Mesopotamian journal of Cybersecurity
Vol.4,No.3, **pp.** 189–202
DOI: https://doi.org/10.58496/MJCS/2024/022; ISSN: 2958-6542
https://mesopotamian.press/journals/index.php/cybersecurity

Research Article

Dental segmentation via enhanced YOLOv8 and image processing techniques

Dhiaa Mohammed Abed [1,2,*], Shuzlina Abdul-Rahman [2], Sofianita Mutalib [2]

[1] *Faculty of Biomedical Engineering, University of Technology, Baghdad, Iraq.*
[2] *Centre of Information Systems Studies, Faculty of Computer and Mathematical Sciences, Universiti Teknologi MARA, 40450 UiTM, Shah Alam, Selangor, Malaysia.*

ARTICLE INFO

Article history
Received 16 Aug 2024
Accepted 15 Nov 2024
Published 08 Dec 2024

Keywords
Dental Segmentation
YOLOv8
Deep Learning
Image processing
Data Privacy

ABSTRACT

By blending computer-aided medical systems with cutting-edge privacy technologies, healthcare providers can deliver more personalized, effective care while maintaining the highest data security standards and patient trust. The challenge of dental segmentation in computer vision, a task focused on accurately outlining dental structures in images, traditional methods, particularly convolution neural networks (CNNs), didn't reach high accuracy in this area due to suboptimal performance and computational inefficiency. The goal of image segmentation is to group pixels on the basis of their visual properties, such as color, texture, intensity, or spatial proximity, to identify and delineate the boundaries of distinct objects or regions within the image. In this paper, You Only Look Once (YOLOv8) algorithm is improved to segment teeth with high accuracy and high execution speed. The increase in the number of layers of YOLOv8 relied upon, as the accuracy of the algorithm segmentation depends on the number of layers used to extract features from the image (backbone) and the number of layers of the head (prediction). In addition, the size of the layers is decreased to increase the execution speed. The novelty of this paper is the proposed YOLOv8 model in addition to the Proposed Activation Function (PAF). The dataset (top view) used was taken from a dental clinic where 526 images were taken of dental and different patients. The best accuracy reached 99.561% when the enhanced YOLOv8 segmentation model was applied to the dental dataset. It can be concluded that the improved model of the YOLOv8 algorithm has increased the accuracy of dental segmentation compared to previous research because it relies on a proposed PAF that increases the difference between the features extracted from the layers of the proposed model which makes it able to distinguish between teeth and surrounding parts significantly.

1. INTRODUCTION

Dental image analysis plays a very important role not only in the oral healthcare domain but also in dental biometrics [1]. There are two types of dental images: color-based intraoral camera images and radiographic images [2]. The process of dental segmentation is one of the important techniques in the field of dentistry, as it helps the dentist in diagnosing diseases and determining the exact locations of tooth decay which is required to build a dental print (Figure 1) [3].

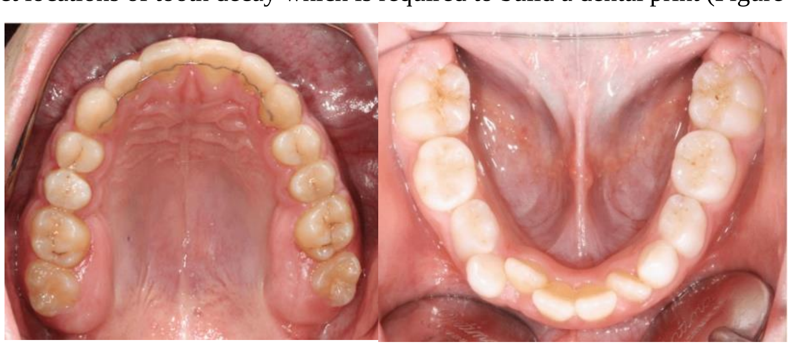

Fig. 1. Dental color images

*Corresponding author. Email: Dhiaa.M.Alfyadh@uotechnology.edu.iq

When dealing with dental images in healthcare applications and integrating YOLO retail and privacy technologies, the same basic principles are applied to maintain the privacy of sensitive patient data, with tools and models customized to suit the nature of dental images. YOLO can be used to detect dental problems (such as cavities, decay, or tooth abnormalities) and analyse dental images (such as dental X-rays or digital oral images) while ensuring that patient privacy is protected. Currently, many studies are investigating the applications of dental image analysis, but the point is which method is appropriate in most cases [4].

The accuracy and efficiency of the dental segmentation is one of the most important challenge facing this study is that needs improvement [5]. The YOLOv8 algorithm is one of the modern algorithms, which is characterized by good accuracy in the segmentation process, but it needs to be improved to increase its accuracy. The second challenge facing this study is the Time-Consuming Process. Where the time of implementation in the medical field is one of the important things greatly as it is related to human treatment.

Two important contributions can be made in the field of dentistry on the one hand and in computer science on the other. The first contribution is the development of a model for the YOLOv8 algorithm to obtain high accuracy in segmenting teeth, segmenting them and isolating them from the rest of the mouth, using a color image of the mouth taken using a micro camera. The second contribution is to increase the speed of the proposed system, that increasing the speed of the proposed system depends on determining the sizes of the layers of the proposed model, where the large sizes of the layers, the greater the execution time and vice versa, and at the same time, the increase in the sizes of layers increases accuracy, and for this it was necessary to reach the number and sizes of layers to reach the lowest execution time without affecting of high accuracy, in addition to the SoftMax equation, which predicts [6].

The image segmentation problem is difficult in image processing and plays an important role in most subsequent image analyses, especially in pattern recognition and image matching [7]. In computer-aided procedures and clinical diagnosis, tooth segmentation plays an important role. It can produce approximate outlines of doubtful regions to provide features that enable distinction between teeth and other tissues [8]. Image segmentation plays a crucial role in various computer vision applications, including object recognition, scene understanding, image editing, medical image analysis, autonomous vehicles, and robotics. It provides a foundational step for extracting meaningful information from images, enabling further analysis and decision-making [9].

2. RELATED WORKS

Previous studies constitute a group of studies on the process of dental segmentation. In 2022, [10] presented a dental detection method that aims to use YOLOv4 to detect and number teeth automatically, and the YOLOv4 algorithm relies on the CNN algorithm in detecting teeth. The reason for adopting the CNN algorithm is that it contains convolution layers that extract the largest number of features from images. Panoramic X-ray images of the teeth have been adopted and trained via the YOLOv4 algorithm, and the model has shown good efficiency in detecting the teeth, as its accuracy reached 92.22%. The proposed CNN methodology exhibited robust and expeditious performance in automating tooth detection and numbering on pediatric panoramic radiographs. The functionality of automated tooth detection has the potential to assist dental practitioners in saving time and serve as a preprocessing tool for identifying dental pathologies.

In 2022, the study [11] concentrated on introducing a postprocessing phase aimed at generating a segmentation map, where objects within an image are distinctly delineated. This technique is applied to tooth instance segmentation via the U-Net network. Postprocessing involves employing grayscale morphological and filtering operations on the sigmoid output of the network before binarization. The overall tooth segmentation yields a Dice overlap score of 95.4±0.3%. The proposed postprocessing stages result in a reduction in the mean error in the tooth count to 6.15%, in contrast to the error of 26.81% without postprocessing. Notably, both the segmentation and tooth counting performances achieve the highest reported values in the literature, given our current knowledge. Furthermore, this accomplishment is realized through the utilization of a comparatively modest training dataset comprising 105 images. While the primary aim of this study is tooth instance segmentation, the introduced method is deemed applicable to analogous challenges in other domains, such as the separation of cell instances.

In 2023, the authors of the study [12] proposed a UNet network structured with a CNN-transformer architecture, strategically combining the strengths of convolutional neural networks (CNNs) and transformers. The CNN component adeptly extracts local features, whereas the transformer captures long-range dependencies. The incorporation of multiple spatial attention modules further enhances the network's proficiency in extracting and representing spatial information. Additionally, a novel masked image modelling method is introduced to simultaneously pretrain the CNN and transformer modules, mitigating limitations arising from a relatively smaller pool of labelled training data. The experimental results underscore the superior performance of the proposed method, with a Dice similarity coefficient (DSC) of 87.12%, an intersection over union (IoU) of 78.90%, a Hausdorff distance 95% (HD95) of 0.525 mm, and an average symmetric surface distance (ASSD) of 0.199

mm. The findings underscore the method's efficiency and effectiveness in automating the precise segmentation of dental CBCT images, highlighting its practical utility in orthodontics and dental implant applications.

In 2023, [13] documented the initial phases of development for an unsupervised, deep learning-driven clinical annotation and segmentation tool, denoted CAST. This tool is designed to autonomously isolate clinically relevant teeth in both intraoral photographs and corresponding oral radiographs. The dataset, comprising 172 intraoral photographs and 424 dental radiographs, was manually annotated by two operators. Through augmentation, the dataset was expanded to include 6258 images for training, 183 for validation, and 98 for testing. The training process incorporated an object detection model, 'YOLOv8', in conjunction with a feature extraction system known as the 'Segment Anything Model.' This combined approach facilitated the automatic annotation and segmentation of tooth-related features and lesions in both types of images without requiring operator intervention. The resulting outputs underwent further refinement via a data relabelling tool, 'X-AnyLabelling', which provided the option to manually reannotate inaccurate data outputs through reinforcement learning. The trained object detection model demonstrated a mean average precision (mAP) of 77.4%, with precision and recall rates of 75.0% and 72.1%, respectively. Compared with radiological images annotated by bounding boxes, the model exhibited a superior ability to segment features from oral images annotated with polygonal boundaries.

In 2024, [14] This study presents an oral-diagnosis framework integrating the YOLOv8 model for precise tooth localization in dental imaging. The dental segmentation and numbering in the right-side bitewing radiographic images were evaluated through comparison of the YOLOv5 and YOLOv8 models, employing confidence thresholds. The dataset comprised 800 training images and 152 testing images, with the YOLOv8 architecture deployed in three variants. Precision, recall, F1-score, and mean average precision (mAP) were evaluated for both models. YOLOv8 demonstrated superior performance over YOLOv5 in precision (0.913 vs. 0.897), F1-score (0.931 vs. 0.920), and mAP (0.96 vs. 0.954). Variations in model dimensions were observed among YOLOv8 S, M, and L variants, with marginal mAP improvements in specific classes. In conclusion, while YOLOv8 did not enhance dental segmentation and numbering tasks across varying architecture sizes, it consistently outperformed YOLOv5, exhibiting superior segmentation and detection abilities.

Through previous research and deep learning algorithms, note that increasing segmentation accuracy depends on several criteria, which are extracting the largest number of features from the image, selecting the most important features and deleting the rest of the features, while the last criterion is the activation function, which greatly affects the accuracy of the segmentation.

The YOLOv8 is the first version of this algorithm that works on segmenting, while previous versions work on detection only [14]. There are many gaps that can be improved in the YOLOv8 algorithm, where the algorithm depends on layer sizes to increase the features extracted from the images, where there are five models of YOLOv8, the first of which is YOLOv8-seg-n, whose layer sizes begin with (16, 32, 64, …..), YOLOv8-seg-l, this model begin with size (32, 64, 128, …..), down to the YOLOv8-seg-x, whose layers are in sizes (80, 160, 320, …) [15][18]. The YOLOv8-seg-x is the most accurate in segmentation than previous models, because the sizes of the image feature extraction layers are larger, but it is the slowest in execution, and the reason is also because of the sizes of its large layers Here the gap can be reached, which is that the accuracy and execution time in the YOLOv8 algorithm are inversely proportional.

Some places where the teeth meet with the rest of the mouth may be unclear, and here begins the role of the power of the activation function in determining whether this part belongs to the teeth or not [16][19]. The activation function of the YOLOv8 algorithm is the SoftMax, and this function predicts all classes within the range of 0 to 1, which leads to some places where the prediction may be close, and here it is necessary to reach an activation function to increase the distance between the prediction values. In addition, the ability of the algorithm to work on more than one dataset at the same time must be tested and therefore made capable of dealing with dental images in different view [17][20].

Previous algorithms such as YOLOv4, Unet, and CNN are good in segmentation, but they lack high accuracy and are slow to execute because these algorithms contain mathematical equations that require a long execution time, especially when dealing with color and big images.

3. MATERIALS AND METHODS

The YOLO algorithm is considered a modern algorithm because it has been shown to be efficient in many areas; for this reason, it was chosen as a proposal for dental segmentation. In this research, YOLOv8 of the algorithm was used and modified to display its increased accuracy in dental segmenting. The modification was made to the layers of the model by increasing the number of layers of the backbone to increase the number of features extracted from the image, as well as the modification in the sizes of the layers to make the proposed model suitable for the process of dental segmenting and isolating it from the rest of the mouth. Figure 2 illustrates the proposed YOLOv8 segmentation model. The limitations in this model is that it deals with color images, and to obtain high accuracy in dental segmentation, the images must be clear and taken with a camera with high specifications through which the teeth are accurately segmented.

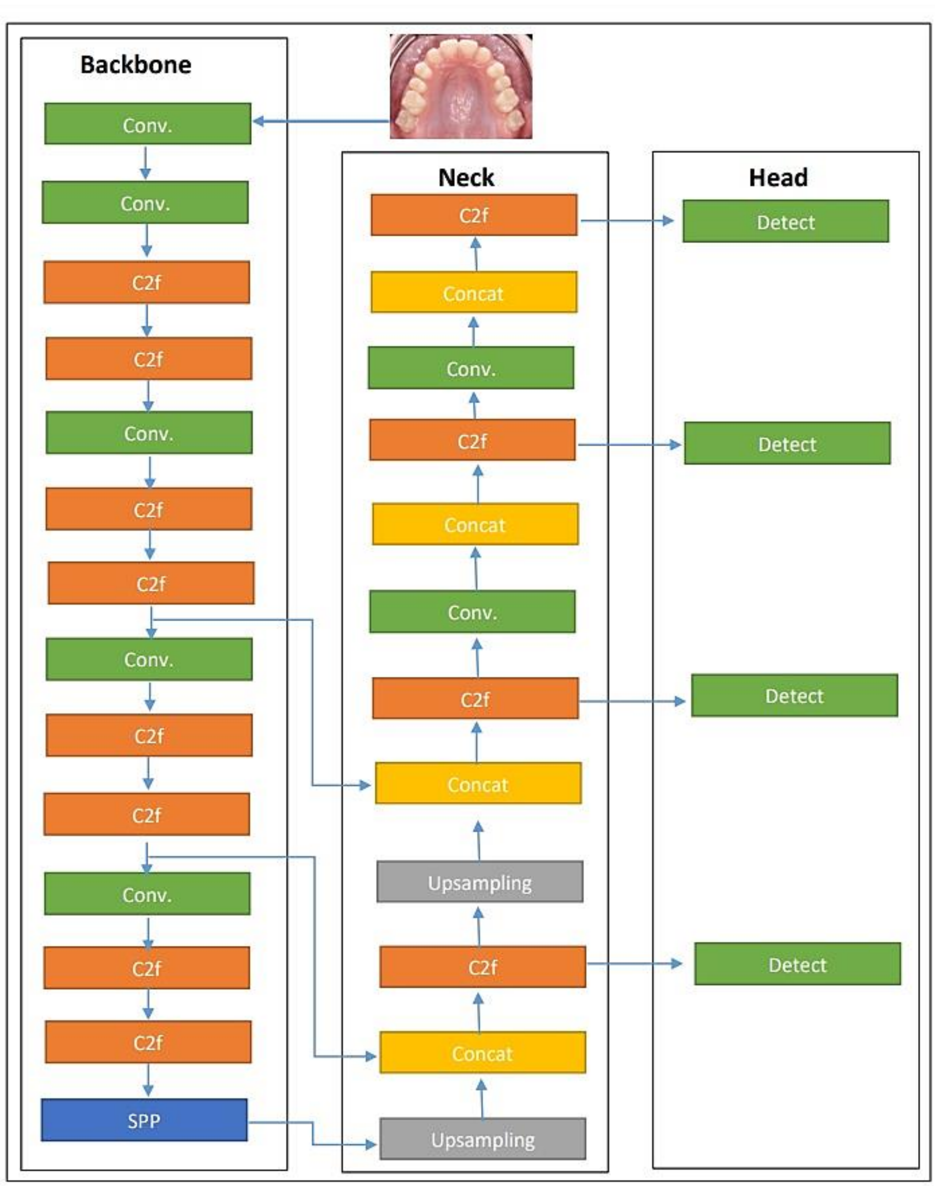

Fig. 2. Proposed YOLOv8 segmentation model

Assume that the input image is of size $W \times H$ (width by height). YOLOv8 divides the image into an $\boldsymbol{S} \times S$ grid, where each grid cell is responsible for predicting object bounding boxes. For a grid size of $\boldsymbol{S} \times S$, each cell predicts B bounding boxes. Each bounding box has the following:

Four parameters are used: (x, y, w, h),, where x and y are the coordinate centers of the bounding box relative to the grid cell, and w and h are the width and height of the box relative to the image size. In this work, image sizes of 640 * 640 pixels were adopted, and the YOLO algorithm divides the image into a grid of 48 * 48 for each cell in the grid. YOLOv8 predicts bounding boxes as offsets relative to the cell $(x_{cell}, y_{cell}), then$ converted to absolute coordinates in the image as follows:

$$x_{absolute} = \frac{1}{S} \left(cell + Predicted_x\right) \tag{1}$$

$$y_{absolute} = \frac{1}{S} \left(cell + Predicted_y\right) \tag{2}$$

Width w and height h are also predicted relative to the whole image, but they are scaled according to the model output:

$$w = PAF(w) \times width_{img} \tag{3}$$

$$h = PAF(h) \times height_{img} \tag{4}$$

where PAF is the proposed activation function to ensure that w and h are between 0 and the image size.

Confidence score: Confidence P_{conf} , which indicates how likely the box is to contain an object. The confidence score for each bounding box is mathematically represented as:

$$P_{conf} = P(obj.) \times IOU_{pred}^{true} \tag{5}$$

where P(obj.) is the probability that there is an object in the bounding box. where IOU_{pred}^{true} is the intersection over union (IOU) between the predicted bounding box and the ground truth box.

$$IOU = \frac{\text{Area of overlap between predicted box and true box}}{\textit{Area of union of the predicted and true boxes}} \tag{6}$$

C class probabilities: Each grid cell also predicts a probability distribution over C object classes. This is computed via the PAF function:

$$P(Class_i) = PAF(z_i, C) \tag{7}$$

where z_i is the score predicted for class i and where C is the total number of classes.

Loss function: YOLOv8 uses a multipart loss function to train the model, which combines

$$Loss_{loc} = \sum_i 1_i^{obj}[(x_i - x_i')^2 + (y_i - y_i')^2 + (w_i - w_i')^2 + (h_i - h_i')^2] \tag{8}$$

The previous equation is localization loss for (x, y, w, h), which are points of the bounding box, and 1_i^{obj} if the grid contains an object and 0_i^{obj} otherwise.

$$Loss_{conf} = \sum_i 1_i^{obj}\left(P_{conf} - 1\right)^2 + \sum_i 1_i^{noobj}\left(P_{conf} - 0\right)^2 \tag{9}$$

The previous equation is confidence loss. The last loss is classification loss.

$$Loss_{class} = \sum_i 1_i^{obj} \sum_c (P_{class} - P'_{class})^2 \tag{10}$$

YOLOv8 of the algorithm was used and modified to display its increased accuracy in segmenting the dentist. The modification was made to the layers of the model by increasing the number of layers of the backbone to increase the number of features extracted from the image, this is because applying more kernels to images leads to an increase in the number of extracted matrices., as well as the modification in the sizes of the layers to make the proposed model suitable for the process of segmenting the dental from the rest of the mouth. Where the backbone contains three types of layers, Conv. C2f and SPP, where the Conv. layer extracts features from images, so increasing the number of this type of layers leads to the extraction of more features. The model's backbone now consists of a C2f module instead of a C3 one. The difference between the two is that in C2f, the model concatenates the output of all bottleneck modules. In contrast, in C3, the model uses the output of the last bottleneck module. the SPP block includes three parallel maximum pooling layers the feature maps of different scales are analyzed separately then the concat operation is applied. The enhanced YOLOv8 segmentation model consists of 30 layers, and the specifications of each layer are accurately determined to suit the process of dental segmentation. Table 1 shows the specifications of the proposed YOLOv8 segmentation model.

TABLE I. SPECIFICATION OF THE PROPOSED YOLOV8 SEGMENTATION MODEL

Layer No.	Part	Layer	Specification
1	Backbone	Conv.	[3, 48, 3, 2]
2		Conv.	[48, 96, 3, 2]
3		C2f	[96, 96, 2, True]
4		C2f	[96, 96, 2, True]
5		Conv.	[96, 192, 3, 2]

6		C2f	[192, 192, 4, True]
7		C2f	[192, 192, 4, True]
8		Conv.	[192, 384, 3, 2]
9		C2f	[384, 384, 4, True]
10		C2f	[384, 384, 4, True]
11		Conv.	[384, 576, 3, 2]
12		C2f	[576, 576, 2, True]
13		C2f	[576, 576, 2, True]
14		SPP	[576, 576, 5]
15	Neck	Upsampling	[None, 2, 'nearest']
16		Concat	[1]
17		C2f	[960, 384, 2]
18		Upsampling	[None, 2, 'nearest']
19		Concat	[1]
20		C2f	[576, 192, 2]
21		Conv.	[192, 192, 3, 2]
22		Concat	[1]
23		C2f	[576, 384, 2]
24		Conv.	[384, 384, 3, 2]
25		Concat	[1]
26		C2f	[960, 576, 2]
27	Head	Detect	[1]
28		Detect	[1]
29		Detect	[1]
30		Detect	[1]

The process of dental segmentation must be accurate, and to increase accuracy, the segmentation activation function must be improved. In the YOLOv8 algorithm, SoftMax was used, and in the proposed system, SoftMax was improved by a sigmoid linear unit (SiLU) as follows:

$$SoftMax = \frac{e^x}{\sum_{i=1}^{N} e^x} \tag{11}$$

$$SiLU = \frac{x}{1 + e^{-x}} \tag{12}$$

Through the previous two equations (Eq. 11 and Eq. 12), it is possible to reach the proposed activation function (PAF) with higher accuracy by substituting the SoftMax equation with the SiLU equation as follows:

$$PAF = \frac{x}{1 + e^{-SoftMax}} \tag{13}$$

$$PAF = \frac{x}{1 + e^{-\left(\frac{e^x}{\sum_{i=1}^{N} e^x}\right)}} \tag{14}$$

The goal of including the Softmax with the SilU formula is that Softmax is used for classification and gives a prediction for all categories, and if these predictions are collected, the result is always 1, meaning that the prediction values are always between 0 and 1, which leads to the prediction values being close and leads to an increase in the percentage of error in the segment. The SilU equation is used with the convolution layer to prevent the features extracted from the images from reaching zero; for this reason, it is suitable for increasing the prediction values and increasing the distances between them, which leads to greater accuracy of the segmentation in the proposed equation than SoftMax. An experiment with a mathematical example of the two equations reveals that the difference in prediction has increased in the proposed equation; therefore, the proposed improvement will increase the accuracy of dental segmentation.

$$\begin{bmatrix} 0.25 \\ 1.23 \\ -0.8 \end{bmatrix} \longrightarrow SoftMax \longrightarrow \begin{bmatrix} 0.249 \\ 0.664 \\ 0.087 \end{bmatrix}$$

$$\begin{bmatrix} 0.25 \\ 1.23 \\ -0.8 \end{bmatrix} \longrightarrow PAF \longrightarrow \begin{bmatrix} 0.141 \\ 0.819 \\ -0.417 \end{bmatrix}$$

The process of building a dataset is highly important for building an intelligent system. AI algorithms need a number and different types of images to obtain high accuracy in training. In this work, top-view dental images were taken from a dental clinic where 526 images were taken of dental and different patients. Any identifying information is removed from the images by hiding the parts containing facial features by cropping techniques to maintain patient data privacy while preserving only the data necessary for dental analysis. Figure 3 shows dental samples from the dataset. The pictures taken from the dental clinic are of different sizes. In this work, a real dental dataset was divided into two parts: 70% for training and 30% for testing. Figure 3 shows samples of the dataset. These images were taken through real cases in a dental clinic in Iraq.

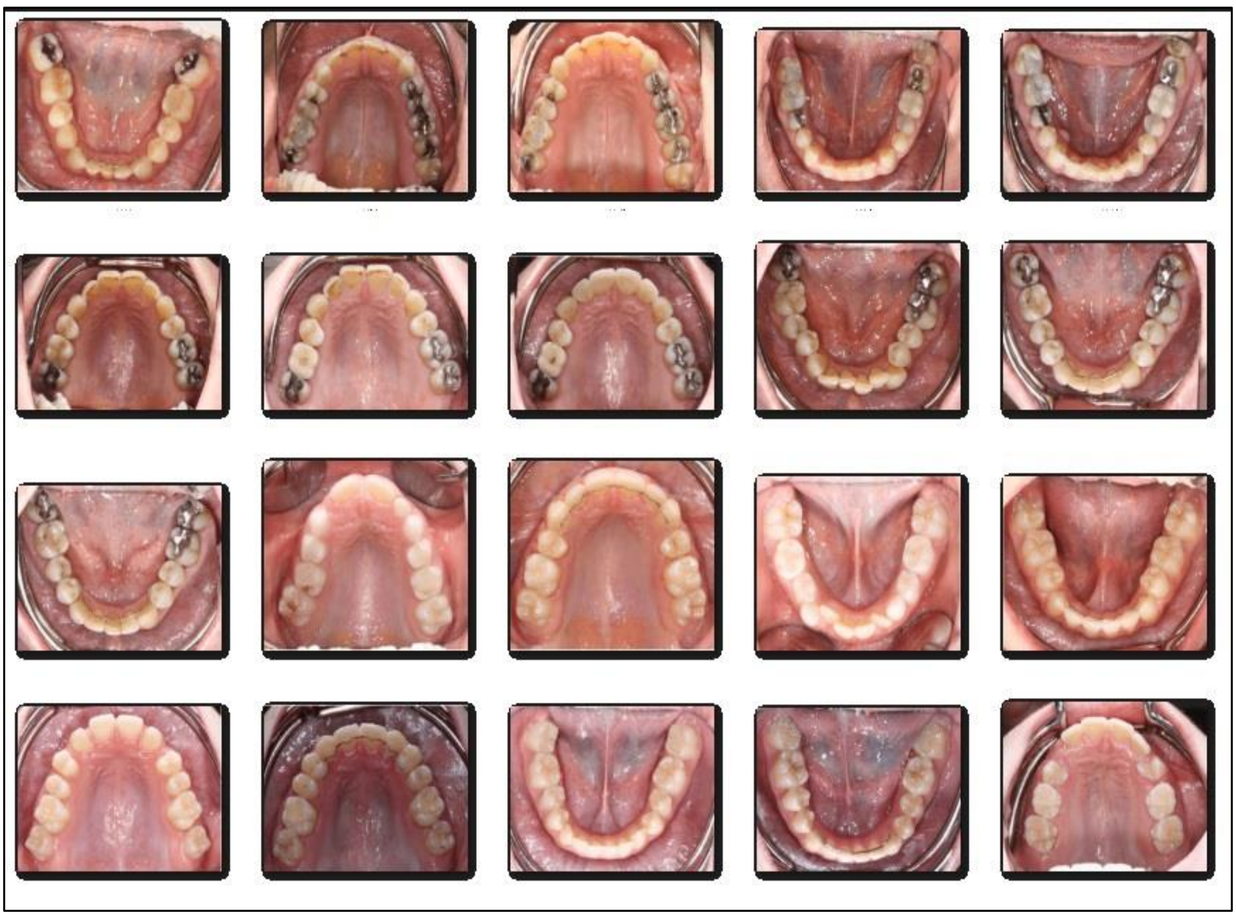

Fig. 3. Samples from the real dental dataset

4. RESULTS

The data used were taken through a dental clinic in Iraq and contained 526 images, and the number of images was increased by augmentation. This step was applied to increase the number of cases in which pictures were taken in different ways by a dentist, and the number of images reached 3156. Several cases have been considered to achieve the best accuracy in segmenting the dentist from the rest of the mouth. The novelty of this paper is proposed YOLOv8 model with specific number and size for layers of segmentation and improve the activation function to increase accuracy of dental segmentation.

First case: In this case, the proposed YOLOv8 segmentation model is trained and tested by using 15 training epochs (Figure 4), which illustrates the results of training and testing the dental dataset during epoch 15.

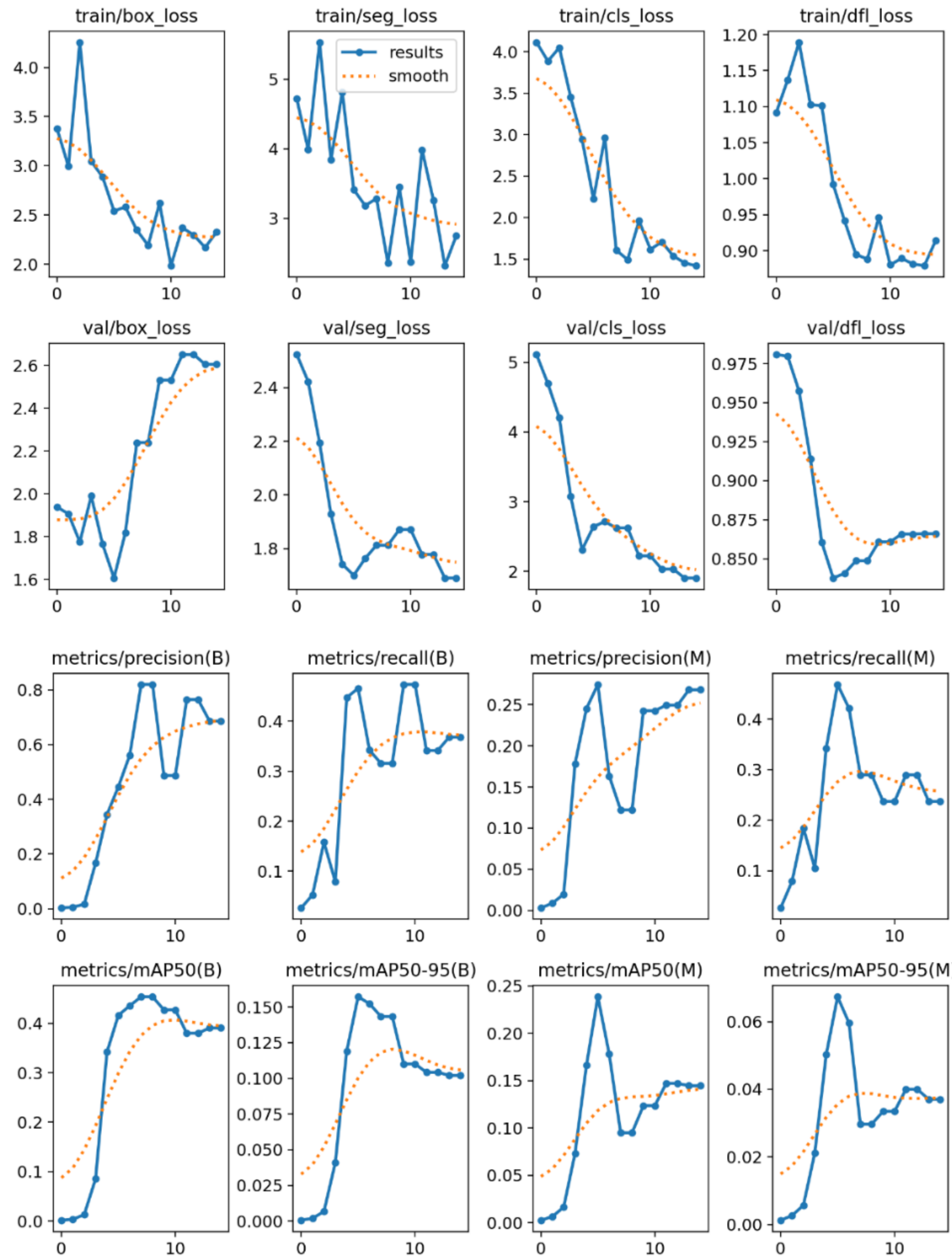

Fig. 4. Results of training and testing the dental dataset during epoch 15

The second case involves training and testing the proposed YOLOv8 segmentation model by using 30 training epochs (Figure 5), which illustrates the results of training and testing the dental dataset during epoch 30.

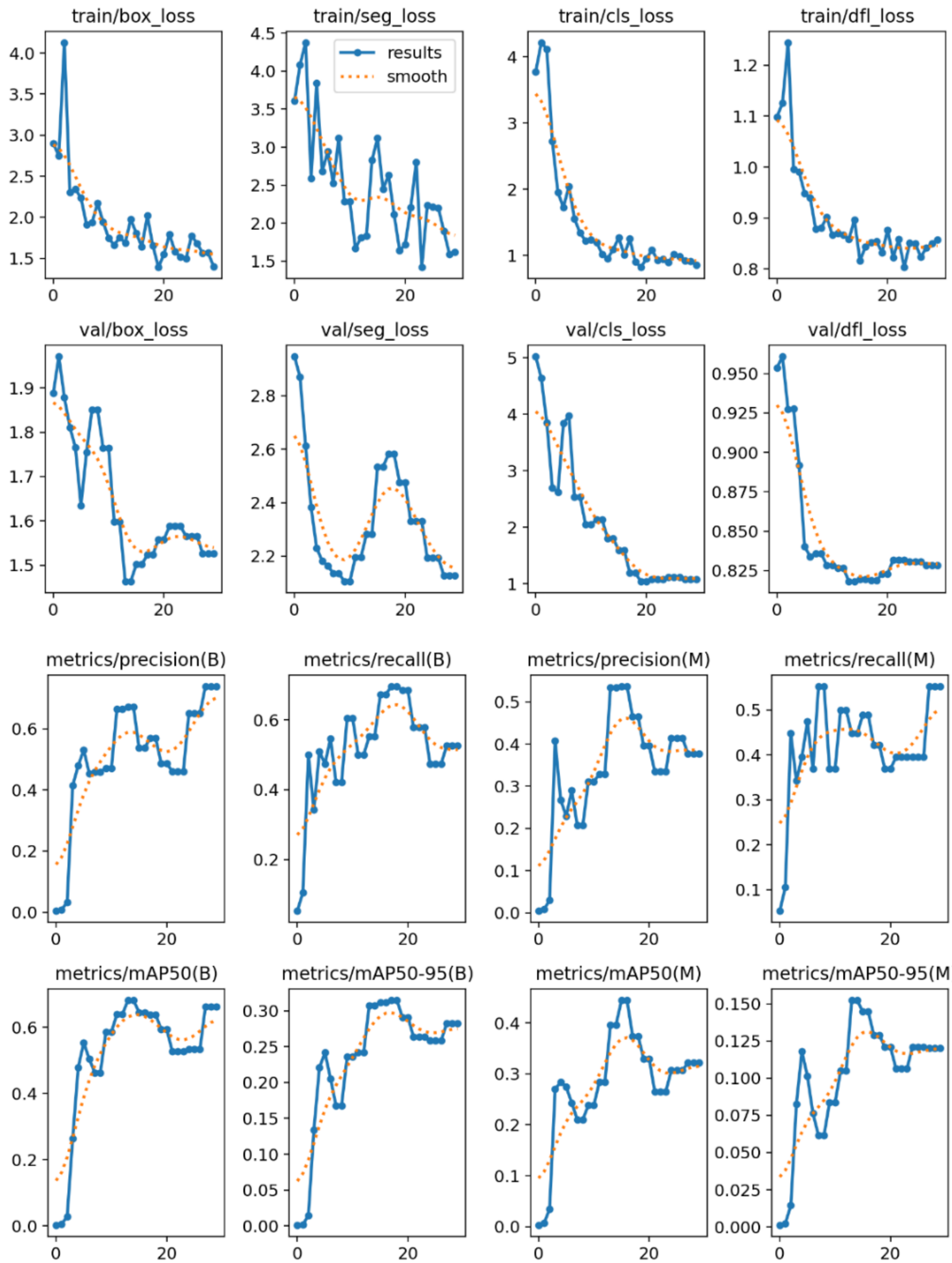

Fig. 5. Results of training and testing on the normal dataset using epoch 30

The third case involves training and testing the proposed YOLOv8 segmentation model by using 50 training epochs (Figure 6), which illustrates the results of training and testing the dental dataset using epoch 50.

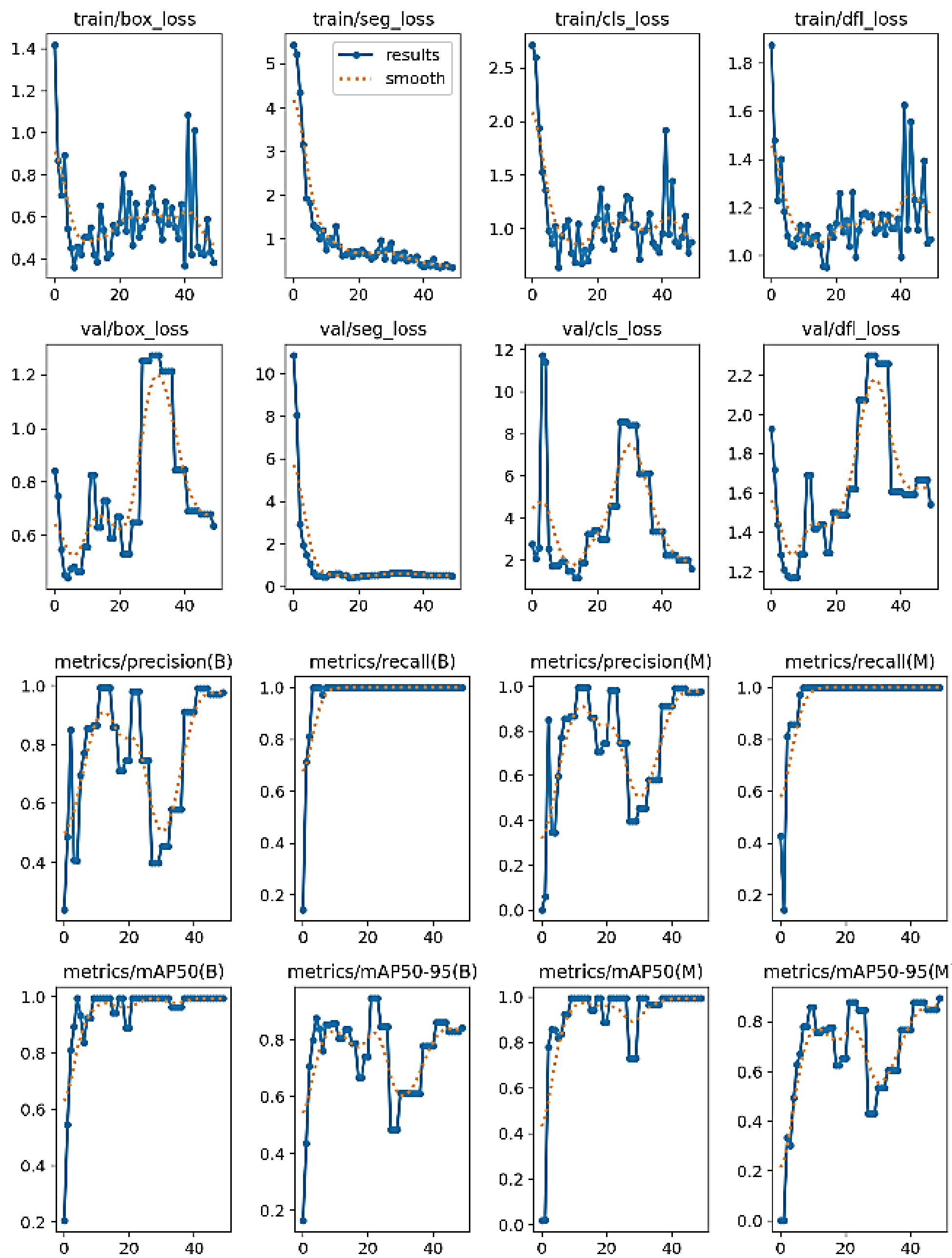

Fig. 6. Results of training and testing the dental dataset using epoch 50

The results of these experiments can be summarized in Table 2, where the best accuracy was reached at 50 epochs, where it reached 99.561% when the proposed YOLOv8 segmentation model was applied to the dental dataset. Where the table summarizes the results of the training of the proposed model in a group of cases to reach the best accuracy of dental segmentation. Where a set of measures were used to inspect the proposed model, these scales determine the prediction accuracy of the set of training and testing images, in addition to loss scales, where the model is tested through the use of loss scales. These scales are calculated by the extent to which the parts that were identified during the dental segmentation match the images that were made of the previously identified dental through supervised, and here the equations of these scales are applied to obtain the results. The metrics/mAP50(B) scale is the dental segmentation scale where notice that its value increased when increased the number of training epochs where its value was 39.073 in the first case, 66.262 in the second case and 99.561 in the third case. The val/seg_loss is loss rate scale for the testing images is 1.6903 in the first case, 2.1272 in the second case and 0.48883 in the third case, this scale represents the amount of data lost during the testing process.

TABLE II. SUMMARY OF THE RESULTS OF THE DENTAL DATASET

Evaluation Measure	First Case	Second Case	Third Case
train/box_loss	2.3257	1.3984	0.38524
train/seg_loss	2.7518	1.6139	0.35353
train/cls_loss	1.4187	0.85871	0.87865
train/dfl_loss	0.91424	0.85774	1.0683
metrics/precision(B)	0.68641	0.73911	0.9761
metrics/recall(B)	0.36842	0.52632	0.99624
metrics/mAP50(B)	0.39073	0.66262	0.99561
metrics/mAP50-95(B)	0.10207	0.28264	0.84186
metrics/precision(M)	0.26819	0.37745	0.9761
metrics/recall(M)	0.23684	0.55263	0.99644
metrics/mAP50(M)	0.14454	0.32163	0.995247
metrics/mAP50-95(M)	0.03703	0.12006	0.8955
val/box_loss	2.6047	1.5259	0.63554
val/seg_loss	1.6903	2.1272	0.48883
val/cls_loss	1.9054	1.0849	1.5975
val/dfl_loss	0.86602	0.82828	1.5415

The process of proposing an equation instead of SoftMax led to a significant increase in the accuracy of dental segmentation, which increased the error rate in Table 3 compared with the prediction between the proposed equation and the results of SoftMax for part of the image. As the PAF aims to increase the distance between the values extracted from the image, because the values in the dental images are close, which confuses the work of SoftMax, and for this reason it was developed. The table contains two columns for prediction label in addition to the correct label When testing 16 values that represent part of testing image, note that the PAF fail in one prediction while the SoftMax fail in three, in addition to that the prediction column gives a clear difference to distinguish between the background and the dental.

TABLE III. COMPARISON BETWEEN PAF AND SOFTMAX PREDICTION

Input Value	PAF			SoftMax		
	Predication	Prediction Label	Label	Predication	Prediction Label	Label
0.17	0.12	Background	Background	0.28	Background	Background
0.38	0.51	Background	Dental	0.58	Background	Dental
0.13	0.10	Background	Background	0.23	Background	Background
1.58	0.95	Dental	Dental	0.77	Dental	Dental
1.33	0.89	Dental	Dental	0.71	Dental	Dental
1.28	0.84	Dental	Dental	0.69	Dental	Dental
1.61	0.91	Dental	Dental	0.82	Dental	Dental
0.58	0.79	Dental	Dental	0.58	Dental	Dental
0.47	0.64	Dental	Dental	0.49	Dental	Background
0.64	0.81	Dental	Dental	0.54	Dental	Dental
0.11	0.19	Background	Background	0.49	Background	Background
0.21	0.34	Background	Background	0.53	Background	Background
0.69	0.83	Dental	Dental	0.48	Dental	Background
0.43	0.67	Dental	Dental	0.44	Dental	Dental
0.22	0.35	Background	Background	0.35	Background	Background
0.34	0.41	Background	Background	0.41	Background	Background

Here, the larger the size of the model (i.e., the sizes of the layers) is, the greater the accuracy, but this leads to a slower implementation; for this reason, the layers of the proposed YOLOv8 segmentation model were determined by obtaining higher accuracy with a lower execution time by increasing the number of layers for extracting features and reducing their sizes. The reason for the increase in the number of layers leads to more accurate features and thus better prediction. The process of reducing the size of the layers decreases the execution time. Table 4 lists the execution times for each model when test sample segmentation is performed. In this table, a different set of previous models of the YOLOv8 algorithm were tested, which contains 5 models with different layers in addition to their different sizes, and the table shows the accelerometer and the accuracy scale to test each model and compare it with the proposed model, which was built using a different number of layers from previous models, as well as using sizes similar to previous models and also different sizes. Note that the smallest model is YOLOv8n, which consists of 23 layers, starting with (16, 32, 64, …) and took 1.47 seconds and an accuracy of 68.444, which is the fastest due to the small number of layers and small sizes, and Model YOLOv8x is the largest, which consists of 25 layers and sizes starting with (80,160,320, ...) and has reached an accuracy of 82.837 and an execution time of 5.68 seconds. As for the proposed model, the number of layers has been increased, which contains 30 layers, which increases accuracy, but to reduce the implementation time, the time is reduced, the sizes of the layers are reduced, starting with (48,96,192, …) where the accuracy reached 99.561 and the execution time is 1.66 seconds.

TABLE IV. COMPARISON OF THE EXECUTION TIMES OF THE YOLOV8 FAMILY WITH THOSE OF THE PROPOSED MODEL

YOLOv8 Model	Number of Backbone Layers	Number of Neck Layers	Number of Head Layers	Start Layer Size	Time Execution	ACC.
YOLOv8n	10	12	1	16, 32, 64, …..	1.47 s	0.68444
YOLOv8s	10	12	1	32, 64, 128, ….	1.98 s	0.54869
YOLOv8m	10	12	3	32, 64, 128, ….	4.64 s	0.73417
YOLOv8l	10	12	3	64, 128, 256, …	5.22 s	0.81692
YOLOv8x	10	12	3	80, 160, 320, …	5.68 s	0.82837
Our Model	14	12	4	48, 96, 192, …	1.66 s	0.99561
	14	12	4	16, 32, 64, …	2.57 s	0.83697
	14	12	4	32, 64, 128, …	5.11 s	0.82687
	14	12	4	64, 128, 256, …	7.54 s	0.96881
	14	12	4	80, 160, 320, …	7.89 s	0.97547

For the purpose of evaluating the proposed model, many images were tested, as the images that were used were of different lighting and sizes, and Figure 7 illustrates the mask detection and segmentation results.

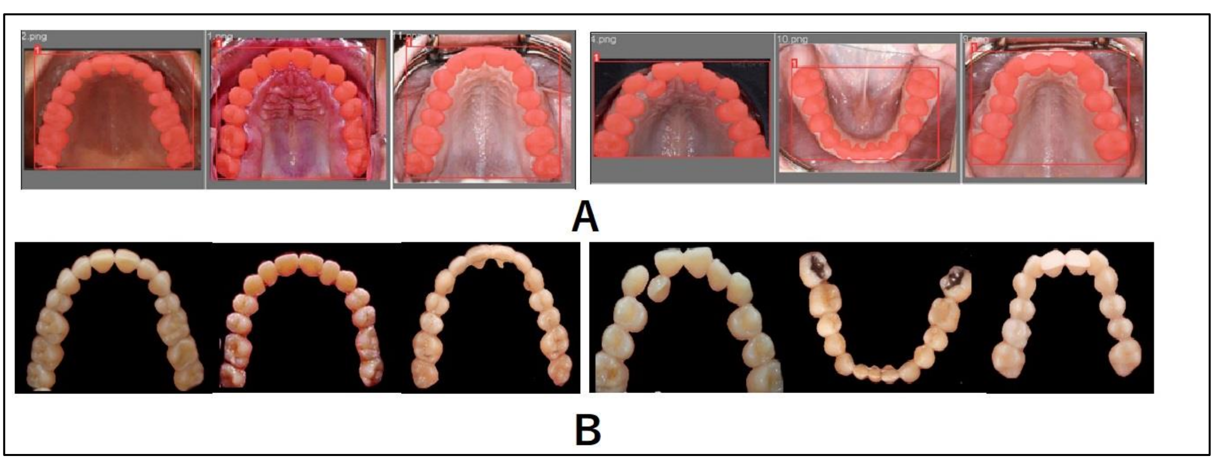

Fig. 7. Mask detection results: (A) dental sample detection and (B) dental sample segmentation

The proposed model is compared with previous methods in the field of dental segmentation. Previous studies have shown different algorithms in this field in addition to applying different types of datasets, such as radiological and X-ray images, which help dentists identify the dentals and diseases that patients suffer from. These results can be summarized in Table 5, where it can be noted that the highest accuracy found in previous research (Gil Jader *et al*, 2020) reached 98% in segmenting the dentist, and compared with the proposed YOLOv8 segmentation model, it reached a higher accuracy of 99.5247%, indicating that the proposed model prefers to reduce the error rate.

TABLE V. COMPARISON WITH PREVIOUS WORKS

Author	Uses Methods	Accuracy
Kaya, Emine, *et al*, 2022 [10]	YOLO V4 and CNN	92.22%
Dhar, Mrinal Kanti, *et al*, 2023 [15]	FUSegNet and PCA	90.37%
da Silva Rocha, Élisson, 2022 [16]	U-Net, DCU-Net, DoubleUNet and Nano-Net	96.57%
HELLİ, Serdar, 2022 [11]	U-Net network	95.40%
Chen, Zeyu *et al*, 2023 [12]	Proposed a CNN-Transformer Architecture UNet network	87.12%
Farook, Taseef H., *et al*, 2023 [13]	YOLOv8	77.40%
The Proposed Work	Enhanced YOLOv8 Segmentation Model	0.996 for detection 0.995 for segmentation

5. CONCLUSION

The integration of YOLO with privacy techniques in dental image analysis helps maintain the confidentiality of personal data while providing accurate and effective analyses of dental health problems via standardized learning. Healthcare providers can provide advanced care without risking patient privacy. When the proposed YOLOv8 segmentation model was trained on a dental dataset, the accuracy was 0.99561 for dental detection, and the accuracy of dental segmentation was 0.995247, as the proposed model takes an implementation time of 1.66 seconds. Compared with the YOLOv8 family, it can be concluded that the best accuracy is reached when the YOLOv8x model is applied, and the larger size of the layers results in high accuracy and high execution time; the best size of layers starts with (48, 96, 192, …) because this size gave a balance between accuracy and speed in performance. The highest accuracy of the previous models when applied to the normal dataset reached 0.98285 for dental detection and 0.82837 for dental segmentation, whereas that of the proposed model reached 0.99561 for dental detection and 0.995247 for dental segmentation. The proposed model greatly helps dentists in determining the teeth for the purpose of diagnosing the disease by the doctor through applied to a different set of dental dataset with different views, provided that these images are colored and can be invested in identifying teeth for the purpose of discovering diseases in future research.

Conflicts of interest

The authors declare that they have no conflicts of interest.

Funding

The author's paper explicitly states that the research project did not receive any funding from institutions or sponsors.

Acknowledgement

The authors is thankful to the institution for their commitment to supporting scholarly endeavors and creating a conducive research environment.

References

[1] A. M. Khan, “Image Segmentation Methods: A Comparative Study,” *Int. J. Soft Comput. Eng.*, vol. 3, no. 4, pp. 84–92, 2013.

[2] T. Bonny *et al.*, “Dental bitewing radiographs segmentation using deep learning-based convolutional neural network algorithms,” *Oral Radiol.*, vol. 40, no. 2, pp. 165–177, 2024, doi: 10.1007/s11282-023-00717-3.

[3] G. Rubiu *et al.*, “Teeth Segmentation in Panoramic Dental X-ray Using Mask Regional Convolutional Neural Network,” *Appl. Sci.*, vol. 13, no. 13, 2023, doi: 10.3390/app13137947.

[4] X. Xu, C. Liu, and Y. Zheng, “3D Tooth Segmentation and Labeling Using Deep Convolutional Neural Networks,” *IEEE Trans. Vis. Comput. Graph.*, vol. PP, p. 1, May 2018, doi: 10.1109/TVCG.2018.2839685.

[5] F. R. S. Teles *et al.*, “Tooth Detection and Numbering in Panoramic Radiographs Using YOLOv8-Based Approach BT - Wireless Mobile Communication and Healthcare,” 2024, pp. 239–253.

[6] A. Fatima *et al.*, “Deep Learning-Based Multiclass Instance Segmentation for Dental Lesion Detection,” *Healthcare*, vol. 11, no. 3. 2023, doi: 10.3390/healthcare11030347.

[7] S. Vinayahalingam *et al.*, “Intra-oral scan segmentation using deep learning,” *BMC Oral Health*, vol. 23, no. 1, p. 643, 2023, doi: 10.1186/s12903-023-03362-8.

[8] Z. Kong *et al.*, “Automated maxillofacial segmentation in panoramic dental x-ray images using an efficient encoder-decoder network,” *IEEE Access*, vol. 8, pp. 207822–207833, 2020, doi: 10.1109/ACCESS.2020.3037677.

[9] E. Shaheen *et al.*, “A novel deep learning system for multi-class tooth segmentation and classification on cone beam computed tomography. A validation study,” *J. Dent.*, vol. 115, p. 103865, 2021, doi: https://doi.org/10.1016/j.jdent.2021.103865.

[10] E. Kaya, H. G. Gunec, S. S. Gokyay, S. Kutal, S. Gulum, and H. F. Ates, “Proposing a CNN Method for Primary and Permanent Tooth Detection and Enumeration on Pediatric Dental Radiographs,” *J. Clin. Pediatr. Dent.*, vol. 46, no. 4, pp. 293–298, 2022, doi: 10.22514/1053-4625-46.4.6.

[11] S. HELLİ and A. HAMAMCI, “Tooth Instance Segmentation on Panoramic Dental Radiographs Using U-Nets and Morphological Processing,” *Düzce Üniversitesi Bilim ve Teknol. Derg.*, vol. 10, no. 1, pp. 39–50, 2022, doi: 10.29130/dubited.950568.

[12] Z. Chen, S. Chen, and F. Hu, “CTA-UNet: CNN-transformer architecture UNet for dental CBCT images segmentation,” *Phys. Med. Biol.*, vol. 68, no. 17, pp. 0–13, 2023, doi: 10.1088/1361-6560/acf026.

[13] T. H. Farook, F. H. Saad, S. Ahmed, and J. Dudley, “Clinical Annotation and Segmentation Tool (CAST) Implementation for Dental Diagnostics,” *Cureus*, vol. 15, no. 11, 2023, doi: 10.7759/cureus.48734.

[14] Deepho, Chutamas, et al. "Toward the Development of an Oral-diagnosis Framework: A Case Study of Teeth Segmentation and Numbering in Bitewing Radiographs via YOLO Models." *2024 IEEE International Conference on Cybernetics and Innovations (ICCI)*. IEEE, 2024.

[15] M. K. Dhar, M. Deb, D. Madhab, and Z. Yu, “A Deep Learning Approach to Teeth Segmentation and Orientation from Panoramic X-rays,” 2023, [Online]. Available: https://github.com/mrinal054/Instance_teeth_segmentation.

[16] 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.

[17] Ali, Ali Abdullah, Mohammed Khaleel Hussein, and Mohammed Ahmed Subhi. "A Classifier-Driven Deep Learning Clustering Approach to Enhance Data Collection in MANETs." Mesopotamian Journal of CyberSecurity 4.3 (2024): 36-45.

[18] W. K. Jummar, A. M. Sagheer, and H. M. Saleh, “Authentication System Based on Fingerprint Using a New Technique for ROI selection ”, Babylonian Journal of Artificial Intelligence, vol. 2024, pp. 102–117, Aug. 2024.

[19] H. M. S. SALEEH, H. Marouane, and A. Fakhfakh, “A Novel Deep Learning Approach for Detecting Types of Attacks in the NSL-KDD Dataset”, BJN, vol. 2024, pp. 171–181, Sep. 2024.

[20] O. M. Hammad, I. Smaoui, A. Fakhfakh, and M. M. Hashim, “Recent advances in digital image masking techniques Future challenges and trends: a review”, *SHIFRA*, vol. 2024, pp. 67–73, May 2024, doi: 10.70470/SHIFRA/2024/008.