applied system innovation

MDPI

Article

Harnessing the Power of an Integrated Artificial Intelligence Model for Enhancing Reliable and Efficient Dental Healthcare Systems

Samar M. Nour [1,*,†], Reem Salah Shehab [1,†], Samar A. Said [2,†] and Islam Tharwat Abdel Halim [3,4,*,†]

[1] Department of Computer and Systems Engineering, Faculty of Engineering & Technology, Badr University in Cairo (BUC), Cairo 11829, Egypt; reemsalah1501@gmail.com
[2] Department of Computer and Systems Engineering, Faculty of Engineering, Helwan University, Cairo 11795, Egypt; samar_said@h-eng.helwan.edu.eg
[3] School of Information Technology and Computer Science (ITCS), Nile University, Giza 12677, Egypt
[4] Center for Informatics Science (CIS), Nile University, 26th of July Corridor, Sheikh Zayed 12677, Egypt
* Correspondence: samar.nour@buc.edu.eg (S.M.N.); ihalim@nu.edu.eg (I.T.A.H.)
† These authors contributed equally to this work.

Abstract: Nowadays, efficient dental healthcare systems are considered significant for upholding oral health. Also, the ability to utilize artificial intelligence for evaluating complex data implies that dental X-ray image recognition is a critical mechanism to enhance dental disease detection. Consequently, integrating deep learning algorithms into dental healthcare systems is considered a promising approach for enhancing the reliability and efficiency of diagnostic processes. In this context, an integrated artificial intelligence model is proposed to enhance model performance and interpretability. The basic idea of the proposed model is to augment the deep learning approach with Ensemble methods to improve the accuracy and robustness of dental healthcare. In the proposed model, a Non-Maximum Suppression (NMS) ensembled technique is employed to improve the accuracy of predictions along with combining outputs from multiple single models (YOLO8 and RT-DETR) to make a final decision. Experimental results on real-world datasets show that the proposed model gives high accuracy in miscellaneous dental diseases. The results show that the proposed model achieves 18% time reductions as well as 30% improvements in accuracy compared with other competitive deep learning algorithms. In addition, the effectiveness of the proposed integrated model, achieved 74% mAP50 and 58% mAP50-90, outperforming existing models. Furthermore, the proposed model grants a high degree of system reliability.

Keywords: deep learning; dental diseases; dental X-ray; ensemble learning; healthcare systems; machine learning; smart dentistry

Academic Editors: Manuel Casal-Guisande and Alberto Comesaña-Campos

Received: 3 November 2024
Revised: 21 December 2024
Accepted: 24 December 2024
Published: 2 January 2025

Citation: Nour, S.M.; Shehab, R.S.; Said, S.A.; Abdel Halim, I.T. Harnessing the Power of an Integrated Artificial Intelligence Model for Enhancing Reliable and Efficient Dental Healthcare Systems. *Appl. Syst. Innov.* **2025**, *8*, 7. https://doi.org/10.3390/asi8010007

1. Introduction

Digitalization and industry development are examples of new innovative technologies and processes in Industry 4.0. Digital technology is increasingly integrated into our daily lives. So, people's lives are more aided by electronic gadgets; the data those devices collect allows artificial intelligence (AI) to more readily use and analyze the data those devices produce. AI is rapidly growing and blooming throughout various sectors. Additionally, AI has been used in the fields of healthcare and dentistry, including drug development, wearable technology, hospital monitoring, robotic and automated decision support, precision and digital medicine, and medical and dental imaging diagnostics [1]. AI can be seen as a

useful tool in many situations to assist healthcare professionals and dentists in lessening their workload [2].

Artificial neural networks (ANNs) can be used as clinical decision support systems (CDSS) to give medical professionals expert support [3]. Deep learning studies that efficiently diagnose dental disease and plan treatments are among the many ways artificial intelligence (AI) is being used in the fields of periodontology, endodontics, orthodontics, restorative dentistry, and oral pathology. Nevertheless, due to restrictions on dataset availability, size, performance, and generalizability, there is very little proof of AI at this time. Throughout the past ten years, artificial intelligence has significantly benefited several dental subfields [4].

Therefore, there is a need to develop automated and accurate tools for detecting and classifying dental diseases using digital images. One such tool is the panoramic dental X-ray, or panorex. This two-dimensional (2-D) dental X-ray examination captures the entire mouth in a single image, including the teeth, upper and lower jaws, and surrounding structures and tissues [5]. In this paper, an integrated deep learning and Ensemble learning model for reliable and efficient dental healthcare is presented. The proposed model uses two state-of-the-art deep learning models, YOLO8 and RT-DETR, to detect dental diseases on panoramic X-ray images.

YOLO8 is an improved version of the You Only Look Once (YOLO) family of models, which are fast and efficient for object detection tasks. The model works by analyzing the input image and predicting bounding boxes and class probabilities directly from the image in one evaluation, hence the name "You Only Look Once". Real-Time Detection Transformer (RT-DETR) is a cutting-edge end-to-end object detector developed by Baidu. It provides real-time performance while maintaining high accuracy. By separating intra-scale interaction and cross-scale fusion, RT-DETR effectively processes multiscale features by utilizing the power of Vision Transformers (ViT). It is highly adaptable, supporting the flexible adjustment of inference speed using different decoder layers without retraining. However, like any deep learning model, the performance of YOLO8 and RT-DETR heavily depends on the quality and diversity of the training data.

We can summarize the major contributions of that paper as follows:

(1) We propose a new integrated artificial intelligence model for reliable and efficient dental healthcare that optimizes and guarantees the effectiveness of the detection of dental diseases, leading to better patient outcomes.
(2) The proposed model helps a professional clinic achieve accurate automated diagnosis in minimal time and involves integrating various advanced technologies and methodologies.
(3) The proposed model redefines the doctor–patient relationship, streamlining communication for easier, more consistent, and highly accurate interactions.
(4) The proposed integrated model is evaluated across several real-world datasets to examine the performance and reliability of the proposed model through real X-ray images.

This paper is organized as follows: Sections 2 and 3 describe the background, motivation, and literature review. Section 4 demonstrates the proposed methodology, including the system model, problem formulation, as well as the implementation details of the proposed integrated model. Section 5 explores the performance evaluation and detailed results. In Section 6, the paper concludes the main findings and the future direction.

2. Background and Motivation

2.1. Dental Health Care System

Dental healthcare system installation is not without its difficulties, despite its many benefits. Due to the potential need for substantial investment in software, hardware, and employee training, initial expenditures may be unaffordable. Furthermore, strong security measures are required to protect sensitive patient data from breaches in order to ensure data privacy. Adoption by practitioners can sometimes be difficult, especially for more established practices that could find the switch to digital systems intimidating [6].

Technical developments will influence dental healthcare systems in the future. Advanced diagnoses and individualized treatment plans will be made possible by AI and machine learning, while 3D printing will transform the production of braces and implants. Blockchain will guarantee safe and easy patient data sharing, while wearable technology will offer real-time oral health monitoring. In addition, to improve effectiveness, accessibility, and patient satisfaction, these systems combine technology, billing, analytics, medical services, and patient management. These developments offer a smoother and superior standard of dental care, which will benefit both patients and providers, despite obstacles including high upfront costs and data privacy issues [7].

2.2. Machine Learning in Dental Health Care System

Artificial intelligence (AI) encompasses a range of subsets, such as speech recognition, fuzzy logic, deep learning, conventional neural networks, and machine learning (ML). Each of these subsets has special abilities and features that might enhance the performance of contemporary medical sciences. These smart solutions make it easier for humans to intervene in clinical diagnosis, medical imaging, and decision-making processes. The most popular AI technique for pattern-based prediction is machine learning (ML). Machine learning can be divided into various types depending on the algorithm's structure and the learning approach used. There are three additional types of learning methods: supervised, unsupervised, and reinforcement learning. Supervised learning involves training the process with input data. When previous data are accessible and can be utilized to forecast potential future events, supervised learning is employed [8].

2.3. Deep Learning in Dental Health Care System

Since these algorithms are trained on historical data, the techniques are more precise and uncomplicated. Regression and classification algorithms are two further categories into which these algorithms can be separated. Algorithms are less accurate and more computationally complex in this direction. Reinforced learning is a technique for learning that relies on rewards and penalties. Algorithms give undesirable impacts a negative value and desired outcomes a positive value [8].

These algorithms take a lot of time and effort to train. A specific kind of machine learning (ML) called deep learning (DL) trains machines to imitate human behavior. Neural networks (NN), which DL employs, demand a lot of processing power to solve complicated problems. However, new developments in data analytics and processing power have made it possible for DL algorithms to watch, learn from, and respond to challenging circumstances. The DL algorithm can use reinforcement learning, unsupervised learning, or supervised learning strategies, depending on the intended use [3]. Deep learning (DL) applications have garnered attention as a means of achieving more advanced and self-sufficient methods for risk assessment, treatment planning, and diagnosis [9,10].

2.4. Ensemble Learning in Dental

Ensemble approaches in machine learning aggregate predictions from several base models to improve model performance beyond what could be accomplished by a single model. The fundamental premise is that various models may identify various patterns or biases in the data and that the Ensemble model can counteract the performance shortcomings of individual models by combining their predictions. Ensembles frequently generate better results overall and with more accuracy [11].

3. Literature Review

Artificial intelligence (AI) incorporation into healthcare systems has created new opportunities to enhance clinical results and operational effectiveness. Large volumes of data may be processed quickly and accurately by AI-powered systems, yielding insights that would be challenging to obtain manually. To help with early diagnosis, AI algorithms, for instance, can examine medical photos to find patterns of particular diseases [12]. Furthermore, according to [13], AI-powered predictive analytics can lessen stress in healthcare facilities, help with resource allocation, and predict patient admissions. Many researchers have studied the influencing factors of medical information technology adoption based on theories like technology adoption behavior to identify the critical elements that influence adopters' adoption of medical information technology and its related applications and to encourage the use and promotion of medical information technology [14–16].

The proposed work in [17] describes the creation and evaluation of YOLOv3, a deep learning model for the automated identification and categorization of four prevalent dental conditions in orthopantomography (OPG) X-ray pictures. Because there were no publicly accessible datasets, the researchers augmented the original 800 OPG photos to build their collection of 1200 images. The study addresses the drawbacks of manual analysis, which can be laborious and prone to human mistakes. It emphasizes the potential of deep learning to increase the effectiveness and accuracy of dental diagnosis.

The research presented in [18] describes the creation and verification of an affordable artificial intelligence system that uses non-standard smartphone photos to identify dental cavities. The researchers used a hybrid model that combined transfer learning with other deep learning models (e.g., VGG16) and Ensemble YOLO object identification methods. Variations in image quality were addressed through the use of data augmentation techniques. By concentrating on non-standardized photos taken using widely accessible technologies, the study overcomes the shortcomings of current AI-based caries detection techniques. There is also a discussion on future research directions.

The authors in [19] analyzed 22 studies that employ machine learning to diagnose oral health problems using X-rays from 2018 to 2022. The study draws attention to both achievements and shortcomings, including sparse datasets and uneven reporting. For more dependable clinical applications, the authors emphasize the necessity of larger, diversified datasets and standardized procedures. Xue, Peng, et al. [20] employed deep learning to identify teeth in dental periapical radiographs. R-FCN (Region-based Fully Convolutional Networks) was combined with a faster R-CNN to detect typical dental issues such as tooth loss, decay, and fillings. Lee et al. [21] used a deep convolutional neural network (CNN) to evaluate radiographs and determine each tooth's radiographic bone loss (RBL). The updated CNN periodontitis categorization's RBL, staging, and presumptive diagnosis were compared to independent examiner findings. However, these algorithms are just concerned with detecting teeth, not identifying each of their faces. Fukuda et al. applied CNNs to detect vertical root fractures (VRF) in panoramic radiography [22]. The CNN was trained using Detect Net five-fold cross-validation, and DIGITS version 5.0 was utilized to improve model reliability. In Ref. [23], this paper describes a novel AI-driven approach for auto-

matically segmenting the PDL (Periodontal Ligament) on CBCT (Cone Beam Computed Tomography) imaging, setting the groundwork for future deep-learning models that can calculate the PDL area or volume directly from CBCT 3D pictures. The study team aims to improve and refine the PDL segmentation system so that periodontists, orthodontists, prosthodontists, and implantologists may diagnose and plan treatments more efficiently and accurately. Namamula et al. [24] proposed an Ensemble learning technique that combines the results of two ML algorithms, 'Edge Detection Instance Preference (EDIP)' and 'Extreme Gradient Boosting (XGboost)'. Voting is used to aggregate precision. Based on these findings, it has been concluded that the suggested learning strategy delivers the maximum level of accuracy when compared to the existing methods. This research presents an autonomous diagnostic strategy for detecting dental cavities in pictures based on deep learning algorithms and Ensemble methodologies. The proposed technique uses a series of pre-trained models, including Xception, VGG16, VGG19, and DenseNet121, to extract fundamental properties from photos and classify them as normal or abnormal [25].

4. Proposed Methodology

This paper introduces a deep neural network model for detecting dental issues. Utilizing an orthopantomography panoramic teeth X-ray dataset, the proposed method aims to identify various dental disorders. The following section outlines the functioning of the proposed architecture.

4.1. The Proposed Integrate Model for Reliable and Efficient Dental Healthcare

This section introduces the proposed model for reliable and efficient dental healthcare. The introduced model utilizes a deep neural network model for detecting dental issues. The suggested model integrates a number of cutting-edge technologies and techniques to assist a professional clinic in quickly and accurately performing an automated diagnostic. It streamlines communication for simpler, more reliable, and highly accurate interactions by redefining the doctor–patient relationship. Figure 1 shows the proposed integrated model of this study. The implementation details of the proposed model are illustrated as follows:

The initial step in the proposed integrated model for reliable and efficient dental healthcare involved collecting the necessary dataset (the input teeth X-ray dataset). Once obtained, the dataset underwent a preprocessing phase. In this phase, the data were further filtered, augmented to images, and annotated, with any images containing irrelevant information excluded. The postprocessing and image analysis techniques were carried out to increase the anticipated mask quality. Following preprocessing, the data were split into training and testing sets, compressed, and uploaded to the drive Then, we apply two deep learning algorithms (YOLO8 and RT-DETR) to those preprocessed input images. The algorithms related to the backbone phase will be discussed in the following subsection. After that, the output from those algorithms became the input in the next phase, named the Ensemble process, by using NMS (Non-Maximum Suppression). Finally, the output of the Ensemble phase detects dental diseases.

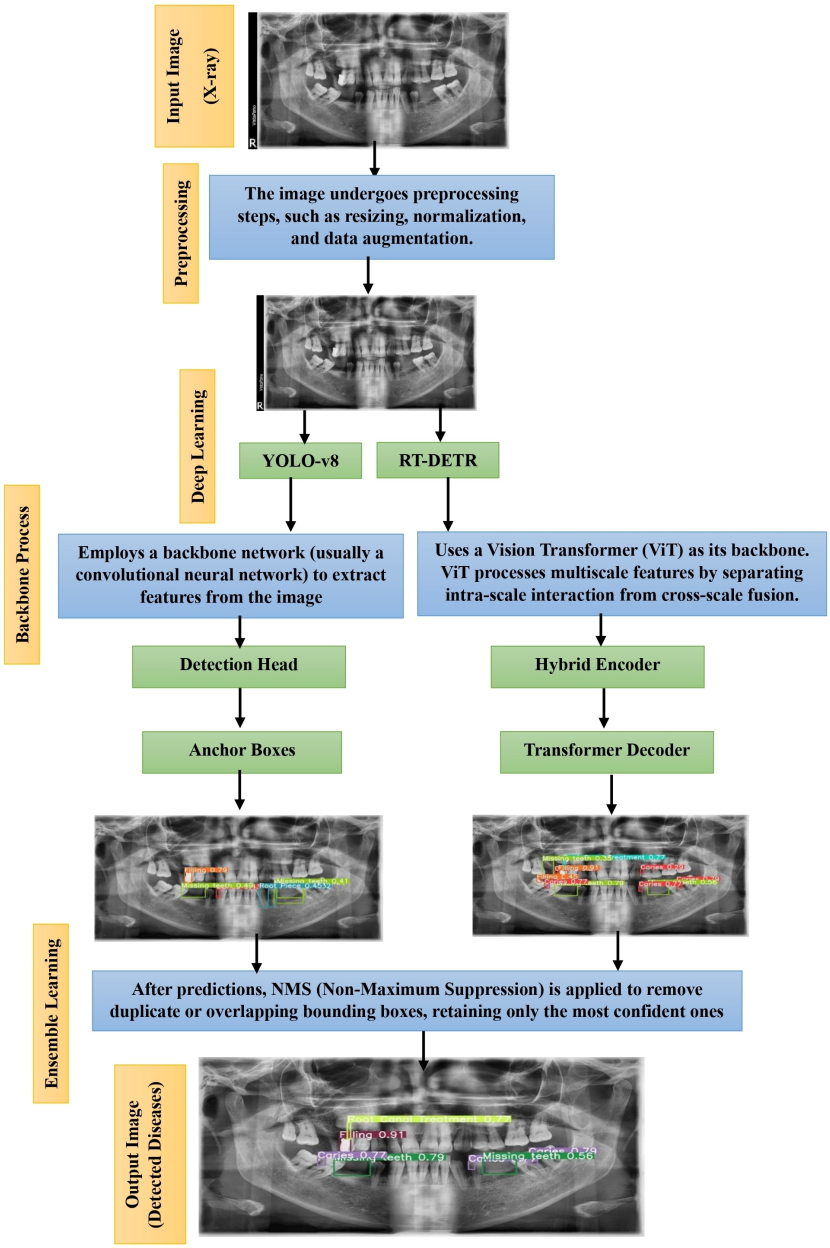


Figure 1. Integrated artificial intelligence model for reliable and efficient dental architecture.

4.2. The Deep Learning Detections

The deep learning model was trained to detect and classify various dental issues from the processed dataset. This involved feeding the neural network with augmented and annotated images. This led to learning as well as identifying patterns associated with different dental disorders. Several metrics were used to assess the model's detection efficacy, resulting in a precise and trustworthy identification of dental issues.

4.2.1. YOLO8

YOLO8 (You Only Look Once version 8) is one of the latest iterations in the YOLO series of object detection models, designed for real-time processing and efficiency. The underlying mathematical model of YOLO is based on deep convolutional neural networks (CNNs) and operates using the following key components [26]:

1. Backbone Network (Feature Extractor):

The backbone is a CNN that extracts features from the input image. YOLO8 uses an efficient architecture to balance speed and accuracy. It transforms the input image into a feature map that captures essential spatial information. The feature map output by the backbone CNN is computed as follows in (1).

$$F = CNN(I) \tag{1}$$

where I is the input image as well as F is the feature map output by the backbone CNN.

2. Detection Head:

YOLO divides the input image into an S × S grid, where each grid cell is responsible for predicting bounding boxes and class probabilities for objects that fall within that region. Each grid cell predicts several bounding boxes, including their coordinates, objectness score (confidence that an object exists in the box), and class probabilities. For each predicted bounding box, YOLO8 outputs the center coordinates of the box, width and height, confidence score, and class probabilities. The bounding box is represented as follows in (2):

$$bbox = (x, y, w, h) \tag{2}$$

where x, y is the center of the bounding box as well as w, h represent the width and height of the box, respectively. The coordinates are normalized relative to the grid cell.

The objectness score CCC is predicted by the model to indicate whether an object exists in the bounding box as follows in (3).

$$C = P(Object) * IOU_{pred}^{truth} \tag{3}$$

where $P(object)$ is the probability of an object in the grid cell. IOU_{pred}^{truth} is the intersection over the union between the predicted box and the ground truth box.

3. Loss Function:

YOLO8 uses a custom loss function to optimize the model during training. The loss function typically includes multiple components. The first component is Bounding Box Regression Loss, which measures how well the predicted bounding boxes match the ground truth boxes as follows in (4)

$$L_{\text{bbox}} = \sum_{i=1}^{n} \left(\text{smooth}_{L_1}(x_i - x_i^{\text{true}}) + \text{smooth}_{L_1}(y_i - y_i^{\text{true}}) + \text{smooth}_{L_1}(w_i - w_i^{\text{true}}) + \text{smooth}_{L_1}(h_i - h_i^{\text{true}})\right) \tag{4}$$

where smooth L_1 is a variation in the standard L_1 loss function.

The second component is confidence loss which measures the objectness score as follows in (5):

$$L_{\text{conf}} = \sum_{i=1}^{n} \left(C_i - C_i^{\text{true}}\right)^2 \tag{5}$$

where C_i typically represents the predicted confidence score for the presence of an object in the ith bounding box.

The third component is class prediction loss, which measures the class probabilities as in (6).

$$L_{\text{class}} = \sum_{i=1}^{n} \sum_{c=1}^{k} \left(P(c_i) - P(c_i^{\text{true}})\right)^2 \tag{6}$$

where $P(c_i)$ presents the predicted probability that the ith bounding box belongs to class c. $P(c_i^{\text{true}})$ represents the true probability that the ith bounding box belongs to class C.

The last component named Total Loss measures the total loss as a weighted sum of the three losses as in (7).

$$L_{\text{total}} = \lambda_{\text{bbox}}\mathcal{L}_{\text{bbox}} + \lambda_{\text{conf}}\mathcal{L}_{\text{conf}} + \lambda_{\text{class}}\mathcal{L}_{\text{class}} \quad (7)$$

where λ_{bbox}, λ_{conf} and λ_{class} are hyperparameters that control the relative importance of each term.

4.2.2. RT-DETR: Real-Time Detection Transformer

RT-DETR is a real-time version of the DETR model. DETR is a transformer-based object detection model that formulates object detection as a set prediction problem, bypassing the need for anchor boxes and allowing for end-to-end training [27]. In this paper, we leverage RT-DETR (Real-Time Detection Transformer) for the real-time detection and classification of dental issues. RT-DETR uses a Vision Transformer (ViT) as its backbone. The ViT processes multiscale features by separating intra-scale interaction from cross-scale fusion. This transformer model allows us to perform high-speed and high-precision detections, making it suited for practical dental diagnostics as shown inFigure 1. The RT-DETR incorporates ViT for decoupling from image input. This allows the model to use ViT's ability to effectively capture global and local features by focusing on itself. The ViT of RT-DETR represents the integration of transformer-based vision models with state-of-the-art object recognition systems, which empowers powerful and efficient real-time recognition capabilities by enabling ViT to be used to capture spatial relationships and global conditions in images. After preparing the dataset as explained in the following subsection, we apply RT-DETR through the steps of Algorithm 1. RT-DETR is a variant of the transformer-based architecture tailored for object detection. Traditional detection methods, such as YOLO, rely on convolutional neural networks (CNNs), but transformers, originally designed for NLP tasks, have shown powerful capabilities in vision tasks as well. However, one of the limitations of early transformer-based detection methods (like DETR) is their slow inference speed. RT-DETR addresses this by optimizing transformers for real-time object detection. The core idea of using transformers in object detection is to leverage self-attention mechanisms to capture long-range dependencies between image features. Transformers excel at handling complex relationships between different parts of an image, making them ideal for learning global context for detection. The Attention Mechanism is computed as in (8)

$$\text{Attention}(Q, K, V) = \text{softmax}\left(\frac{QK^T}{\sqrt{d_k}}\right)V \quad (8)$$

where Q (Query), K (Key), and V (Value) are projections of the input image feature embeddings; dk is the dimension of the key vector.

Similar to other detection architectures, RT-DETR uses a backbone network (often a CNN or a vision transformer variant) to extract feature maps from the input image. The feature extractor reduces the spatial dimensions of the image while retaining important details for the detection task. In the transformer encoder–decoder stage, the encoder takes the feature map from the backbone and applies a series of multi-head attention and feedforward layers to process the features. Then, the decoder generates object queries that are responsible for detecting objects in the image. Instead of sliding windows (as in CNN-based approaches), RT-DETR uses object queries to predict the bounding boxes directly from the image's global context. In the learnable object queries stage, RT-DETR, like DETR, uses a fixed number of learnable object queries, each one responsible for detecting a specific object in the image. These queries are passed through the decoder and used to predict bounding boxes and class labels. One of the critical innovations in RT-DETR is the efficiency improvements in the architecture that allow it to run in real-time, a significant improvement over early transformer-based models like DETR, which are computationally

expensive. This efficiency comes from reducing the number of transformer layers and optimizing operations like self-attention, as well as parallelization of computations. After that, RT-DETR predicts bounding boxes in a manner similar to DETR, using a regression head to output the bounding box coordinates (x, y, w, h) and a classification head to predict the object class as in (9) [28]:

$$\text{bbox} = (x, y, w, h) \tag{9}$$

where x, y represent the center of the bounding box; w, h represent the width and height of the box. The model uses a combination of two main loss functions to train the detection model. The first is named Bounding Box Regression Loss, which measures the difference between predicted and ground truth bounding boxes, as in (10).

Algorithm 1: YOLO8 and RT-DETR

```
Define the main function:
    Load the object detection dataset
    Preprocess the dataset (cleaning, normalization, etc.)
    Split the dataset into training and testing sets
    Define the object detection model (YOLO8 & RT-DETR)
    Define the loss function (Cross-Entropy Loss)
    Define the optimizer (Adam, SGD)
for each epoch in total number of epochs do
    for each batch in training data do
        Forward pass the batch through the model
        Compute the loss
        Backward pass to compute gradients
        Update the model parameters using the optimizer
    end
    Evaluate the model on the testing set
    Print the performance metrics
end
End of the main function
Call the main function
```

* **Cross-Entropy Loss:** is a measure used in machine learning and specifically in classification tasks to quantify the difference between two probability distributions.
* **Adam:** Adaptive Moment Estimation is an advanced optimization algorithm used in training deep learning models.
* **SGD:** Stochastic Gradient Descent is a core optimization algorithm used extensively in training machine learning models.

$$\mathcal{L}_{\text{bbox}} = \sum_{i} (smooth_{L_1}(\hat{b}_i - b_i)) \tag{10}$$

where $\hat{b}_i$ is the predicted bounding box, and b_i is the ground truth. The latter is named classification loss which measures the difference between predicted class scores and true class labels (often using a cross-entropy loss). The Total Loss is computed as follows in (11)

$$\mathcal{L}_{\text{total}} = \lambda_{\text{class}} \mathcal{L}_{\text{class}} + \lambda_{\text{bbox}} \mathcal{L}_{\text{bbox}} \tag{11}$$

where λ_{class} and λ_{bbox} are weighting factors.

Finally, RT-DETR used Hungarian matching to assign ground truth objects to predicted boxes. This bipartite matching process ensures that each ground truth object is matched to

only one predicted object, which helps avoid duplicate predictions and enforces one-to-one matching.

4.3. Ensemble Learning

In order to improve model performance beyond what could be accomplished by a single model, Ensemble techniques in machine learning aggregate predictions from several base models. The fundamental idea is that various models may identify various biases or trends in the data, and by combining their predictions, the Ensemble model can make up for the shortcomings of each individual model. Ensembles often achieve higher accuracy and produce better overall [29], as shown in Figure 1. NMS (Non-Maximum Suppression) is a post-processing step in object detection algorithms where highly overlapping bounding boxes are pruned to keep only the most confident one. It helps in eliminating redundant detections. The bounding boxes that the object detection algorithms generate are not always accurate representations of the things in the image; ideally, we should have exactly one bounding box for each object. Among the several predicted bounding boxes, these object identification algorithms employ non-max suppression to choose the optimal bounding box. By using this method, the less probable bounding boxes are "suppressed", leaving only the best one. The optimum bounding box for an item is chosen via non-max suppression, which rejects or "suppresses" any alternative bounding boxes. Up to convergence, the algorithm examines overlaps, iteratively chooses the optimum bounding box, and eliminates unnecessary boxes. Two considerations are made by the NMS. The model first provides the objectiveness score. The bounding boxes overlap, or IOU, as explained in Algorithm 2.

Algorithm 2: Proposed Ensemble NMS (Non-Maximum Suppression)

Input : A list of Proposal boxes B, corresponding confidence scores S, and overlap threshold N.
Output: A list of filtered proposals D.
The proposed Algorithm:
Select the proposal with the highest confidence score.
Remove it from B and add it to the final proposal list D (Initially D is empty).
Now compare this proposal with all the proposals.
for *D to all B list* **do**
  Calculate the IOU (Intersection over Union) of this proposal D with every other proposal B.
  if *the IOU is greater than the threshold N* **then**
    1. Remove that proposal from B.
    2. Repeat: take the proposal with the highest confidence from the remaining proposals in B and remove it from B and add it to D.
  else
    Once again calculate the IOU of this proposal with all the proposals in B and eliminate the boxes which have higher IOU than the threshold.
  end
  This process is repeated until there are no more proposals left in B.
end
End for.

* **Boxes** B**:** refers to a collection of bounding boxes used in object detection and localization tasks, crucial for accurately identifying and localizing objects within images.
* **IOU (Intersection over Union):** is a metric used to measure the overlap between two bounding boxes or regions of interest (ROIs) in object detection and image segmentation tasks.

Non-Maximum Suppression (NMS)

After generating multiple bounding box predictions, YOLO applies Non-Maximum Suppression (NMS) to eliminate redundant boxes for the same object as shown in Algorithm 2. NMS selects the bounding box with the highest confidence score and removes other boxes that have a high IoU with the selected box as follows in (12).

$$\text{IoU} = \frac{\text{Area of Overlap}}{\text{Area of Union}} \tag{12}$$

where the input is considered the image is divided into grid cells. For each grid cell, YOLO predicts bounding boxes, confidence scores, and class probabilities that consider the output. The loss function consists of bounding box regression, object confidence, and classification losses.

NMS suppresses boxes with an IoU above a certain threshold as in subsequent steps.

1. Sort bounding boxes by a confidence score.
2. For each box, calculate IoU with boxes ranked below it.
3. If the IoU is greater than a predefined threshold (e.g., 0.5), discard the box with the lower score.
4. Repeat until no more boxes are left to compare.

Where NMS uses the following practical considerations:

- IoU Threshold: A lower threshold (e.g., 0.3) leads to more aggressive suppression, whereas a higher threshold (e.g., 0.7) allows more overlap.
- Speed-Accuracy Tradeoff: While NMS helps remove redundant detections, it may also suppress valid ones in certain cases, especially with small IoU thresholds. Balancing precision and recall is important.
- Class-Specific NMS: In some models, NMS is applied separately for each class, preventing boxes from being suppressed across different object categories.

YOLO8 improves upon previous versions by optimizing these components for better accuracy and speed. It is considered one of the latest versions of the popular YOLO (You Only Look Once) family of algorithms. YOLO8 has no official paper but boasts higher accuracy and faster speed. It employs a backbone network (usually a convolutional neural network, CNN) to extract features from the image. The backbone, based on CSPDarknet53 (cross-Stage Partial network), has 53 layers that talk to each other better for a clearer understanding of images. This part helps focus on different parts of a picture and figure out which features matter the most for finding objects. CSPDarknet53 acts as a backbone network, providing a robust and efficient method for extracting objects from image input. The integration in YOLO8 helps strike a balance between computational efficiency and detection accuracy, making it suitable for real-time object recognition applications. YOLO8 is also good at spotting objects of different sizes in pictures, thanks to its feature pyramid network as shown in Figure 1 phase Backbone Process [30]. Also, object detection localizes an object within an image by drawing bounding boxes. The implementation needs you to declare the model as yolo8n.pt for object detection using the nano variant. You do not need to add any suffix to use YOLO8 for detection. YOLO8 uses anchor-free detection, where the model directly predicts an object's center rather than calculating the offset from a predefined anchor box.

Anchor boxes are predefined boxes with specific heights and widths, designed to detect object classes at particular scales and aspect ratios. During detection, they are distributed throughout the image and chosen according to object sizes in the training dataset [30]. After preparing the dataset as explained in the following subsection, we apply YOLO8 through the steps of Algorithm 1.

5. Performance Evaluation

The proposed integrated model is tested on a different set of experiments for validation. Many random experiments with real-time datasets have been conducted. This section presents the platform, how to set up the investigation, the results, and the discussion. For comparison, the proposed model is compared with three competitive algorithms.

5.1. Panoramic X-ray Dataset (Collect Data)

According to [31] and various other sources, the detector model's label data format should have a particular structure as shown in Table 1. Each object's coordinates and class index are included in the data; these are normalized according to the image's width and height. For an 864 × 1188 image, each item is represented by a single line in the text file, which has the format $(x1/864, y1/1188, x2/864, y2/1188)$. The class index for the "caries" class is one, whereas the class index for the "no-caries" class is zero. By ensuring that the coordinates are scaled about the image dimensions, this normalization guarantees that the detector model can handle them. Label data, often referred to as ground truth data, provide annotations that specify the presence and location of objects of interest in an image .

Table 1. The versions of datasets used in the proposed model.

vzrad2 Dataset (V4)	**vzrad2 Dataset (V5)**	**vzrad2 Dataset (V6)**
Train Set 62% by 5773 Images Valid Set 21% by 1956 Images Test Set 16% by 1519 Images	Train Set 56% by 3617 Images Valid Set 25% by 1612 Images Test Set 20% by 1268 Images	Train Set 57% by 4774 Images Valid Set 25% by 2074 Images Test Set 19% by 1581 Images
Preprocessing: • Auto-Orient: Applied • Resize: Stretch to 640*640 • Modify Classes: 14 remapped, 33 dropped • Filter Null: Require all images to contain annotations.	**Preprocessing:** • Auto-Orient: Applied • Resize: Stretch to 640*640 • Modify Classes: 10 remapped, 38 dropped • Filter Null: Require all images to contain annotations	**Preprocessing:** • Auto-Orient: Applied • Resize: Stretch to 640*640 • Modify Classes: 14 remapped, 34 dropped • Filter Null: Require all images to contain annotations
Augmentations: • Outputs per training example: 3 • Grayscale: Apply to 25% of images • Noise: Up to 5% of pixels • Mosaic: Applied • Bounding Box: • Rotation: Between −15° and +15°	**Augmentations:** • Outputs per training example: 3 • Grayscale: Apply to 25% of images • Noise: Up to 5% of pixels • Mosaic: Applied • Bounding Box: • Rotation: Between −15° and +15°	**Augmentations:** • Outputs per training example: 2 • Grayscale: Apply to 25% of images • Hue: Between −6° and +6° • Brightness: Between −15% and +15% • Blur: Up to 0.7px

5.2. Performance Metrics

The performance of the image detection models was evaluated using accuracy as in (13).

$$\text{The Total Accuracy} = \frac{\text{Total Number of Correctly Detection}}{\text{Total Number of images}} \tag{13}$$

The confusion matrix is considered another performance metric that is used to evaluate the proposed model. It explains the expected classifications are either positive (the patient

is sick) or negative (the patient is healthy). The proposed model is run on 1000 patients, and the model predictions are entered. The recall metric known as sensitivity tells us how many of the sick patients our model incorrectly diagnosed as well [32]. It also shows how many times the model incorrectly diagnosed a sick patient as negative. In this paper detection recall is considered another metric that was calculated as in (14).

$$\text{Detection Recall} = \frac{\text{Number of successful match}}{\text{Total Number of standard bounding box}} \tag{14}$$

Precision (also known as specificity) is the opposite of recall. It tells us how many of the healthy patients our model incorrectly diagnosed as sick. In other words, how many times did the model incorrectly diagnose a healthy patient as positive? Precision is calculated as in (15).

$$\text{Detection Precision} = \frac{\text{Number of true positive}}{\text{Number of detected bounding box}} \tag{15}$$

F-score summarizes the performance of a classifier with a single metric that represents both recall and precision, as in (16).

$$F\text{-score} = 2 \cdot \frac{\text{Detection Precision} \cdot \text{Detection Recall}}{\text{Detection Precision} + \text{Detection Recall}} \tag{16}$$

The mean average precision (mAP) metric means simply the average of AP of all classes. It helps to assess the balance between precision and recall for an object detection model across various probability thresholds. A higher mAP value indicates better performance of the object detection model, as in (17).

$$mAP = \frac{1}{n} \sum_{k=1}^{k=n} APK \tag{17}$$

where n is the number of classes and APk is the average of the precision of class k.

In the proposed model, mAP50-95 is calculated. It refers to the mean Average Precision (mAP) calculated at multiple Intersection over Union (IoU) thresholds, ranging from 0.5 to 0.95 with a step size of 0.05. This provides a more comprehensive evaluation of the model's performance across different IoU thresholds. mAP50 is computed in the proposed model [33,34]. It refers to the mAP calculated at an IoU threshold of 0.5. IoU is a measure of the overlap between the predicted bounding box and the ground truth bounding box. An IoU of 0.5 means the predicted bounding box and the ground truth bounding box overlap by 50%. So, mAP50 gives the average precision of the model at this IoU threshold.

5.3. Result

Loss functions are employed in the proposed model in the training and validation phases. It is used to predict object positions, classifications, and other parameters with accuracy. Usually, the loss function is composed of several parts that assess various facets of prediction accuracy. In the proposed model are measured three losses, which are bounding regression loss, class prediction loss, and distribution focal loss [35]. Figure 2 shows train/valid loss for YOLO8. The train/box_loss is considered the loss incurred in predicting the bounding boxes of the objects during the training phase. A lower value indicates a better fit of the model to the training data. The proposed model, given the minimum train/box_loss is 1.2476 through 40 epochs. At the same time, the train/cls_loss is the classification loss during training. It measures how well the model is performing at classifying the objects. So, the proposed model given the minimum train/cls_loss is 0.69987 through 40 epochs. Finally, in the training phase, we measure the train/dfl_loss, where the loss is related to the distribution-focused learning during training. It is a measure of how

well the model is learning the data distribution and is given 0.9827 [36]. In addition, we measure the loss in the validation phase (val/box_loss is 1.4334, val/cls_loss is 0.90363, and val/dfl_loss is 1.0912). The lower value indicates that the model generalizes well to unseen data, as shown in Figure 2.

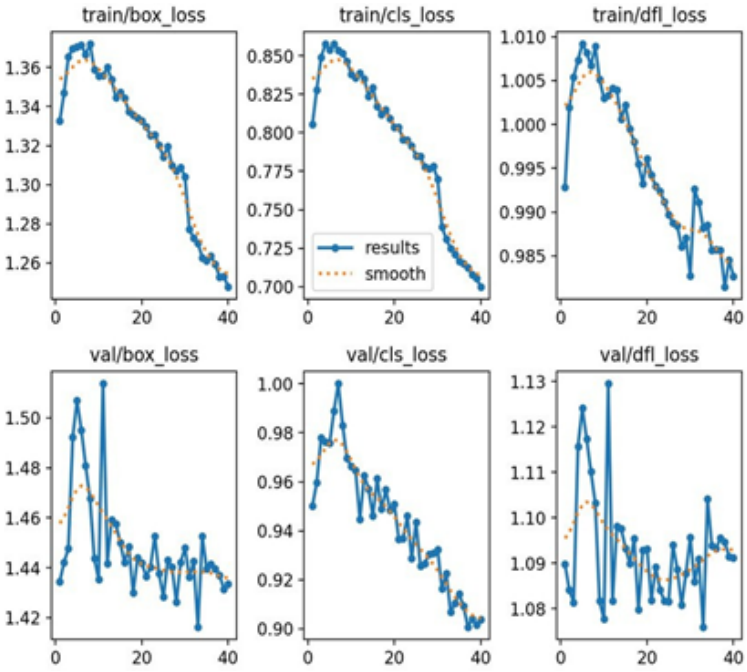


Figure 2. Train and valid loss for Yolo8.

Figure 3 shows train/valid loss for RT-DETR. The train/box_loss is 0.4146 through 15 epochs; train/cls_loss is 0.39282 and train/dfl_loss is 0.05091. The loss in the validation phase is where val/box_loss is 0.56202, val/cls_loss is 0.52314, and val/dfl_loss is 0.09694. Comparing Figures 1 and 2 explains the loss in RT-DETR is less than YOLO 8, despite the number of epochs in RT-DETR being the least. This is due to RT-DETR model complexity, while YOLO8 is simplicity and speed.

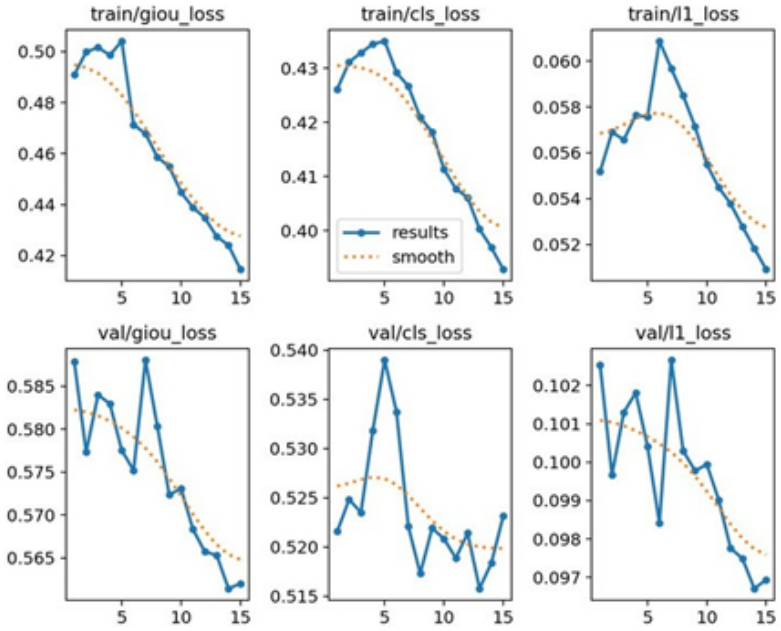


Figure 3. Training and validation loss for RT-DETR.

Optimizing the model's loss function and architecture resulted in an 18% reduction in training and inference time. Using a more efficient loss function allows the model to converge faster, lowering the number of epochs required for training. This not only speeds up training but also improves inference times, resulting in better overall performance.

Figure 4a shows the Precision Metric and Recall Metric for YOLO8. The proposed model has a precision of 0.74206, which means it is quite good at not labeling a negative sample as positive. Also, Recall Metric has a recall of 0.54332, indicating its ability to find all the positive samples.

Figure 4b shows precision metrics and recall metrics for RT-DETR. The proposed model has a precision of 0.8051, which means it is quite good at not labeling a negative sample as positive. Also, Recall Metric has a recall of 0.74484, indicating its ability to find all the positive samples.

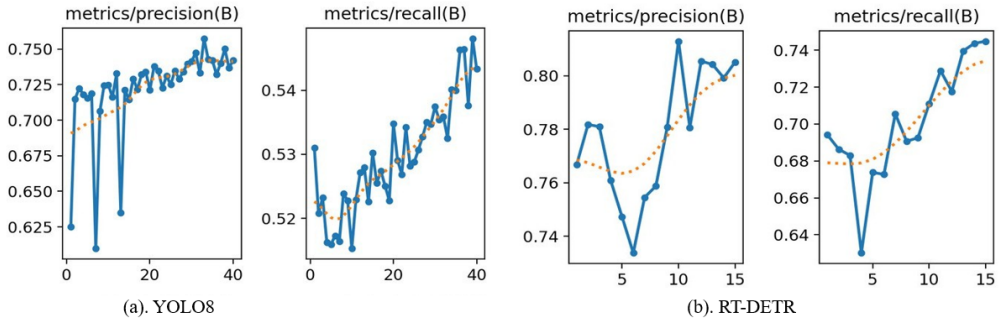


Figure 4. Precision and recall metrics for YOLO8 and RT-DETR.

Figure 5a shows mAP50 and mAP50-95 metrics for YOLO8. The proposed model has an mAP50 of 0.57447, indicating a moderate ability to detect objects with a large overlap with ground truth and an mAP50-95 of 0.32494, indicating its performance across different levels of IoU thresholds is fair.

Figure 5b shows mAP50 and mAP50-95 metrics for RT-DETR. The proposed model has an mAP50 of 0.73544, indicating a moderate ability to detect objects with a large overlap with ground truth. While the mAP50-95 of 0.4595 indicates its performance across different levels of IoU thresholds is fair.

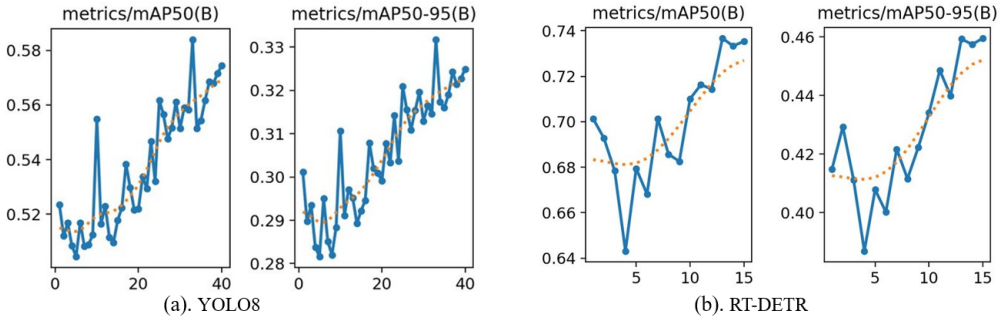


Figure 5. mAP50 and mAP50-95 metrics for YOLO8 and RT-DETR.

Figure 6 shows the overall performance for YOLO8, where the curve labeled "all classes 0.584 mAP0.5" shows the average performance across all classes, providing a general sense of the model's effectiveness. Figure 7 helps to understand the effectiveness of a classifier across different thresholds and for various conditions. This type of analysis is crucial in fields such as medical imaging, where accurate and reliable predictions can significantly impact patient outcomes.

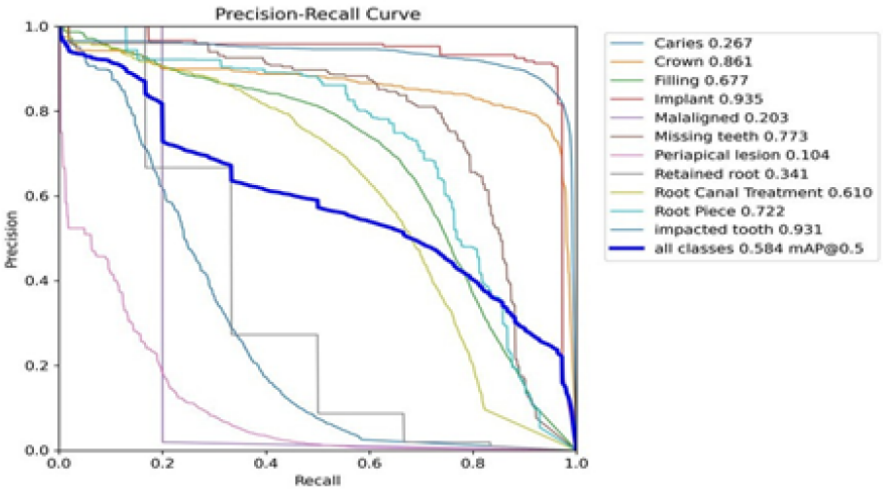


Figure 6. Precision-recall curve for YOLO8.

In the RT-DETR, the precision-recall curves help in understanding which classes are being detected more accurately and reliably and which ones may require further training or adjustment in the model parameters. The overall performance across all classes (indicated by the blue line labeled "all classes 0.736 mAP0.5") suggests a moderate level of accuracy when considering all detections collectively [37]. This chart is essential for evaluating and refining the RT-DETR, ensuring it is optimized for the specific detection tasks it is intended to perform, as shown in Figure 7.

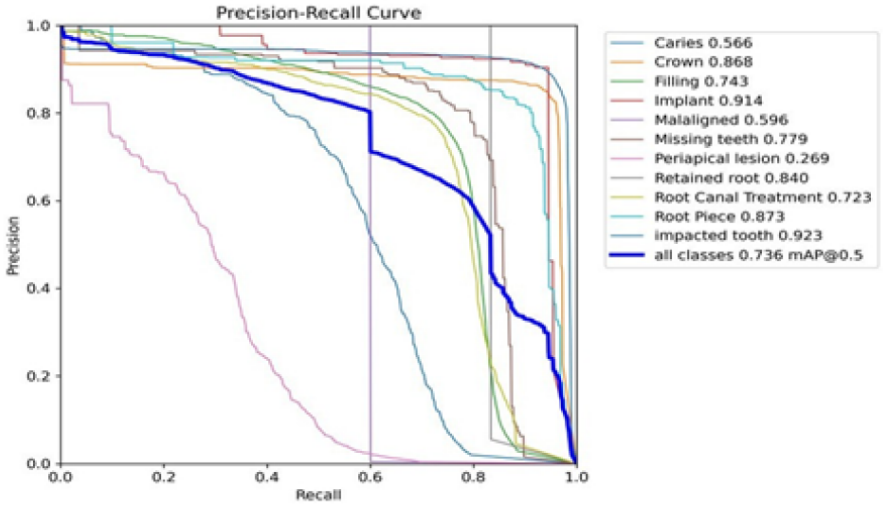
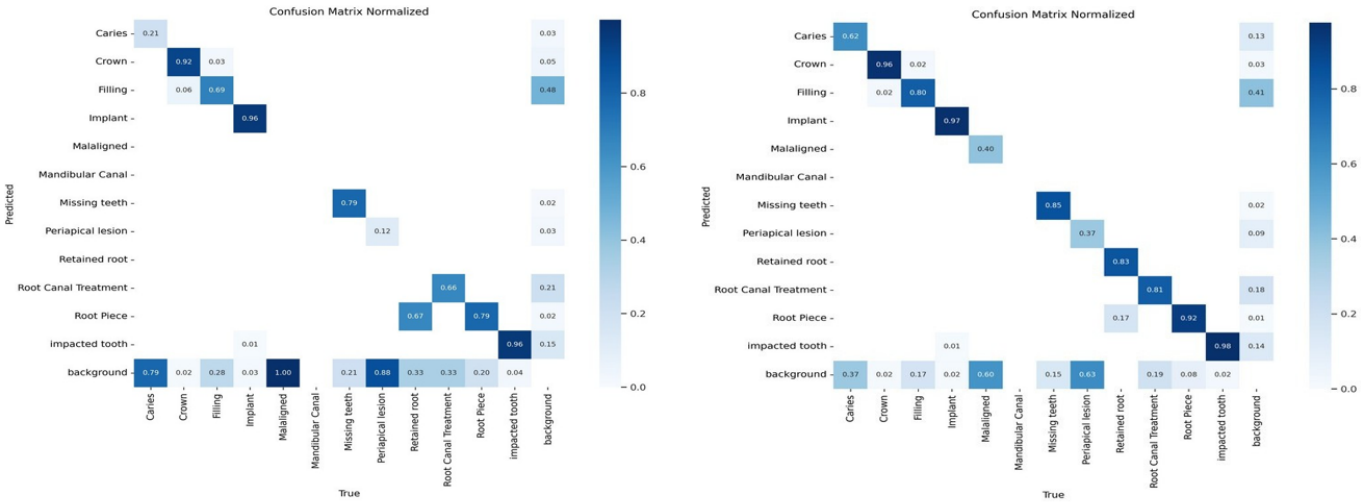


Figure 7. Precision-recall curve for RT-DETR.

Figure 8a shows the Confusion Matrix Normalized for YOLO8, which is a visualization tool typically used to measure the performance of a classification model. Each cell in the matrix represents the proportion of predictions made by the model. For instance, the cell at the intersection of the "Crown" row and the "Crown" column shows 0.92, which means 92% of the actual Crown cases were correctly predicted as a Crown by the model. Off-diagonal cells indicate incorrect predictions. For example, the 0.21 in the "Caries" row and "Crown" column indicates that 21% of the cases that were actually Caries were incorrectly predicted as Crown. High Accuracy Categories like "Implant" (0.96) and "Crown" (0.92) show very high diagonal values, indicating that the model is highly accurate in predicting these conditions. Misclassifications such as "Caries" being mistaken for "Crown" (0.21) and "Root Canal Treatment" being confused with "Root Piece" (0.21). The performance of the model varies across different categories. Some categories, like "Implant", are predicted

with high accuracy, while others, like "Periapical Lesion", have lower correct prediction rates (0.12).



(a). YOLO8 (b). RT-DETR

Figure 8. Normalized confusion matrix for YOLO8 and RT-DETR.

Figure 8b shows the normalized confusion matrix for an RT-DETR. Diagonal values represent the proportion of correctly predicted instances for each class. For example, "Crown" has a value of 0.96, indicating that 96% of the crown predictions were correct. Off-diagonal values show the proportion of incorrect predictions, where the model predicted the row label, but the true label was the column label. For example, under "Caries" (row) and "Crown" (column), there is a value of 0.02, indicating that 2% of the true "Crown" instances were incorrectly predicted as "Caries". High performers like "Crown" (0.96), "Implant" (0.97), and "Root Piece" (0.92) have high predictive accuracy, indicating effective recognition by the model. Misclassifications such as "mandibular canal" are often misclassified as "missing teeth" (0.15) and "peripical lesion" (0.37). "Root Canal Treatment" is sometimes confused with "Retained Root" (0.17) and "Root Piece" (0.18).

Figure 9a shows a filling located in the upper left molar region. This indicates the presence of a dental filling with a confidence score of 0.79. Missing teeth in two areas are marked with this label, one on the left side and another on the right side of the lower jaw. This indicates missing teeth with a confidence score of 0.49. The root piece is located in the lower central region. This indicates the presence of a root piece with a confidence score of 0.4532. Missing teeth are located in the lower right molar region. This indicates missing teeth with a confidence score of 0.41.

(a). YOLO8 (b). RT-DETR

Figure 9. The sample output image for YOLO8 and RT-DETR for the same image.

Figure 9b shows a root canal treatment located in the upper central region. This indicates the presence of a root canal treatment with a confidence score of 0.77. The filling

is located in the upper left molar region. This indicates the presence of a dental filling with a confidence score of 0.91. Caries are located in the lower left molar region. This indicates the presence of dental caries (tooth decay) with a confidence score of 0.77. Teeth are located in the lower left central region. This likely indicates the presence of a tooth or teeth with a confidence score of 0.79. Missing teeth located in the lower right molar region. This indicates missing teeth with a confidence score of 0.56. Caries are located in the upper right molar region. This indicates the presence of dental caries (tooth decay) with a confidence score of 0.79.

The provided Figure 10 is a bar chart comparing the counts of different dental disease classes as detected by three different methods: Ensemble, RT-DETR, and YOLO8. Here, is the analysis of the chart: Key takeaways like filling is the most frequently detected class across all methods, with the Ensemble method detecting the highest number. Also, implant and root pieces have the lowest counts across all methods. For periapical lesions, RT-DETR detects cases while the other two methods do not. Overall, the Ensemble method tends to detect higher counts across most classes compared to RT-DETR and YOLO8.

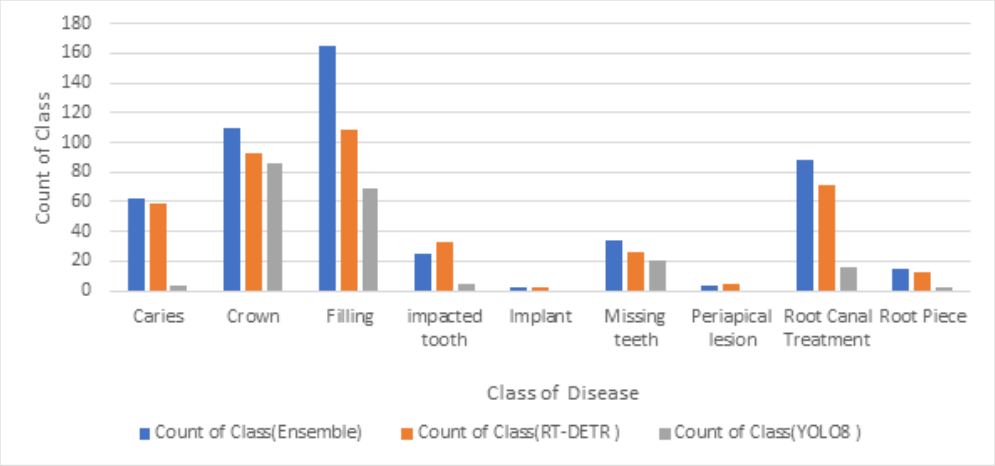


Figure 10. Comparing the counts of different dental disease classes as detected by three models.

Figure 11 provides a comparative analysis of the effectiveness and sensitivity of three different detection methods in identifying various dental disease classes. Figure 11 is a bar graph that appears to be comparing the performance of three different deep-learning models for a dental health classification task. Here is a breakdown of the information in the image. Where axis: The x-axis labels show the different classes of dental diseases. These include "Caries", "Crown Filling", "Impacted Tooth", "Implant", "Missing Teeth", "Periapical Lesion", "Root Canal", and "Root Piece". Y-axis: The y-axis label is "Average of Score (Ensemble)", though it likely refers to a percentage as the values range from 0.00% to 100.00%. This axis likely represents how well each model performed in classifying each disease class. Higher scores likely indicate better performance. There are three data series plotted in the graph, each representing the performance of a different deep-learning model. The models are labeled "Ensemble", "RT-DETR", and "YOLO8". It appears that the "Ensemble" model achieved the highest average scores across all disease classifications. It is important to note that the data for the "Median After-tax Income in Canada by Ethnicity" seems irrelevant and appears to be from a separate dataset unintentionally included in the image. Overall, the graph shows that the deep learning models were able to classify various dental diseases with some level of accuracy, with the "Ensemble" model performing the best. However, without knowing the specifics of the scoring system and the ground truth data used for evaluation, it is difficult to say definitively how well these models performed.

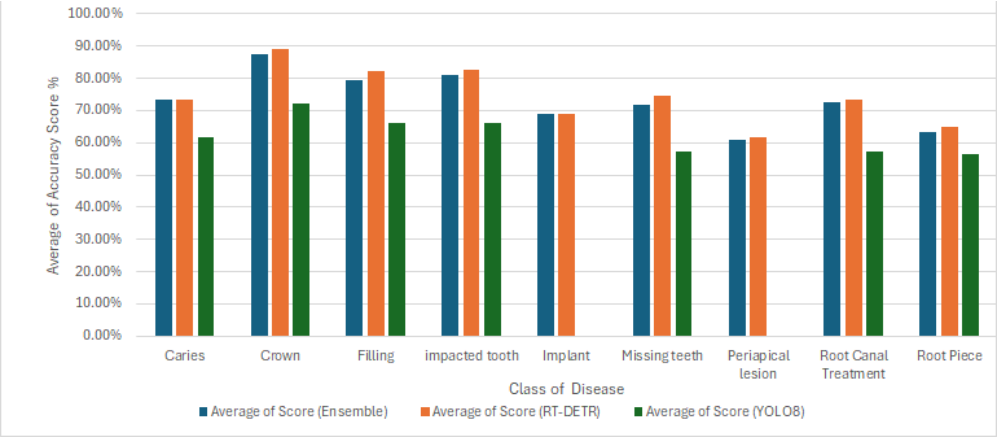


Figure 11. Comparative analysis of the effectiveness and sensitivity of three models.

Figure 12 shows the performance comparison of three object detection models (YOLO 8, RT-DETR, and Ensemble) on dental X-ray images for identifying various dental conditions. Here is the detailed analysis in Table 2. YOLO 8 provides reasonable detections but with lower confidence scores, especially for "missing teeth" and "caries". However, RT-DETR shows higher confidence scores for detections such as "filling" and "impaired tooth", providing more detailed annotations for "caries". Also, the Ensemble consistently shows high confidence scores for multiple conditions, including "root canal treatment", "Caries", and "missing teeth", making it the most comprehensive and accurate among the three models. Finally, this comparison indicates that the Ensemble model tends to provide more confident and detailed detections in dental X-ray images than both YOLO 8 and RT-DETR.

Table 2. Compares object detection models (YOLO8, RT-DETR) to real samples.

Position of Image	YOLO8	RT-DETR	Ensemble
Top Row	- "Filling" with a confidence score of 0.79. - "Missing teeth" with confidence scores of 0.49 and 0.41. - "Root piece" with a confidence score of 0.52.	- "Filling" with confidence scores of 0.91 and 0.43. - "Missing teeth" with a confidence score of 0.35. - "Caries" with confidence scores of 0.29, 0.79, and 0.69.	- "Filling" with a confidence score of 0.91. - "Missing teeth" with a confidence score of 0.62. - "Root canal treatment" with a confidence score of 0.72. "Caries" with confidence scores of 0.79 and 0.79.
Middle Row	- "Filling" with a confidence score of 0.32	- "Impacted tooth" with a confidence score of 0.73.	- "Root canal treatment" with confidence scores of 0.73 and 0.78. - "Missing teeth" with confidence scores of 0.65 and 0.68.
Bottom Row	- "Caries" with a confidence score of 0.51. - "Missing teeth" with a confidence score of 0.41.	- "Caries" with confidence scores of 0.85, 0.81, and 0.82. - "Caries" with a confidence score of 0.32. - "Filling" with a confidence score of 0.91.	- "Caries" with confidence scores of 0.85 and 0.81. - "Missing teeth" with a confidence score of 0.63.

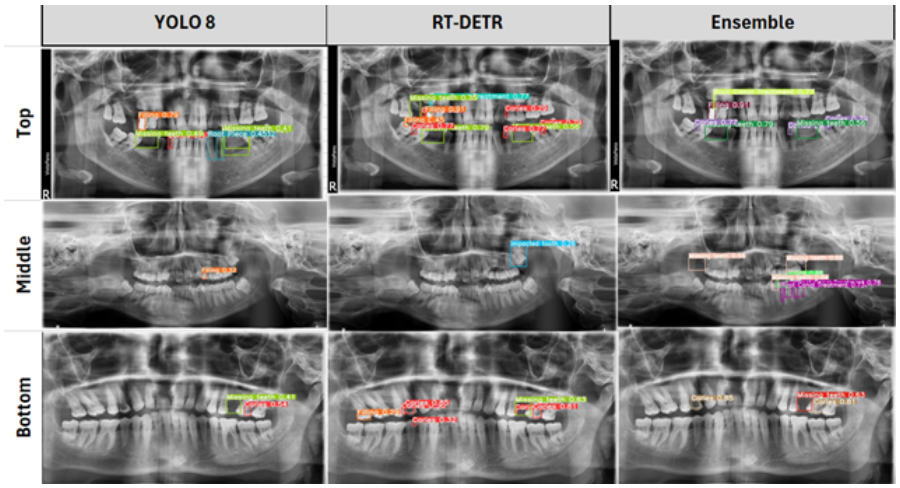


Figure 12. The samples of output image for YOLO8, RT-DETR, and Ensemble for the same images.

5.4. Discussion

Table 3, titled "Overall Performance", provides a comprehensive comparison of the performance metrics of various object detection models, specifically focusing on the YOLO8 and RT-DETR models augmenting their Ensemble versions. The table is organized into distinct columns representing the different iterations of these datasets (V4, V5, V6), with an additional column for an ensemble that integrates the outputs of both models. Evaluative metrics include precision, recall, mean average precision at 50% IoU (mAP50), mean average precision averaged from 50% to 90% IoU (mAP50-90), and overall accuracy. Key observations reveal that while the precision of YOLO8 shows an increase from version V4 to V5 (from 50% to 74%), there is a decline in V6 (68%). Conversely, RT-DETR demonstrates a high precision peak in V5 at 81%, in addition to superior recall metrics. Notably, the Ensemble model in V6 outperformed all individual models, achieving remarkable metrics: 90% precision, 79% recall, and 92% accuracy. This analysis underscores the enhanced efficacy of Ensemble methods by capitalizing on the strengths of both object detection architectures, reflecting a more consistent performance trajectory, which is particularly evident in the improvements leading up to version V6. Thus, the accuracy improved by about 30% as the minimum value Yolo8 (V4) and maximum value Ensemble (V6).

Table 3. Overall Performance.

Metrics	YOLO8 (V4)	YOLO8 (V5)	YOLO8 (V6)	RT-DETR (V4)	RT-DETR (V5)	RT-DETR (V6)	Ensemble (V4)	Ensemble (V5)	Ensemble (V6)
Precision(B)	50%	74%	68%	66%	81%	78%	85%	88%	90%
Recall(B)	50%	54%	53%	68%	74%	68%	74%	77%	79%
mAP50(B)	40%	57%	47%	65%	73%	69%	71%	74%	74%
mAP50-90(B)	30%	32%	31%	50%	46%	51%	51%	55%	58%
Accuracy	62%	71%	65%	66%	75%	72%	71%	84%	92%

Finally, Table 4 presents a comparative examination of the performance indicators between the proposed Ensemble model and two other individual models, demonstrating the progressive improvements gained by Ensemble learning techniques. The individual models in [17,18] provide a basic assessment of prediction effectiveness, but the Ensemble model (V6) demonstrates the advantages of combining many predictors. The results from the comparative table show that the Ensemble model regularly outperforms both separate models across all metrics. The Ensemble model obtains an astounding 90%, surpassing the 69% and 77% values of models [17,18]. This notable improvement indicates a significant

reduction in false positive occurrences, demonstrating the Ensemble model's dependability in making true positive predictions. The accuracy statistic highlights the Ensemble model's efficacy, with a noteworthy 92% correctness in predictions compared to the individual models' 77% and 83% accuracy, respectively. This significant difference demonstrates the benefits of Ensemble techniques, which combine the capabilities of multiple models to improve overall performance.

Table 4. Comparison of the proposed model with other models.

Metric	[17]	[18]	Ensemble(V6)
Precision(B)	69%	77%	***90***%
Recall(B)	56%	71%	***79***%
mAP50(B)	54%	72%	***74***%
mAP50-90(B)	33%	35%	***58***%
Accuracy	77%	83%	***92***%

The Ensemble model demonstrates improved precision, recall, and overall accuracy, making it a superior option to individual models [17,18]. These findings call for a larger use of Ensemble learning approaches in predictive modeling, particularly in domains where high precision and reliability are crucial. The Ensemble model's excellent metrics not only represent improved performance but also indicate promising potential for future developments within machine learning frameworks.

6. Conclusions

The integration of various artificial intelligence models into dental healthcare systems increases the accuracy and robustness of the diagnosis of dental diseases. The proposed integrated model demonstrates significant improvements in diagnostic accuracy as well as efficiency. It also simplified initial interventions as well as improved patient outcomes out of Ensemble Non-Maximum Suppression (NMS). Experimental results reveal the high performance of the proposed model, with an average 30% improvement in accuracy and 18% reduction in execution time compared to other deep learning algorithms.

In addition, this research has been practically implemented and validated by specialized dental professionals, ensuring its relevance and applicability to real-world scenarios. Unlike purely theoretical studies, this work has been integrated into practical settings, demonstrating its effectiveness and value in addressing actual challenges faced by experts in the field. These improvements not only enhance the reliability of the system within dental teams but also simplify the workflow for dental professionals, enabling them to handle and manage more complex tasks with greater efficiency and ease.

Author Contributions: Conceptualization, S.M.N., S.A.S. and I.T.A.H.; Methodology, S.M.N., S.A.S. and I.T.A.H.; Software, S.M.N. and R.S.S.; Validation, S.M.N., R.S.S., S.A.S. and I.T.A.H.; Formal analysis, S.M.N., S.A.S. and I.T.A.H.; Investigation, R.S.S., S.A.S. and I.T.A.H.; Resources, S.M.N., R.S.S., S.A.S. and I.T.A.H.; Data curation, S.M.N., R.S.S. and S.A.S.; Writing—original draft, S.M.N., R.S.S. and S.A.S.; Writing—review & editing, S.M.N., S.A.S. and I.T.A.H.; Visualization, S.M.N., R.S.S. and S.A.S.; Supervision, S.M.N.; Project administration, S.M.N.; Funding acquisition, S.M.N. and I.T.A.H.; All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Data Availability Statement: The data presented in this study are available on request from the corresponding author.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Rath, K.C.; Khang, A.; Rath, S.K.; Satapathy, N.; Satapathy, S.K.; Kar, S. Artificial Intelligence (AI)-Enabled Technology in Medicine—Advancing Holistic Healthcare Monitoring and Control Systems. In *Computer Vision and AI-Integrated IoT Technologies in the Medical Ecosystem*; CRC Press: Boca Raton, FL, USA, 2024; pp. 87–108.
2. Daniyal, M.; Qureshi, M.; Marzo, R.R.; Aljuaid, M.; Shahid, D. Exploring Clinical Specialists' Perspectives on the Future Role of AI: Evaluating Replacement Perceptions, Benefits, and Drawbacks. *BMC Health Serv. Res.* **2024**, *24*, 587. [CrossRef]
3. Park, W.J.; Park, J.B. History and Application of Artificial Neural Networks in Dentistry. *Eur. J. Dent.* **2018**, *12*, 594–601. [CrossRef]
4. Fatima, A.; Shafi, I.; Afzal, H.; Díez, I.D.; Lourdes, D.R.; Breñosa, J.; Espinosa, J.C.; Ashraf, I. Advancements in Dentistry with Artificial Intelligence: Current Clinical Applications and Future Perspectives. *Healthcare* **2022**, *10*, 11. [CrossRef]
5. Nour, S.M. Artificial Intelligence (AI) for Improving Performance at the Cutting Edge of Medical Imaging. In Proceedings of the 2023 5th Novel Intelligent and Leading Emerging Sciences Conference (NILES), Giza, Egypt, 21–23 October 2023.
6. Ghoneim, A.; D'Souza, V.; Ebnahmady, A.; Kaura Parbhakar, K.; He, H.; Gerbig, M.; Laporte, A.; Hancock Howard, R.; Gomaa, N.; Quiñonez, C.; et al. The Impact of Dental Care Programs on Individuals and Their Families: A Scoping Review. *Dent. J.* **2023**, *11*, 33. [CrossRef]
7. Kashwani, R.; Ahuja, G.; Narula, V.; Jose, A.T.; Kulkarni, V.; Hajong, R.; Gupta, S. Future of Dental Care: Integrating AI, Metaverse, AR/VR, Teledentistry, CAD & 3D Printing, Blockchain, and CRISPR Innovations. *Dent. J.* **2023**, *21*, 123–137.
8. Manickam, P.; Mariappan, S.A.; Murugesan, S.M.; Hansda, S.; Kaushik, A.; Shinde, R.; Thipperudraswamy, S.P. Artificial intelligence (AI) and internet of medical things (IoMT) assisted biomedical systems for intelligent healthcare. *Biosensors* **2022**, *12*, 562. [CrossRef]
9. Nour, S.M.; Said, S.A. Harnessing the Power of AI for Effective Cybersecurity Defense., The 6th International Conference on Computing and Informatics (ICCI 2024) be held at Future University in Egypt, Cairo, Egypt, 6–7 March 2024; IEEE: Piscataway, NJ, USA, 2024.
10. Stead, W.W. Clinical Implications and Challenges of Artificial Intelligence and Deep Learning. *JAMA* **2018**, *320*, 1107–1108. [CrossRef]
11. Géron, A. *Hands-on Machine Learning with Scikit-Learn, Keras, and TensorFlow*; O'Reilly Media, Inc.: Sebastopol, CA, USA, 2022.
12. Esteva, A.; Feng, J.; van der Wal, D.; Huang, S.C.; Simko, J.P.; DeVries, S.; Chen, E.; Schaeffer, E.M.; Morgan, T.M.; Sun, Y.; et al. Prostate Cancer Therapy Personalization via Multi-Modal Deep Learning on Randomized Phase III Clinical Trials. *NPJ Digit. Med.* **2022**, *5*, 71. [CrossRef]
13. Rajkomar, A.; Oren, E.; Chen, K.; Dai, A.M.; Hajaj, N.; Hardt, M.; Liu, P.J.; Liu, X.; Marcus, J.; Sun, M.; et al. Scalable and Accurate Deep Learning with Electronic Health Records. *NPJ Digit. Med.* **2018**, *1*, 1–10. [CrossRef]
14. Nour, S.M.; Said, S.A. Deep Learning Performance Evaluation Model for Enhancing Network Intrusion Detection Systems. In Proceedings of the 2024 6th Novel Intelligent and Leading Emerging Sciences Conference (NILES), Cairo, Egypt, 19–21 October 2024; pp. 61–65.
15. Liu, C.; Li, K.; Liang, J.; Li, K. Service Reliability in an HC: Considering from the Perspective of Scheduling with Load-Dependent Machine Reliability. *IEEE Trans. Reliab.* **2019**, *68*, 476–495. [CrossRef]
16. Lee, J.-H.; Han, S.-S.; Kim, Y.-H.; Lee, C.; Kim, I. Application of a Fully Deep Convolutional Neural Network to the Automation of Tooth Segmentation on Panoramic Radiographs. *Oral Surg. Oral Med. Oral Pathol. Oral Radiol.* **2020**, *129*, 635–642. [CrossRef] [PubMed]
17. Almalki, Y.E.; Din, A.I.; Ramzan, M.; Irfan, M.; Aamir, K.M.; Almalki, A.; Alotaibi, S.; Alaglan, G.; Alshamrani, H.A.; Rahman, S. Deep Learning Models for Classification of Dental Diseases Using Orthopantomography X-ray OPG Images. *Sensors* **2022**, *22*, 7370. [CrossRef] [PubMed]
18. Tareq, A.; Faisal, M.I.; Islam, M.S.; Rafa, N.S.; Chowdhury, T.; Ahmed, S.; Farook, T.H.; Mohammed, N.; Dudley, J. 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* **2023**, *20*, 5351. [CrossRef] [PubMed]
19. Martins, M.V.; Baptista, L.; Luís, H.; Assunção, V.; Araújo, M.R.; Realinho, V. Machine Learning in X-ray Diagnosis for Oral Health: A Review of Recent Progress. *Computation* **2023**, *11*, 115. [CrossRef]
20. Xue, P.; Wang, J.; Qin, D.; Yan, H.; Qu, Y.; Seery, S.; Jiang, Y.; Qiao, Y. Deep Learning in Image-Based Breast and Cervical Cancer Detection: A Systematic Review and Meta-Analysis. *NPJ Digit. Med.* **2022**, *5*, 19. [CrossRef]
21. Lee, C.-T.; Kabir, T.; Nelson, J.; Sheng, S.; Meng, H.-W.; Van Dyke, T.E.; Walji, M.F.; Jiang, X.; Shams, S. Use of the Deep Learning Approach to Measure Alveolar Bone Level. *J. Clin. Periodontol.* **2022**, *49*, 260–269. [CrossRef]
22. Fukuda, M.; Inamoto, K.; Shibata, N.; Ariji, Y.; Yanashita, Y.; Kutsuna, S.; Nakata, K.; Katsumata, A.; Fujita, H.; Ariji, E. Evaluation of an Artificial Intelligence System for Detecting Vertical Root Fracture on Panoramic Radiography. *Oral Radiol.* **2020**, *36*, 337–343. [CrossRef]
23. Su, S.; Jia, X.; Zhan, L.; Gao, S.; Zhang, Q.; Huang, X. Automatic Tooth Periodontal Ligament Segmentation of Cone Beam Computed Tomography Based on Instance Segmentation Network. *Heliyon* **2024**, *10*, e11858. [CrossRef]

24. Namamula, L.R.; Chaytor, D. Effective Ensemble Learning Approach for Large-Scale Medical Data Analytics. *Int. J. Syst. Assur. Eng. Manag.* **2024**, *15*, 13–20. [CrossRef]
25. AlSayyed, A.; Taqateq, A.; Al-Sayyed, R.; Suleiman, D.; Shukri, S.; Alhenawi, E.; Albsheish, A. Employing CNN Ensemble Models in Classifying Dental Caries Using Oral Photographs. *Int. J. Data Netw. Sci.* **2023**, *7*, 1535–1550. [CrossRef]
26. Li, P.; Xu, J.; Liu, S. Solid Waste Detection Using Enhanced YOLOv8 Lightweight Convolutional Neural Networks. *Mathematics* **2024**, *12*, 2185. [CrossRef]
27. Zhao, Y.; Lv, W.; Xu, S.; Wei, J.; Wang, G.; Dang, Q.; Liu, Y.; Chen, J. DETRs Beat YOLOs on Real-Time Object Detection. *arXiv* **2023**, arXiv:2304.08069.
28. Liu, J.-W.; Yang, D.; Feng, T.-W.; Fu, J.-J. MDFD2-DETR: A Real-Time Complex Road Object Detection Model Based on Multi-Domain Feature Decomposition and De-Redundancy. *IEEE Trans. Intell. Veh.* **2024**. [CrossRef]
29. Wang, C.; Sun, Y.; Wang, W.; Liu, H.; Wang, B. Hybrid Intrusion Detection System Based on Combination of Random Forest and Autoencoder. *Symmetry* **2023**, *15*, 568. [CrossRef]
30. Terven, J.; Córdova-Esparza, D.-M.; Romero-González, J.-A. A Comprehensive Review of YOLO Architectures in Computer Vision: From YOLOv1 to YOLO8 and YOLO-NAS. *Mach. Learn. Knowl. Extr.* **2023**, *5*, 1680–1716. [CrossRef]
31. Arshs Workspace Radio. vzrad2 Dataset. Roboflow Universe, 2024. Available online: https://universe.roboflow.com/arshs-workspace-radio/vzrad2 (accessed on 10 December 2024).
32. Pappireddy, A.R. Predictive Analytics in Dental Health: Leveraging Data for Early Detection and Prevention. *J. Sci. Technol.* **2024**, *15*, 121–137.
33. Chauhan, R.B.; Shah, T.V.; Shah, D.H.; Gohil, T.J.; Oza, A.D.; Jajal, B.; Saxena, K.K. An Overview of Image Processing for Dental Diagnosis. *Innov. Emerg. Technol.* **2023**, *10*, 2330001. [CrossRef]
34. Anil, S.; Porwal, P.; Porwal, A. Transforming Dental Caries Diagnosis through Artificial Intelligence-Based Techniques. *Cureus* **2023**, *15*, e42490. [CrossRef]
35. Kobi, J.; Nchaw Nchaw, A.; Otieno, B. Big Data-Driven Insights for Equitable Healthcare Access and Quality for US Immigrants. *Int. J. Res. Trends Innov.* **2024**, *9*, 392–408.
36. Dhopte, A.; Bagde, H. Smart Smile: Revolutionizing Dentistry with Artificial Intelligence. *Cureus* **2023**, *15*, e41735. [CrossRef]
37. Lubbad, M.A.H.; Kurtulus, I.L.; Karaboga, D.; Kilic, K.; Basturk, A.; Akay, B.; Nalbantoglu, O.U.; Yilmaz, O.M.; Ayata, M.; Yilmaz, S.; et al. A Comparative Analysis of Deep Learning-Based Approaches for Classifying Dental Implants Decision Support System. *J. Imaging Inform. Med.* **2024**, *37*, 2559–2580. [CrossRef]