

Article

Automatic Tooth Detection and Numbering Using a Combination of a CNN and Heuristic Algorithm

Changgyun Kim ^{1,†}, Donghyun Kim ^{2,†}, HoGul Jeong ², Suk-Ja Yoon ³ and Sekyoung Youm ^{1,*}¹ Department of Industrial and Systems Engineering, Dongguk University, Seoul 04620, Korea; tiocdrbs@dongguk.edu² Medipartner Co., Ltd., Seoul 06135, Korea; psds075@yonsei.ac.kr (D.K.); rari98@yonsei.ac.kr (H.J.)³ Department of Oral and Maxillofacial Radiology, School of Dentistry, Chonnam National University, Gwangju 61186, Korea; yoonfr@chonnam.ac.kr

* Correspondence: skyoum@dongguk.edu

† The authors are co-first authors.

Received: 17 July 2020; Accepted: 10 August 2020; Published: 13 August 2020



Abstract: Dental panoramic radiography (DPR) is a method commonly used in dentistry for patient diagnosis. This study presents a new technique that combines a regional convolutional neural network (RCNN), Single Shot Multibox Detector, and heuristic methods to detect and number the teeth and implants with only fixtures in a DPR image. This technology is highly significant in providing statistical information and personal identification based on DPR and separating the images of individual teeth, which serve as basic data for various DPR-based AI algorithms. As a result, the mAP(@IOU = 0.5) of the tooth, implant fixture, and crown detection using the RCNN algorithm were obtained at rates of 96.7%, 45.1%, and 60.9%, respectively. Further, the sensitivity, specificity, and accuracy of the tooth numbering algorithm using a convolutional neural network and heuristics were 84.2%, 75.5%, and 84.5%, respectively. Techniques to analyze DPR images, including implants and bridges, were developed, enabling the possibility of applying AI to orthodontic or implant DPR images of patients.

Keywords: tooth detection; tooth numbering; panoramic radiography; implant detection; radiology AI

1. Introduction

Dental panoramic radiography (DPR) is an examination that uses an extremely small dose of ionizing radiation to capture a single image of the entire mouth. This technique is commonly applied by dentists and oral surgeons in their everyday practice and has significant potential in the planning of treatment involving dentures, braces, extractions, and implants.

DPR is a commonly used imaging method applied to an overall evaluation of the jaw bones and teeth. Compared to intraoral radiographs, it has the advantages of a shorter imaging time, lower exposure dose, and the ability to approximate the real size and location of the major anatomical structures in the oral and maxillofacial region more accurately. Therefore, DPR not only complements an oral examination consisting of questionnaires and visual examinations, but also serves as an indicator of permanent and objective records of teeth and hard tissues [1].

Although DPR is part of a basic examination, only professionally trained doctors can read a DPR image. As a result, several attempts to aid in the diagnosis by an automatic reading of panoramic images have remained at the early stage of development owing to problems such as diagnostic accuracy [2].

The recent development of big data and cloud technology has vastly augmented the availability of learnable medical information and algorithms. Algorithms for panoramic images using artificial

intelligence (AI) have been recently applied in various fields, including maxillary bone, tooth age, and osteoporosis determination [3].

Tooth and implant detection and tooth numbering are fundamental concepts that have been used in various studies applying AI to the analysis of DPR images. Tooth detection involves determining whether a tooth is a prosthesis. Tooth numbering refers to numbering the prostheses and teeth by locating a divided prosthesis and obtaining its relationship with the surrounding teeth.

The tooth detection and numbering from a DPR image can be utilized as primary statistical data and identification information because teeth are the hardest tissues in the human body and are composed of enamel, dentin, and chalk [4].

Tooth numbering information can also be applied to the development of various other algorithms. The location information of the panorama can be linked to the existing readings, upon which new algorithms can be developed.

Although algorithms for automatically detecting and diagnosing the condition or number of a tooth using panoramic information have been occasionally used, the methods applied in existing studies do not support an automatic tooth detection within a specific location and have a drawback of requiring the user to manually set the tooth position first [5–10]. Although the accuracy of tooth detection and numbering has increased in recent studies [11,12], a high detection rate for images containing implant fixtures and crowns has not yet been achieved [13]. Thus, in this study, an algorithm is proposed for modeling, detecting, and numbering implant fixtures, crowns, and normal teeth using complex panoramic images that include information on various dental treatments.

2. Prior Research

2.1. Object Detection and CNN Algorithm

The first technology used to detect and recognize objects and people involved object detection using a convolutional neural network (CNN). However, the simple use of a CNN generally requires a substantial amount of training data and a significantly long training time. However, owing to the development of object detection techniques for classifying objects through a bounding-box, the accuracy of the algorithms used for object detection has increased significantly [14]. With the R-CNN technique, which is the basis of object detection, the potential regions of interest (ROIs) are first extracted, and detection is conducted according to the given algorithm. The main advantage of an R-CNN is that it can rapidly extract the location of regions with a relatively high accuracy even if there are few datasets. After the development of an R-CNN algorithm in 2013, various algorithms, including the fast R-CNN, Faster R-CNN, YOLO, and RefineDet, have been developed as object detection techniques [15].

A study on object detection conducted in 2014 used the Sports-1M dataset, which is composed of YouTube video data, to classify the labels into 487 classes. In this study, each video class was classified using a fusion CNN, which analyzes the fusion of two or more spatial and temporal dimensions in a single frame for video classification. In another study proposed in 2017, the object regions of the image data were divided, and training on the regions was conducted using a CNN to send the feature extraction results to the LSTM model for object detection [16]. In this study, as the analysis data, the Chinese University of Hong Kong square dataset, related to walking, and MIT traffic data, related to vehicle detection, were used. Using the CNN model specialized for image analysis and the LSTM model specialized for long-term memory, the feature results of images extracted through a CNN were trained using LSTM. The subsequent memorization of the patterns on a long-term basis enabled object detection through the extraction of the object labels.

In a study conducted in 2018, the Detection with Enriched Semantics (DES) model, which detects objects using a single image, was used to verify the detection data of PASCAL VOC and MS COCO. The DES model was trained using six consecutive activation maps, where each activation map provides values for multi-box detection. The integration of the six activation maps enables detecting objects in a single image when using the trained model [17]. Object detection techniques have also been widely

used in various areas including self-driving cars, CCTV, surveillance, and sporting events. In this study, object detection was implemented using faster R-CNN frameworks. Although a faster R-CNN is known to be slightly slower than the recently developed RefineDet and YOLO, it is known to perform at a similar or even higher level of accuracy [14].

2.2. Dental Radiology

Various studies on tooth detection and numbering have been conducted. Early tooth numbering studies were primarily performed on periapical images. Studies conducted in 2005 [18,19] used pattern recognition classification to classify teeth as molars and premolars and assigned numbers to each tooth according to its position.

Studies on tooth detection using CNNs began in 2017. Automated tooth detection and numbering were conducted using a CNN, and a heuristic method was applied to tooth detection [12]. To detect an object as a tooth and process the tooth number, a heuristic method was used to detect the number corresponding to each tooth [20].

In 2018, the segmentation of each tooth was carried out using a DPR image. ResNet101, which is the basis of a Mask R-CNN, was used in these tooth studies. With this approach, the categories are divided into ten different classes. During learning, only images of teeth without an implant are trained, and the learning result for tooth detection was found to achieve an accuracy of 95% [9].

In 2019, tooth X-ray images were evaluated using a VGG16 CNN model for tooth detection and numbering. However, this study has limited applications i.e., periapical images and DPR images without implants. Therefore, it was found to have limited use in the real world [10].

In 2020, Researchers in Japan conducted research on implant fixtures and developed an algorithm to classify implants by implant company and model. This algorithm judges which implant fixture is the implant fixture when there is a defined ROI area, and there is a lack of locating the implant [21].

The Table 1 shows that recent studies on the tooth detection, segmentation and numbering algorithm through CNN. The high accuracy of tooth detection and numbering has achieved in the recent studies, implant fixtures and crowns didn't included. Our algorithm is proposed for modeling, detecting, and numbering implant fixtures, crowns, and normal teeth using complex panoramic images that include information on various dental treatments.

Table 1. Comparative overview table for machine learning article in the field of tooth detection.

Author	Architecture	Object	Evaluation	Pros (+)/Cons (–)	Ref.	Year
Oktay	CNN	Incisors, Premolars, Molars	Over 90%	(+) CNN Approach (–) Few Dataset (–) Manual ROI Selection	17	2017
Jader et al.	Mask R-CNN	Tooth	94%	(+) Region Segmentation (–) Without Implants	18	2018
Tuzoff et al.	R-CNN	Tooth	99%	(+) R-CNN Approach (+) Many Data (–) Without Implants	19	2019
Sukegawa et al.	CNN	Implants	Over 90%	(+) CNN Approach (+) Implant (–) Only Implants	26	2020

In our study, an algorithm is proposed to detect teeth, implant fixtures, and crowns, and subsequently number the teeth and implants in a DPR image based on the FDI two-digit notation [22]. This algorithm can be used universally, including cases of DPR images of implanted teeth. The notation consists of two digits, where the first digit indicates one of the four tooth quadrants, and the second digit indicates the detailed shape of the tooth.

3. Materials and Methods

3.1. Study Design

Based on previous research results, we propose an algorithm that can recognize the regions of the teeth and classify them using DPR images during the training process. To achieve this, we also propose methods to obtain images for training purposes. The images obtained were labeled, and the objects within the images were detected using a Single Shot Multibox Detector (SSD) and a regional convolutional neural network (RCNN) algorithm. These images were classified again using a heuristic algorithm [23].

The dataset used in this study is composed of 303 anonymized DPR images. The equipment models used for acquiring the DPR images were Vatech PaX-i, HDX Will, and Rayscan. Dislocated teeth and late residual teeth are not covered in this study. The reason for this is that we did not have sufficient data for these two categories of teeth, and hence, classification would be difficult. Therefore, images containing dislocated teeth and late residual teeth were excluded from this study. The overall research design flow is presented in Figure 1. The DPR images were divided into 253 training sets and 50 test sets. Each data entry was labeled with dental objects (implant fixture, crown, tooth), and numbered by three dentists. We then developed two algorithms using a CNN and a faster RCNN, object detection, and tooth classification algorithms, respectively.

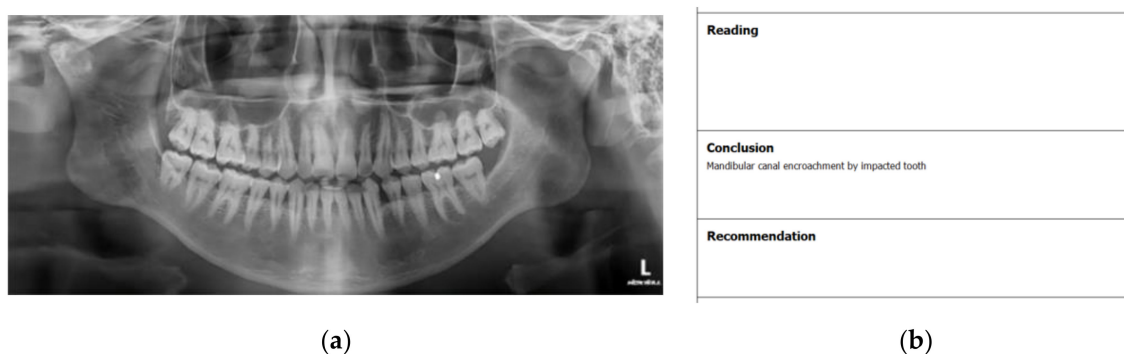


Figure 1. (a) Sample DPR image and (b) study report.

Figure 2 shows the flow of this study for implementing a heuristic algorithm for tooth and implant detection and numbering. The primary task is to locate areas related to the teeth or implant fixtures in the given X-ray image. To achieve this, an object detection technique, which is frequently used in the field of image recognition, was used. The technique applied to conduct the training process is called a faster RCNN algorithm. Among the various learning models, the learning model we used was the Faster RCNN Inception v3 architecture developed by Google. During the training process, the objects were divided into three labels: teeth, crowns, and implant fixtures.

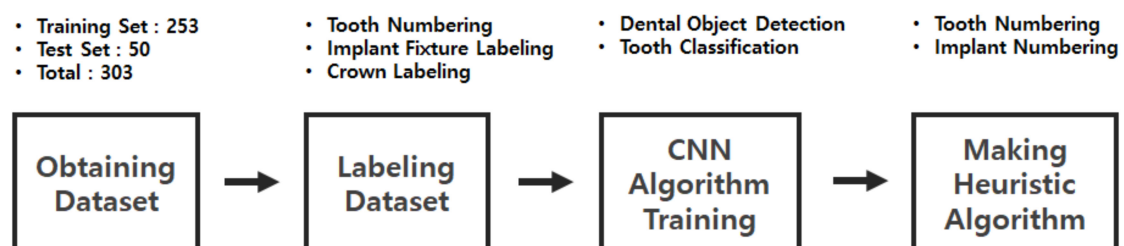


Figure 2. Study flow of teeth and implant detection.

Faster RCNN is a method that derives better accuracy than existing object detection algorithms by extracting image features and minimizing noise for image analysis. Faster-RCNN is composed of a

convolution feature map and ROI feature vector. The convolution feature map delivers images to the convolution and max-pooling layers, and the received information is placed as features in the ROI feature vector map. It is converted to a map with various features and moved to fully connected layers (FCs) to determine the object value of the image for the object of the K class. In this process, through the loss function, the multi-task loss is minimized, and the learning accuracy is increased. In Equation (1), i represents the index value of the anchor, and p_i is an object or a value predicting whether it is a background. p_i^* is 1 if the value is an object, and 0 if it is an object. Here, the object is detected through the loss function L_{cls} and smooth L1 loss function values L_{reg} and normalization mini-batch values N_{cls} normalization anchor location value N_{reg} values between classes [24].

$$L(\{p_i\}, \{t_i\}) = \frac{1}{N_{cls}} \sum_i L_{cls}(p_i, p_i^*) + \lambda \frac{1}{N_{reg}} \sum_i p_i^* L_{reg}(t_i, t_i^*) \quad (1)$$

An SSD simultaneously performs the bounding box and class prediction while processing an image in a single shot. The SSD works well on low-resolution images, derives output through each multi-feature map, and predicts bounding box and class scores through appropriate convolution operations on each feature map. Therefore, the SSD is represented by the predicted value for the class and the predicted value for the bounding box.

$$L(x, c, l, g) = \frac{1}{N} (L_{conf}(x, c) + \alpha (L_{loc}(x, l, g))) \quad (2)$$

$$L_{loc}(x, c) = - \sum_{i \in Pos}^N x_{ij}^p \log(\hat{c}_i^0) - \sum_{i \in Neg} \log(\hat{c}_i^0) \quad (3)$$

$$L_{loc}(x, l, g) = \sum_{i \in Pos}^N \sum_{m \in \{cx, cy, w, h\}} x_{ij}^k smooth_{L1}(l_i^m - \hat{g}_j^m) \quad (4)$$

The object recognition index in category p of the i default box and the j ground truth box. p is 1 if the IOU between the object's j ground truth and the i default box is 0.5 or more, 0 is N , and N is the number matching the default boxes; i is the predicted box; g is the ground truth box; values other than g is the default box; cx , cy are the box x , y coordinates; w and h are the box width and height, respectively; alpha is 1 predicted box cx to predict, cy , w , and h values predict $\hat{g}$ values. Therefore, the object value is predicted through object [25].

Next, once the regions of the teeth were recognized through images and categorized, an algorithm was designed to predict the number of teeth. To distinguish the position of a given tooth, another algorithm was designed that uses the position and shape information of the tooth.

Figure 3 shows how the proposed algorithm conducts tooth and implant detection and tooth numbering. To simultaneously secure both the numbering and object detection information, numbering and object detection labeled data were collected separately. By collecting the data separately, the algorithm was generated to independently classify the position of a given tooth through both the tooth position and individual tooth shape information.

3.2. Dataset and Labeling Dataset

A total of 303 panoramic patient data were collected from the Medipartner Dental Network Hospital after obtaining patient agreements. Each image was anonymized and converted into a 1600 pixel $\times$ 900 pixel image in JPG format. Three dentists labeled each panoramic tooth image, using Label Box as the labeling tool. Figure 4 shows the labeling results on an image from the Label Box tool used by the dentists. Only labeled dental implants with fixtures were included in the data set because we did not have enough abutments and prostheses images on the dental panorama images.

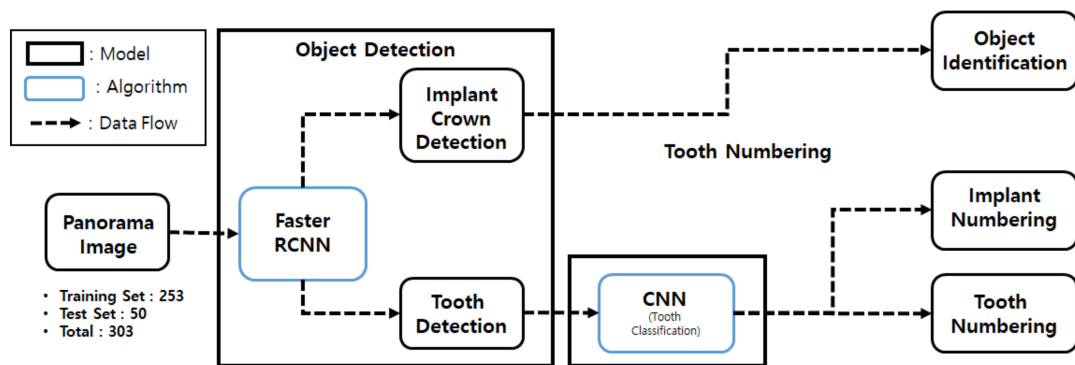


Figure 3. Algorithm for object detection and tooth numbering.

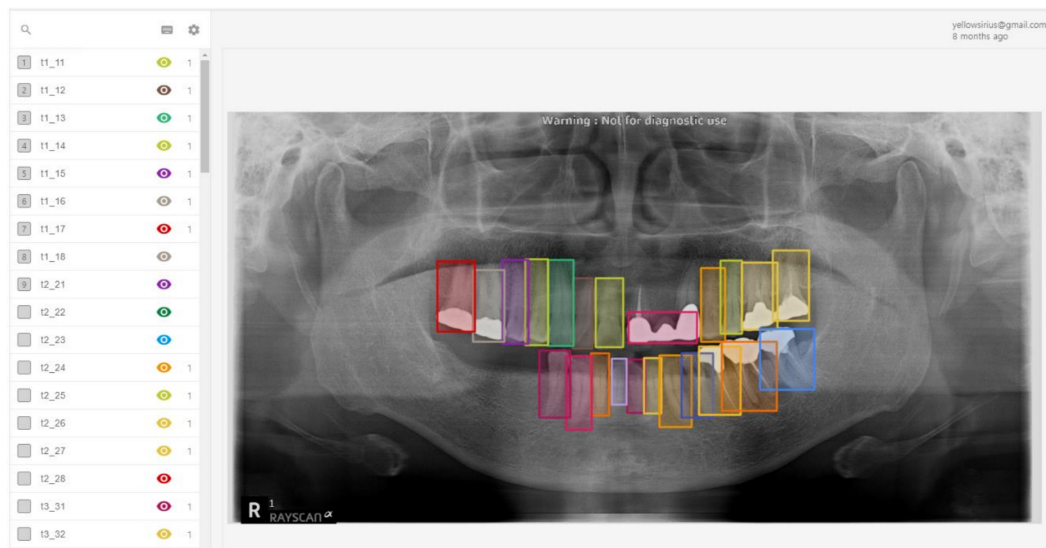


Figure 4. Labeling tool used to generate training and test sets.

For the tooth detection, teeth and implants were detected and parts of the root of one tooth and parts of two individuals were considered to be in different classes. The detailed execution method is shown in Figure 5. The objects in the panoramic image are mainly composed of implant fixtures, crowns, and teeth. Here, specific numbers are assigned to each tooth. Using such images, a total of 253 training sets and 50 test sets were obtained.

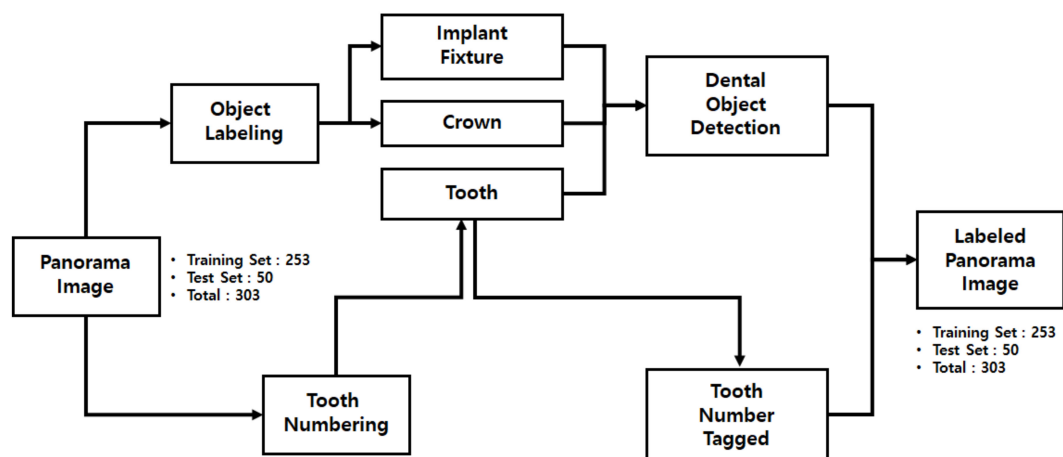


Figure 5. DPR labeling method.

The labeled data were categorized as dental object detection information or tooth numbering information. The collected data are presented in Figure 6. In the image on the top-right of Figure 6, blue parts indicate the labeling data on the implant fixtures, and the orange parts indicate the data on the crowns. The tooth labeling information presented in the lower-right image in Figure 6 indicates the information used to label the individual tooth numbers.

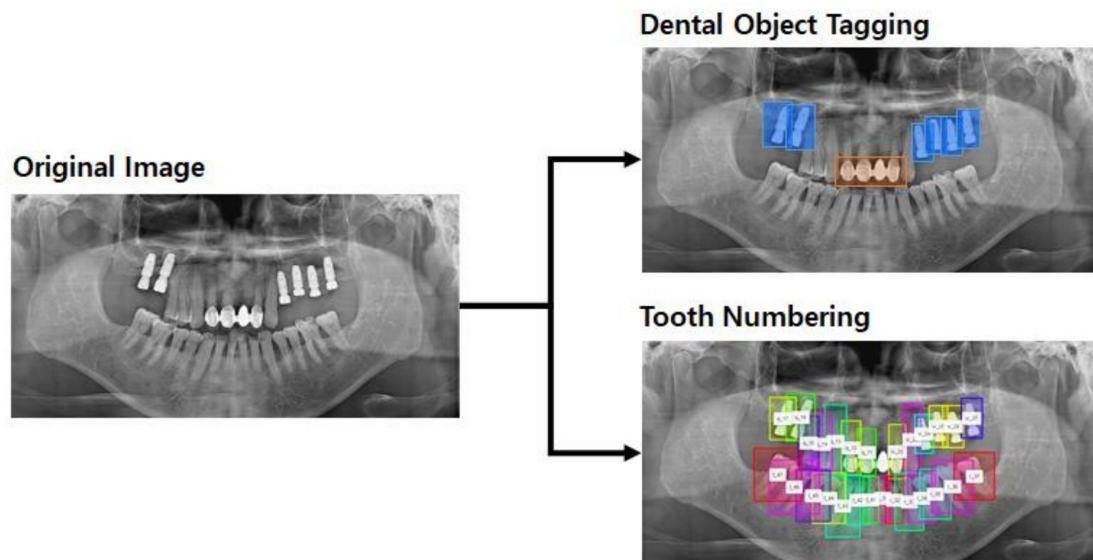


Figure 6. Example of labeled data.

3.3. Dental Object Detection Modeling and Training

The tooth training set was composed of 6446 teeth, excluding missing teeth. This set was obtained from a total of 253 panoramic tooth images. These teeth were categorized into classes, including teeth, implant fixtures, and crowns. There were 402 implants with fixtures only and 205 crowns. The Tensorflow Slim Library was used for learning, and the Faster RCNN and Inception V3 neural networks were used as described above. An Intel Xeon and GeForce RTX 2080TI were used as the learning equipment. The tooth, implant-fixture, and crown detection were conducted 42,000, 24,000, and 70,000 times, respectively, for each of the three learning models used during the training. The average loss value was terminated between 0.05 and 0.1. To prevent overfitting, the loss value was less than 0.02, and each model was trained less than 100,000 times.

The faster RCNN inception model combines a region proposal network, a region proposal method that captures the features using deep learning, and an inception model that reduces the number of computations and improves the speed and accuracy [22]. Figure 7 shows the model process for detecting teeth. Through this process, classes such as tooth number, implant, and crown are derived and evaluated by groups belonging to each tooth class.

3.4. Tooth Numbering Modeling and Training

After detecting dental objects, tooth classification and numbering algorithms were trained to determine the numbering of the individual teeth. To identify the number of a specific tooth, the combined positional values of each tooth are required. The method is outlined in Figure 8.

An RCNN was constructed to combine the extracted tooth information and position data and classify the objects into each number. Subsequently, an algorithm was designed that classifies each tooth type based on the constructed model. Using this algorithm, individual teeth were predicted and numbered. The predicted values were input as the second digit of the tooth ranging from 1–7, and the training was conducted in such a manner that the model can classify the given tooth as an incisor, canine, or molar according to its shape.

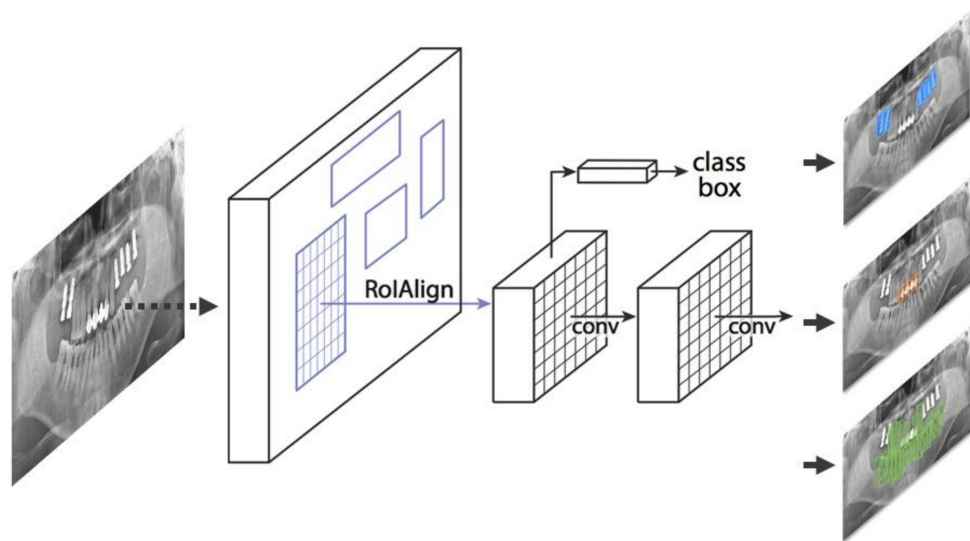


Figure 7. Algorithm model for dental object detection.

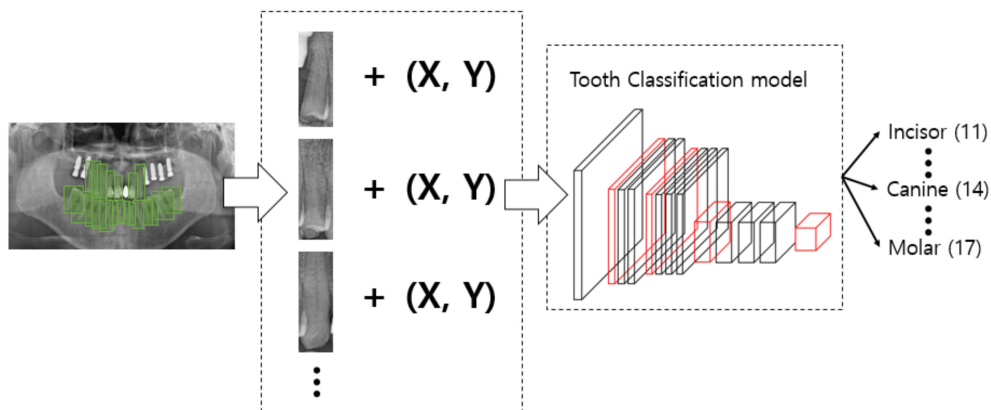


Figure 8. Tooth numbering algorithm.

4. Results

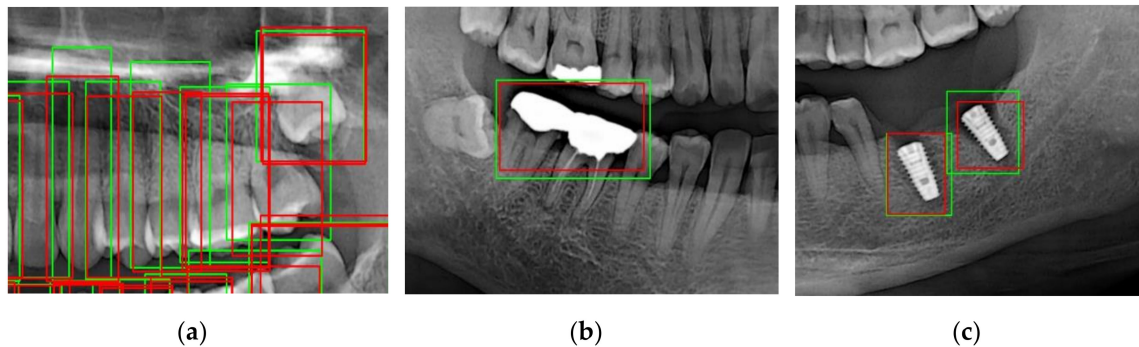
4.1. Dental Object Detection Model

The performance of the dental object detection model was evaluated based on the standards of the intersection over union (IOU). In the case of mAP, when the IOU value is 0.5, a successful prediction is considered to have an accuracy of 50% or higher. In the case of tooth detection, a significantly high accuracy of 96.7% was obtained when $\text{mAP@IOU} = 0.5$. Even at $\text{mAP@IOU} = 0.7$, a high accuracy of 75.4% was obtained, suggesting that the detection algorithm can determine the location of the teeth with high accuracy.

By contrast, for the case of implant fixtures and crowns, the accuracies were 45.1% and 50.9%, respectively, as shown in Table 2 when $\text{IOU} = 0.5$. Further, at $\text{IOU} = 0.7$, low accuracies of 26.6% and 40.8% were obtained for implant fixtures and crowns, respectively. These results indicate that the shapes of the implant fixtures and crowns are detected less accurately than expected. One possible reason for this is that the crowns and implant fixtures have various unstructured shapes, and hence the model may be unable to accurately detect the shapes when compared to normal teeth. Nevertheless, the results obtained are significant because the implant fixtures and crowns were still detected through the panoramic images. Therefore, as shown in Figure 9, each tooth, implant, and crown can be detected. As described above, the numbering of the teeth, the implant, and the crown is determined by detecting the teeth.

Table 2. Results of Dental Object Detection on AP@0.5, 0.7.

Category	mAP @ IOU = 0.5	mAP @ IOU = 0.7
Tooth	96.7%	75.4%
Implant Fixture	45.1%	26.6%
Crown	60.9%	40.8%

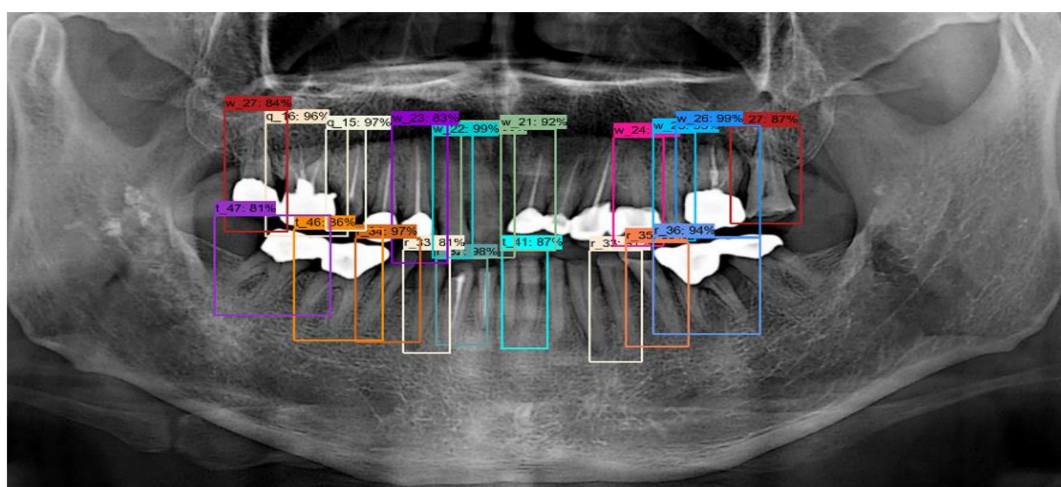
**Figure 9.** (a) Example of tooth, (b) crown, and (c) implant fixture detection.

4.2. Tooth Numbering Model

Table 3 shows that the probability of a tooth actually existing in the location indicated by the RCNN algorithm was 84.2%, with a sensitivity of 75.5%, and a precision of 84.5%. In addition, for tooth numbering, an accuracy of 77.4% was consistently obtained between the location of the actual tooth and the location indicated by the algorithms. The results of tooth detection are shown in Figure 10. As shown in this figure, all information about the position of the tooth can be detected correctly. However, it can be seen that the small tooth and crown cannot be detected for the tooth whose shape is not recognized.

Table 3. Comparison between the proposed algorithm and RCNN.

Category	Sensitivity	Specificity	Precision
Proposed Algorithm (RCNN + Heuristics)	75.5%	80.4%	84.5%
Conventional Algorithm (RCNN)	60.2%	76.7%	72.4%
Single Shot Multibox Detector (SSD)	56.7%	70.4%	64.7%

**Figure 10.** Results of tooth detection.

In the existing model composed of only an RCNN, because functions for detecting the object and classifying the tooth are combined, the accuracy of the model is relatively low in its ability to search and detect a tooth compared to our proposed approach.

4.3. Practical Application for the Algorithm

We confirmed that the detection of the dental object and the implant fixture were successful through previous results. As mentioned earlier, the tooth detection algorithm has several research advantages. One example is the identification of an individual through tooth detection. Another example is to use it for further research, such as classifying caries in the teeth. In the case of the implant detection algorithm, it would be possible to separate manufacturers and models in the future by classifying the implants. In addition, it is possible to generate new labeling information related to the information from existing patient charts.

Dental caries is one of the research fields that can be explored through representative object detection. It is possible to combine caries diagnosis information in the patient chart with panoramic images and diagnose caries progression of individual teeth. In the case of implants, it is thought that it will be possible to develop an algorithm for detecting abnormalities in the area of the gums as a follow-up study by identifying the abnormalities of the gums at the location where the implant.

5. Conclusions

In this study, we presented an algorithm that can detect dental objects in a DPR image and assign a number to each tooth based on its shape and location as obtained using the RCNN and CNN algorithms. The results confirm that the numbering of the teeth and implants is possible in a DPR image. Based on the analysis results, an RCNN + heuristics algorithm, which exhibited the best performance in dental detection, was adopted. As a result, precision values of 84.5%, sensitivity values of 75.5%, and specificity values of 80.4% were achieved. Thus, the proposed algorithm was found to yield the best performance in detecting teeth.

The panoramic images derived through the tooth numbering and implant fixture and crown analysis methods applied in this study share the following three purposes: to show and explain the panoramic image results to patients and non-professional personnel, statistical relevance, and an implementation in other algorithms. Although interpreting panoramic images is easy for dentists, patients face difficulties in such an interpretation. When a dentist needs to show and describe a panoramic tooth image to a patient, providing the tooth numbers on the crowns, implant fixtures, and teeth will help the patient easily understand the diagnostic information of the panoramic image. In addition, panoramic images can be used to identify a large number of teeth and dental objects and understand their statistical significance. Statistical information, including the number of patients with crowns and implant treatments, and specific tooth information based on the tooth number, can be used as an important tool in health-related dental research. Finally, panoramic images applying an automatic detection of the tooth number, as well as the position of the implant fixtures and crowns, can contribute to the development of radiology-related algorithms. Because existing methods generally provide tooth and implant information verbally, they have not been proven helpful in image training. Therefore, the proposed algorithm is expected to contribute to the development of image detection algorithms by providing additional tooth position information to existing methods.

To develop the algorithm proposed in this study, we used a total of 253 teeth images and extracted 6446 teeth, implant, and crown data. Specifically, the detection processes for teeth, implant fixtures, and crowns were conducted first. Subsequently, we derived an algorithm that can number each tooth. The experimental results showed a high accuracy in tooth numbering, although the accuracy in terms of the numbering of the implant fixtures and crowns was lacking. The main reason for such accuracy is the smaller amount of training data for implant fixtures and crowns compared to that of teeth.

Further, in the case of implant fixtures, the images of the training data had similar implant sizes and shapes, and hence the accuracy for detecting the implants was higher than that for the crowns.

By contrast, the sizes and shapes of the crowns in the training data appeared to be different in each image, yielding a much lower detection rate during the training process. Accordingly, we expect to derive better results by training the model using a larger amount of teeth data in the future. In addition, to further increase the accuracy of the tooth numbering, the algorithm should be supplemented by improving the wisdom tooth detection rate, which we also expect to resolve by collecting additional data in the future.

Author Contributions: Conceptualization, C.K., D.K. and S.Y.; data, D.H., S.-J.Y., and H.J.; methodology, C.K., H.J. and S.-J.Y.; validation, C.K. and D.K.; formal analysis, C.K. and D.K.; investigation, C.K. and S.Y.; writing—original draft preparation, C.K.; writing—review and editing S.Y. and H.J.; supervision, S.Y. All authors have read and agreed to the published version of the manuscript.

Funding: This work was supported by the Dongguk University Research Fund of 2016 and the Gerontechnology Research Center.

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

References

1. Yim, J.H.; Ryu, D.M.; Lee, B.S.; Kwon, Y.D. Analysis of digitalized panorama and cone beam computed tomographic image distortion for the diagnosis of dental implant surgery. *J. Craniofacial Surg.* **2011**, *22*, 669–673. [[CrossRef](#)] [[PubMed](#)]
2. Kalinowska, I.R. Artificial intelligence in dentomaxillofacial radiology: Hype or future? *J. Oral Maxillofac. Radiol.* **2018**, *6*, 1. [[CrossRef](#)]
3. Joon, H.J.; Hoa, J.Y.; Hae, C.B.; Suk, H.M. An overview of deep learning in the field of dentistry. *Imaging Sci. Dent.* **2019**, *49*, 1. [[CrossRef](#)]
4. Bucky, T. Kriminalistische Feststellungen durch Röntgenstrahlen. *Arztl. Sachverst.* **1992**, *28*, 166–170.
5. Miki, Y.; Muramatsu, C.; Hayashi, T.; Zhou, X.; Hara, T.; Katsumata, A.; Fujita, H. Classification of teeth in cone-beam CT using deep convolutional neural network. *Comput. Biol. Med.* **2017**, *80*, 24–29. [[CrossRef](#)] [[PubMed](#)]
6. Cui, Z.; Li, C.; Wang, W. ToothNet: Automatic tooth instance segmentation and identification from cone beam CT images. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Long Beach, CA, USA, 16–20 June 2019; pp. 6368–6377.
7. Koch, T.L.; Perslev, M.; Igel, C.; Brandt, S.S. Accurate segmentation of dental panoramic radiographs with U-NETS. In Proceedings of the IEEE 16th International Symposium on Biomedical Imaging, Venice, Italy, 8–11 April 2019; pp. 15–19.
8. Lamecker, H.; Kainmueller, D.; Zachow, S. Automatic detection and classification of teeth in CT data. In *International Conference on Medical Image Computing and Computer-Assisted Intervention*; Springer: Berlin/Heidelberg, Germany, 2012; pp. 609–616.
9. Jader, G.; Fontineli, J.; Ruiz, M.; Abdalla, K.; Pithon, M.; Oliveira, L. Deep instance segmentation of teeth in panoramic X-Ray images. In Proceedings of the 31st SIBGRAPI Conference on Graphics, Patterns and Images (SIBGRAPI), Parana, Brazil, 29 October–1 November 2018; pp. 400–407.
10. Tuzoff, D.V.; Tuzova, L.N.; Bornstein, M.M.; Krasnov, A.S.; Kharchenko, M.A.; Nikolenko, S.I.; Sveshnikov, M.M.; Bednenko, G.B. Tooth detection and numbering in panoramic radiographs using convolutional neural networks. *Dentomaxillofac. Radiol.* **2019**, *48*, 20180051. [[CrossRef](#)] [[PubMed](#)]
11. Tuzoff, D.V.; Tuzova, L.N.; Kharchenko, M.A. *Report on Tooth Detection and Numbering in Panoramic Radiographs Using CNNs*; MIDL: Amsterdam, The Netherlands, 2018.
12. Betul, O. Tooth detection with convolutional neural networks. In Proceedings of the Medical Technologies National Congress (TIPTEKNO), Trabzon, Turkey, 12–14 October 2017; pp. 1–4.
13. Türp, J.C.; Alt, K.W. Designating teeth: The advantages of the FDI's two-digit system. *Quintessence Int.* **1995**, *26*, 501–504. [[PubMed](#)]
14. 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, Columbus, OH, USA, 24–27 June 2014; pp. 580–587.

15. A 2019 Guide to Object Detection. Available online: <https://heartbeat.fritz.ai/a-2019-guide-to-object-detection-9509987954c3A> (accessed on 11 June 2020).
16. Li, X.; Ye, M.; Liu, Y.; Zhang, F.; Liu, D.; Tang, S. Accurate object detection using memory-based models in surveillance scenes. *Pattern Recognit.* **2017**, *67*, 73–84. [[CrossRef](#)]
17. Harris, E.F. Tooth-coding systems in the clinical dental setting. *Dent. Anthr. J.* **2018**, *18*, 43–49. [[CrossRef](#)]
18. Mahoor, M.H.; Mottaleb, M.A. Classification and numbering of teeth in dental bitewing images. *Pattern Recognit.* **2005**, *38*, 577–586. [[CrossRef](#)]
19. Chen, H.; Zhang, K.; Lyu, P.; Li, H.; Zhang, L.; Wu, J.; Lee, C.H. A deep learning approach to automatic teeth detection and numbering based on object detection in dental periapical films. *Sci. Rep.* **2019**, *9*, 3840. [[CrossRef](#)] [[PubMed](#)]
20. Sukegawa, S.; Yoshii, K.; Hara, T.; Yamashita, K.; Nakano, K.; Yamamoto, N.; Nagatsuka, H.; Furuki, Y. Deep neural networks for dental implant system classification. *Biomolecules* **2020**, *10*, 984. [[CrossRef](#)] [[PubMed](#)]
21. *Dentistry—Designation System for Teeth and Areas of the Oral Cavity*; ISO 3950:2009; ISO: Geneva, Switzerland, 2009.
22. Zhang, Z.; Qiao, S.; Xie, C.; Shen, W.; Wang, B.; Yuille, A.L. Single-shot object detection with enriched semantics. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Salt Lake City, UT, USA, 18–22 June 2018; pp. 5813–5821.
23. 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*; MIT Press: San Diego, CA, USA, 2015; pp. 91–99.
24. Liu, W.; Anguelov, D.; Erhan, D.; Szegedy, C.; Reed, S.; Fu, C.Y.; Berg, A.C. Ssd: Single shot multibox detector. In *European Conference on Computer Vision*; Springer: Cham, Switzerland, 2016; pp. 21–37.
25. Szegedy, C.; Ioffe, S.; Vanhoucke, V.; Alemi, A.A. Inception-v4, inception-resnet and the impact of residual connections on learning. In Proceedings of the Thirty-First AAAI Conference on Artificial Intelligence, San Francisco, CA, USA, 4–9 February 2017.



© 2020 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).