Proposing a CNN Method for Primary and Permanent Tooth Detection and Enumeration on Pediatric Dental Radiographs

Emine Kaya*/Huseyin Gurkan Gunec**/Sitki Selcuk Gokyay***/Secilay Kutal****/
Semih Gulum*****/Hasan Fehmi Ates******

Objective: *In this paper, we aimed to evaluate the performance of a deep learning system for automated tooth detection and numbering on pediatric panoramic radiographs.* ***Study Design:*** *YOLO V4, a CNN (Convolutional Neural Networks) based object detection model was used for automated tooth detection and numbering. 4545 pediatric panoramic X-ray images, processed in labelImg, were trained and tested in the Yolo algorithm.* ***Results and Conclusions:*** *The model was successful in detecting and numbering both primary and permanent teeth on pediatric panoramic radiographs with the mean average precision (mAP) value of 92.22 %, mean average recall (mAR) value of 94.44% and weighted-F1 score of 0.91. The proposed CNN method yielded high and fast performance for automated tooth detection and numbering on pediatric panoramic radiographs. Automatic tooth detection could help dental practitioners to save time and also use it as a pre-processing tool for detection of dental pathologies.*

Keywords: *Deep learning; Tooth enumeration; Panoramic radiograph*

*Emine Kaya, Assistant Professor, Department of Pediatric Dentistry, Faculty of Dentistry, Istanbul Okan University, Istanbul, Turkey.
**Huseyin Gurkan Gunec, Assistant Professor, Department of Endodontics, Faculty of Dentistry, Atlas University, Istanbul, Turkey.
***Sitki Selcuk Gokyay, Research Assistant, Department of Endodontics, Faculty of Dentistry, Istanbul University, Istanbul, Turkey.
****Secilay Kutal, Undergraduate Student, Mechatronics Engineering, Faculty of Technology, Marmara University.
*****Semih Gulum, Undergraduate Student, Mechatronics Engineering, Faculty of Technology, Marmara University.
******Hasan Fehmi Ates, Professor , Computer Engineering, School of Engineering and Natural Sciences, Istanbul Medipol University.

Corresponding author:
Emine Kaya, Department of Pediatric Dentistry, Faculty of Dentistry, Istanbul Okan University, Tuzla/Istanbul, Turkey.
Phone: 00553 3583870
E-mail: eminetass@gmail.com

INTRODUCTION

Evaluation of dental radiographs is one of the most crucial parts of oral examination in clinical dental practice. Panoramic radiography has been used routinely for dental examination with the advantages like minimal patient discomfort, fast and easy application and limited radiation dose[1]. Interpretation of panoramic radiographs by a dental practitioner as an important step of the diagnosis involves teeth detection and numbering. Although correct detection and numbering of each tooth on panoramic images help dentists in making better diagnosis, the manual process is time-consuming and depends on the qualification of practitioners [2]. Automated tooth detection and numbering may help dentists perform more effective treatment options by reducing fatigue-related errors and saving time. Furthermore, automated tooth detection may also be useful to observe the positions of roots for orthodontic treatment [3] and to enhance dental forensic by identifying dental records [4].

Tooth identification is a basis of automated complicated detection systems that dental diseases are then determined and assigned to the identified tooth in a next step. Many researchers have investigated Artificial intelligence (AI) based methods for detection and numbering of teeth on panoramic radiographs [2,5-7]. Deep convolutional neural networks (CNNs) have demonstrated promising results for tooth detection and numbering [5-7]. CNNs, a standard architecture class for deep feedforward neural networks, are used for various computer vision tasks as a state-of-the-art approach [8]. There have been many kinds of neural networks such as region-based convolutional neural network (R-CNN)[9], faster region proposal with convolutional neural network features (fast R-CNN)[10] and faster R-CNN[11] developed for object detection and adopted for dental research successfully. Silva et al[12] introduced mask R-CNN and reported that although carrying out tooth detection is more difficult because of the characteristics of teeth on panoramic radiographs, CNN-based methods may ease this process. Recent studies have reported that the success level of CNN algorithms on tooth detection and numbering is close to the level of dental experts [2,5-7]. Although different CNN algorithms were evaluated, only adult panoramic radiographs were used in these studies. In the literature, there is only one study detecting and numbering the primary teeth on pediatric panoramic images [13]. To the best of our knowledge, this is the first study that a CNN algorithm was proposed for automated detection and numbering of both primary and permanent teeth on pediatric panoramic radiographs.

MATERIALS AND METHOD

Image Data Set

A total of 4545 panoramic radiographs of pediatric patients aged 5-13 were collected from Istanbul Okan University, Faculty of Dentistry for this study. No additional information, such as age or gender was revealed, since the radiographs were collected anonymously. Approval for this study was obtained from the institutional ethical committee (13.01.2021/132). Each panoramic radiograph was anonymized and saved as different image formats like .png, .bmp, .jpg etc. Annotation of each permanent and primary tooth in the maxillae and the mandible was manually performed by using labeling of bounding box. Detecting the location of a tooth was carried out by the bounding box and all teeth were labeled according to FDI tooth numbering system.

20 classes for primary teeth and 28 classes for permanent teeth were labeled on each image. We excluded 18,28,38 and 48 classes, since third molars are not erupted in this age range. Among 4545 radiographs, 62538 primary teeth and 43321 permanent teeth were labeled. (Figure 1)

The distribution of the number of data according to each of the 48 classes in the model was presented in Figure 2. General distribution of the data set was not regular. Although the irregularity of the distribution was a compelling factor for training the model, this situation was not considered as a disadvantage, since this irregular distribution may show similarity to the numbering distribution of the pediatric patients that the model encounters in real life.

Labeling Process

Labeling is the process carried out to make dataset meaningful for the predetermined objects for the machine learning algorithm. LabelImg program was used for labelling the permanent and primary teeth. The tooth to be labeled with LabelImg was selected as a rectangle and the name of the relevant class was assigned to represent the selected area. In this way, a label file containing various information such as the coordinates of all labeled areas on the images, the class name and the size of the image was saved. The format of the label files, which were originally kept in XML format, was later changed to TXT in order to be fed as an input to the model. On average, 24 teeth were labeled on a panoramic radiograph. The distribution of the number of labels on an image was presented in Figure 3.

Deep Learning Method

CNNs, one of the most popular architectures of deep learning, is commonly used for object recognition. Object detection techniques are classified as the one-stage detectors and the two-stage detectors. YOLO algorithm, as the most noticeable model of the one-stage detectors can detect and classify objects in a single image.[14] YOLO is a real-time object recognition algorithm that detects multiple objects and draws bounding boxes around each object to indicate the area of the detection.[15] We used YOLO V4 because of extreme speed and accuracy for object detection. The training data set with 4045 images was used to train the model, and a randomly selected testing group with 500 images was used to evaluate the performance of the model.

The images were resized as 608x608 pixels for the training of the model. The training was carried out on a server with Nvidia RTX2080 Ti (11 GB RAM) graphics card, 192 GB RAM and

Figure 1: Labelling permanent and primary teeth on labelImg

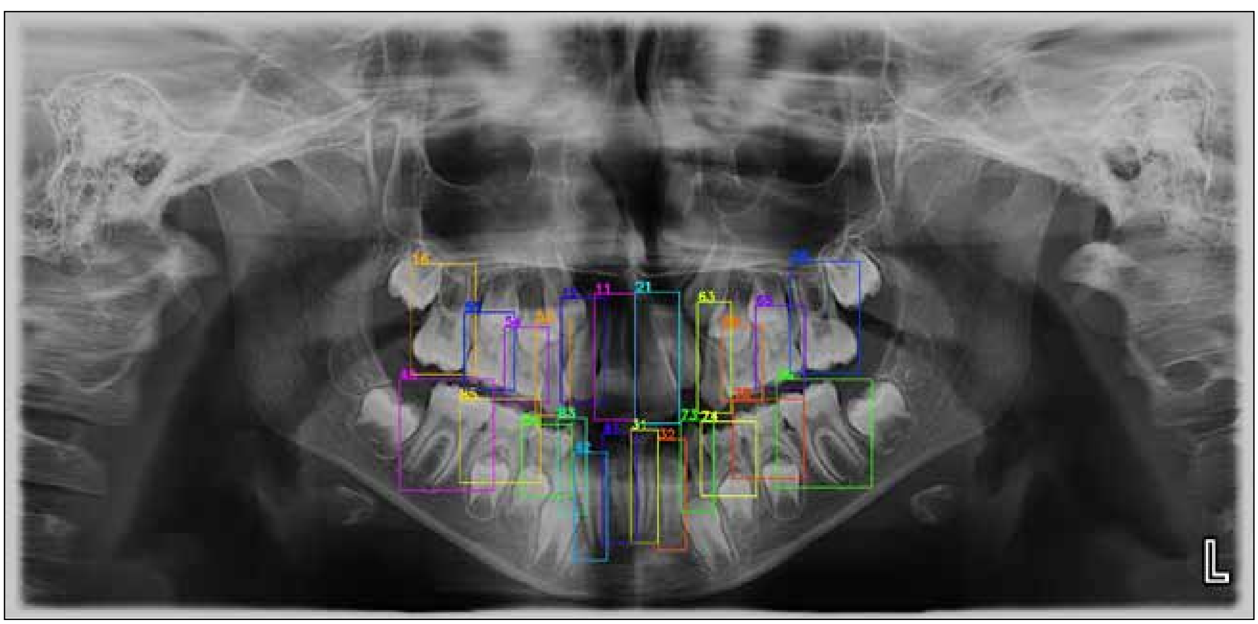


Figure 2: Distribution of the number of data by classes

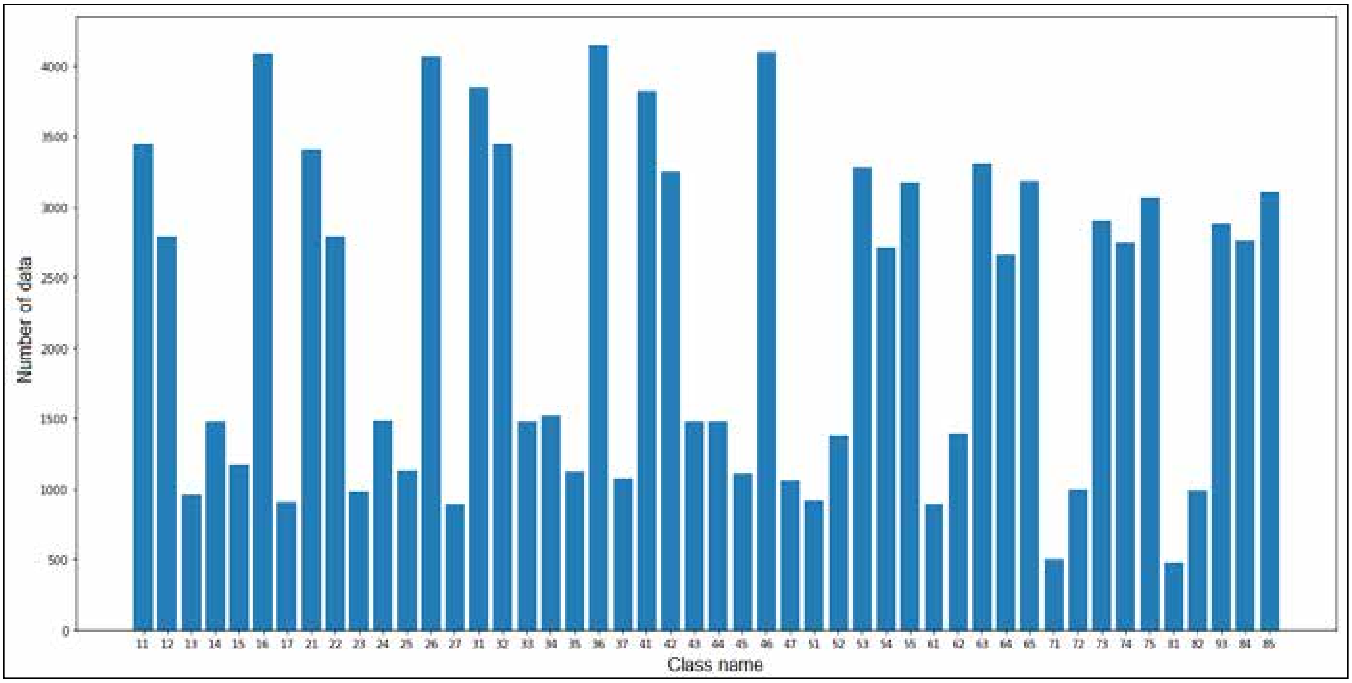


Figure 3: Distribution of number of labels on an image

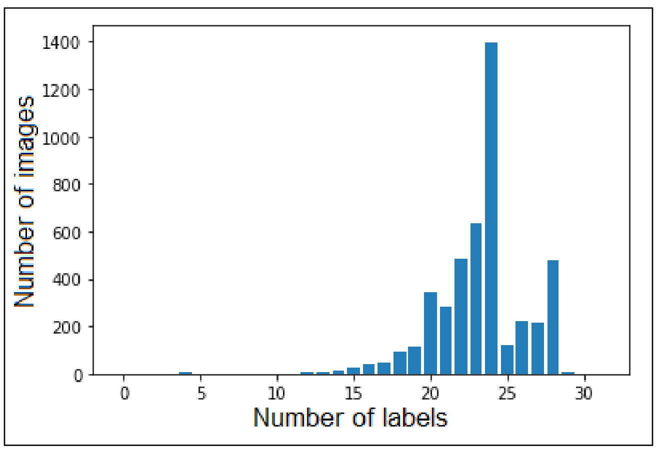


Table 1: Confusion Matrix.[22]

		Labeling	
Predicted labeling	Total	Positive Labeling	Negative Labeling
	Positive Prediction	True Positive (TP)	False Positive (FP)
	Negative Prediction	False Negative (FN)	True Negative (TN)

trained for 30 epochs. During the training process, the learning rate value was taken as 0.00261. While the batch was given as 16, the subdivision value was used as 4. Approximately 90% of the dataset was used in training and the remaining 10% was reserved for testing the performance of the model. Results of permanent and primary tooth detection and numbering in YOLO V4 were presented in figure 4 a-b.

Performance Metrics

The detection and classification accuracy of YOLO V4 model is based on three standard metrics. These metrics are: true positive (TP), false positive (FP), and false negative (FN). As indicated in Table 1, a correctly detected and numbered tooth indicates a TP; a correctly detected but incorrectly numbered tooth is FP; a missing detection indicates a FN.

Precision (Equation 2) represents the percentage of all positive detections that are truly positive. Recall (Equation 3) is the percentage of positive detections among all samples of the given class. F1 score (Equation 4) is a metric calculated by taking the harmonic average of precision and recall values, which is an alternative metric preferred to show the performance in unevenly distributed datasets.

Average precision (Equation 5) value, which is another metric included in the study during the evaluation phase of the model, is basically calculated by the area under the recall-precision curve. The mAP value (Equation 6) is a metric obtained by averaging this calculated class-based value.

The object detection model produces per-class probabilities for each detected tooth. Precision and recall values can be adjusted by thresholding these output probabilities. Detections with probabilities above the threshold are accepted as true detections, and detections with probabilities below the threshold are accepted as false detections. A low threshold generates too many detections, which is likely to increase recall but reduce precision. By changing the threshold value, a precision- recall curve can be generated to highlight the trade-off between the two metrics.

In object detection, average precision (AP) is commonly used as the evaluation metric. AP (Equation 5) value, is basically calculated from the area under the precision- recall curve. The mean AP (mAP) value (Equation 6) is a metric obtained by averaging these calculated class-based AP values over all distinct object classes.

 doi 10.22514/1053-4625-46.4.6

Figure 4-a: Results of permanent and primary tooth detection and numbering in YOLO V4. Numbering outputs are written based on FDI notation.

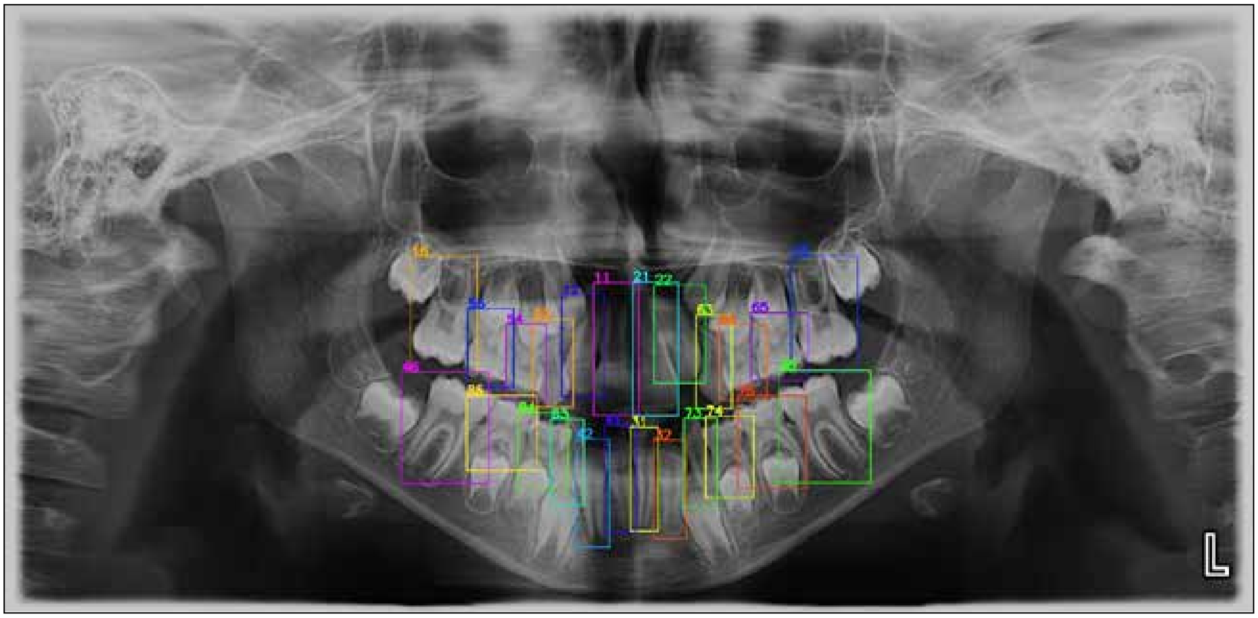
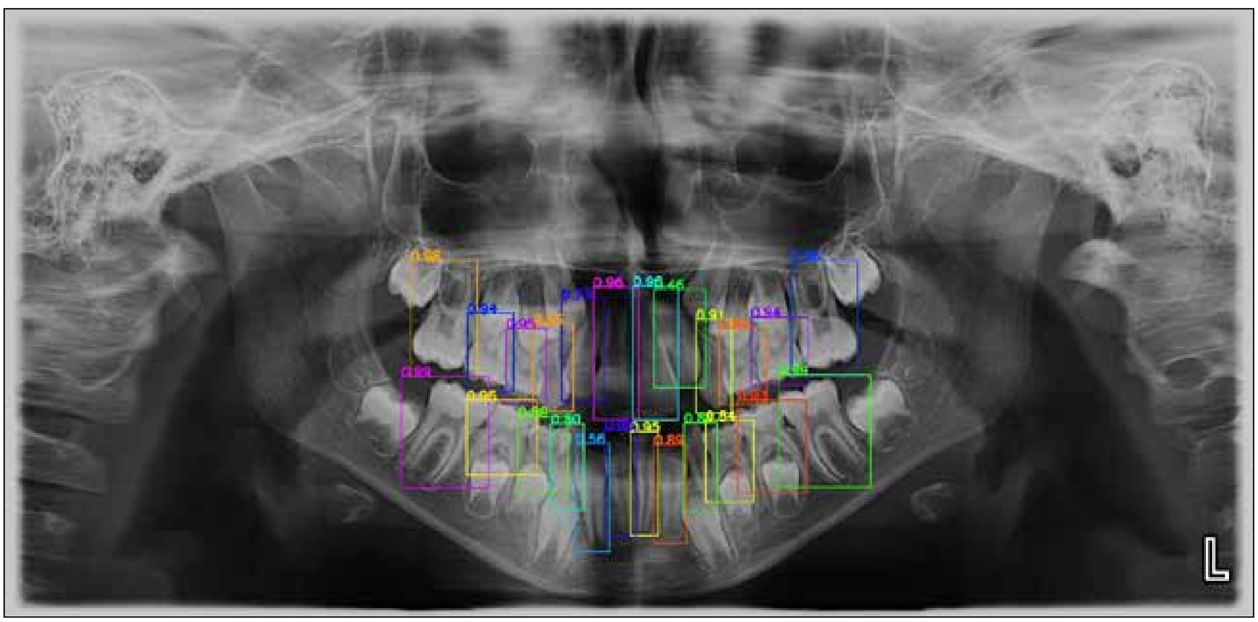


Figure 4-b: Results of permanent and primary tooth detection and numbering in YOLO V4. Detection confidence score is provided on top of each detected tooth.



$$\text{Accuracy}: \frac{TP+TN}{TP+TN+FP+FN} \quad (1)$$

$$\text{Precision}: \frac{TP}{TP+FP} \quad (2)$$

$$\text{Recall}: \frac{TP}{TP+FN} \quad (3)$$

$$\text{F1 Score}: 2.\frac{Precision.Recall}{Precision+Recall} \quad (4)$$

$$\text{AP(average precision) Value}: \int_0^1 P(R)dR \quad (5)$$

$$\text{mAP(mean average precision) Value}: \frac{1}{n}\sum_{i=1}^{n} AP_i \quad (6)$$

When evaluating the detections of the model, the bounding boxes of generated detections are compared against the bounding boxes of ground truth samples in the test images. Intersection over union (IoU) metric is computed by the ratio of the area of intersection area of the two bounding boxes to the area of their unions. The higher the IoU the more precise the detection is. Typically, IoU>50% is required to accept a detection as a positive detection. mAP @ IoU=0.50 means that mean average precision is computed under the requirement of a minimum 50% IoU for the detected bounding boxes. In this present study, mean (AP, AR) at 0.50 IoU: 92.22%, 94.44% were calculated. TP, FP, FN, Precision, Recall and F1 Score for each tooth were presented in table 2.

DISCUSSION

Tooth detection and enumeration, providing basis for the automatic diagnosis of dental diseases on radiographs, is the first step of radiological examination. Although tooth enumeration on radiographs is challenging to make into a learning task, automation of tooth detection and enumeration may make easier the daily practice of dentists [2]. Most research on deep learning performance of dental diagnosis has used intraoral radiography that provides more detailed information about the relevant region. [16-19]. However, panoramic radiography has gain of high research interest due to the allowing a single annotated image that include all teeth, vast number of anatomical structures and possible pathologies [20]. On the other hand, positioning artifacts that are commonly occurred due to the superposing of anatomical structures and teeth on panoramic images of children prevents to obtain more detailed information. Therefore, the number of studies on automated tooth detection and numbering using pediatric panoramic images is insufficient [13].

Tuzoff *et al* [7] used Faster R-CNN architecture on 1352 randomly chosen panoramic radiographs of adults for tooth detection and numbering tasks and reported that the sensitivity and precision for the system and experts demonstrated similarity. The authors also recommended that these systems could be enhanced by applying additional techniques, such as advanced image augmentation, and using recent CNN algorithms. Similarly, Lee et al [5] proposed a R-CNN method for tooth enumeration using individual annotation on 30 panoramic radiographs of adults and F1 score of 0.875 (precision: 0.858, recall: 0.893) and a mean IoU of 0.877 were obtained for automated tooth enumeration. In this present study, weighted F1 score: 91.1% (precision:0.89, recall:0.90) at IoU=0.50 were obtained for automated tooth enumeration of primary and permanent teeth. However, the results cannot be directly compared because of the differences in the datasets and applied algorithms.

Table 2: TP, FP,FN, Precision, Recall and F1 score for each tooth.

Tooth Number (class names)	TP	FP	FN	Precision	Recall	F1 Score
11	359	8	18	0.98	0.95	0.97
12	266	26	23	0.91	0.92	0.92
13	78	12	9	0.87	0.90	0.88
14	132	23	14	0.85	0.90	0.88
15	105	12	10	0.90	0.91	0.91
16	428	13	15	0.97	0.97	0.97
17	77	10	19	0.89	0.80	0.84
21	355	13	12	0.96	0.97	0.97
22	264	20	31	0.93	0.89	0.91
23	85	16	9	0.84	0.90	0.87
24	134	24	12	0.85	0.92	0.88
25	106	12	6	0.90	0.95	0.92
26	419	20	17	0.95	0.96	0.96
27	77	13	18	0.86	0.81	0.83
31	337	61	85	0.85	0.80	0.82
32	335	34	38	0.91	0.90	0.90
33	127	19	14	0.87	0.90	0.89
34	133	11	11	0.92	0.92	0.92
35	99	5	9	0.95	0.92	0.93
36	440	20	17	0.96	0.96	0.96
37	93	19	14	0.83	0.87	0.85
41	297	50	120	0.86	0.71	0.78
42	285	52	58	0.85	0.83	0.84
43	126	10	12	0.93	0.91	0.92
44	133	14	7	0.90	0.95	0.93
45	103	10	8	0.91	0.93	0.92
46	436	15	12	0.97	0.97	0.97
47	95	17	9	0.85	0.91	0.88
51	91	13	9	0.88	0.91	0.89
52	157	29	17	0.84	0.90	0.87
53	357	39	27	0.90	0.93	0.92
54	293	44	22	0.87	0.93	0.90
55	356	25	8	0.93	0.98	0.96
61	94	19	8	0.83	0.92	0.87
62	155	16	12	0.91	0.93	0.92
63	363	37	16	0.91	0.96	0.93
64	297	38	22	0.89	0.93	0.91
65	348	21	8	0.94	0.98	0.96
71	43	15	20	0.74	0.68	0.71
72	96	30	19	0.76	0.83	0.80
73	325	32	21	0.91	0.94	0.92
74	303	22	21	0.93	0.94	0.93
75	352	21	12	0.94	0.97	0.96
81	36	11	25	0.77	0.59	0.67
82	100	31	17	0.76	0.85	0.81
83	315	56	26	0.85	0.92	0.88
84	318	29	14	0.92	0.96	0.94
85	338	19	16	0.95	0.95	0.95

Kilic *et al*[13] evaluated a Faster R-CNN method for detecting and numbering of primary teeth on 421 pediatric panoramic radiographs and reported that the sensitivity, precision, and F1 score were 0.9804, 0.9571, and 0.9686, respectively. Although this study showed similarity with our study due to the use of pediatric panoramic images and the high success performance of the CNN method on detecting and numbering according to the FDI notation, the number of images was low and only primary teeth were detected and numbered. This present study was the first to use a CNN algorithm for automated detection and numbering of both primary and permanent teeth on pediatric panoramic radiographs.

All the aforementioned studies used two stage detectors like Mask-RCNN[12], R-CNN[5] and Faster R-CNN[7,13] that achieved great accomplishments in object detection. Although the two-stage detectors are usually more precise, this approach is slower than the one-stage detectors and requires high computation time. YOLO is one example of the one stage detectors that is used for detection and classification of objects with extreme speed and high accuracy. Furthermore, the feature of YOLO as performing real time object detection with its overall good performance in various object classes average values distinguishes it from other CNN algorithms. [14] We used YOLO V4 for tooth detection and numbering because of its speed and high accuracy level for object detection. In a study of Yuksel *et al* [21], YOLO was used to detect five different dental therapy options and number the teeth according to the FDI notation on 1005 panoramic radiographs of adults. Although this study had a significant drawback as a small dataset, the model used for tooth numbering showed satisfactory results with an AP score of 89.1%. Similarly, the mean AP score of our model was 92.22% showing that the model was very successful in automated tooth enumeration. Since we used pediatric panoramic radiographs to train and test the model for primary and permanent tooth numbering and detection, it is not possible to make a full comparison with this study.

CONCLUSION

In this study, we proposed a deep learning algorithm for detecting and numbering the primary and permanent dentition on pediatric panoramic radiographs. Our results showed that the performance of the proposed model was fast and accurate. To the best of our knowledge, this is the first study that used a CNN algorithm for detecting and numbering both primary and permanent teeth on pediatric panoramic images. Combination of deep learning-based models with the practice of dental experts may provide better treatment outcomes and accurate diagnoses of diseases in less time.

REFERENCES

1. Choi JW. Assessment of panoramic radiography as a national oral examination tool: review of the literature. *Imaging science in dentistry.* 2011;41(1):1-6.
2. Leite AF, Gerven AV, Willems H, et al. Artificial intelligence-driven novel tool for tooth detection and segmentation on panoramic radiographs. *Clinical oral investigations.* 2021;25(4):2257-67.
3. Xie X, Wang L, Wang A. Artificial neural network modeling for deciding if extractions are necessary prior to orthodontic treatment. *The Angle orthodontist.* 2010;80(2):262-6.
4. De Tobel J, Radesh P, Vandermeulen D, Thevissen PW. An automated technique to stage lower third molar development on panoramic radiographs for age estimation: a pilot study. *The Journal of forensic odonto-stomatology.* 2017;35(2):42-54.
5. Lee JH, Han SS, Kim YH, Lee C, Kim I. Application of a fully deep convolutional neural network to the automation of tooth segmentation on panoramic radiographs. *Oral surgery, oral medicine, oral pathology and oral radiology.* 2020;129(6):635-42.
6. Mahdi FP, Motoki K, Kobashi S. Optimization technique combined with deep learning method for teeth recognition in dental panoramic radiographs. *Scientific reports.* 2020;10(1):19261.
7. Tuzoff DV, Tuzova LN, Bornstein MM, et al. Tooth detection and numbering in panoramic radiographs using convolutional neural networks. *Dento maxillo facial radiology.* 2019;48(4):20180051.
8. LeCun Y, Bengio Y, Hinton G. Deep learning. *Nature.* 2015;521(7553):436-444.
9. Girshick R, Donahue, J., Darrell, T. & Malik, J. Rich feature hierarchies for accurate object detection and semantic segmentation. *In Proceedings of the IEEE conference on computer vision and pattern recognition.* 2014; 580–587.
10. He K, Zhang, X., Ren, S. & Sun, J. Spatial pyramid pooling in deep convolutional networks for visual recognition. *IEEE transactions on pattern analysis and machine intelligence* 2015;37:1904–1916.
11. Ren S, He, K., Girshick, R. & Sun, J. Faster r-cnn: Towards real-time object detection with region proposal networks. *In Advances in neural information processing systems (NIPS).* 2015; 91–99 (2015).
12. Silva G, Oliveira, L., Pithon, M. . Automatic segmenting teeth in X-ray images: trends, a novel data set, benchmarking and future perspectives. . *Expert Syst Appl.* 2018;107:15-31.
13. Kilic MC, Bayrakdar IS, Celik O, et al. Artificial intelligence system for automatic deciduous tooth detection and numbering in panoramic radiographs. *Dento maxillo facial radiology.* 2021;20200172.
14. Redmon J, Farhadi, A. Yolov3: An incremental improvement. *arXiv preprint arXiv: 1804. 02767.* 2018.
15. Chung YL, Lin C.K. Application of a Model that Combines the YOLOv3 Object Detection Algorithm and Canny Edge Detection Algorithm to Detect Highway Accidents. *Symmetry* 2020;12(11):1875.
16. Chen H, Zhang K, Lyu P, et al. A deep learning approach to automatic teeth detection and numbering based on object detection in dental periapical films. *Scientific reports.* 2019;9(1):3840.
17. Ekert T, Krois J, Meinhold L, et al. Deep Learning for the Radiographic Detection of Apical Lesions. *Journal of endodontics.* 2019;45(7):917-922 e915.
18. Valizadeh S, Goodini M, Ehsani S, Mohseni H, Azimi F, Bakhshandeh H. Designing of a Computer Software for Detection of Approximal Caries in Posterior Teeth. *Iranian journal of radiology : a quarterly journal published by the Iranian Radiological Society.* 2015;12(4):e16242.
19. Yasa Y, Celik O, Bayrakdar IS, et al. An artificial intelligence proposal to automatic teeth detection and numbering in dental bite-wing radiographs. *Acta odontologica Scandinavica.* 2021;79(4):275-281.
20. Schwendicke F, Golla T, Dreher M, Krois J. Convolutional neural networks for dental image diagnostics: A scoping review. *Journal of dentistry.* 2019;91:103226.
21. Yuksel AE, Gultekin S, Simsar E, et al. Dental enumeration and multiple treatment detection on panoramic X-rays using deep learning. *Scientific reports.* 2021;11(1):12342.
22. Goutte CGE. A Probabilistic Interpretation of Precision, Recall and F-Score, with Implication for Evaluation. *Paper presented at the Proceedings of the 27th European conference on Advances in Information Retrieval Research.* 2005.