Dental Image (X-ray) Tooth Detection and Labeling of Teeth

By

Joe Chudzik,

Olivia Fenwick,

Priyanka Karekar,

Jeff Xu

Supervisor: Utku Pamuksuz

A Capstone Project

Submitted to the University of Chicago in partial fulfillment of the requirements of the degree of

Master of Science in Analytics

Division of Physical Sciences

December 2021

Abstract

Panoramic extra-oral dental x-rays provide an overall view of the mouth that is instrumental in helping dentists monitor tooth health, however, they can be challenging to interpret and require a high level of training. Identifying individual teeth is the first step in dental artificial intelligence research but this can be particularly difficult to do on panoramic dental x-rays and manual tooth detection can be very time-consuming and labor-intensive. This work aims to help solve those issues by training convolutional neural networks to detect and classify teeth and combine it with a rule-based algorithm to label existing and missing teeth in panoramic dental x-rays. The tooth detection and classification model yielded a recall score of 0.9448 and mAP score of 0.9613 on the testing set and the rule-based tooth numbering algorithm had an accuracy of 91.48%.

Keywords: deep learning, dental AI, YOLOv5, convolutional neural networks, tooth detection, tooth numbering

Executive Summary

Dental x-rays are essential in monitoring oral health; however, highly trained professionals are required for them to be accurately read and interpreted. Dental artificial intelligence (AI) can support dental diagnosis by reducing mistakes and increasing efficiency. Panoramic x-rays are simple to perform yet provide great value for diagnosing tooth problems as they provide an overall view of the mouth. Identifying individual teeth is the first step in dental AI research but manual tooth detection and tooth type classification can be very time-consuming and labor-intensive. This work aims to help solve that by using convolutional neural networks to detect and classify teeth and a rule-based algorithm to number teeth and identify missing teeth.

The model was trained using 2328 panoramic dental x-rays from two different data sources which helped to reduce bias in the model. There were three distinct stages to the process; detect teeth in the x-ray, classify tooth type, and number the teeth and label missing teeth. The first two steps were achieved using YOLOv5 models and the final step used a rule-based approach.

The model performed strongly to identify and classify teeth with a recall score of 0.9448 and mAP score of 0.9613 on the testing set. The rule-based algorithm had an accuracy of 91.48%. This approach, combined with a related tooth defect detection and classification model could be used by dentists to increase their diagnosis efficiency and accuracy.

Table of Contents

Commented [OF1]: Make sure to double check page numbers at the end

List of Figures

Commented [OF2]: At the end check the figure numbers and titles are correct

List of Tables

Introduction

Problem Statement

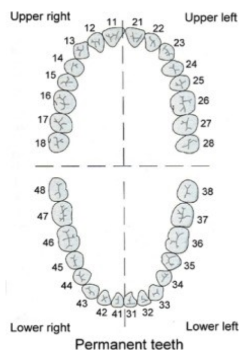
In dentistry, radiological examinations help dentists by visualizing the structure of teeth, jaw bones and surrounding structures with the goal of screening embedded teeth, bone abnormalities, cysts, tumors, infections, fractures, problems in the temporomandibular regions (Jader et al., 2019). Sometimes, relying solely on the dentist's opinion can bring differences in diagnosis and can lead to prescribing a faulty treatment plan, thereby hindering proper treatment. According to Chen et al. (2019), a senior dentist has a diagnostic accuracy of about 70%. This indicates that there is a considerable scope for improving decision support systems to improve dental diagnosis.

Artificial intelligence (AI) based solutions for medical diagnosis have been a growing field of image pattern recognition research. Deep learning-based solutions have the potential to greatly enhance decision support, such as with the detection of teeth and related components in x-ray images. Inference Analytics Neural Network (IANN) is interested in developing a solution using deep learning approaches to improve and expedite the x-ray screening process. Such a product will significantly reduce time and effort required for screening panoramic x-rays, thereby increasing the number of patients that can be seen at a time, while also allowing the dentist more time to formulate appropriate treatment plans. This would also serve as a validation tool for dentists to reduce human error. This could potentially also increase revenue for dental clinics by improving patient and staff scheduling.

Analysis Goals

The goal of this project is to detect teeth on panoramic dental x-rays, classify the tooth into its proper tooth type category, correctly label individual teeth numbers according to standard dental convention, and to identify missing teeth and their locations. The final model should be able to correctly identify and label adult teeth on panoramic dental x-ray image, with an accuracy of more than 90%, and label existing teeth locations with an accuracy more than 90%. The novelty of this work is to refine and speed up the tooth detection process, and to correctly identify and label existing and missing teeth according to the ISO 3950 standard (*ISO 3950:2016*, n.d.), as shown in Figure 1.

Figure 1. *Tooth numbering according to ISO 3950 standard.*



This work on tooth detection is an essential first step for developing dental AI, as it serves as a backbone for automatic charting, problem detection, etc. The output of the tooth detection model from this study will be used in subsequent research to develop AI solutions to identify dental problems in individual teeth detections obtained from panoramic x-rays. These works when combined would be able to alert a dentist to potential problems relating to missing teeth, dental carries, tumors etc. This could save a lot of time and resources required for initial dental screening, and lead to more effective treatment planning.

Scope

Due to the complexity of the overall project at IANN and time constraints, this work is focused mainly on identifying and labeling teeth in patients with no primary teeth. As a result, the tooth identification model will only be applicable for patients aged 14 years and older, specifically those who do not have their primary teeth

Background

IANN is a health-technology start-up with the goal of creating products to help improve efficiency and reduce mistakes in healthcare (*Inference Analytics*, n.d.). They use technology such as deep learning combined with transfer learning to mimic how the brain works and improve traditional NLP practices. Their software enables organizations to include predictive and descriptive analytics in their decision-making process and to help them make data driven decisions. Already, IANN's platform has been shown to provide many benefits to the healthcare community; a recent clinical study showed that radiologists were able to reduce the time spent on each report by around 20% while using IANN technology. It has also been shown to help catch mistakes and reduce human error to lead to improved quality of care while also aiding physicians in improving their efficiency. Through this research and in conjunction with another similar project which builds upon these results, IANN aims to enter the dental AI industry by creating a product that will use AI to assist dentists with faster screening of the dental x-rays and detecting abnormalities.

Currently, most of the available products in dental AI focus on supporting routine tasks such as classifying issues with scanned preps, automating administrative tasks, and identifying coding and billing errors (*Overjet*, n.d.). However, there are companies, such as Denti.AI, Pearl and ORCA Dental AI, with products specifically designed to analyze dental x-rays and provide a second opinion for dentists (*Denti.AI*, n.d.; *Hello Pearl*, n.d.; *ORCA Dental AI*, n.d.). Diagnostic dental AI is still a developing field, and most companies are still waiting for approval from regulatory agencies before it is widely available in the United States. However, Overjet recently received FDA approval earlier in 2021 to be used as a medical device for aiding dental tooth screening based on x-rays (Overjet, 2021).

Literature Review

Tooth Detection

Most of the current research in dental imaging relies on deep learning-based approaches as these approaches most effectively solve complicated tooth segmentation challenges as further described later. Mask Regional-based Convolutional Networks (Mask R-CNNs) have achieved instance segmentation of teeth with 98% accuracy, 94% precision, and 84% recall (Jader et al., 2019). Since then, Mask R-CNN has become a popular architecture for CNN based tooth segmentation, with many further studies in this field using this architecture (Gurses & Oktay, 2020; Tuzoff et al., 2019; Zhu et al., 2020). Although Masked R-CNN-based detection methods show impressive results, high accuracy may not be guaranteed in all the categories of x-ray tooth images, especially in images with dental implants or missing teeth (Jader et al., 2019; Tuzoff et al., 2019).

Deleted:

You Only Look Once v5 (YOLOv5) is an object detection algorithm that provides a novel approach to solve the challenges in tooth detection. Dwivedi (2020) demonstrated that YOLOv5 outperforms Faster R-CNN in terms of inference speed, detection of small objects, reduced overlapping boxes, and the detection of crowded objects in MOTChallenge. Pretrained on the MS-COCO dataset, transfer learning can be effectively utilized to train a YOLOv5 model for the required context. Although there are no published applications of YOLOv5 in the medical industry, this work is fundamental to not only further introduce various AI frameworks to the medical industry but also to improve x-ray image-based tooth detection.

Tooth Numbering

Maximum numbering is a common method used to label teeth into distinct categories and report the missing tooth type (Tuzoff et al., 2019). This approach uses bounding box outputs from a CNN and applies tooth location rules to estimate a tooth number. Although this approach is highly accurate when most teeth are present on the x-ray, it struggles to locate missing teeth without adjacent teeth.

More recently, point-based localization approaches have been used for tooth detection and numbering (Chung et al., 2020). This approach attempts to regress points to delineate objects such as the center, left-top, and right bottom points of the teeth. This improved the average precision (AP) of teeth detection by an average of 15.71% compared to previous methods using R-CNNs. This approach also does not require a second CNN to classify teeth, but instead, it uses the center points estimates for all possible 32 teeth regardless of the existence of the teeth and matches the center points of each tooth on a given panoramic x-ray image with the estimated center points. The major advantage of this approach is that it saves a significant amount of computing time while finding both existing and missing teeth locations.

Performance Analysis

Region of interest (ROI), mean average precision (mAP), and mean squared error (MSE) are commonly used metrics to evaluate the performance of tooth detection models. To capture as many teeth as possible, it is common for tooth detection models to have higher recall than precision (Jader et al., 2019). Within a dental context, higher recall is preferred as it is better to detect all the teeth with some non-tooth objects incorrectly detected than not detect some of the teeth.

Data

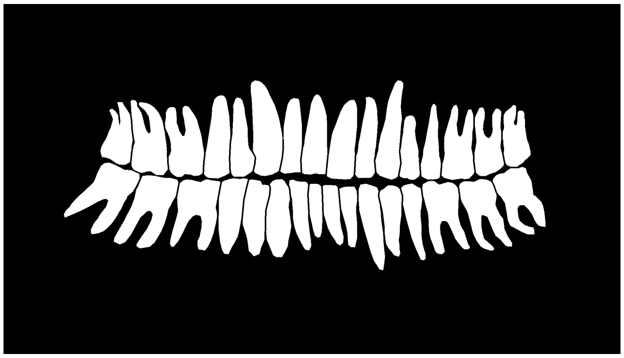
Data Sources

The most used radiographic images in dentistry include panoramic, bitewing, periapical and cephalometric x-rays (Schwendicke et al., 2019). This work focused on panoramic dental x-rays which are 2D images that capture the entire mouth structure in a single image, by producing a flattened image of the curved jaw. They are extra-oral and simple to perform, making them a valuable and essential tool for dental diagnosis (*Panoramic Dental X-Ray*, n.d.). Two sources of panoramic dental x-rays were used in this work: UFBA-UESC dental images dataset and Great Lake Dental Partners (GLDP) dataset.

The UFBA-UESC dental images were first used by Silva, et al. (2018) to analyze the performance of different segmentation algorithms and have become a standard for dental research. The images from GLDP were collected from about 40 clinics under their umbrella and assembled specifically for IANN's dental AI research. Images from the UFBA-UESC dental images dataset were used for initial analysis and later combined with panoramic dental x-rays from GLDP to train the final model.

The UFBA-UESC Dental Images dataset contains 1500 panoramic dental x-rays that fall into various categories depending on if there are missing, extra, or all the teeth and if there are any dental implants, appliances, or restorations. 1330 images from the dataset were utilized, as the remaining 170 x-ray images had extra teeth including primary teeth, as shown in Table 1, which is beyond the scope of this work. Segmented dental masks were available for 276 of the images, an example of which can be seen below in Figure 2.

Figure 2. *Dental mask from the UFBA-UESC Dental Image Dataset*



The data from GLDP contains over 2 million panoramic dental x-rays, 998 that were used for training and testing and the rest for validation. All of the 998 images were from people aged 14 and older so the model is only trained on adult teeth. Contrary to the UFBA-UESC dental x-ray dataset, none of the data we received from GLDP had corresponding dental masks or labels segmenting the teeth.

Descriptive Analysis

Only four of the categories in the UFBA-UESC Dental Images dataset have images containing all 32 teeth, 57% of the images have missing teeth, 58% have restorations, 19% have dental appliances, and only 8% have dental implants. Labelled data for tooth type is imbalanced, as each set of 32 teeth has 12 molar, 8 pre-molars and incisors, but only 4 canines. Table 1 further summarizes the details of the categories the UFBA-UESC Dental Images dataset was divided into as well as the number of images in each category.

Table 1. *Description of x-ray categories for UFBA-UESC Dental Images dataset*

Category	Description	Images	Average Number of Teeth
1	Images with all the teeth, containing teeth with restoration and with dental appliance	73	32
2	Images with all the teeth, containing teeth with restoration and without dental appliance	220	32
3	Images with all the teeth, containing teeth without restoration and with dental appliance	45	32
4	Images with all the teeth, 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 missing teeth, containing teeth with restoration and dental appliance	115	27
8	Images missing teeth, containing teeth with restoration and without dental appliance	457	29
9	Images missing teeth, containing teeth without restoration and with dental appliance	45	28
10	Images missing teeth, containing teeth without restoration and without dental appliance	115	28

Data Limitations

There were several challenges in the data to consider as panoramic dental x-ray images can often contain inherent bias. According to Dhillon et al. (2012), 90% of panoramic dental x-rays have some form of error, for example changes in patient position or movement during exposure. Differences in image quality can also arise from acquiring images from different technologies, however, this is not significant if they meet a minimum standard as defined by the FDA for use as a medical diagnostic tool (FDA, 2007). Sampling from a limited source can exacerbate these challenges but the variety of data sources used in this work negated that concern.

Methodology

Feature Engineering

There was little feature engineering to be performed since the dataset is image data. Each image was normalized as not all x-ray machines are created equally; for example, a particular x-ray machine may produce an image of higher or lower contrast than others. Additionally, x-ray images are prone to user error- the images could be off-centered or too far away from the patient's teeth. Normalizing all the x-ray images helped to ensure they were as similar as possible.

The UFBA-UESC dental images dataset included dental masks for a portion of the images, however, YOLOv5 requires specifically formatted labels for training. Therefore, to train the model, bounding box coordinates were extracted from the dental masks and converted into YOLOv5 label format. Finally, as YOLOv5 requires its training labels to be normalized, all the YOLOv5 labels and corresponding images were normalized.

Modeling Framework

Mask R-CNN model development

The first modeling framework developed and tested was Mask R-CNN. Mask R-CNN produced sufficient and acceptable results as the initial model for instance segmentation of panoramic dental x-rays. Mask R-CNN, an extension of Faster R-CNN, includes a branch of convolutional networks that performs instance segmentation. Ultimately, this framework outputs fixed size features which each feature is first classified as a tooth or background, localized by regressing the bounding box coordinates, and finally segmented per-pixel by the fully convolutional network in each of the detected tooth bounding boxes (Jader et al., 2019). This model was trained using the UFBA-UESC dental images dataset with the output goal of detecting teeth within an x-ray. For the initial model, the same methodologies and modeling parameters were reproduced as in the aforementioned paper; the main adjustment was changing the data that this model was originally trained with. With the same methodologies and using our curated dataset, we were able to produce a strong, initial tooth detection model.

Mask R-CNN returns model metrics that are based on pixel-wise comparisons between the training masks and the detections output by the model. It was more relevant for this work to evaluate performance based on individual tooth regions detected by the model and complete manual evaluation. These model metrics proved useful later on to give an idea of how this model performed compared to other approaches.

You Only Look Once v5 (YOLOv5) model development

The second modeling framework developed and tested was YOLOv5. YOLOv5's initial weights, trained on the MS-COCO dataset, gave a strong start to building this model. This object detection algorithm divides a given image into a grid system, with each cell in the grid being responsible for detecting objects within itself; this means that this modeling framework is performing detections based on the global context. Using the UFBA-UESC dental images dataset in conjunction with panoramic dental x-rays from GLDP, the YOLOv5 tooth detection model was trained with 1500 epochs to obtain the detected objects (teeth) for each x-ray as bounding boxes. The bounding boxes of the detected objects were recorded along with their locations in the image.

The YOLOv5 model developed for tooth detection was extended to classify teeth as one of four types: molars, pre-molars, incisors, or canines. The outputted detections of the YOLOv5 tooth detection model were manually placed into its respective tooth type category. The weights from the final YOLOv5 tooth detection model were used to initialize this tooth classification

Deleted:

Commented [AS3]: You need more details on your implementation of Mask R-CNN. Any parameterization? Was it pre-trained, and did you use transfer learning? Did you have a custom header? How many epochs? When you say "predicting teeth", do you mean "predicting bounding boxes around teeth"?

Deleted: evaluation. This analysis was performed on a random 30% of the testing images to calculate recall and precision. …

Deleted: general,

Deleted: relative

Deleted: s

Deleted: ¶

Deleted: You Only Look Once v5 (YOLOv5). The YOLOv5 object detection algorithm outperformed Mask R-CNN primarily in both its speed and accuracy.

Deleted: YOLOv5 is a group of object detection architectures and models developed by Ultralytics (2020). We were able to achieve competitive results to the Mask R-CNN model through transfer learning. We utilized model with a strong foundation. Using the newly adjusted data, the tooth type classification YOLOv5 model had the same steps as the tooth detection YOLOv5 model and was retrained and optimized through non-maximum suppression.

Tooth numbering model development

Finally, three different tooth numbering algorithms were tested using results from the YOLOv5 tooth detection and type classification model to number teeth according to the dental ISO standard 3950 and identify missing teeth. The three modelling approaches were a rule-based algorithm, Bayesian neural network, and minimum Euclidean distance. In all three approaches, the center point of the x-ray was used to divide the mouth into quadrants. The rule-based approach then used various rules to number teeth per quadrant. These rules consisted of, but are not limited to, checking number of teeth per tooth type per quadrant, checking average y-axis location of teeth in each quadrant, and finding gaps between teeth in a quadrant to detect missing teeth. The second approach used the x-center and y-center of the tooth bounding boxes and tooth categories to train a Bayesian neural network to number teeth based on cross-entropy loss and kweight loss. The third approach matched the YOLOv5 bounding boxes with estimated tooth location, by calculating the minimum Euclidean distance between the bounding box and all estimated teeth. A list of missing teeth was generated based on these three methods and approximate locations for tooth number labels that are indicated on the x-ray with the appropriate numbering of gaps between existing teeth. Finally, a set of 175 randomly sampled images was used to calculate accuracy of the algorithm.

Optimization Techniques

There were two main optimization techniques utilized for developing the YOLOv5 model: early stopping (patience) and non-maximum suppression. Both of these methods proved

Formatted: Complex Script Font: Not Bold, Not Italic, English (US)

Commented [AS10]: All paragraphs (except for abstract) should be indented.

Commented [OF11]: The literature review talks about maximum numbering, should we connect the two?

Deleted: first

Deleted:

Deleted: using Bayesian neural-network

Deleted: consisted of using

Deleted: /

Commented [JX13R12]: sure, I am on it

Deleted: …

Deleted: uses …

Deleted: s

Deleted: euclidean

Commented [PK14]: Same here

Deleted: with

Deleted: when

Deleted: are available

Deleted: A

Commented [OF15]: Should we mention this if we aren't going into the random 30% for m rcnn?

to be useful in resource management and accuracy improvement. Firstly, patience is a built-in parameter to the YOLOv5 modeling framework. This parameter allows the training process to be stopped once a monitored metric has stopped improving. The early stopping parameter (patience) was set to monitor the mean average precision (mAP) of the model. Lastly, non-maximum suppression was applied to the detections during testing. YOLOv5 is unique in that each grid cell it is looking in to detect an object, YOLOv5 can propose two bounding boxes. Non-maximum suppression aims to eliminate this problem by comparing the confidence scores and overlap threshold of a list of proposal detections. This method ensured only one detection was obtained for the detected object in question.

Deleted: In our case, we chose our early stopping parameter to be…

Deleted: be

Deleted: based on

Deleted: ing

Deleted: we employed

Deleted: allows us to

Deleted: s

Deleted: obtain

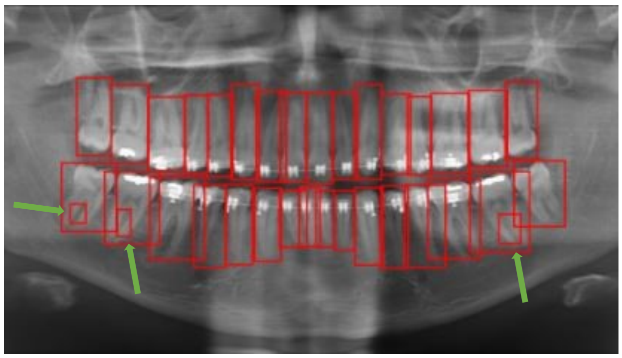
Deleted: is

Deleted: one

Findings

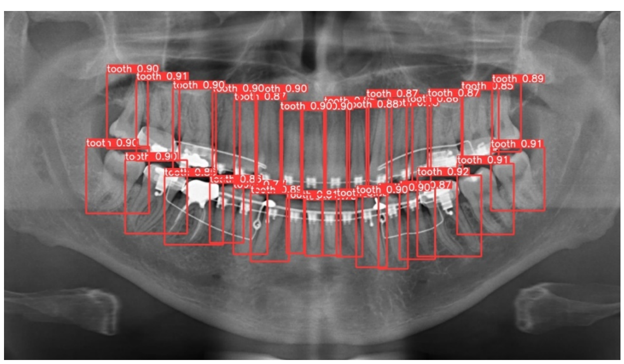
The Mask R-CNN model performed well with a precision score of 0.9372 and a recall score of 0.9664 on the test data, however, it often identified non-tooth objects as teeth which is reflected in the lower precision score. It was observed the inaccuracies fell into one of several categories: identifying a root as a separate tooth instance, identifying a foreign object (e.g. dental implants) as a separate tooth instance, grouping multiple teeth as one, identifying a gap from a missing tooth as a separate tooth, or not identifying a tooth entirely. Nearly 53% of the errors can be explained by either a root or a foreign object being identified as a tooth. Figure 3 demonstrates an example of multiple instances of incorrect bounding boxes on one x-ray; in this case three instances of tooth roots are detected as teeth.

Figure 3. *X-ray image with teeth detected by Mask R-CNN model highlighting errors*



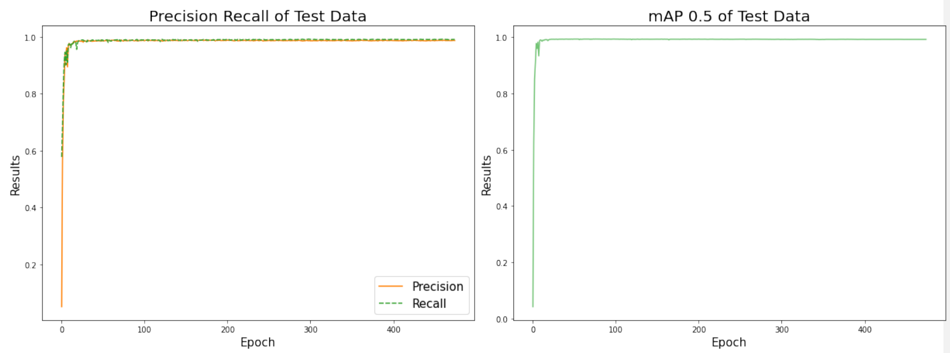
The initial YOLOv5 tooth detection model reliably identified individual teeth and segmented them with bounding boxes. As can be seen in Figure 4, each of the teeth in the example x-ray are identified with a red bounding box and the model has high confidence (greater than 0.85) for each one.

Figure 4. *X-ray with teeth detected by initial YOLOv5 tooth detection model*



The model was set to train for 1500 epochs but experienced early stopping after only 474 epochs once the performance stopped improving for more than 100 epochs. It had a recall score of 0.9908 and a mAP score of 0.9924 on the test data. Figure 5 shows how the model performance improved over training; the performance quickly stabilized within the first 100 epochs.

Figure 5. *Graphs of initial YOLOv5 tooth detection model evaluation metrics*



The secondary YOLOv5 model, trained to classify tooth type, was also set to train for 1500 epochs and experienced early stopping after 506 epochs. As can be seen in Figure 6, the model showed slower improvement than the initial YOLOv5 tooth detection model and stabilized after 300 epochs. The model achieved a recall score of 0.9448 and mAP score of 0.9613.

Figure 6. *Graphs of secondary YOLOv5 tooth classification model evaluation metrics*

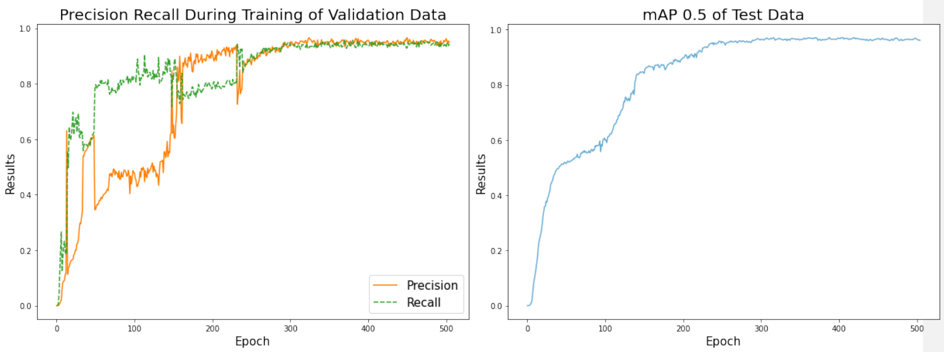


Figure 7 shows two examples of x-rays with the teeth detected and classified by the model, the confidence is high for each tooth (greater than 0.80) and the model performs well even when many of the teeth are missing.

Figure 7. *X-rays with teeth detected and classified by the secondary YOLOv5 model*

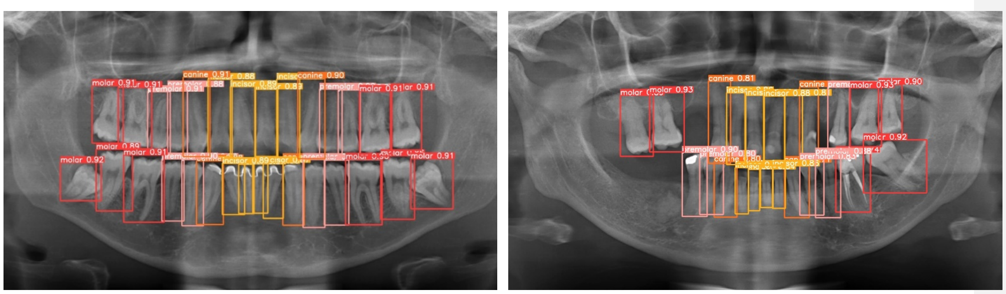


Figure 8 show the precision-recall curve for each tooth type. The model performs well for each type with AP of 0.94 or above. Canines were detected and classified the poorest, however, this was expected as out of 32 possible teeth only 4 are canines so there is a lot less data available to train the model.

Figure 8. *Precision-recall curve of each tooth type for secondary YOLOv5 model*

Deleted:

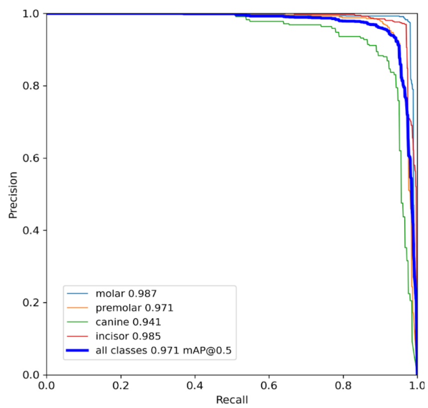


Table 2 shows a comparison between the performance of the different models that were developed. YOLOv5 had better performance than Mask R-CNN. Most notably, the precision score is significantly higher in the tooth detection model meaning instances of non-tooth objects being detected as teeth happened far less. Mask R-CNN performance struggled with images containing dental implant where it had a precision of 0.65 and recall of 0.90. Comparatively YOLOv5 had a precision of 0.96 and recall of 0.98 for those x-ray types.

Deleted: ¶

Table 2. *Comparison of metrics between CNN models*

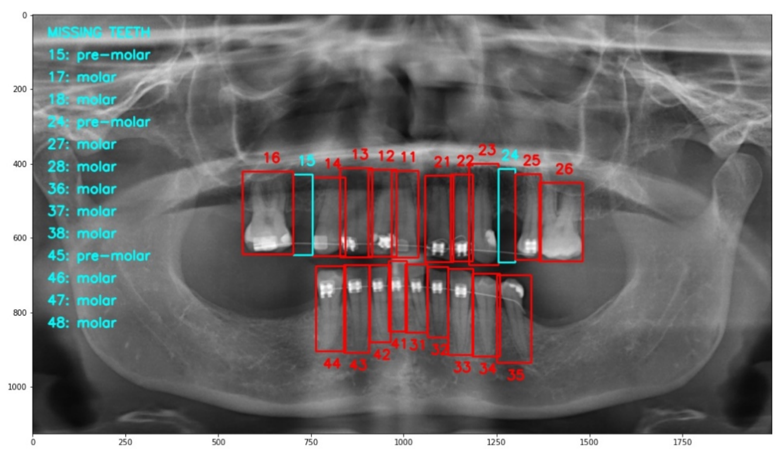
Model	Precision	Recall
Mask R-CNN	0.9372	0.9664
Initial YOLOv5 (tooth detection)	0.9870	0.9908
Secondary YOLOv5 (tooth classification)	0.9513	0.9448

In conjunction with the trained YOLOv5 models, three different tooth labelling approaches were tested to number existing tooth segments and identify missing teeth. As shown in Table 3, the rule-based approach performed the best out of the three algorithms, with an accuracy score of 91.48%. A sample image is shown in Figure 9, where the red boxes mark and label teeth numbers of teeth present in the x-ray image, while the cyan boxes show gaps in the teeth with corresponding missing tooth numbers. A complete list of missing teeth numbers is listed on the upper-left corner of the x-ray.

Table 3. *Comparison of accuracy across different tooth numbering models*

Model	Accuracy
Rule-based	0.9148
Bayesian neural network	0.7342
Minimum Euclidean distance-based	0.6460

Figure 9. *X-ray labelled with rule-based tooth numbering by ISO 3950 standard*



Discussion

The primary goal of this research was to train a model that can accurately detect teeth and classify tooth type in an extra-oral panoramic dental x-ray of an adult which was achieved with a YOLOv5 model. The secondary goal of numbering teeth and identifying missing teeth was accomplished by using the YOLOv5 model output combined with a rule-based algorithm.

Although Mask R-CNN was originally believed to be the solution, this work shows that YOLOv5 produced better results. Mask R-CNN also has other challenges such as requiring dental masks to be created for the x-rays by highly trained dentists to train the model, which reduced the number of x-ray images that it was possible to use to train the model. YOLOv5 uses a text file containing bounding box data that could be generated using available open-source annotation tools, and further edited with user relevance feedback, so it was possible to quickly generate labelled data for the GLDP dataset and combine it with the UFBA-UESC dental images dataset to improve the training dataset.

As Mask R-CNN performs pixel-wise comparisons, it struggles when objects are overlapping in an image, this frequently occurs in dental x-rays, especially panoramic ones since they are a 3-D image converted to 2-D, as discussed previously. The detection methods used by YOLOv5 enabled it to overcome this issue experienced by Mask R-CNN, and likely accounts for some of the difference in performance. YOLOv5 informs its detections of the global context surrounding a tooth, as it is performing its detections in a grid, cell by cell. With this a measure of recall is lost and more localization error is gained, but accuracy and generalizability through region-based detections is also obtained. One limitation of the YOLOv5 model to note is it will not work as well on x-rays containing infant teeth, as it was only trained using permanent teeth.

The rule-based algorithm for tooth numbering, while performing the best out of the three approaches tested, has certain limitations because of the non-uniformity of tooth size and structure across different patients. It performs well in most cases where most of the teeth are present in the x-ray, however, it sometimes struggles to accurately label teeth and suggest missing teeth locations where there are multiple consecutive teeth missing. Another source of errors arise from the results of the YOLOv5 model for tooth classification. Any errors generated

Deleted: ;

Deleted: s

Deleted: previous

by this model get propagated forward with the rule-based approach. Combining panoramic and bitewing images together would reduce this problem and increase numbering accuracy, as it would increase the accuracy of tooth classification.

Dental AI is newly developing and there are a very limited number of products currently available in the field for tooth detection and classification. Despite the constraints of this work, this approach can be used as a benchmark going forward to improve upon existing limitations, while also opening doors to other possibilities of further expansion into products such as automatic dental charting, that would not be possible without tooth detection and numbering models.

Conclusion

This work explains and showcases a final model that uses a combination of a custom trained YOLOv5 model and a rule-based algorithm to detect and classify teeth then identify existing and missing teeth within a panoramic dental x-ray. The next step in the pipeline is combining our model's detection outputs with a closely related project to detect and classify anomalies, and producing a GUI where dentists could submit a panoramic dental x-ray and instantly receive back suggested tooth locations with dental anomalies. This will help dentists to catch mistakes, increase efficiency, and provide a second opinion to validate their work.

There are many possible extensions to this work including expanding to other types of dental x-rays. Bitewing images show a zoomed in view of the upper and lower teeth of one part of the mouth. They can show a clearer view of specific teeth and the gums than panoramic x-rays and therefore help to make more accurate diagnosis. Combined with the work already finalized, this could potentially revolutionize the dental field.

Deleted:

Deleted: ¶

References

Commented [PK16]: Check formatting for this section before submitting

Chung, M., Lee, J., Park, S., Lee, M., Lee, C. E., Lee, J., & Shin, Y. G. (2020). Individual Tooth Detection and Identification from Dental Panoramic X-Ray Images via Point-wise Localization and Distance Regularization. *Artificial Intelligence in Medicine*, *111*. https://doi.org/10.1016/j.artmed.2020.101996

Denti.AI. (n.d.). Retrieved May 1, 2021, from https://www.denti.ai/

Dhillon, M., Raju, S. M., Verma, S., Tomar, D., Mohan, R. S., Lakhanpal, M., & Krishnamoorthy, B. (2012). Positioning errors and quality assessment in panoramic radiography. *Imaging Science in Dentistry*, *42*(4), 207–212. https://doi.org/10.5624/isd.2012.42.4.207

FDA. (2007). *Performance Standard for Diagnostic X-Ray Systems and Their Major Components*. FDA Guidance Document. https://www.fda.gov/regulatory-information/search-fda-guidance-documents/performance-standard-diagnostic-x-ray-systems-and-their-major-components-21cfr-102030-102031-102032

Gurses, A., & Oktay, A. B. (2020). Human identification with panoramic dental images using mask R-CNN and SURF. *5th International Conference on Computer Science and Engineering, UBMK 2020*, 232–237. https://doi.org/10.1109/UBMK50275.2020.9219513

Hello Pearl. (n.d.). Retrieved May 1, 2021, from https://www.hellopearl.com/

Inference Analytics. (n.d.). Retrieved May 1, 2021, from https://www.inferenceanalytics.com/

ISO 3950:2016. (n.d.). Retrieved May 27, 2021, from https://www.iso.org/standard/68292.html

Jader, G., Fontineli, J., Ruiz, M., Abdalla, K., Pithon, M., & Oliveira, L. (2019). Deep Instance Segmentation of Teeth in Panoramic X-Ray Images. *Proceedings - 31st Conference on Graphics, Patterns and Images, SIBGRAPI 2018*, 400–407. https://doi.org/10.1109/SIBGRAPI.2018.00058

Formatted: Italian

Formatted: Italian

Formatted: Finnish

ORCA Dental AI. (n.d.). Retrieved May 1, 2021, from https://www.orca-ai.com/

Overjet. (n.d.). Retrieved May 1, 2021, from https://www.overjet.ai/

Overjet. (2021). *In a Dental Industry First, Overjet's Dental Assist Receives FDA Clearance to Bring AI to the Dental Practice*. Cision - PR Newswire. https://www.prnewswire.com/news-releases/in-a-dental-industry-first-overjets-dental-assist-receives-fda-clearance-to-bring-ai-to-the-dental-practice-301295927.html

Panoramic Dental X-ray. (n.d.). Retrieved May 1, 2021, from https://www.radiologyinfo.org/en/info/panoramic-xray

Schwendicke, F., Golla, T., Dreher, M., & Krois, J. (2019). Convolutional neural networks for dental image diagnostics: A scoping review. In *Journal of Dentistry* (Vol. 91). Elsevier Ltd. https://doi.org/10.1016/j.jdent.2019.103226

Silva, G., Oliveira, L., & Pithon, M. (2018). Automatic segmenting teeth in X-ray images: Trends, a novel data set, benchmarking and future perspectives. *Expert Systems with Applications*, *107*, 15–31. https://doi.org/10.1016/j.eswa.2018.04.001

Tuzoff, D. V., Tuzova, L. N., Bornstein, M. M., Krasnov, A. S., Kharchenko, M. A., Nikolenko, S. I., Sveshnikov, M. M., & Bednenko, G. B. (2019). Tooth detection and numbering in panoramic radiographs using convolutional neural networks. *Dentomaxillofacial Radiology*, *48*(4). https://doi.org/10.1259/dmfr.20180051

Ultralytics. (2020). *YOLOv5 Documentation*. https://docs.ultralytics.com/

Zhu, G., Piao, Z., & Kim, S. C. (2020). Tooth Detection and Segmentation with Mask R-CNN. *Undefined*, 070–072. https://doi.org/10.1109/ICAIIC48513.2020.9065216

Formatted: French

Formatted: German