

# Segmentation on Panoramic Dental X- Ray Images Using U-Net

**Prof. Rakesh Pandey, Akash Reddy, Arya Waghmare, Avdhesh Kamojjalwar, Chirag Rathi,  
Devanshu Ghuse, Mansi Kuranjekar**

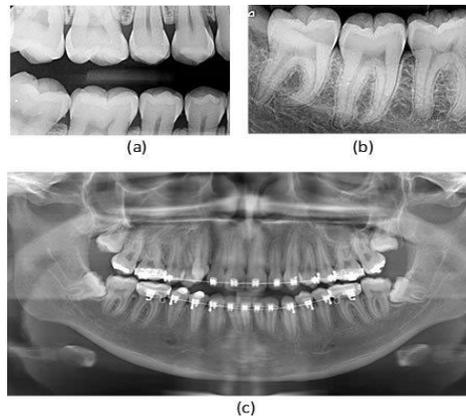
Department of Artificial Intelligence Engineering  
G. H Raisoni Institute of Engineering and Technology, Nagpur, India

**Abstract:** Radiological examinations in dentistry help professionals by targeting teeth for dental implants, bone defects, cysts, abscesses, infections, tumors, problems in the temporomandibular regions, to name a few some names, the skeletal system indicated. Although there is not yet a need for fully automated diagnostic tools, image pattern recognition has progressed towards decision support, primarily from the identification of teeth and their contents in X-ray images. Following a new direction, this paper proposes the investigation of a deep learning technique, such as tooth segmentation. To our knowledge, this is the first system to recognize and classify each tooth in panoramic X- ray images. It should be noted that this image is the most robust for dental separation because it shows other parts of the patient's body (e.g. mandible, pelvis, and jaw) We propose a classification system based on a mask area-based complex neural networks so do not pattern classification. Performance was evaluated from 1500 robust image datasets, which had high diversity and included 10 groups of different buccal images The proposed ancient system used only 193 facial images with an average of 1500 images. there are 32 teeth, using the transfer learning methods, we obtained 98% accuracy, 88 % F1-score, 94% accuracy; More than 1224 unrecognized images achieved 84% recall and 99% specificity, better results than the other 10 unsupervised methods

**Keywords:** Radiological examinations

## I. INTRODUCTION

From the X-ray images, dentists can evaluate the overall structure of the tooth and (if necessary) plan the patient's treatment. In fact, X- ray imaging is a tool used in dentistry to evaluate the structure of teeth, gums, and jawbones in the mouth to diagnose oral problems. The ray image is taken from inside the mouth), and extraoral, with the patient between the radiographic film and the x-ray source (the x-ray image is taken from the back of the patient's mouth). There are three types of dental x-rays in these two groups that are commonly used in dental examinations: Extraoral panoramic radiography - also known as panoramic x-ray or orthopantomography intraoral bitewing radiography - or bitewing Xray and periapical of the mouth



**Fig. 1.** Types of X-ray images: (a) Bitewing X- ray; (b) Periapical X-ray; (c) Panoramic X-ray

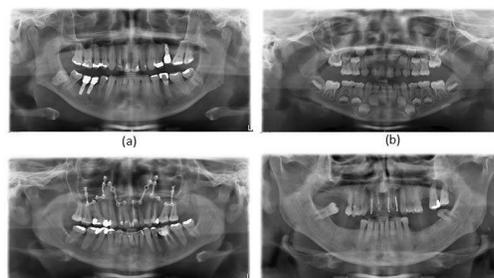
Radiography or periapical x-ray only. Figure 1 shows examples of these X-ray images. Panoramic x-rays are particularly useful examinations to aid the clinical diagnosis of dental diseases (caries or endodontic diseases) Such examinations may include dental and jaw irregularities, such as: appendages, bones internal defects, cysts, tumors, cancer, infection, fractures after accidents, temporomandibular joint disease of the ear, Areas causing pain in the face, neck and head. Dentists often request a closer look at oral and orthopedic surgeries in the temporomandibular region as a preoperative evaluation of the teeth

**A. Training Cover R-CNN**

It is not limited to isolated teeth, as is the case with endofacial radiographs, including the joints between the jaws, including details such as the skull and maxillary vertebrae arising from the bones of the jaw and facial regions. such as teeth types from patient to patient, restorative and prosthetic dental products, proximity areas of common interests, location of a missing tooth, and access limitations. Figure 2 shows some examples of these data. In short, the analysis of panoramic X-ray images depends on the careful work of an expert who does not have technical support. An automated classification method for filtering the features of spatial x-ray images could therefore be a starting point to help dentists diagnose their disease

**TABLE 1: CATEGORIZATION OF THE DATA SET IMAGES AND AVERAGE NUMBER OF TEETH PER CATEGORY**

Number	Category	Images	Average number of teeth
1	Images with <b>all the teeth</b> , containing teeth with restoration and with dental appliance	73	32
2	Images with <b>all the teeth</b> , containing teeth with restoration and without dental appliance	220	32
3	Images with <b>all the teeth</b> , containing teeth without restoration and with dental appliance	45	32
4	Images with <b>all the teeth</b> , containing teeth without restoration and without dental appliance	140	32
5	Images containing dental implant	120	18
6	Images containing more than 32 teeth	170	37
7	Images <b>missing teeth</b> , containing teeth with restoration and dental appliance	115	27
8	Images <b>missing teeth</b> , containing teeth with restoration and without dental appliance	457	29
9	Images <b>missing teeth</b> , containing teeth without restoration and with dental appliance	45	28
10	Images missing teeth, containing teeth without restoration and without dental appliance	115	28



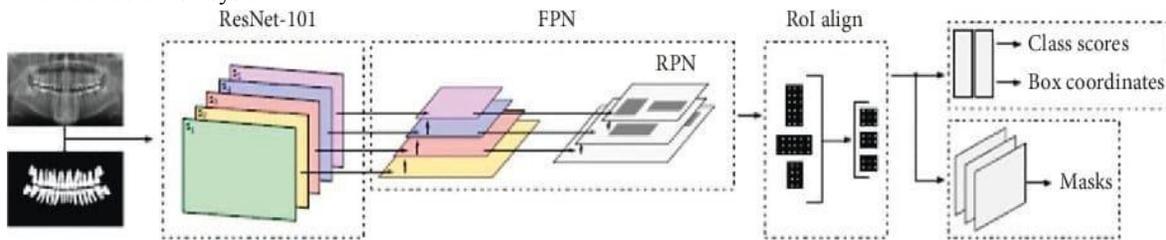
**Fig. 2.** Examples of some problems detected in panoramic X-ray images: (a) Dental implant, restored teeth, apex of some teeth; (b) supernumerary teeth; (c) devices for mandibular trauma; (d) missing and broken teeth

**B. Related work**

Unsupervised tooth segmentation in X- ray images: Until now, most dental image studies have relied on pixel-wise unsupervised tooth segmentation [3]. These works including bitewing images [2], [4-15], periapical [16-29 ], or panoramic [1], . The methods run in [30-37]. According to [3] , 80% of the studied projects use nasal images such as bitten or periapical wings for their tests. Some of these use more than one model in their experiments: Bitewing and periapical [38]; caught with birds [39]; periapical and panoramic [40]; touch, peripeak, and bird [41]. Most papers work with small data sets, typically from 1 to 100 images [1], [4-5], [8-9], [12], [14-17], [19- 20] . , [ 22-24 ] , [ 23 ] , [26-30], [32-34], [37], [39] and their associates. [38] was the only one [3] (out of 630 independent images) to make extensive use of images in the reviewed works. All the programs evaluated the performance of their proposed programs by considering small changes in the dataset.

Dental classification in panoramic images: As far as we know, our work is the first to use instance segmentation based on depth learning from panoramic X-ray images and the rest of the work follows the classification method a they do not support it afterwards, e.g. , based on the boundary

Other types of medical image classification: In, authors trained end-to-end convolutional networks to classify neuronal structures in electronic microcells In [43] authors proposed a spatially constrained convolutional neural network (SC-CNN) for identifying and classifying nuclei in routine colon cancer histology images from lung segmentation Working with cancer detection system



**Fig. 3.** Training process of the segmentation system. From left to right: X-ray images and annotation masks as inputs, ResNet101 backbone with 5-stage feature extractor (from S1 to S5), where the output of each ResNet stage, but S1, forms a layer in the feature pyramid network (FPN); anchors are determined over FPN, and regions of interest (RoI) are computed (defining the region proposal network (RPN)) and, finally, aligned (RoI aligned). Outputs are the class scores and box coordinates, given by full connected network, and masks, given by a fully convolutional network

**C. Contributions**

This study represents a significant advance in dental classification, especially in dental image analysis using deep learning techniques. Historical difficulties in accurately classifying teeth, combined with limitations of visibility and the severity of traditional images, have hindered progress in this field for a long time Work a previous studies have focused mainly on images of the posterior or anterior aspects, which often presented difficulties in adequately capturing the upper teeth

In contrast, this study is a pioneer in the explicit application of deep learning techniques, in particular, exploiting the capabilities of convolutional neural networks (CNNs), especially masked R-CNNs, to analyze X-1. spatial ray images thus extending the analysis beyond the limitations posed by traditional imaging Using spatial images combined with multiple perspectives allows for better identification of teeth with bones the surroundings interrupt less, making the classification system more accurate and robust

The dataset used in this study includes several methods, including 1500 images, which greatly contributes to the generalizability and applicability of the proposed method Through data properly collected and refined, the dataset facilitates advanced training of the deep learning mode. The highlight of this work was a remarkable performance with minimal training issues. Despite the small size of the training set, the deep connectivity-based masked R-CNN model showed exceptional accuracy, with 98% accuracy, 88% F1-score, and 94% accuracy. For example, a tooth fracture on x-rays. The objective of the object detection task is to locate and classify individual objects. The

goal of semantic classification is to classify each pixel of an object into known classes, without distinguishing between object instances. Instance segmentation combines these two classical computer vision tasks (recognition and semantic segmentation), where each detected object is classified, spatially, and segmented Here is our proposed algorithm using Mask R-CNN [1]. [47] is used, for example, to classify teeth in X-ray images.

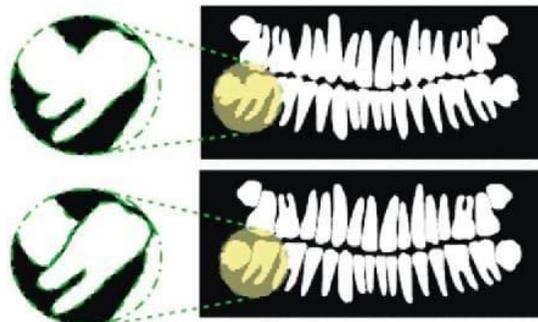
**Deep network architecture details**

Masked R-CNN architecture is shown in Figure 3. As an extension of the fast R-CNN [48], masked R-CNN incorporates a convolutional network branch to perform instance segmentation task Extracting features from ResNet101, these features compose a feature pyramid network (FPN), where finally anchors are defined and Regions of interest (RoIs) are selected These two steps (FPN + anchors) form the Region Proposal Network (RPN) introduced by [48]. The RoIs are subsequently sorted to achieve the same size. Finally, each object of a fixed size is: i) classified as a tooth or jaw (numerical square); ii) localized by regressing the bounding box coordinates; and iii) each pixel segmented by a fully convolutional network (FCN) [49] within each reported tooth boundary box (mask). Only 193 buccal images were annotated in our data set. This number of descriptors is insufficient to train them from scratch due to the large number of free parameters in the deep learning network. To overcome the lack of annotations, the pre-trained weights were taken from the MSCOCO dataset [46], which contains, for example, 80 annotated features for the classification task We include only the pre-trained weights were used on the Mask-RCNN network backbone (ResNet 101). Only the weights of the top layers (RPN, etc.) were initialized with our data set

**TABLE 2: SUMMARY OF THE QUANTITATIVE RESULTS.**

Category	Accuracy	Specificity	Precision	Recall	F1-score
#1 (73)	0.98	0.99	0.92	0.91	0.92
#2 (60)	0.98	0.99	0.94	0.91	0.92
#3 (2)	0.97	0.99	0.96	0.77	0.85
#4 (67)	0.98	0.99	0.96	0.89	0.93
#5 (120)	0.98	0.99	0.94	0.82	0.87
#6 (170)	0.97	0.99	0.94	0.83	0.88
#7 (115)	0.97	0.98	0.80	0.90	0.84
#8 (457)	0.97	0.99	0.97	0.80	0.87
#9 (45)	0.98	0.99	0.96	0.88	0.92
#10 (115)	0.99	0.99	0.93	0.89	0.91
<b>Weighted sum:</b>	1195.83	1215.73	1146.63	1032.57	1082.11
<b>Average±STD:</b>	0.98 ±0.008	0.99±0.006	0.94±0.06	0.84±0.07	0.88±0.05

Mask training hyperparameters (e.g., number of studies, number of epochs) were defined empirically using observed training and validation data In this step we divided 193 annotated images into two types for training segmentation networks, validate results, and tune hyperparameters , the results of which are described in Section III.



**Fig. 4.** Process of separating the teeth in the data set proposed by [3]. Zoomed tooth in the top is as it was in [3], and in the bottom is as it was changed here.

In our data set, the training step was performed in two steps. In the first step, the Adam optimizer [50] with  $\alpha = 10^{-3}$ ,  $\beta_1 = 0.9$ ,  $\beta_2 = 0.999$ , and  $\epsilon = 10^{-8}$  was used. The head layers were weighted by training with 100 epochs. The ADAM optimizer was used to obtain fast results from adjusting the weights for the new classification task. In the second stage, a stochastic gradient descent (SGD) optimizer was used to optimize the load, without specifying motion, with a learning rate of  $10^{-6}$ . In this last stage, weights of columns 4 and 5 are from ResNet 101 (see Fig. 3), as well as the head layers were considered as each part of ResNet 101 trained corresponds to a series of diffraction layers with the same feature map size. The training was performed with an error rate of 106. The mesh weights resulting from this training step were used in the analysis and comparison of the results with other methods.

## II. EXPERIMENTAL ANALYSIS

The references of the 276 images in the first four classes (1, 2, 3, 4) of the original dataset [3] were modified by splitting the teeth according to Figure 4. Those groups were chosen because they image varieties with 32 teeth gave, having more or less the expected positions; 193 images (se 6987) were used for training, while 83 images (se 3040) were used as validation sets for optimizing the deep mesh.

The method of splitting teeth in each image was developed only to train Mask R-CNN with more samples (now, the objects are teeth, not all the teeth) than the originally collected dataset. It was shown that this method is effective, even considering that 1224 test images were used as a dataset with annotations of all dental arches. After training Mask R-CNN with 6987 tooth images (from image 193), and optimizing the mesh parameters from 3040 tooth images (from image 83), Mask R-CNN (from image 83). about basic description use by).

### Quantitative analysis

The following metrics were used to evaluate the performance of the segmenter:

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

$\text{Accuracy} = \frac{TP}{(TP + FN)}$ ,  $\text{Accuracy} = \frac{TN}{(TN + FP)}$ ,  $\text{Precision} = \frac{TP}{(TP + FP)}$ ,  $\text{Recall} = \frac{TP}{(TP + FN)}$ , and  $F1\text{-score} = 2 * \text{Precision} * \text{Recall} / (\text{Precision} + \text{Recall})$ , where TP, TN, FN and FP represent true positive, true negative, false negative and false positive, respectively; in any case. These metrics were applied pixel-wise.

Following the method proposed by [3], weighted averages were calculated considering the number of images in each class (weighted), and the sum of the resulting metrics for each image, all divided by test over the number of images. Table II quantitatively summarizes the results obtained by our system. All standard deviations were found to be very small, indicating that all individual results were all close to the mean. This fact indicates that, although the data set was complex, the proposed algorithm achieves good generalization and consistency in the results.

Briefly, the results showed a good balance between true negative/false negative (precision, accuracy) and true positive/false positive (precision, accuracy and recall) rates, with ultimately an F1-score is verified, where harmonic means are calculated between Recall and accuracy.

Compared with the unsupervised methods investigated in [3], the R-CNN overlay showed the best results. Table III presents the results of the MASK R-CNN in comparison with the results of the unsupervised classification methods reviewed in [3]. Although the number of test images is different between the two classes, the differences (1500 and 1224) are not so large as to hinder the comparison. Table III highlights the best results from unsupervised and from MASK R-1. CNN results confirmed. Except in specific, which resulted in the deep learning approach approaching split/merge, when other metrics are considered, mask R-CNNs are clearly superior. Also, it is worth mentioning that no other unsupervised method presents all the metrics as consistently and robustly as the deep learning method does.

### B. Qualitative analysis

Figure 5 shows a mosaic of good results, while Figure 6 shows the worst results, in each metric. An image mosaic (see Fig. 5) is obtained by nearly perfect results in which the original X-ray images of each tooth are plotted from left to right columns, ground truth, tooth classification, and model classification, respectively, except for recall and F1 scores. For the worst classification models requiring higher levels (see Fig. 6), accuracy and specificity also reached higher values however, the F1 score, accuracy, and recall obtained yielded worse results. This result was expected due to artifacts not detected in the training set after visual inspection. This is similar to such images in rows two, three, and four (top down) in Figure 6: Note that two and four panoramic X-ray images are equivalent, where the two metrics, F1-score

and recall, failed to get a negative result, when in row three, we can see that one denture is counted as a tooth, while only one tooth is detected by Mask R-CNN (that because of the accuracy of 48%) with a relatively low standard deviation (see Table II) indicates that the bad separations were such they are rarely observed

### III. CONCLUSION

Dental classification on dental x-rays has been studied for many years, mainly relying on unsupervised methods. Although many methods had been proposed and tested, results were still distant. For more complex tasks in decision support systems, it is mandatory to classify teeth in oral images. This is a first step to detect not only teeth and their attachments, but also artifacts (e.g. dentures), dental problems, and missing teeth. Considering our proposed deep learning algorithm showed promising results on a complex dataset, future work on model classification of each aspect segment edges and teeth, and reports all these features with the aim of providing itself treatment result

**TABLE 3: COMPARISON OF THE UNSUPERVISED METHODS STUDIED IN [3] AND MASK R-CNN.**

Method	Accuracy	Specificity	Precision	Recall	F1-score
<b>Region growing [12]</b>	0.68	0.69	0.35	0.63	0.44
<b>Splitting/merging [51]</b>	0.81	0.99	0.81	0.08	0.14
<b>Global thresholding [21]</b>	0.79	0.81	0.52	0.69	0.56
<b>Niblack method [34]</b>	0.81	0.81	0.51	0.82	0.61
<b>Fuzzy C-means [30]</b>	0.82	0.91	0.61	0.45	0.49
<b>Canny [35]</b>	0.79	0.96	0.45	0.11	0.17
<b>Sobel [35]</b>	0.80	0.99	0.66	0.03	0.06
<b>Active contours without edges [13]</b>	0.80	0.85	0.51	0.57	0.52
<b>Level set method [5]</b>	0.76	0.78	0.48	0.68	0.52
<b>Watershed [14]</b>	0.77	0.75	0.48	0.82	0.58
<b>Mask R-CNN</b>	0.98	0.99	0.94	0.84	0.88

### REFERENCES

- [1]. Y. Y. Amer and M. J. Aqel, "An efficient segmentation algorithm for panoramic dental images," *Procedia Computer Science*, vol. 65, pp. 718–725, 2015.
- [2]. C. W. Wang, C. T. Huang, J. H. Lee, C. H. Li, S. W. Chang, M. J. Siao, T. M. Lai, B. Ibragimov, T. Vrtovec, O. Ronneberger, P. Fischer, T. F. Cootes, and C. Lindner, "A benchmark for comparison of dental radiography analysis algorithms," *Medical Image Analysis*, vol. 31, pp. 63–76, 2016.
- [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.
- [4]. A. K. Jain and H. Chen, "Matching of dental x-ray images for human identification," *Pattern Recognition*, vol. 37, no. 7, pp. 1519–1532, 2004.
- [5]. Ehsani Rad, M. S. Mohd Rahim, and A. Norouzi, "Digital Dental X-Ray Image Segmentation and Feature Extraction," *TELKOMNIKA Indonesian Journal of Electrical Engineering*, vol. 11, no. 6, pp. 3109–3114, 2013.
- [6]. N. Senthilkumaran, "Genetic Algorithm Approach to Edge Detection for Dental X-ray Image Segmentation," *International Journal of Advanced Research in Computer Science and Electronics Engineering*, vol. 1, no. 7, pp. 5236–5238, 2012.
- [7]. O. Nomir and M. Abdel-Mottaleb, "Fusion of matching algorithms for human identification using dental x-ray radiographs," *IEEE Transactions on Information Forensics and Security*, vol. 3, no. 2, pp. 223–233, 2008.
- [8]. P. L. Lin, Y. H. Lai, and P. W. Huang, "An effective classification and numbering system for dental bitewing radiographs using teeth region and contour information," *Pattern Recognition*, vol. 43, no. 4, pp. 1380–1392, 2010.
- [9]. P. L. Lin and Y. H. Lai, "Dental biometrics: Human identification based on teeth and dental works in bitewing radiographs," *Pattern Recognition*, vol. 45, no. 3, pp. 934–946, 2012.

- [10]. F. Keshkar and W. Gueaieb, "Segmentation of dental radiographs using a swarm intelligence approach," in Canadian Conference on Electrical and Computer Engineering, 2007, pp. 328–331.
- [11]. O. Nomir and M. Abdel-Mottaleb, "Hierarchical contour matching for dental X-ray radiographs," Pattern Recognition, vol. 41, no. 1, pp. 130–138, 2008.
- [12]. K. Modi and N. P. Desai, "A simple and novel algorithm for automatic selection of ROI for dental radiograph segmentation," in Canadian Conference on Electrical and Computer Engineering, 2011, pp. 000 504–000 507.
- [13]. R. B. Ali, R. Ejbali, and M. Zaied, "GPU- based Segmentation of Dental X-ray Images using Active Contours Without Edges," in International Conference on Intelligent Systems Design and Applications, vol. 1, 2015, pp. 505–510.
- [14]. S. Srinivasan, S. K. Sannapureddy, and S. J. H. M. M. Hattab, "Deep learning approach for automated tooth identification in dental panoramic radiographs," Journal of Imaging, vol. 6, no. 10, p. 110, 2020.
- [15]. A. Estai, L. Huang, M. C. Z. J. W. J. Aranda, and J. R. D. C. R. Z. L. C. E. Kruger, "Comparing the diagnostic outcomes of bacterial compositions from dental implant fixtures and teeth using pyrosequencing: a systematic review and meta-analysis," Journal of Periodontal Research, vol. 56, no. 4, pp. 614–629, 2021.
- [16]. E. T. Puertolas, L. Rodriguez, J. C. R. D. L. V. I. Y. A. M. L. M. A. J. A. D. N. M. Fernandez, "Effect of orthodontic treatment on periodontal health: a systematic review and meta-analysis," Journal of Clinical Periodontology, vol. 48, no. 2, pp. 246–262, 2021.
- [17]. Cardoso, L. O. F. C. E. A. L. C. M. A. Rodrigues, and R. F. C. A. S. M. D. R. F. A. C. L. B. P. Machado, "Machine learning approaches for caries detection and diagnosis in dental imaging: a systematic review," Journal of Dentistry, vol. 105, p. 103576, 2021.
- [18]. S. Ahmed, M. M. Islam, S. Paul, K. A. S. M. A. G. M. Shams, and M. R. D. R. S. Paul, "A deep learning approach for classification of dental periapical lesions on radiographs," Diagnostics, vol. 11, no. 3, p. 393, 2021.
- [19]. M. Nascimento, M. S. Machado, D. S. Silva, F. M. D. R. R. da Silva, J. F. R. da Silva, R. F. V. M. d. Oliveira, M. d. S. Leao, and S. F. F. L. C.
- [20]. D. B. H. Vasconcelos, "Accuracy of different methods for dental age estimation in Brazilian children and adolescents," International Journal of Legal Medicine, vol. 135, no. 1, pp. 151–160, 2021.
- [21]. N. J. da Costa, R. V. S. G. Goncalves, M. d. S. d. B. Tucunduva Neto, M. d. C. Z. K. A. D. T. M. Oliveira, A. G. d. Oliveira, and G. d. M. L. G. Silva, "Effect of orthodontic debonding and residual adhesive removal techniques on enamel surface roughness," Journal of Dentistry, vol. 98, p. 103403, 2020