

Panoramic Dental X-Ray Restorative Elements Segmentation using Hybrid Deep Learning

MOUSA MOHAMMED KHUBRANI, FATHE JERIBI*, ALI TAHIR,
ABDULNASSER ABDULWAKIL METWALLY
College of Engineering and Computer Science,
Jazan University,
Jazan 45142,
SAUDI ARABIA

**Corresponding Author*

Abstract: - Panoramic radiography is a commonly used imaging technique for dental X-rays, it is used as a diagnostics tool in dentistry. The study introduced a hybrid deep learning approach for detecting and segmenting dental restorative elements from panoramic dental X-rays. By integrating the You Look Only Once (YOLO v8) model for object detection and the Segment Anything Model (SAM) for segmentation, the aim is to enhance the identification of different dental restorative elements such as dental implants, crowns, fillings, and root canals. The datasets of the study comprised 1290 dental X-ray images. The YOLO model effectively recognizes regions of interest and generates bounding boxes and then for achieving precise segmentation SAM is utilized. The results demonstrate high accuracy for classification rates of 95% for fillings, 88% for crowns, 93% for root canals, and 97% for implants and the Intersection over Union (IoU) metrics results also improve systems accuracy. The results show significant improvement in accuracy and highlight the effectiveness of the hybrid approach in refining diagnostic precision and enhancing efficiency in dental imaging.

Key-Words: - Segmentation, Dental X-rays, YOLO, SAM, Hybrid model, Artificial Intelligence, Panoramic Radiography.

Received: April 12, 2024. Revised: November 5, 2024. Accepted: December 4, 2024. Published: December 31, 2024.

1 Introduction

Panoramic radiography is a commonly used imaging technique for dental X-rays, [1]. Dentists rely on panoramic dental X-rays to visualize the oral health of the patient. It can identify dental diseases including caries, cavities, gum diseases, restorations, implants, broken teeth, and oral cancer using panoramic X-rays, [2]. Without using the X-rays technique, the dentist cannot identify the problem until the disease reaches the severe stage. Clinical practicians examine dental X-rays for identifying and classifying tooth diseases, for detecting anatomical features, and segmenting tooth structures, [3].

Recent advancements in deep learning show significant promise in improving the accuracy and efficiency of medical image analysis. Tooth caries detection from panoramic dental X-ray improved classification over traditional rule-based systems and manual diagnosis by doctors, [4]. Utilizing Radio Visiography x-ray images for accurately classifying dental diseases while achieving promising accuracy, [5].

Accurate segmentation is essential for distinguishing between natural tooth structures and restorative materials, which is crucial for effective treatment planning and assessment. Segmentation of an image is a computer vision technique that basically involves the partitioning of an image into distinct regions or segments, each corresponding to different objects. Segmentation of an image gives more precision for diagnostic radiographs such as X-rays, Cone Beam Computer Tomography (CBCT), and Multi-Slice Computer Tomography (MSCT), [6], [7].

In the field of computing, Artificial intelligence mimics human intelligence in terms of reasoning, learning, language understanding, and problem-solving. In the past few years, AI made significant progress in enabling machines to automatically analyze and sort complex data, [8]. AI has significantly transformed dental tasks by enhancing the speed and accuracy of tasks, also reducing errors as compared to human performance, [9]. The deep learning architectures are specialized artificial neural networks that outperform for learning in hierarchical data representations. The deep learning technique is proficient in performing tasks such as prediction, object detection, and classification. The Convolutional Neural Networks (CNN) are efficient for image classification tasks because of their ability to recognize complex. The CNN models learn directly from the data and identify patterns for classification images avoiding the necessity of manually extracting features.

The aim of the study is to propose a hybrid deep learning approach for identifying different dental restorative elements such as root canals, dental implants, filling, and crowns in panoramic 2D dental X-rays.

2 Literature Review

In previous studies, a deep-learning algorithm has been developed for classifying dental implants using approximately 156,000 radiographic images. The proposed algorithm achieved an accuracy of 88.53% and gave reliable performance with both panoramic and peripheral images, [10].

This research focus on identification of different dental implant brands from panoramic radiographs. 25 deep models such as ResNet-50, EfficientNet, VGG16, and ConvNeXt are employed and compare on the base of accuracy. The ConvNeXt model used achieve a highest accuracy of 95.74% with strong precision while also maintaining efficiency with fewer parameters, [11].

In study is use for classification of tooth by using panoramic radiographs. Deep learning models YOLO-V4 and R-CNN are employ and compare. Results show that YOLO-V4 has superior performance over R-CNN achieving 99.90 accuracy and 99.18% recall along with a faster detection rate and superior performance for the identification of impacted third molars. By improving accuracy and efficiency, YOLO-V4 enhances the diagnosis and classification of teeth, [12].

A study was introduced for the classification of dental implant brands and treatment stages using a multi-task Deep Learning Model. This research utilizes dental panoramic radiograph images gathered from Kagawa Prefectural Central Hospital, Japan. This dataset contains 9,767 images containing 12 different implant brands treatment. Five deep learning models such as CNN model ResNet18, ResNet34, ResNet50, ResNet101, and ResNet152 are evaluated on this dataset. The findings show that the multi-task learning model offers high validity, for detecting different dental implant brands and treatment stages in dentistry and improving overall analysis accuracy, [13].

Gil Jader et all introduced an approach using Mask Region-Based CNN and transfer learning. Their model can accurately segment each tooth in panoramic X-rays. The model trained on 193 annotated images of 32 teeth, achieved 98% accuracy and 94% precision and also obtained 84% recall and 99% specificity on 1,224 images not previously encountered, [14].

The study proposed an approach to predict dental decay. Three different types of dental decay treatments such as fluoride, filling, and root canal are predicted in this research from X-ray images. 200 X-ray images of dental patients were utilized to train the model and achieved an overall accuracy of 87%, with the highest accuracy for fluoride treatment (98%), followed by root canal (88%), and the lowest for filling treatment (77%), [15].

In a study, a deep learning-based method is presented for computer-aided dental caries diagnosis. The dataset used in this study consists of 533 panoramic dental X-ray images, 229 caries teeth images, and 304 non-caries teeth images. The proposed approach used a dual approach for feature extraction, combining deep learning and mathematical modeling. Then five different classification models Naïve Bayes, SVM, KNN, Decision Tree, and Random Forest are employed for the detection of caries. The study results show 91.70% accuracy, 90.43% sensitivity, and 92.67% specificity, [16].

Another study is proposed to recognize periodontitis and dental caries. Pre-trained deep model EfficientNet-B0 model is used for the classification of periodontitis and dental caries in dental X-rays.YOLOv7 is applied for single-tooth image detection from dental X-ray. The dataset used in this study comprised of 1525 periapical dental X-rays. Results are evaluated using the area under the curve and Receiver Operating Characteristic curve which are 98.67% and 98.31% for recognition of periodontitis and dental caries, [17].

In this research, a deep learning model is presented to identify and detect dental implant manufacturers and kinds. The dataset consisted of 1,574 panoramic radiographs of which 3,675 were implant radiographs. The YOLOv7 model employed in the study identifies the implant locations and classifies the manufacturers. Subsequently, the EfficientNet model was employed to classify various dental implant types. YOLOv7 outperformed and achieved exceptional performance, with 1.00 recall, 0.979 precision, and 0.989 of an F1 score for implant detection. EfficientNet is also performed well, with classification metrics exceeding 0.92, [18].

In previous studies, the segmentation technique is used rarely and is not used in combination with the YOLO model. Our research combines YOLOv8 and SAM models for dental restorative element detection.

3 Proposed Work

The study introduced a hybrid approach by utilizing the You Only Look Once (YOLO v8) model for dental restorative elements detection and segmentation using the Segment Anything Model (SAM). The process involves detecting dental restorative elements in X-rays using YOLO v8 and refining the detection with detailed segmentation through SAM. The final output is an image where the accurate boundaries are detected for restorative elements which offer specific localization and comprehensive segmentation. The proposed methodology diagram is shown in Figure 1 (Appendix).

3.1 YOLO Model for ROI Detection

YOLO is an object detection model; it is prominent due to its speed and accuracy. YOLO excels in high-speed performance, it treats object detection as a single regression problem and the model predicts all important information (class labels, bounding box coordinates) in a single step. YOLO is more efficient as compared to the other conventional object detection models that apply a sliding window over the image. YOLO model enables real-time processing of dental X-rays with minimal latency, making it highly suitable for clinical settings where quick decision-making is crucial. The YOLO bounding box is a rectangular box, which is used to locate the position of an object in an image. YOLO breaks the input image into a grid and allocates each grid cell the task of prediction for objects whose centers are located within the cell.

YOLO splits the input image into a S X S grid. Every grid cell is accountable for detecting bounding boxes and their corresponding score. For every bounding box, YOLO has five parameters in which, (x,y) shows center coordinates for the bounding box, w and h are the height and weight of the bounding box and c is the confidence score, which represents what is the probability of object lies inside the bounding box.

YOLO predicts probability distribution over C classes for every grid cell.

$$\mathrm{C} = P(\text{ object }) \times \mathrm{IoU}_{\text{pred}}^{\text{truth}} \times P(\text{ class } \mid \text{ object }) \quad (1)$$

Here P(object) is the probability of an object lying in a bounding box, $\mathrm{IoU}^{\text{truth}}{}_{\text{pred}}$ represents Intersection over union relative to the ground truth and the predicted bounding box and P (" class "|" object ") shows the probability of the detected object fitting into a specific class. This process allows YOLO v8 to perform more effectively for dental implant detection.

3.2 Bounding Boxes Extraction

Following the YOLO processing of the input image, it generates the output of multiple bounding boxes in conjunction with their confidence score and class labels. For further processing extract these bounding boxes, utilizing Non-Maximum Suppression (NMS) which is used to handle multiple overlapping boxes for the same dental restorative elements, retaining only those boxes that have the highest score. The final bounding boxes are extracted and presented as the output of the YOLO model, with their class labels and corresponding scores.

3.3 Segment Restorative Elements

The next step of the proposed methodology is the segmentation of dental X-rays, utilizing the SAM model which gives high-precision accuracy, detailed segmentation of objects within the image, and provides flexibility with input prompts such as including points, bounding boxes, and masks. Bounding boxes extracted from the YOLO model are used as input for the SAM model.

In the SAM model, segmentation encompasses the process of assigning labels to each pixel in an image, and every pixel with the same label relates to the same object.

The segmentation process can be modeled by a function S which transforms the input image and bounding box into a corresponding segmentation mask.

$$\mathcal{S}: (I, B_i) \rightarrow M_i \quad (2)$$

Here let I donate the input image, let $B_i=\{B_1,B_2,........,B_n\}$ that represents bounding boxes from the YOLO model. For every bounding box $B_i \in B_1$ the SAM model evaluates segmentation masks M_i, where M_i is the binary mask for the same size of the bounded box B_i. $M_i(x,y)=1$, it represents the pixel (x,y) related to the object lies in the bounding box Bi, and when Mi(x,y)=0, it represents the pixel (x,y) is not related to the object lies in the bounding box Bi. The binary mask M_i helps to isolate objects in an image.

SAM uses deep learning techniques, creating binary masks and precisely depicting the boundaries of the objects in an image. SAM describes segmentation mask Mi for each bounding box given by YOLO as a prompt for SAM and these masks capture accurate shape and boundaries of object which lies in the bounding boxes.

3.4 Overlaying Segmentation Masks on the Original Image

The final step of the methodology is to overlap the segmented mask constructed by SAM with the original image. This step reveals the accurate boundaries of the detected objects.

The final output image I_{final} is constructed by:

$$I_{\text{final}}(x,y) = \begin{cases} I(x,y) & \text{if } M_i(x,y) = 0 \text{ for all } i \\ \alpha \cdot M_i(x,y) + (1-\alpha) \cdot I(x,y) & \text{if } M_i(x,y) = 1 \text{ for some } i \end{cases} \quad (3)$$

Here α is the blending factor that shows the transparency of mask overlay, which controls the visibility of the segmentation masks relative to the original image. I(x,y) represents pixel value at position (x,y) in the original position. M_i (x,y) denotes the pixel values at (x,y) position in the segmentation mask M_i. When M_i (x,y)=0" ", it indicates the pixel does not belong to any object and this shows the final output image retains the original pixel value I(x, y), and M_i(x,y)=1 it represents a pixel (x,y) is part of an object according to the segmentation mask M_i, the final pixel value is a blend of the mask and the original image, adjusted by the blending factor α, which means it shows the final output image blending the SAM object mask with the original image. The final output image I_{final} provides a visual representation for different dental restorative elements.

4 Results and Experiments

This section discussed the efficiency evaluation of the proposed method, emphasizing both quantitative metrics and qualitative observations. The dataset used for this study consists of panoramic X-ray images, obtained from publically available datasets, [19]. The dataset is made by only taking filling, crown, implant, and root canal X-ray images. The study conducted multiple experiments to evaluate the performance of the proposed method. The result of different experiments is evaluated by different evaluation metrics, IoU is used for segmented results, and precision, recall, and F1 score are utilized for classification.

In the first experiment, 80% of label data is used for training purposes and 20% of the data is used for validation and testing. The results of the YOLO model show positive trends for both testing and validation over time, with decreasing losses and enhancing precision, recall, and Mean Average Precision (mAP) metrics. The training graph is shown in Figure 2 (Appendix).

The second experiment is focused on detecting and classifying dental implants, crowns, fillings, and root canals. The results are represented in terms of percentage for correct classifications and misclassifications for four different dental conditions. The accurate classification rates for filling, crown, root canal, and implant are 95%, 88%, 93%, and 97% respectively. Accuracies of restorative elements are described through the confusion matrix in Figure 3 (Appendix).

The third experiment is focused on measuring the average Intersection over Union (IoU) for each class. IoU is a metric that is used for object detection to estimate the accuracy of projected bounding boxes against the ground truth. The IoU values for filling, crowns, root canals, and implants are 95.98%, 83.21%, 90.21%, and 95.20% respectively as shown in Table 1.

Table 1. The average Intersection over Union (IoU) for each class

Class	IOU (%)	Precision (%)	Recall (%)	F1-Score (%)
Filling	95.98	96.5	94.7	95.59
Crown	83.21	85.3	81.5	83.36
Root Canal	90.21	91	89.4	90.19
Implant	95.2	96.1	94.2	95.14

In the fourth experiment, individual images are analyzed for the model's performance, while focusing on different metrics such as precision, recall, F1-score, and IoU. The results are shown in Table 2 for ten individual images, with each metric being calculated separately:

Table 2. The average Intersection over Union (IoU) for each class

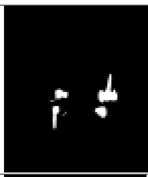
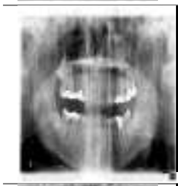
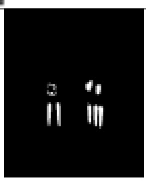
Images	Precision (%)	Recall (%)	F1-Score (%)	IoU (%)
Image1	96.0	97.0	96.5	92.8
Image2	91.0	93.0	92.0	85.8
Image3	92.0	90.0	91.0	84.2
Image4	89.0	87.0	88.0	80.5
Image5	88.0	90.0	89.0	81.2
Image6	87.0	85.0	86.0	79.3
Image7	90.0	88.0	89.0	82.3
Image8	94.0	95.0	94.5	89.8
Image9	93.0	91.0	92.0	85.3
Image10	95.0	96.0	95.5	91.5

The comparison of our proposed model with different existing base detection and segmentation models is shown in Table 3.

Table 3. Comparison with existing base models

Class	Model	Precision (%)	Recall (%)	F1-Score (%)	IoU (%)
Filling	ResNet	94.8	93.5	94.14	93
	UNet	95	93.5	94.24	94
	YOLO + SAM	96.5	94.7	95.59	95.98
Crown	ResNet	82.5	79.7	81.07	80.5
	UNet	84	80.5	82.15	82.1
	YOLO + SAM	85.3	81.5	83.36	83.21
Root Canal	ResNet	90.3	88.5	89.39	88.2
	UNet	90	88	89	89
	YOLO + SAM	91	89.4	90.19	90.21
Implant	ResNet	95	93.8	94.39	94.1
	UNet	95	94	94.67	94.5
	YOLO + SAM	96.1	94.2	95.14	95.2

Table 4. Visual comparison of segmented results of restorative elements with ground truth

Input image	Ground Truth	SAM Segmented	Class
			Filling/ Root Canal
			Filling
			Filling
			Implant /Filling
			Implant /Filling

This study presents a visual comparison of segmented restorative dental elements with their corresponding ground truth as shown in Table 4. It includes input images, ground truth annotations of the images, segmentation results from the SAM, and the classified restorative element types, such as fillings, root canals, and implants.

5 Conclusion

The study successfully demonstrates the efficiency and accuracy of a hybrid deep learning approach by integrating the YOLO and SAM models for the detection and segmentation of dental restorative elements in panoramic dental X-rays. The automated detection and segmentation process reduces the reliance on manual interpretation, minimizing human error and variability by reducing the need for manual interpretation. The proposed method can assist dentists in accurately identifying dental restorative elements, such as fillings, crowns, and implants, from X-rays. Also automate routine X-ray analyses in dental clinics, reducing the time dentists spend manually reviewing images The proposed hybrid model faced challenges, when the YOLO model fails to correct the localization of dental restorative elements, SAM also inherits these errors resulting in incorrect or incomplete segmentations. For future work, generative AI techniques can be explored to enhance the current method, which improves accuracy for rare conditions, also integrating this approach with 3D imaging modalities like Cone Beam Computed Tomography (CBCT) for more comprehensive diagnostic insights. Problems, leading to inaccurate segmentation. The proposed study can be extended to develop intelligent dental diagnostic systems by integrating Generative AI with dental X-ray analysis and a conversational dental chatbot. This dental chatbot provides explanations of detected issues, recommends treatment options, and answers patient queries in real-time, which enhances dental care.

References:

[1] F. Schwendicke, T. Golla, M. Dreher, and J. Krois, "Convolutional neural networks for dental image diagnostics: A scoping review," *Journal of Dentistry*, vol. 91, p. 103226, 2019. DOI: 10.1016/j.jdent.2019.103226.

[2] C. W. Wang, C. T. Huang, J. H. Lee, C. H. Li, S. W. Chang, M. J. Siao, and C. Lindner, "A benchmark for comparison of dental radiography analysis algorithms," *Medical*

Image Analysis, vol. 31, pp. 63-76, 2016. DOI: 10.1016/j.media.2016.02.004.

[3] G. Silva, L. Oliveira, and M. Pithon, "Automatic segmenting teeth in X-ray images: Trends, a novel data set, benchmarking and future perspectives," *Expert Systems with Applications*, vol. 107, pp. 15-31, 2018. DOI: 10.1016/j.eswa.2018.04.001.

[4] A. A. ALbahbah, H. M. El-Bakry, and S. Abd-Elgahany, "Detection of caries in panoramic dental X-ray images using back-propagation neural network," *International Journal of Electronics, Communication, and Computer Engineering*, vol. 7, no. 5, pp. 250-256, 2016.

[5] S. A. Prajapati, R. Nagaraj, and S. Mitra, "Classification of dental diseases using CNN and transfer learning," in *Proc. 2017 5th International Symposium on Computational and Business Intelligence (ISCBI)*, Dubai, UAE, Aug. 2017, pp. 70-74. DOI: 10.1109/ISCBI.2017.8053547.

[6] N. Patanachai, N. Covavisaruch, and C. Sinthanayothin, "Wavelet transformation for dental X-ray radiographs segmentation technique," in *Proc. 2010 Eighth Int. Conf. ICT Knowledge Eng.,* Nov. 2010, pp. 103-106. DOI: 10.1109/ICTKE.2010.5692904.

[7] Sepehrian, M., Deylami, A. M., & Zoroofi, R. A. (2013, December). Individual teeth segmentation in CBCT and MSCT dental images using watershed. In *2013 20th Iranian Conference on Biomedical Engineering (ICBME)* (pp. 27-30). IEEE. DOI: 10.1109/ICBME.2013.6782187.

[8] K. Suzuki, "Overview of deep learning in medical imaging," *Radiological Physics and Technology*, vol. 10, no. 3, pp. 257–273, Sep. 2017. DOI: 10.1007/s12194-017-0406-5.

[9] Y. W. Chen, K. Stanley, and W. Att, "Artificial intelligence in dentistry: current applications and future perspectives," *Quintessence International*, vol. 51, no. 3, pp. 248-257, 2020. DOI: 10.3290/j.qi.a43952.

[10] W. Park, J. K. Huh, and J. H. Lee, "Automated deep learning for classification of dental implant radiographs using a large multi-center dataset," *Scientific Reports*, vol. 13, no. 1, p. 4862, 2023.

[11] I. L. Kurtulus, M. Lubbad, O. M. D. Yilmaz, K. Kilic, D. Karaboga, A. Basturk, and I. Pacal, "A robust deep learning model for the classification of dental implant brands," *Journal of Stomatology, Oral and Maxillofacial Surgery*, vol. 125, p. 101818, 2024. DOI: 10.1016/j.jormas.2024.101818.

[12] S. Yilmaz, M. Tasyurek, M. Amuk, M. Celik, and E. M. Canger, "Developing deep learning methods for classification of teeth in dental panoramic radiography," *Oral Surgery, Oral Medicine, Oral Pathology and Oral Radiology*, vol. 138, no. 1, pp. 118-127, 2024. DOI: 10.1016/j.oooo.2023.02.021.

[13] S. Sukegawa, K. Yoshii, T. Hara, T. Matsuyama, K. Yamashita, K. Nakano, K. Taka-batake, H. Kawai, H. Nagatsuka, and Y. Furuki, "Multi-Task Deep Learning Model for Clas-sification of Dental Implant Brand and Treatment Stage Using Dental Panoramic Radio-graph Images," *Biomolecules,* vol. 11, no. 5, p. 815, 2021. DOI: 10.3390/biom11060815.

[14] G. Jader, J. Fontineli, M. Ruiz, K. Abdalla, M. Pithon, and L. Oliveira, "Deep instance segmentation of teeth in panoramic X-ray images," in *2018 31st SIBGRAPI Conference on Graphics, Patterns and Images (SIBGRAPI)*, Paraná, Brazil, Oct. 2018, pp. 400-407. DOI: 10.1109/SIBGRAPI.2018.00058.

[15] M. Bouchahma, S. B. Hammouda, S. Kouki, M. Alshemaili, and K. Samara, "An automatic dental decay treatment prediction using a deep convolutional neural network on X-ray images," in *2019 IEEE/ACS 16th International Conference on Computer Systems and Applications (AICCSA)*, Nov. 2019, pp. 1-4. DOI: 10.1109/AICCSA47632.2019.9035278.

[16] T. H. Bui, K. Hamamoto, and M. P. J. A. S. Paing, "Deep fusion feature extraction for caries detection on dental panoramic radiographs," *Applied Sciences*, vol. 11, no. 5, p. 2005, 2021. DOI: 10.3390/app11052005.

[17] I. D. S. Chen, C.-M. Yang, M.-J. Chen, M.-C. Chen, R.-M. Weng, and C.-H. Yeh, "Deep Learning-Based Recognition of Periodontitis and Dental Caries in Dental X-ray Images," *Bioengineering*, vol. 10, no. 8, Art. no. 911, Aug. 2023. DOI: 10.3390/bioengineering10080911.

[18] Y. Ariji et al., "Two-step deep learning models for detection and identification of the manufacturers and types of dental implants on panoramic radiographs," *Odontology*, Aug. 2024. DOI: 10.1007/s10266-024-00989-z.

[19] A. H. Abdi, S. Kasaei, and M. Mehdizadeh, "Automatic segmentation of mandible in panoramic x-ray," *Journal of Medical Imaging*, vol. 2, no. 4, pp. 1-9, 2015. DOI: 10.1117/1.JMI.2.4.044003.

APPENDIX

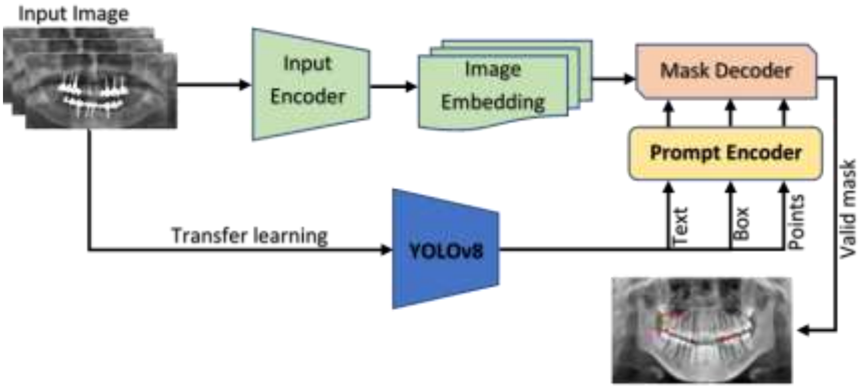


Fig. 1: Proposed system diagram of restorative elements detection and segmentation

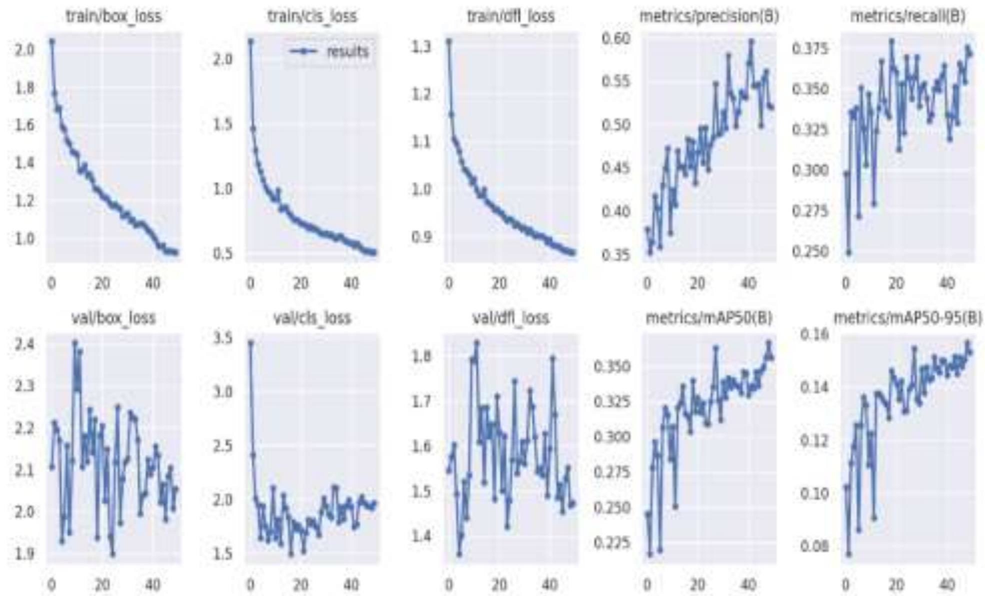


Fig. 2: Training graph for bounding box detection

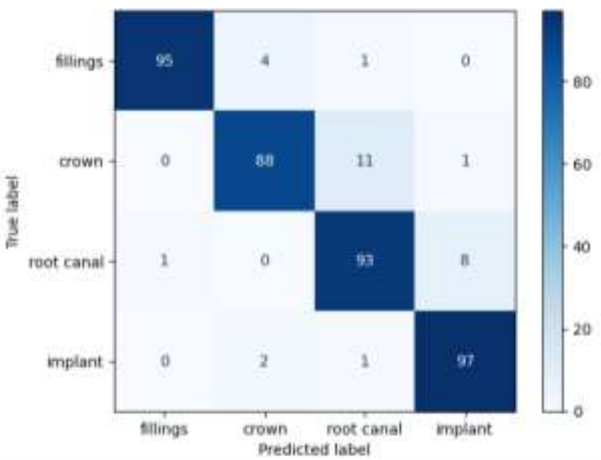


Fig. 3: Confusion matrix of restorative elements accuracy

Contribution of Individual Authors to the Creation of a Scientific Article (Ghostwriting Policy)

- Mousa Mohammed Khubrani provided the problem statement and proposed the model.
- Fathe Jeribi contributed to technical writing and analytical writing.
- Ali Tahir worked on implementation and result analysis.
- -Abdulnasser Abdulwakil Metwally performed technical proofreading and result gathering.

Sources of Funding for Research Presented in a Scientific Article or Scientific Article Itself

No funding was received for conducting this study.

Conflict of Interest

The authors have no conflicts of interest; the data is publicly available on internet.