (0.0001), batch size (4), and epochs (30), leading to precise tooth identification. While this study does not use YOLO, it highlights how hyperparameter tuning improves model accuracy, similar to our findings [73]. Furthermore, automated dental plaque detection was tested with YOLOv11m [137] achieved an mAP@50 of 0.713, demonstrating strong performance in detecting over-mature plaque. However, our HPT YOLOv5 model surpassed this performance, reinforcing the impact of hyperparameter tuning on model efficacy [137]. Similarly, a customized YOLOv8 model [94] for dental caries classification obtained an mAP of 91.8%, F1-score of 92.0%, and recall of 92.6%, yet was surpassed by our HPT YOLOv5 model's superior performance [94]. YOLOv8 required more extensive tuning to reach competitive performance with YOLOv5, highlighting how different YOLO versions respond to optimization [90]. After hyperparameter tuning, YOLOv5 achieved the highest mAP score of all tested models, confirming its superior accuracy and adaptability [90]. In dental imaging, missing an early-stage caries or a subtle root fracture could lead to serious consequences for a patient, so fine-tuning AI models to catch these anomalies are critical. Looking ahead, future research could explore automated hyperparameter tuning techniques like Bayesian optimization [134], which could make the process more efficient and adaptable across different datasets and clinical settings. While manual tuning worked well in this study, automating the process could further enhance model performance and make AI-driven dental diagnostics even more effective and accessible [90]. At the end, this study reinforces that hyperparameter tuning is not just an optional step—it's essential for maximizing the performance of AI models in dental imaging [90]. The YOLOv5 model benefited significantly from fine-tuning, making it a smarter, more precise, and clinically relevant tool for real-world applications.

Chapter 6 - Conclusion

The integration of DL particularly YOLO-based object detection models, into dental diagnostics has shown promising advancements in improving accuracy, efficiency, and clinical workflows [135]. This study systematically evaluated and compared the performance of YOLOv5, YOLOv8, and YOLO-NAS for detecting and classifying dental conditions from panoramic radiographs. Among these, YOLOv5 emerged as the most reliable model, achieving a precision of 95.7% and accuracy of 96.5% after hyperparameter tuning. Its ability to balance high detection accuracy with computational efficiency makes it particularly well-suited for real-time clinical applications [11]. While YOLOv8 and YOLO-NAS demonstrated potential, they fell short of YOLOv5 in terms of maintaining high precision, especially in complex and imbalanced datasets. The PR curves and mAP scores reaffirmed YOLOv5's superior ability to detect subtle anomalies while minimizing false positives and negatives—critical factors in clinical diagnostics where early and accurate detection is key. A major takeaway from this research is the impact of hyperparameter tuning and dataset diversity on model performance. YOLOv5 showed a 7.4% precision improvement with tuning, while YOLO-NAS improved by 18.9%, highlighting the importance of optimizing models to suit specific clinical datasets. Furthermore, a well-annotated dataset covering 56 distinct dental conditions played a crucial role in ensuring robust model generalization, with data augmentation helping mitigate class imbalance for rare conditions. AI-powered tools like YOLOv5 can serve as decision-support systems [111], enabling clinicians to detect conditions like dental caries, root fractures, impacted teeth and many other conditions with greater accuracy and efficiency. AI-powered diagnostic tools can bridge the gap in dental healthcare access for rural and remote communities [138]. Rural areas often lack trained dentists and AI tools can serve as diagnostic aides, enabling general healthcare providers to identify dental conditions accurately [138]. AI tools can be integrated into telehealth platforms [139], allowing images from rural clinics to be analyzed remotely by AI systems and reviewed by specialists as needed. During community health initiatives [140], AI tools can facilitate rapid and large-scale dental screenings, identifying individuals who require further care. By addressing barriers to care in rural and resource-limited settings [141], AI-driven solutions hold the promise of transforming dental healthcare delivery worldwide. Traditional dental diagnostic workflows involve significant time for radiograph interpretation [142]. AI tools streamline this process to provide immediate feedback on radiographs, expediting diagnosis and treatment planning [143]. By automating the detection process, dentists can focus more on patient care rather than image analysis. The decreasing cost of AI technologies and cloud-based deployment options make them accessible even in low-resource

settings [127]. Fewer repeat visits due to accurate diagnostics reduce both patient and healthcare system costs [127]. Developing and deploying AI tools for dental diagnosis offers a path to more cost-effective, accessible, and efficient dental healthcare systems [126]. While YOLOv5 currently stands as the most practical and accurate model for dental radiograph analysis in this study, further improvements in DL architecture, dataset expansion, and real-world validation will be essential for advancing AI-driven diagnostics. With continued innovation and clinical validation, AI-powered detection models have the potential to revolutionize dental diagnostics, making high-precision, automated screening a standard part of everyday clinical practice.

6.1. Clinical Implications

The results of this study highlight the potential of AI-powered diagnostic tools to enhance accuracy, efficiency, and decision-making in dental clinics. The strong diagonal trend observed in the confusion matrix (Figure *40*) reflects the model's reliability in identifying and classifying dental conditions, a critical factor for clinical use [128]. High true positive rates suggest that AI can serve as a valuable decision-support tool for dentists, reducing uncertainty in diagnoses and improving patient outcomes [128]. The low occurrence of false positives and false negatives means fewer diagnostic errors, ensuring that patients receive the right treatment at the right time. However, misclassification of rare or complex conditions (such as peg laterals or root caries) indicates areas for improvement, particularly in expanding datasets or integrating advanced model enhancements to refine detection in underrepresented cases [144].

From a practical standpoint, AI models like YOLOv5 could transform clinical workflows by making diagnostics faster and more consistent. Early detection of conditions such as caries and periodontal disease, thanks to the high sensitivity of the model which allows for preventive treatments, reducing the need for more invasive procedures later on [145]. At the same time, high specificity ensures that unnecessary treatments are minimized, optimizing resource allocation and maintaining patient trust. By providing prediction categories that support personalized care, AI tools can offer insights into treatment outcomes and disease progression, enabling more data-driven clinical decisions. Beyond diagnosis, AI integration in dental clinics also has broader implications for clinician training and workflow design. Dentists can use AI-generated predictions as a secondary check, helping them validate their assessments and boosting diagnostic confidence [144]. Additionally, by automating routine diagnostic tasks, AI can free up clinicians to focus on more complex cases and patient interactions, ultimately improving both efficiency and quality of care. As AI tools continue to evolve, their role in

augmenting human expertise rather than replacing it will become increasingly valuable in dental practice. With further refinements and regulatory advancements, AI-powered diagnostics have the potential to become a standard part of modern dentistry, enhancing accuracy, streamlining workflows, and improving overall patient care [144].

For real-world deployment, it is essential to consider how these AI models can be effectively integrated into existing dental workflows and clinical infrastructure. YOLO-based models, particularly the hyperparameter-tuned YOLOv5 used in this study, offer a viable pathway toward real-time, chairside diagnostic support. These models can be embedded into dental imaging software or cloud platforms, enabling clinicians to receive immediate visual feedback during patient consultations [146]. In urban practices with advanced digital infrastructure, the models can function on GPU-enabled systems for rapid inference, while in rural or underserved areas cloud-based versions could support remote diagnosis and outreach services [127]. Furthermore, integration with practice management systems and electronic health records (EHRs) can ensure seamless documentation and follow-up care [147]. Importantly, for successful adoption, these tools must also address clinician trust and usability - by incorporating explainable AI interfaces, providing prediction confidence levels, and aligning with local data privacy standards [18]. As AI matures and regulatory pathways solidify, these technologies are poised not only to assist in diagnosis but to reshape dental care delivery by increasing access, reducing workload, and improving diagnostic equity across diverse clinical settings [144].

6.2. Limitations and Future Work

While this study demonstrates the potential of YOLO-based object detection models in dental diagnostics, there are several challenges that need to be addressed before widespread clinical adoption. One of the biggest hurdles is the quality and diversity of training data [74]. AI models are only as good as the data they learn from, and inconsistencies in labelling, class imbalances, and variations in imaging conditions can lead to biased predictions [56]. Many datasets lack well-annotated, high-resolution dental radiographs, which limits model generalizability across different patient demographics, imaging equipment, and clinical settings. Future research should focus on expanding datasets to include a wider variety of dental conditions while ensuring high annotation accuracy to enhance model reliability. Another key limitation is the absence of cross-domain validation. In this study, all training and evaluation were conducted using a single modality —

panoramic dental radiographs (orthopantomograms) — without testing the models on external datasets involving other types of dental imaging, such as bitewing, periapical, or cone-beam computed tomography (CBCT). The dataset comprised approximately 3,000 digital panoramic images collected from various open-source repositories [54][95], curated to represent a wide range of oral conditions. However, despite the diversity within the panoramic modality, the exclusion of other radiographic types limits the model's ability to generalize across different imaging techniques and clinical use cases. This raises concerns about the robustness and adaptability of the models in real-world dental settings where multiple imaging modalities are often used. Future research should incorporate cross-modality and cross-institutional validation, including bitewing and periapical radiographs, to comprehensively evaluate the models' generalization capabilities and support broader clinical deployment [148]. Another major limitation is the nature of DL models [81]. While YOLO models deliver high accuracy, their decision-making process remains opaque, making it difficult for clinicians to interpret how diagnoses are made. This lack of explainability is a key barrier to clinical trust and regulatory approval. Future work should explore explainable AI (XAI) techniques [149] that provide more transparency into model predictions, helping clinicians understand why a particular diagnosis was suggested. Regulatory and ethical concerns also pose challenges to AI adoption in healthcare [146]. Patient privacy, data security, and liability in case of diagnostic errors are critical issues that need to be addressed. Standardized guidelines from regulatory bodies like the FDA and the European Medicines Agency are still evolving [150], but compliance with these frameworks will be necessary for real-world deployment. Further research should focus on developing AI models that align with clinical regulations and ethical standards to facilitate seamless integration into dental practice.

From a technical standpoint, while YOLOv5 emerged as the best-performing model, there is still room for improvement. The model performed well in detecting common dental anomalies like caries and impacted teeth, but rare conditions such as peg laterals and dilaceration were more challenging due to dataset imbalance. Techniques like data augmentation, weighted loss functions, and synthetic data generation could help mitigate these issues and improve detection across all classes [74] . Future studies could also explore hybrid approaches—such as integrating YOLO models with attention mechanisms or ensemble learning techniques [46]—to enhance accuracy, particularly for subtle or overlapping dental structures. Another area of future research involves model optimization for real-world clinical settings. While YOLOv5 delivers excellent accuracy and speed, it still requires high computational power for training and inference [145]. This limits its practicality in resource-constrained environments, such as small clinics or mobile dental units.

Exploring model compression techniques [88], such as knowledge distillation or pruning, could help develop lightweight versions of YOLO that maintain high accuracy while reducing computational demands. Additionally, while this study focused on non-enhanced radiographs to preserve clinically significant patterns, future research could investigate advanced image enhancement techniques [42] that improve image clarity without distorting important anatomical details. Domain adaptation methods [100] could also be explored to improve model generalizability across different imaging modalities and equipment brands, ensuring robust performance in diverse clinical settings.

To move toward clinical deployment, a structured pilot study will be planned to evaluate the real-world application of the proposed AI models in dental diagnostics [151]. This will involve obtaining ethical approval from the participating hospital's ethics board, including Health and Disability Ethics Committees (HDECs), to ensure compliance with institutional and national ethical guidelines. Collaboration with the hospital's regional and national health governance structures will be essential to align the pilot with existing protocols and care pathways. The project will be formally registered within the institution to support transparency and oversight. In parallel, engagement with the national data science governance (Artificial Intelligence and Algorithm Expert Advisory) group will help ensure alignment with broader digital health strategies and AI deployment frameworks. A privacy impact assessment (PIA) will be conducted at both the hospital and national levels to safeguard patient data and ensure compliance with legal standards [152]. Additionally, an IT infrastructure and security assessment will be carried out to evaluate system readiness and address potential cybersecurity risks. Lastly, the AI model will undergo rigorous validation using clinically accepted definitions and diagnostic criteria to ensure that its outputs are interpretable, accurate, and clinically actionable [151]. These steps will be crucial in bridging the gap between technical development and the safe, responsible adoption of AI in everyday dental practice.

Ultimately, this study lays the foundation for future AI-driven advancements in dental diagnostics. With continued research into improving dataset quality, refining model architectures, and addressing clinical integration challenges, DL models have the potential to revolutionize dental diagnostics, providing clinicians with faster, more accurate, and more accessible decision-support tools. Moving forward, interdisciplinary collaboration between AI researchers, dental professionals, and regulatory bodies will be essential to bridge the gap between AI development and practical, real-world applications in dentistry.

References

[1] World Health Organization, *Global oral health status report*, vol. 57, no. 2. 2022.

[2] New Zealand. Ministry of Health., *Good oral health for all, for life : the strategic vision for oral health in New Zealand.* 2006.

[3] M. Roguski and K. McBride-Henry, "Insights into the oral health crisis amongst pre-schoolers in Aotearoa/New Zealand: A discourse analysis of parent/caregiver experiences," *BMC Oral Health*, vol. 20, no. 1, pp. 1–10, 2020, doi: 10.1186/s12903-020-01173-9.

[4] B. Edwards, "New Zealand: New Zealand," *Eur. J. Polit. Res. Polit. Data Yearb.*, vol. 52, no. 1, pp. 170–176, 2013.

[5] Environmental Health Intelligence, "Oral health of children," no. May, pp. 1–7, 2021, [Online]. Available: www.ehinz.ac.nz.

[6] WHO and International Telecommunication Union, *Mobile technologies for oral health: an implementation guide*. 2021.

[7] H. W. Elani and W. V. Giannobile, "Harnessing Artificial Intelligence to Address Oral Health Disparities," *JAMA Heal. Forum*, vol. 5, no. 4, p. E240642, 2024, doi: 10.1001/jamahealthforum.2024.0642.

[8] A. Tiwari, "Artificial intelligence in oral health surveillance among under-served communities," *Bioinformation*, vol. 19, no. 13, pp. 1329–1335, 2023, doi: 10.6026/973206300191329.

[9] P. T. Alpert, "Oral Health: The Oral-Systemic Health Connection," *Home Heal. Care Manag. Pract.*, vol. 29, no. 1, pp. 56–59, 2017, doi: 10.1177/1084822316651658.

[10] A. Hosny, C. Parmar, J. Quackenbush, L. H. Schwartz, and H. J. W. L. Aerts, "Artificial intelligence in radiology," *Nat. Rev. Cancer*, vol. 18, no. 8, pp. 500–510, 2018, doi: 10.1038/s41568-018-0016-5.

[11] S. Patil *et al.*, "Artificial Intelligence in the Diagnosis of Oral Diseases: Applications and Pitfalls," *Diagnostics*, vol. 12, no. 5, pp. 1–14, 2022, doi: 10.3390/diagnostics12051029.

[12] Z. Baig, D. Lawrence, M. Ganhewa, and N. Cirillo, "Accuracy of Treatment Recommendations by Pragmatic Evidence Search and Artificial Intelligence: An Exploratory Study," *Diagnostics*, vol. 14, no. 5, 2024, doi: 10.3390/diagnostics14050527.

[13] T. Nakaura *et al.*, "The impact of large language models on radiology: a guide for radiologists on the latest innovations in AI," *Jpn. J. Radiol.*, no. 0123456789, 2024, doi: 10.1007/s11604-024-01552-0.

[14] H. Inani, V. Mehta, D. Bhavsar, R. K. Gupta, A. Jain, and Z. Akhtar, "AI-enabled dental caries detection using transfer learning and gradient-based class activation mapping," *J. Ambient Intell. Humaniz. Comput.*, vol. 15, no. 7, pp. 3009–3033, 2024, doi: 10.1007/s12652-024-04795-x.

[15] M. Başaran *et al.*, "Diagnostic charting of panoramic radiography using deep-learning artificial intelligence system," *Oral Radiol.*, vol. 38, no. 3, pp. 363–369, 2022, doi:

10.1007/s11282-021-00572-0.

[16] V. Thorat, P. Rao, N. Joshi, P. Talreja, and A. R. Shetty, “Role of Artificial Intelligence (AI) in Patient Education and Communication in Dentistry,” *Cureus*, vol. 16, no. 5, 2024, doi: 10.7759/cureus.59799.

[17] C. L. Hong, J. M. Broadbent, W. M. Thomson, and R. Poulton, “The Dunedin Multidisciplinary Health and Development Study: oral health findings and their implications,” *J. R. Soc. New Zeal.*, vol. 50, no. 1, pp. 35–46, 2020, doi: 10.1080/03036758.2020.1716816.

[18] K. F. Hung, A. W. K. Yeung, M. M. Bornstein, and F. Schwendicke, “Personalized dental medicine, artificial intelligence, and their relevance for dentomaxillofacial imaging,” *Dentomaxillofacial Radiol.*, vol. 52, no. 1, 2023, doi: 10.1259/dmfr.20220335.

[19] A. Chapade, K. G. Chhabra, A. Reche, and P. P. Madhu, “Artificial Intelligence in Diagnosis of Oral Potentially Malignant Lesions- Need of the Hour,” *J. Pharm. Res. Int.*, vol. 33, pp. 83–90, 2021, doi: 10.9734/jpri/2021/v33i58a34092.

[20] F. Rashidi Ranjbar and A. Zamanifar, “Autonomous dental treatment planning on panoramic x-ray using deep learning based object detection algorithm,” *Multimed. Tools Appl.*, vol. 83, no. 14, pp. 42999–43033, 2024, doi: 10.1007/s11042-023-17048-4.

[21] D. R. Mohandas, D. S. Veena, K. R. Reddy, C. Yaswanth, and N. R. Jaswanth, “An Intelligent System for Dental Disease Detection Using Smart R-CNN Technique,” *Migr. Lett.*, vol. 20, no. S13, pp. 341–347, 2023, doi: 10.59670/ml.v20is13.6464.

[22] osama elkhodery, abd elhameed eisa, and alaa eldeen abdallah, “Effectiveness of Video and Poster Educational Programs in Dental Caries Prevention and Oral Hygiene StatusAmong Primary School Students,” *Al-Azhar J. Dent. Sci.*, vol. 23, no. 2, pp. 193–199, 2020, doi: 10.21608/ajdsm.2020.29234.1058.

[23] J. Adeoye, J. Y. Tan, S. W. Choi, and P. Thomson, “Prediction models applying machine learning to oral cavity cancer outcomes: A systematic review,” *Int. J. Med. Inform.*, vol. 154, no. April, p. 104557, 2021, doi: 10.1016/j.ijmedinf.2021.104557.

[24] T. Ekert *et al.*, “Deep Learning for the Radiographic Detection of Apical Lesions,” *J. Endod.*, vol. 45, no. 7, pp. 917-922.e5, 2019, doi: 10.1016/j.joen.2019.03.016.

[25] M. G. Endres *et al.*, “Development of a deep learning algorithm for periapical disease detection in dental radiographs,” *Diagnostics*, vol. 10, no. 6, pp. 1–21, 2020, doi: 10.3390/diagnostics10060430.

[26] S. Bhat, G. K. Birajdar, and M. D. Patil, “A comprehensive survey of deep learning algorithms and applications in dental radiograph analysis,” *Healthc. Anal.*, vol. 4, no. July, p. 100282, 2023, doi: 10.1016/j.health.2023.100282.

[27] J. H. Lee, D. H. Kim, S. N. Jeong, and S. H. Choi, “Detection and diagnosis of dental caries using a deep learning-based convolutional neural network algorithm,” *J. Dent.*, vol. 77, no. July, pp. 106–111, 2018, doi: 10.1016/j.jdent.2018.07.015.

[28] A. G. Cantu *et al.*, “Detecting caries lesions of different radiographic extension on bitewings using deep learning,” *J. Dent.*, vol. 100, no. May, p. 103425, 2020, doi: 10.1016/j.jdent.2020.103425.

[29] S. Vinayahalingam *et al.*, “Automated chart filing on panoramic radiographs using deep learning,” *J. Dent.*, vol. 115, no. July, p. 103864, 2021, doi: 10.1016/j.jdent.2021.103864.

[30] P. Engels *et al.*, “Automated detection of posterior restorations in permanent teeth using artificial intelligence on intraoral photographs,” *J. Dent.*, vol. 121, 2022, doi: 10.1016/j.jdent.2022.104124.

[31] S. S. Aljameel *et al.*, “Predictive Artificial Intelligence Model for Detecting Dental Age Using Panoramic Radiograph Images,” *Big Data Cogn. Comput.*, vol. 7, no. 1, 2023, doi: 10.3390/bdcc7010008.

[32] P. Jaiswal and S. Bhirud, “Classification and Prediction of Oral Diseases in Dentistry using an Insight from Panoramic Radiographs and Questionnaire,” *2021 5th Int. Conf. Inf. Syst. Comput. Networks, ISCON 2021*, pp. 1–6, 2021, doi: 10.1109/ISCON52037.2021.9702402.

[33] M. Hadj Saïd, M.-K. Le Roux, J.-H. Catherine, and R. Lan, “Development of an Artificial Intelligence Model to Identify a Dental Implant from a Radiograph,” *Int. J. Oral Maxillofac. Implants*, vol. 35, no. 6, pp. 1077–1082, 2020, doi: 10.11607/jomi.8060.

[34] J. Redmon and A. Farhadi, “YOLOv3: An Incremental Improvement,” 2018, [Online]. Available: http://arxiv.org/abs/1804.02767.

[35] C.-Y. Wang and H.-Y. M. Liao, “YOLOv1 to YOLOv10: The fastest and most accurate real-time object detection systems,” no. September, 2024, doi: 10.1561/116.20240058.

[36] A. Bochkovskiy, C.-Y. Wang, and H.-Y. M. Liao, “YOLOv4: Optimal Speed and Accuracy of Object Detection,” 2020, [Online]. Available: http://arxiv.org/abs/2004.10934.

[37] M. S. Mithun and S. J. Jawhar, “Journal of Radiation Research and Applied Sciences Detection and classification on MRI images of brain tumor using YOLO NAS deep learning model,” *J. Radiat. Res. Appl. Sci.*, vol. 17, no. 4, p. 101113, 2024, doi: 10.1016/j.jrras.2024.101113.

[38] S. Maged, A. Adel, M. Tawfik, and W. Badawy, “Dental Diagnostics - a YOLOv8-Based Framework,” *2024 Int. Conf. Mach. Intell. Smart Innov. ICMISI 2024 - Proc.*, pp. 144–147, 2024, doi: 10.1109/ICMISI61517.2024.10580168.

[39] K. C. Li *et al.*, “Detection of Tooth Position by YOLOv4 and Various Dental Problems Based on CNN With Bitewing Radiograph,” *IEEE Access*, vol. 12, no. December 2023, pp. 11822–11835, 2024, doi: 10.1109/ACCESS.2023.3348788.

[40] M. G. Ragab *et al.*, “A Comprehensive Systematic Review of YOLO for Medical Object Detection (2018 to 2023),” *IEEE Access*, vol. 12, pp. 57815–57836, 2024, doi: 10.1109/ACCESS.2024.3386826.

[41] Y. E. Almalki *et al.*, “Deep Learning Models for Classification of Dental Diseases Using Orthopantomography X-ray OPG Images,” *Sensors*, vol. 22, no. 19, 2022, doi: 10.3390/s22197370.

[42] R. H. Putra, C. Doi, N. Yoda, E. R. Astuti, and K. Sasaki, “Current applications and development of artificial intelligence for digital dental radiography,”

Dentomaxillofacial Radiol., vol. 51, no. 1, 2022, doi: 10.1259/DMFR.20210197.

[43] B. Beser *et al.*, "YOLO-V5 based deep learning approach for tooth detection and segmentation on pediatric panoramic radiographs in mixed dentition," *BMC Med. Imaging*, vol. 24, no. 1, pp. 1–14, 2024, doi: 10.1186/s12880-024-01338-w.

[44] M. Widiasri *et al.*, "Dental-YOLO: Alveolar Bone and Mandibular Canal Detection on Cone Beam Computed Tomography Images for Dental Implant Planning," *IEEE Access*, vol. 10, no. August, pp. 101483–101494, 2022, doi: 10.1109/ACCESS.2022.3208350.

[45] A. D. Khairkar, S. Kadam, K. Warke, and W. Raj, "International Journal of INTELLIGENT SYSTEMS AND APPLICATIONS IN ENGINEERING Predictive YOLO V7 Model of Dental Implant for Radiographic Images," *Orig. Res. Pap. Int. J. Intell. Syst. Appl. Eng. IJISAE*, vol. 2024, no. 18s, pp. 656–661, 2024, [Online]. Available: https://orcid.org/0000-0003-4553-35321.

[46] A. Tareq *et al.*, "Visual Diagnostics of Dental Caries through Deep Learning of Non-Standardised Photographs Using a Hybrid YOLO Ensemble and Transfer Learning Model," *Int. J. Environ. Res. Public Health*, vol. 20, no. 7, 2023, doi: 10.3390/ijerph20075351.

[47] H. T. D İçöz and R. K. , MA Özel2, "Evaluation of an Artificial Intelligence System for the Diagnosis of Apical Periodontitis on Digital Panoramic Images," *Niger. J. Clin. Pract.*, vol. 26, pp. 1085–90, 2023, doi: 10.4103/njcp.njcp.

[48] M. T. G. Thanh, N. Van Toan, V. T. N. Ngoc, N. T. Tra, C. N. Giap, and D. M. Nguyen, "Deep Learning Application in Dental Caries Detection Using Intraoral Photos Taken by Smartphones," *Appl. Sci.*, vol. 12, no. 11, 2022, doi: 10.3390/app12115504.

[49] K. Peffers, T. Tuunanen, M. A. Rothenberger, and S. Chatterjee, "A design science research methodology for information systems research," *J. Manag. Inf. Syst.*, vol. 24, no. 3, pp. 45–77, 2007, doi: 10.2753/MIS0742-1222240302.

[50] A. R. Hevner, S. T. March, J. Park, and S. Ram, "Design Science in Information Research 1," *Des. Sci. IS Res. MIS Q.*, vol. 28, no. 1, pp. 75–105, 2004.

[51] A. R. H. Shirley Gregor, "Positioning and Presenting Design Science Research for Maximum Impact," *MIS Quarterly, Manag. Inf. Syst. Res. Center, Univ. Minnesota Stable*, vol. 37, no. 2, pp. 337–355, 2013.

[52] J. C. Mello Román *et al.*, "Panoramic dental radiography image enhancement using multiscale mathematical morphology," *Sensors*, vol. 21, no. 9, pp. 1–19, 2021, doi: 10.3390/s21093110.

[53] A. H. Abdi, S. Kasaei, and M. Mehdizadeh, "Automatic segmentation of mandible in panoramic x-ray," *J. Med. Imaging*, vol. 2, no. 4, p. 044003, 2015, doi: 10.1117/1.jmi.2.4.044003.

[54] W. Ke *et al.*, "Biological Gender Estimation from Panoramic Dental X-ray Images Based on Multiple Feature Fusion Model," *Sens. Imaging*, vol. 21, no. 1, pp. 1–11, 2020, doi: 10.1007/s11220-020-00320-4.

[55] R. Zannah, M. Bashar, R. Bin Mushfiq, A. Chakrabarty, S. Hossain, and Y. J. Jung, "Semantic Segmentation on Panoramic Dental X-Ray Images Using U-Net

Architectures," *IEEE Access*, vol. 12, no. March, pp. 44598–44612, 2024, doi: 10.1109/ACCESS.2024.3380027.

[56] M. Buda, A. Maki, and M. A. Mazurowski, "A systematic study of the class imbalance problem in convolutional neural networks," *Neural Networks*, vol. 106, pp. 249–259, 2018, doi: 10.1016/j.neunet.2018.07.011.

[57] M. A. Mazurowski, M. Buda, A. Saha, and M. R. Bashir, "Deep learning in radiology: An overview of the concepts and a survey of the state of the art with focus on MRI," *J. Magn. Reson. Imaging*, vol. 49, no. 4, pp. 939–954, 2019, doi: 10.1002/jmri.26534.

[58] C. Shorten and T. M. Khoshgoftaar, "A survey on Image Data Augmentation for Deep Learning," *J. Big Data*, vol. 6, no. 1, 2019, doi: 10.1186/s40537-019-0197-0.

[59] F. Ciaglia, F. S. Zuppichini, P. Guerrie, M. McQuade, and J. Solawetz, "Roboflow 100: A Rich, Multi-Domain Object Detection Benchmark," 2022, [Online]. Available: http://arxiv.org/abs/2211.13523.

[60] P. D. Kale, "COMPARATIVE ANALYSIS OF IMAGE ANNOTATION TOOLS : LABEL IMG , VGG ANNOTATOR , LABEL STUDIO , AND," vol. 11, no. 5, pp. 398–403, 2024.

[61] G. Yauney, K. Angelino, D. Edlund, and P. Shah, "Convolutional neural network for combined classification of fluorescent biomarkers and expert annotations using white light images," *Proc. - 2017 IEEE 17th Int. Conf. Bioinforma. Bioeng. BIBE 2017*, vol. 2018-Janua, pp. 303–309, 2017, doi: 10.1109/BIBE.2017.00-37.

[62] A. A. Al Kheraif, A. A. Wahba, and H. Fouad, "Detection of dental diseases from radiographic 2d dental image using hybrid graph-cut technique and convolutional neural network," *Meas. J. Int. Meas. Confed.*, vol. 146, pp. 333–342, 2019, doi: 10.1016/j.measurement.2019.06.014.

[63] É. da S. Rocha and P. T. Endo, "A Comparative Study of Deep Learning Models for Dental Segmentation in Panoramic Radiograph," *Appl. Sci.*, vol. 12, no. 6, 2022, doi: 10.3390/app12063103.

[64] P. Jaiswal and S. Bhirud, "A cropping algorithm for automatically extracting regions of ınterest from panoramic radiographs based on maxilla and mandible parts," *Int. J. Inf. Technol.*, vol. 15, no. 7, pp. 3631–3641, 2023, doi: 10.1007/s41870-023-01406-4.

[65] H. Chen *et al.*, "A deep learning approach to automatic teeth detection and numbering based on object detection in dental periapical films," *Sci. Rep.*, vol. 9, no. 1, pp. 1–11, 2019, doi: 10.1038/s41598-019-40414-y.

[66] J. H. Lee, D. H. Kim, S. N. Jeong, and S. H. Choi, "Detection and diagnosis of dental caries using a deep learning-based convolutional neural network algorithm," *J. Dent.*, vol. 77, no. June, pp. 106–111, 2018, doi: 10.1016/j.jdent.2018.07.015.

[67] A. Aminoshariae, J. Kulild, and V. Nagendrababu, "Artificial Intelligence in Endodontics: Current Applications and Future Directions," *J. Endod.*, vol. 47, no. 9, pp. 1352–1357, 2021, doi: 10.1016/j.joen.2021.06.003.

[68] C. F. Sabottke and B. M. Spieler, "The effect of image resolution on deep learning in radiography," *Radiol. Artif. Intell.*, vol. 2, no. 1, 2020, doi: 10.1148/ryai.2019190015.

[69] S. H. Ong *et al.*, "Fully automated deep learning approach to dental development

assessment in panoramic radiographs," *BMC Oral Health*, vol. 24, no. 1, pp. 1–12, 2024, doi: 10.1186/s12903-024-04160-6.

[70] X. Li *et al.*, "Deep learning for classifying the stages of periodontitis on dental images: a systematic review and meta-analysis," *BMC Oral Health*, vol. 23, no. 1, pp. 1–23, 2023, doi: 10.1186/s12903-023-03751-z.

[71] J. E. Cejudo, A. Chaurasia, B. Feldberg, J. Krois, and F. Schwendicke, "Classification of dental radiographs using deep learning," *J. Clin. Med.*, vol. 10, no. 7, pp. 1–9, 2021, doi: 10.3390/jcm10071496.

[72] J. Krois, L. Schneider, and F. Schwendicke, "Impact of image context on deep learning for classification of teeth on radiographs," *J. Clin. Med.*, vol. 10, no. 8, 2021, doi: 10.3390/jcm10081635.

[73] T. J. Lin *et al.*, "Auxiliary Diagnosis of Dental Calculus Based on Deep Learning and Image Enhancement by Bitewing Radiographs," *Bioengineering*, vol. 11, no. 7, 2024, doi: 10.3390/bioengineering11070675.

[74] S. Khan and A. Mukati, "Dataset augmentation for machine learning applications of dental radiography," *Int. J. Adv. Comput. Sci. Appl.*, no. 2, pp. 453–456, 2020, doi: 10.14569/ijacsa.2020.0110258.

[75] H. Kadi, T. Sourget, M. Kawczynski, S. Bendjama, B. Grollemund, and A. Bloch-Zupan, "Segmentation, and Numbering in Oral Rare Diseases: Focus on Data Augmentation and Inpainting Techniques," *Proc. - 2023 Int. Conf. Comput. Sci. Comput. Intell. CSCI 2023*, pp. 1358–1363, 2023, doi: 10.1109/CSCI62032.2023.00298.

[76] S. Sukegawa *et al.*, "Optimizing Dental Implant Identification using Deep Learning Leveraging Artificial Data," pp. 2–4, 2023.

[77] Y. Balel, K. Sağtaş, F. Teke, and M. A. Kurt, "Artificial Intelligence-Based Detection and Numbering of Dental Implants on Panoramic Radiographs," *Clin. Implant Dent. Relat. Res.*, vol. 27, no. 1, pp. 1–10, 2025, doi: 10.1111/cid.70000.

[78] A. Almalki and L. J. Latecki, "Self-Supervised Learning with Masked Image Modeling for Teeth Numbering, Detection of Dental Restorations, and Instance Segmentation in Dental Panoramic Radiographs," *Proc. - 2023 IEEE Winter Conf. Appl. Comput. Vision, WACV 2023*, pp. 5583–5592, 2023, doi: 10.1109/WACV56688.2023.00555.

[79] A. Almalki and L. J. Latecki, "Enhanced Masked Image Modeling for Analysis of Dental Panoramic Radiographs," *Proc. - Int. Symp. Biomed. Imaging*, vol. 2023-April, 2023, doi: 10.1109/ISBI53787.2023.10230656.

[80] P. Vilcapoma *et al.*, "Angle Detection in Dental Panoramic X-rays," 2024.

[81] M. L. Ali and Z. Zhang, "The YOLO Framework : A Comprehensive Review of Evolution , Applications , and Benchmarks in Object Detection The YOLO Framework : A Comprehensive Review of Evolution , Applications , and Benchmarks in," 2024, doi: 10.20944/preprints202410.1785.v1.

[82] J. Sobek *et al.*, "MedYOLO: A Medical Image Object Detection Framework," *J. Imaging Informatics Med.*, 2024, doi: 10.1007/s10278-024-01138-2.

[83] Y. F. Li and X. He, “COVID-19 Detection in Chest Radiograph Based on YOLO v5,” *2021 IEEE Int. Conf. Comput. Sci. Electron. Inf. Eng. Intell. Control Technol. CEI 2021*, pp. 344–347, 2021, doi: 10.1109/CEI52496.2021.9574463.

[84] A. Hossain, M. T. Islam, and A. F. Almutairi, “A deep learning model to classify and detect brain abnormalities in portable microwave based imaging system,” *Sci. Rep.*, vol. 12, no. 1, pp. 1–27, 2022, doi: 10.1038/s41598-022-10309-6.

[85] F. Prinzi, M. Insalaco, A. Orlando, S. Gaglio, and S. Vitabile, “A Yolo-Based Model for Breast Cancer Detection in Mammograms,” *Cognit. Comput.*, vol. 16, no. 1, pp. 107–120, 2024, doi: 10.1007/s12559-023-10189-6.

[86] Y. Luo, Y. Zhang, X. Sun, H. Dai, and X. Chen, “Intelligent Solutions in Chest Abnormality Detection Based on YOLOv5 and ResNet50,” *J. Healthc. Eng.*, vol. 2021, 2021, doi: 10.1155/2021/2267635.

[87] N. Akgül, C. Yilmaz, E. Bilgir, Ö. Çelik, O. Baydar, and İ. Ş. Bayrakdar, “A YOLO-V5 approach for the evaluation of normal fillings and overhanging fillings: an artificial intelligence study,” *Braz. Oral Res.*, vol. 38, p. e098, 2024, doi: 10.1590/1807-3107bor-2024.vol38.0098.

[88] X. Xu, X. Zhang, and T. Zhang, “Lite-YOLOv5: A Lightweight Deep Learning Detector for On-Board Ship Detection in Large-Scene Sentinel-1 SAR Images,” *Remote Sens.*, vol. 14, no. 4, 2022, doi: 10.3390/rs14041018.

[89] A. Korkmaz, M. T. Agdas, S. Kosunalp, T. Iliev, and I. Stoyanov, “Detection of Threats to Farm Animals Using Deep Learning Models: A Comparative Study,” *Appl. Sci.*, vol. 14, no. 14, 2024, doi: 10.3390/app14146098.

[90] O. G. Ajayi, P. O. Ibrahim, and O. S. Adegboyega, “Effect of Hyperparameter Tuning on the Performance of YOLOv8 for Multi Crop Classification on UAV Images,” *Appl. Sci.*, vol. 14, no. 13, 2024, doi: 10.3390/app14135708.

[91] J. Howard and S. Gugger, “Fastai: A layered api for deep learning,” *Inf.*, vol. 11, no. 2, pp. 1–26, 2020, doi: 10.3390/info11020108.

[92] P. D. Tsung-Yi Lin, Priya Goyal, Ross Girshick, Kaiming He, “Focal Loss for Dense Object Detection Tsung-Yi,” *Proc. IEEE Int. Conf. Comput. Vis.*, pp. 2980–2988, 2017, doi: 10.1109/ICAICTA49861.2020.9428882.

[93] A. Juyal, H. Tiwari, U. K. Singh, N. Kumar, and S. Kumar, “Dental Caries Detection Using Faster R-CNN and YOLO V3,” *ITM Web Conf.*, vol. 53, p. 02005, 2023, doi: 10.1051/itmconf/20235302005.

[94] E. S. M. Zia Uddin, D. M. I. Aslam, and D. M. Moinuddin, “Dental Carries Classification Using Yolo V8,” *J. Popul. Ther. Clin. Pharmacol.*, vol. 31, no. 6, pp. 2570–2586, 2024, doi: 10.53555/jptcp.v31i6.6983.

[95] Y. He, Y. Su, X. Wang, J. Yu, and Y. Luo, “An improved method MSS-YOLOv5 for object detection with balancing speed-accuracy,” *Front. Phys.*, vol. 10, no. January, pp. 1–13, 2023, doi: 10.3389/fphy.2022.1101923.

[96] “Ultralytics YOLO Hyperparameter Tuning Guide.” https://docs.ultralytics.com/guides/hyperparameter-tuning (accessed Dec. 20, 2024).

[97] “Hyperparameter evolution.”

https://docs.ultralytics.com/yolov5/tutorials/hyperparameter_evolution/ (accessed Dec. 15, 2024).

[98] I. Loshchilov and F. Hutter, “SGDR: Stochastic gradient descent with warm restarts,” *5th Int. Conf. Learn. Represent. ICLR 2017 - Conf. Track Proc.*, pp. 1–16, 2017.

[99] D. P. Kingma and J. L. Ba, “Adam: A method for stochastic optimization,” *3rd Int. Conf. Learn. Represent. ICLR 2015 - Conf. Track Proc.*, pp. 1–15, 2015.

[100] H. Guan and M. Liu, “Domain Adaptation for Medical Image Analysis: A Survey,” *IEEE Trans. Biomed. Eng.*, vol. 69, no. 3, pp. 1173–1185, 2022, doi: 10.1109/TBME.2021.3117407.

[101] G. Jocher, “YOLOv5 by Ultralytics.” .

[102] D. L. Duong *et al.*, “diagnostics Proof-of-Concept Study on an Automatic Computational System in Detecting and Classifying Occlusal Caries Lesions from Smartphone Color Images of Unrestored Extracted Teeth,” 2021, doi: 10.3390/diagnostics11071136.

[103] Y. Skaik, “Understanding and using sensitivity, specificity and predictive values,” *Indian J. Ophthalmol.*, vol. 56, no. 4, p. 341, 2008, doi: 10.4103/0301-4738.41424.

[104] D. M. W. Powers, “EVALUATION: FROM PRECISION, RECALL AND F-MEASURE TO ROC, INFORMEDNESS, MARKEDNESS & CORRELATION,” *J. Mach. Learn. Technol.*, vol. 2, no. 1, pp. 37–63, 2011.

[105] “Precision-Recall.” https://scikit-learn.org/stable/auto_examples/model_selection/plot_precision_recall.html (accessed Jan. 10, 2025).

[106] R. Padilla, W. L. Passos, T. L. B. Dias, S. L. Netto, and E. A. B. Da Silva, “A comparative analysis of object detection metrics with a companion open-source toolkit,” *Electron.*, vol. 10, no. 3, pp. 1–28, 2021, doi: 10.3390/electronics10030279.

[107] S. Schmid, H. Wei, and C. U. Grosse, “On the uncertainty in the segmentation of ultrasound images reconstructed with the total focusing method,” *SN Appl. Sci.*, vol. 5, no. 4, 2023, doi: 10.1007/s42452-023-05342-7.

[108] L. T. Arsiwala-Scheppach, A. Chaurasia, A. Müller, J. Krois, and F. Schwendicke, “Machine Learning in Dentistry: A Scoping Review,” *J. Clin. Med.*, vol. 12, no. 3, pp. 1–23, 2023, doi: 10.3390/jcm12030937.

[109] D. Albano *et al.*, “Artificial intelligence for radiographic imaging detection of caries lesions: a systematic review,” *BMC Oral Health*, vol. 24, no. 1, pp. 1–10, 2024, doi: 10.1186/s12903-024-04046-7.

[110] S. Datta and N. Chaki, “Detection of dental caries lesion at early stage based on image analysis technique,” *2015 IEEE Int. Conf. Comput. Graph. Vis. Inf. Secur. CGVIS 2015*, pp. 89–93, 2016, doi: 10.1109/CGVIS.2015.7449899.

[111] T. Shan, F. R. Tay, and L. Gu, “Application of Artificial Intelligence in Dentistry,” *J. Dent. Res.*, vol. 100, no. 3, pp. 232–244, 2021, doi: 10.1177/0022034520969115.

[112] M. K. Alam *et al.*, “Applications of artificial intelligence in the utilisation of imaging modalities in dentistry: A systematic review and meta-analysis of in-vitro studies,”

Heliyon, vol. 10, no. 3, p. e24221, 2024, doi: 10.1016/j.heliyon.2024.e24221.

[113] A. Chaurasia, A. Namachivayam, R. B. Koca-Ünsal, and J. H. Lee, "Deep-learning performance in identifying and classifying dental implant systems from dental imaging: a systematic review and meta-analysis," *J. Periodontal Implant Sci.*, vol. 54, no. 1, pp. 3–12, 2024, doi: 10.5051/jpis.2300160008.

[114] K. Potdar, T. S., and C. D., "A Comparative Study of Categorical Variable Encoding Techniques for Neural Network Classifiers," *Int. J. Comput. Appl.*, vol. 175, no. 4, pp. 7–9, 2017, doi: 10.5120/ijca2017915495.

[115] N. S. Alghamdi, M. Zakariah, and H. Karamti, *A deep CNN-based acoustic model for the identification of lung diseases utilizing extracted MFCC features from respiratory sounds*, no. 0123456789. Springer US, 2024.

[116] H. Zhang and Y. Qie, "Applying Deep Learning to Medical Imaging: A Review," *Appl. Sci.*, vol. 13, no. 18, 2023, doi: 10.3390/app131810521.

[117] H. Den, J. Ito, and A. Kokaze, "Diagnostic accuracy of a deep learning model using YOLOv5 for detecting developmental dysplasia of the hip on radiography images," *Sci. Rep.*, vol. 13, no. 1, pp. 1–10, 2023, doi: 10.1038/s41598-023-33860-2.

[118] T. Ahmad *et al.*, "Object Detection through Modified YOLO Neural Network," *Sci. Program.*, vol. 2020, 2020, doi: 10.1155/2020/8403262.

[119] R. Ren, H. Luo, C. Su, Y. Yao, and W. Liao, "Machine learning in dental, oral and craniofacial imaging: A review of recent progress," *PeerJ*, vol. 9, no. Ml, pp. 1–35, 2021, doi: 10.7717/peerj.11451.

[120] Y. Yang, H. Zhang, J. W. Gichoya, D. Katabi, and M. Ghassemi, "The limits of fair medical imaging AI in real-world generalization," *Nat. Med.*, vol. 30, no. October, 2024, doi: 10.1038/s41591-024-03113-4.

[121] J. Terven, D. M. Córdova-Esparza, and J. A. Romero-González, "A Comprehensive Review of YOLO Architectures in Computer Vision: From YOLOv1 to YOLOv8 and YOLO-NAS," *Mach. Learn. Knowl. Extr.*, vol. 5, no. 4, pp. 1680–1716, 2023, doi: 10.3390/make5040083.

[122] M. Tassoker, M. Ü. Öziç, and F. Yuce, "Performance evaluation of a deep learning model for automatic detection and localization of idiopathic osteosclerosis on dental panoramic radiographs," *Sci. Rep.*, vol. 14, no. 1, pp. 1–14, 2024, doi: 10.1038/s41598-024-55109-2.

[123] S. AbuSalim, N. Zakaria, M. R. Islam, G. Kumar, N. Mokhtar, and S. J. Abdulkadir, "Analysis of Deep Learning Techniques for Dental Informatics: A Systematic Literature Review," *Healthc.*, vol. 10, no. 10, pp. 1–30, 2022, doi: 10.3390/healthcare10101892.

[124] P. Shah and K. Zalavadia, "Recent Advances in Dental Imaging," *Sch. J. Dent. Sci.*, vol. 8, no. 6, pp. 184–188, 2021, doi: 10.36347/sjds.2021.v08i06.004.

[125] F. Schwendicke, S. Mertens, A. G. Cantu, A. Chaurasia, H. Meyer-Lueckel, and J. Krois, "Cost-effectiveness of AI for caries detection: randomized trial," *J. Dent.*, vol. 119, no. March, p. 104080, 2022, doi: 10.1016/j.jdent.2022.104080.

[126] N. N. Khanna *et al.*, "Economics of Artificial Intelligence in Healthcare: Diagnosis vs.

Treatment," *Healthc.*, vol. 10, no. 12, 2022, doi: 10.3390/healthcare10122493.

[127] X. Wu *et al.*, "Cost-effectiveness and cost-utility of a digital technology-driven hierarchical healthcare screening pattern in China," *Nat. Commun.*, vol. 15, no. 1, 2024, doi: 10.1038/s41467-024-47211-w.

[128] B. Vandenberghe, "The digital patient – Imaging science in dentistry," *J. Dent.*, vol. 74, no. April, pp. S21–S26, 2018, doi: 10.1016/j.jdent.2018.04.019.

[129] A. Katsumata, "Deep learning and artificial intelligence in dental diagnostic imaging," *Jpn. Dent. Sci. Rev.*, vol. 59, no. September, pp. 329–333, 2023, doi: 10.1016/j.jdsr.2023.09.004.

[130] M. Hussain, "YOLO-v1 to YOLO-v8, the Rise of YOLO and Its Complementary Nature toward Digital Manufacturing and Industrial Defect Detection," *Machines*, vol. 11, no. 7, 2023, doi: 10.3390/machines11070677.

[131] T. J. Lin *et al.*, "Evaluation of the Alveolar Crest and Cemento-Enamel Junction in Periodontitis Using Object Detection on Periapical Radiographs," *Diagnostics*, vol. 14, no. 15, 2024, doi: 10.3390/diagnostics14151687.

[132] H. H. and S. J. Maya Fitria, Yasmina Elma, Maulisa Oktiana, Khairun Saddami, Rizki Novita, Rizkika Putri, Handika Rahayu, "THE DEEP LEARNING MODEL FOR DECAYED- MISSING-FILLED TEETH DETECTION: A COMPARISON BETWEEN YOLOV5 AND YOLOV8," *JJCIT*, vol. 10, no. 3, pp. 335–348, 2024, [Online]. Available: http://hliangzhao.me/slides/synergy.pdf.

[133] B. Wu, C. Pang, X. Zeng, and X. Hu, "ME-YOLO: Improved YOLOv5 for Detecting Medical Personal Protective Equipment," *Appl. Sci.*, vol. 12, no. 23, 2022, doi: 10.3390/app122311978.

[134] N. Dahiya *et al.*, "Hyper-parameter tuned deep learning approach for effective human monkeypox disease detection," *Sci. Rep.*, vol. 13, no. 1, pp. 1–18, 2023, doi: 10.1038/s41598-023-43236-1.

[135] M. L. Ali and Z. Zhang, "The YOLO Framework: A Comprehensive Review of Evolution, Applications, and Benchmarks in Object Detection," *Computers*, vol. 13, no. 12, 2024, doi: 10.3390/computers13120336.

[136] R. B. Peker, "Evaluation of the Performance of a YOLOv10-Based Deep Learning Model for Tooth Detection and Numbering on Panoramic Radiographs of Patients in the Mixed Dentition Period," 2025.

[137] A. Ramírez-Pedraza *et al.*, "Deep Learning in Oral Hygiene: Automated Dental Plaque Detection via YOLO Frameworks and Quantification Using the O'Leary Index," *Diagnostics*, vol. 15, no. 2, pp. 1–20, 2025, doi: 10.3390/diagnostics15020231.

[138] K. Perez, D. Wisniewski, A. Ari, K. Lee, and C. Lieneck, "Investigation into Application of AI and Telemedicine in Rural Communities : A Systematic Literature Review," pp. 1–54, 2025.

[139] K. Reddy, A. Taksande, and B. Kurian, "Harnessing the Power of Mobile Phone Technology: Screening and Identifying Autism Spectrum Disorder With Smartphone Apps," *Cureus*, vol. 16, no. 2, 2024, doi: 10.7759/cureus.55004.

[140] F. Carrillo-Perez *et al.*, "Applications of artificial intelligence in dentistry: A

comprehensive review," *J. Esthet. Restor. Dent.*, vol. 34, no. 1, pp. 259–280, 2022, doi: 10.1111/jerd.12844.

[141] Y. Wang, F. Sibaii, K. Lee, M. J. Gill, and J. L. Hatch, "Leveraging AI Tools to Bridge the Healthcare Gap in Rural Areas in India," *medRxiv*, vol. 1, no. 165, pp. 1–13, 2021.

[142] L. Liu, J. Xu, Y. Huan, Z. Zou, S. C. Yeh, and L. R. Zheng, "A Smart Dental Health-IoT Platform Based on Intelligent Hardware, Deep Learning, and Mobile Terminal," *IEEE J. Biomed. Heal. Informatics*, vol. 24, no. 3, pp. 898–906, 2020, doi: 10.1109/JBHI.2019.2919916.

[143] N. Alzaid *et al.*, "Revolutionizing Dental Care: A Comprehensive Review of Artificial Intelligence Applications Among Various Dental Specialties," *Cureus*, vol. 15, no. 10, 2023, doi: 10.7759/cureus.47033.

[144] D. N. K. Dr. Anjum Younus, Dr Abdul Aleem, Dr Zubia Waqar, Dr Samra Bokhari, Dr Sana Adeeba Islam, "INVESTIGATING THE ROLE OF ARTIFICIAL INTELLIGENCE IN DENTAL DIAGNOSTICS AND TREATMENT PLANNING – A COMPARATIVE STUDY," vol. 31, no. 09, pp. 2362–2372, 2024, doi: 10.53555/q833js02.

[145] R. Pauwels, "A brief introduction to concepts and applications of artificial intelligence in dental imaging," *Oral Radiol.*, vol. 37, no. 1, pp. 153–160, 2021, doi: 10.1007/s11282-020-00468-5.

[146] I. Shafi *et al.*, "A Comprehensive Review of Recent Advances in Artificial Intelligence for Dentistry E-Health," *Diagnostics*, vol. 13, no. 13, pp. 1–30, 2023, doi: 10.3390/diagnostics13132196.

[147] H. Chennamsetty, S. Chalasani, and D. Riley, "Predictive analytics on Electronic Health Records (EHRs) using Hadoop and Hive," *Proc. 2015 IEEE Int. Conf. Electr. Comput. Commun. Technol. ICECCT 2015*, pp. 1–5, 2015, doi: 10.1109/ICECCT.2015.7226129.

[148] R. B. Chauhan *et al.*, "An overview of image processing for dental diagnosis," *Innov. Emerg. Technol.*, vol. 10, no. September 2022, 2023, doi: 10.1142/s2737599423300015.

[149] A. Nascita, A. Montieri, G. Aceto, D. Ciuonzo, V. Persico, and A. Pescape, "XAI Meets Mobile Traffic Classification: Understanding and Improving Multimodal Deep Learning Architectures," *IEEE Trans. Netw. Serv. Manag.*, vol. 18, no. 4, pp. 4225–4246, 2021, doi: 10.1109/TNSM.2021.3098157.

[150] D. Cawthorne and A. Robbins-van Wynsberghe, "An Ethical Framework for the Design, Development, Implementation, and Assessment of Drones Used in Public Healthcare," *Sci. Eng. Ethics*, vol. 26, no. 5, pp. 2867–2891, 2020, doi: 10.1007/s11948-020-00233-1.

[151] C. A. Gomez-Cabello, S. Borna, S. Pressman, S. A. Haider, C. R. Haider, and A. J. Forte, "Artificial-Intelligence-Based Clinical Decision Support Systems in Primary Care: A Scoping Review of Current Clinical Implementations," *Eur. J. Investig. Heal. Psychol. Educ.*, vol. 14, no. 3, pp. 685–698, 2024, doi: 10.3390/ejihpe14030045.

[152] X. Liu *et al.*, "A comparison of deep learning performance against health-care

professionals in detecting diseases from medical imaging: a systematic review and meta-analysis," *Lancet Digit. Heal.*, vol. 1, no. 6, pp. e271–e297, 2019, doi: 10.1016/S2589-7500(19)30123-2.

[153] Y. Yu, "Machine Learning for Dental Image Analysis," 2016, [Online]. Available: http://arxiv.org/abs/1611.09958.

[154] M. C. Kılıc *et al.*, "Artificial intelligence system for automatic deciduous tooth detection and numbering in panoramic radiographs," *Dentomaxillofacial Radiol.*, vol. 50, no. 6, 2021, doi: 10.1259/dmfr.20200172.

[155] Z. Zheng, H. Yan, F. C. Setzer, K. J. Shi, M. Mupparapu, and J. Li, "Anatomically Constrained Deep Learning for Automating Dental CBCT Segmentation and Lesion Detection," *IEEE Trans. Autom. Sci. Eng.*, vol. 18, no. 2, pp. 603–614, 2021, doi: 10.1109/TASE.2020.3025871.

[156] K. Giannakopoulos, A. Kavadella, A. A. Salim, V. Stamatopoulos, and E. G. Kaklamanos, "Evaluation of the Performance of Generative AI Large Language Models ChatGPT, Google Bard, and Microsoft Bing Chat in Supporting Evidence-Based Dentistry: Comparative Mixed Methods Study," *J. Med. Internet Res.*, vol. 25, no. 1, 2023, doi: 10.2196/51580.

[157] M. E. Celik, "Deep Learning Based Detection Tool for Impacted Mandibular Third Molar Teeth," *Diagnostics*, vol. 12, no. 4, 2022, doi: 10.3390/diagnostics12040942.

[158] G. Jader, J. Fontineli, M. Ruiz, K. Abdalla, M. Pithon, and L. Oliveira, "Deep Instance Segmentation of Teeth in Panoramic X-Ray Images," *Proc. - 31st Conf. Graph. Patterns Images, SIBGRAPI 2018*, pp. 400–407, 2018, doi: 10.1109/SIBGRAPI.2018.00058.

[159] R. A. Welikala *et al.*, "Automated Detection and Classification of Oral Lesions Using Deep Learning for Early Detection of Oral Cancer," *IEEE Access*, vol. 8, pp. 132677–132693, 2020, doi: 10.1109/ACCESS.2020.3010180.

[160] F. P. Mahdi, N. Yagi, and S. Kobashi, "Automatic Teeth Recognition in Dental X-Ray Images Using Transfer Learning Based Faster R-CNN," *Proc. Int. Symp. Mult. Log.*, vol. 2020-Novem, pp. 16–21, 2020, doi: 10.1109/ISMVL49045.2020.00-36.

[161] H. Chen, H. Li, Y. Zhao, J. Zhao, and Y. Wang, "Dental disease detection on periapical radiographs based on deep convolutional neural networks," *Int. J. Comput. Assist. Radiol. Surg.*, vol. 16, no. 4, pp. 649–661, 2021, doi: 10.1007/s11548-021-02319-y.

[162] K. Singh and M. Abrol, "To Use a Quick R-CNN Algorithm - An Automated Strategy for Tooth Diagnostics," *Proc. 2022 11th Int. Conf. Syst. Model. Adv. Res. Trends, SMART 2022*, pp. 1160–1164, 2022, doi: 10.1109/SMART55829.2022.10047030.

[163] K. Hung, A. W. K. Yeung, R. Tanaka, and M. M. Bornstein, "Current applications, opportunities, and limitations of AI for 3D imaging in dental research and practice," *Int. J. Environ. Res. Public Health*, vol. 17, no. 12, pp. 1–18, 2020, doi: 10.3390/ijerph17124424.

Appendix Table 1

Table 8. Scoping review extracted information.

S.no.	Authors	Focus area	AI/ML model used	Study design	Key findings
1.	Yu et al. [153]	The use of AI in dental imaging, specifically leveraging ML models such as CNNs to interpret and analyze dental radiographs [153].	CNNs, including a 4-layer and a 16-layer model, were utilized. Additionally, other ML techniques such as Bag-of-Words models, the k-nearest neighbors algorithm, and SVMs were also considered.	The study involved experiments using panoramic dental radiographs of patients from Pusan National University Hospital. AI models were trained and tested for two tasks: teeth classification and sex classification from the radiographs.	CNNs were effective in classifying teeth with a high degree of accuracy, especially when using a 16-layer model which achieved higher accuracy rates over multiple training iterations. The 4-layer CNN model performed well in sex classification tasks, achieving up to 97.5% accuracy. Bag-of-Words models, k-NN, and SVMs also showed good performance but were generally less effective compared to CNNs.
2.	Kilic et al.[154]	The application of AI, particularly DL, for the automated identification and numbering of deciduous teeth in children using panoramic radiographs [154].	The model employed is Faster R-CNN Inception v2 (COCO), a DL framework built on CNNs and optimized for object detection tasks.	The research involves training and testing an AI algorithm on a dataset comprising 421 panoramic images of children's teeth. The AI system's	The AI system exhibited strong performance in detecting and numbering deciduous teeth, achieving a sensitivity of 98.04%, precision of

				performance was assessed using a confusion matrix to evaluate its sensitivity, precision, and F1 score.	95.71%, and an F1 score of 96.86%. These findings indicate that DL -based AI models hold significant potential for the automated charting of children's panoramic dental radiographs.
3.	Zheng et al. [155]	The study aims to enhance dental care automation through AI, utilizing anatomically constrained DL for automated segmentation and lesion detection in dental CBCT [155].	The research employs an anatomically constrained Dense U-Net, a modified version of the Dense U-Net that integrates anatomical knowledge as constraints within the DL model for CBCT segmentation and lesion detection.	The research integrates oral anatomical knowledge with DL to improve CBCT segmentation and lesion detection. It introduces a novel regularized optimization framework that employs mean-field variational approximation for enhanced computational efficiency. The system is evaluated using a dataset of CBCT images, with its performance compared to standard Dense U-Net models under different loss functions.	The anatomically constrained Dense U-Net surpasses the standard Dense U-Net in both lesion detection accuracy and voxel-matching Dice coefficient (DICE) accuracy for each label. The system exhibits notable improvements, particularly in identifying challenging labels such as lesions and materials, highlighting the advantages of incorporating anatomical knowledge into DL models.
4.	Giannakopoulos et al. [156]	The study examines the performance of	The discussed AI models are	A comparative mixed-methods	ChatGPT-4 performed the

		generative AI LLMs, including ChatGPT (versions 3.5 and 4), Google Bard, and Microsoft Bing Chat, in facilitating evidence-based dentistry. It evaluates their accuracy in responding to clinically relevant questions within the field of dentistry [156].	generative LLMs, including ChatGPT-3.5, ChatGPT-4, Google Bard, and Microsoft Bing Chat. Built on the transformer architecture, these models are specifically designed for natural language processing, comprehension, and generation.	study in which each LLM was tested with 20 clinical dentistry-related questions. Responses were evaluated on a scale of 0 to 10 based on strong scientific evidence, including guidelines and consensus statements, by two experienced faculty members. Statistical analyses were conducted to compare the scores and determine the highest-performing model.	best among the models, showing statistically significant better results compared to other LLMs. However, all models occasionally displayed inaccuracies and lacked detailed source citations, raising concerns about their use in clinical decision-making without proper oversight.
5.	Rashidi et al. [20]	The implementation of a DL based object detection algorithm specifically designed for autonomous dental treatment planning using panoramic X-ray images [20].	The model used is YOLOv7, a CNN model part of the YOLO family, known for its high accuracy and quick performance on small datasets.	The research utilized a dataset of 1,025 panoramic dental X-ray images, annotated with the guidance of a dentist with over 25 years of experience. These images were employed to train and evaluate the YOLOv7 model for the autonomous detection of damaged teeth and the recommendation	The YOLOv7 model achieved a precision of 100%, an F1 score of 76%, a recall of 96%, and a mAP of 81.9%. These results indicate excellent potential for clinical use, suggesting that the model is highly effective in identifying damaged teeth and recommending appropriate treatments.

				of treatment plans.	
6.	Celik et al., [157]	The development of a computer-assisted detection system utilizing deep CNNs to identify impacted mandibular third molar teeth using various architectures on panoramic radiographs[157].	The study employs DL models such as Faster R-CNN with a ResNet50 backbone, AlexNet, VGG16, and YOLOv3 to detect impacted third molar teeth in panoramic radiographs.	The research compares two detector techniques, Faster R-CNN and YOLOv3, applied to 440 panoramic radiographs from 300 patients to assess their effectiveness in detecting impacted third molar teeth. The evaluation is based on performance metrics such mAP, recall, and precision.	YOLOv3 showed the highest detection efficacy with an mAP@0.5 of 0.96, and Faster RCNN with a ResNet50 backbone yielded a mAP@0.5 rate of 0.91. The study concluded that YOLOv3 had excellent performance for detecting impacted mandibular third molar teeth.
7.	Jader et al. [158]	DL method for instance segmentation of teeth in panoramic X-ray images. It specifically addresses the challenges of isolating teeth in these images, which also display other parts of the patient's body [158].	The study employs a Mask R-CNN for instance segmentation of the teeth.	The study introduces a segmentation system designed to detect and segment individual teeth in panoramic X-ray images. Its performance was evaluated using a diverse dataset of 1,500 images, encompassing 10 categories of various buccal image types.	The proposed Mask R-CNN system achieved 98% accuracy, 88% F1-score, 94% precision, 84% recall, and 99% specificity over 1224 unseen images. These results were superior to 10 unsupervised methods previously studied, highlighting the effectiveness of the DL approach in this complex imaging scenario.
8.	Welikala et al. [159]	The automated detection and	The AI models discussed	This research utilized a dataset	The study found that DL

		classification of oral lesions using DL to facilitate early detection of oral cancer. It emphasizes the development and application of AI technologies for analyzing images captured through mobile devices as part of a telemedicine approach [159].	include ResNet-101 for image classification and Faster R-CNN for object detection. These models were employed to analyze oral cavity images for the presence of lesions and other indicators of potential malignancies.	of 2155 oral cavity images from 1085 individuals, employing DL to classify and detect oral lesions. The study evaluated two different computer vision approaches: image classification and object detection, to assess their effectiveness in the early detection of oral cancer.	models, particularly ResNet-101 and Faster R-CNN, are effective tools for detecting and classifying oral lesions. ResNet-101 showed a high F1 score in identifying images that contain lesions or require referrals, while Faster R-CNN was effective in detecting and classifying the location and type of lesions.
9.	Mahdi et al. [160]	The automatic recognition of teeth in dental X-ray images using a transfer learning-based Faster R-CNN approach, following the universal tooth numbering system [160].	The model used is a Faster R-CNN, with transfer learning from a pre-trained ResNet-50 model. This approach is chosen for its efficiency in automatic feature extraction and object detection.	The research utilizes 900 panoramic dental X-ray images, split into a training set of 810 images and a test set of 90 images. The system is trained to recognize and number teeth based on their position in the X-ray images.	The system achieved mAP of 0.942, indicating high reliability and accuracy in the automatic recognition of teeth. This demonstrates the model's potential utility in aiding dental professionals.
10.	Hu Chen et al. [161]	Developing an auxiliary diagnosis system for detecting dental diseases in periapical radiographs using deep CNNs. It specifically targets decay, periapical periodontitis, and periodontitis, and	The research utilizes Faster R-CNN, a DL model for object detection. The Faster R-CNN is equipped with region proposal techniques to	The study involved training CNNs on a dataset of 2,900 periapical radiographs, using different training strategies to handle disease categories and	The CNNs were effective in detecting lesions, particularly more severe conditions. Training strategies that customized the CNNs for

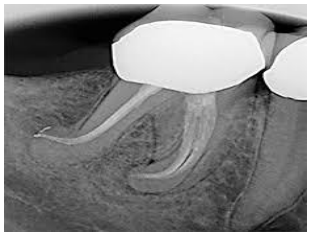
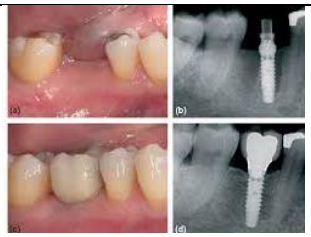
		explores their detection across different severity levels [65].	identify and classify dental lesions in periapical radiographs.	severity levels. Fivefold cross-validation was used to evaluate model performance with metrics such as IoU, precision, recall, and AP.	specific diseases or severity levels slightly improved performance in certain scenarios. The influence of training strategy, disease category, and severity level on detection performance was statistically significant.
11.	Kushdeep et al. [162]	The use of a Quick R-CNN algorithm for automated tooth diagnostics. The study involves detecting various dental issues like cavities and oral diseases through an automated diagnostic model trained on clinical dental images [162].	The model used is a Quick R-CNN, which is applied for object detection in dental diagnostic imaging.	The study involved collecting clinical dental images from various clinics, upon which the Quick R-CNN was trained. The images were used to train and construct an automated diagnostic model capable of identifying different types of dental cavities and oral disorders.	The Quick R-CNN model provided high diagnostic accuracies, sensitivity, and specificity in identifying dental issues. It successfully differentiated between normal teeth and those affected by specific dental conditions.
12.	Hung et al. [163]	Current applications, opportunities, and limitations of AI for 3D imaging in dental research and practice, focusing particularly on dentomaxillofacial radiology and intraoral and facial scanning [163].	The review discusses various ML techniques, including CNNs, SVMs, and other DL models. These models are used for automated diagnosis, treatment	It discusses the methodology and findings of studies that have applied AI to 3D imaging in dental practice, highlighting the benefits and limitations of these technologies.	AI applications in dental 3D imaging show promise in improving diagnostic accuracy, automating the localization of anatomical landmarks, and enhancing image quality.

			planning, and enhancing image quality in dental imaging.		However, challenges remain, including the need for large and diverse datasets to train the models effectively and the integration of these systems into clinical workflows.
13.	Nakaura et al. [13]	The impact of LLMs on the field of radiology, exploring how these models can enhance various aspects of the radiology workflow, including report generation, diagnostics, and patient care, while also addressing potential challenges and limitations [13].	The use of LLMs, particularly those based on the Transformer architecture, such as GPT and BERT.	The paper outlines how LLMs have been integrated into radiological practices, assessing both their advantages and the complexities they introduce.	LLMs offer substantial potential to improve the efficiency and accuracy of radiological processes through automation and advanced data handling capabilities. However, challenges such as data hallucination, bias, and the need for substantial training datasets are significant and need addressing to ensure clinical reliability.
14.	Ruben Pauwels[145]	The concepts and applications of AI in dental imaging, covering diagnostics, forensics, image processing, and image reconstruction. It aims to provide a	The use of ML and DL, particularly CNNs, in dental imaging applications. These AI models are used for tasks like diagnosing	A narrative review provides an overview of AI concepts and applications in dental imaging, including ML and DL methodologies. The study	AI technologies, particularly DL, are proving valuable in dental imaging for tasks such as lesion detection, forensic

		non-exhaustive overview of AI's role in enhancing dental radiology through various ML and DL approaches [145].	pathologies, forensic analysis, and image processing.	summarizes various AI applications in diagnostics, forensics, image processing, and reconstruction rather than conducting a structured experiment or comparative analysis.	analysis, and automatic image processing. These technologies can perform at or above the level of human experts in some cases, offering significant time savings and potential accuracy improvements in diagnostic processes.
15.	Putra et al. [42]	The current applications and development of AI in digital dental radiography, exploring its use across various diagnostic and treatment planning tasks, including detection of dental caries, periodontal bone loss, cephalometric analysis, and enhancements in image quality [42].	The application of several AI models, particularly DL methods like CNNs, for tasks such as image classification, detection, and segmentation in dental radiography.	It includes a literature search up to December 2020, categorizing AI applications by their purposes and discussing methodological developments and challenges in each area.	AI has shown significant potential in enhancing dental radiographic assessments, providing more accurate and reproducible results compared to traditional methods. The review highlights various successful applications but also notes ongoing challenges such as the need for larger and more diverse datasets and standardization in reporting and methodological approaches.

Appendix Table 2

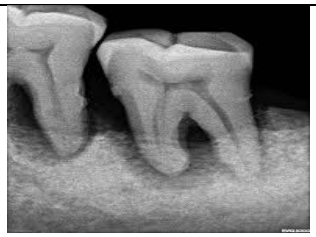
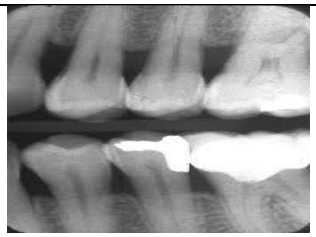
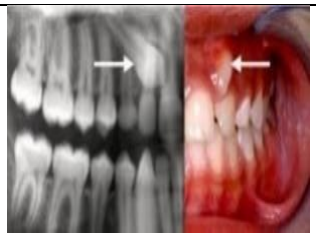
Table 9. Dental annotation classes and their clinical features [50]

S.No.	Class Annotated	Clinical Feature	Clinical Image	Dental Disease
1.	Curved root	Root shape anomaly		Root anomaly
2.	Dental Filling	Dental restoration		Restoration status
3.	Dental caries	Tooth decay		Caries
4.	Dental implant	Prosthetic replacement		Tooth replacement
5.	Impacted teeth	Tooth eruption anomaly		Impaction
6.	Missing teeth	Absence of teeth		Edentulism

7.	RCT Treated teeth	Endodontic treatment		Endodontic treatment outcome
8.	RCT Treated tooth	Endodontic treatment		Endodontic treatment outcome
9.	Root stumps	Residual root fragments		Residual infection risk
10.	Attrition	Tooth wear		Attrition
11.	Completely edentulous mandible	Complete tooth loss in mandible		Edentulism
12.	Completely edentulous maxilla	Complete tooth loss in maxilla		Edentulism
13.	Conical tooth	Abnormal tooth shape		Developmental anomaly

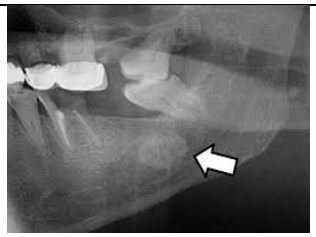
14.	Dental crown	Tooth restoration		Restoration status
15.	Dental post	Post used in restoration		Endodontic restoration
16.	Dental post and crown	Post and crown restoration		Endodontic restoration
17.	Developing permanent tooth	Tooth development		Tooth development
18.	Developing tooth bud	Developing tooth structure		Tooth development
19.	Dilaceration	Root abnormality		Root anomaly

20.	Edentulous ridge	Residual alveolar ridge		Edentulism
21.	Erupting permanent tooth	Tooth eruption		Tooth eruption status
22.	Exfoliating primary tooth	Shedding primary tooth		Primary tooth shedding
23.	Extensive dental caries	Severe tooth decay		Severe caries
24.	Fixed partial denture	Prosthetic appliance		Restoration status
25.	Fractured tooth	Tooth fracture		Trauma

26.	Furcation bone loss	Advanced periodontal disease		Periodontitis
27.	Fused roots	Root anomaly		Root anomaly
28.	Fusion of teeth	Tooth fusion anomaly		Fusion anomaly
29.	Horizontal bone loss	Bone loss in periodontium		Periodontitis
30.	Horizontal impaction	Impacted tooth in horizontal position		Impaction
31.	Impacted canine	Impacted canine		Impaction

32.	Impacted premolar	Impacted premolar		Impaction
33.	Impacted tooth	Tooth impaction		Impaction
34.	Mesially tilted tooth	Tooth tilting anomaly		Malocclusion
35.	Metal wire	Orthodontic material		Orthodontic treatment
36.	Mini implant	Prosthetic mini-implant		Prosthetic replacement
37.	Missing filling	Lost filling		Restoration failure

38.	Missing laterals	Missing lateral incisor		Developmental anomaly
39.	Open apex	Root development anomaly		Root development
40.	Orthodontic brackets with wire	Orthodontic treatment		Orthodontic treatment
41.	Para molar	Extra tooth		Developmental anomaly
42.	Partially impacted tooth	Partially erupted tooth		Impaction
43.	Peg laterals	Small lateral incisor		Developmental anomaly

44.	Periapical lesion	Apical infection		Periapical pathology
45.	Primary tooth	Primary dentition		Primary tooth
46.	Proximal dental caries	Decay near tooth contact		Caries
47.	Radio opaque structure	Opaque structure in radiograph		Radiographic finding
48.	Recently extracted socket	Healing socket after extraction		Post-extraction healing
49.	Retained deciduous tooth	Deciduous tooth retained		Developmental anomaly

50.	Root canal filling	Endodontic treatment		Endodontic treatment
51.	Root caries	Decay on root surface		Root decay
52.	Secondary caries	Recurrent decay		Recurrent decay
53.	Spacing	Abnormal tooth spacing		Malocclusion
54.	Tooth bud	Developing tooth structure		Tooth development
55.	Veneers	Tooth veneers		Cosmetic restoration

56.	Vertical bone loss	Bone loss in periodontium		Periodontitis