Instance Segmentation for Detecting Dental Caries in Panoramic X-rays using Detectron2

MSc Research Project
Data Analytics

Kajol Daiya
Student ID: x19216831

School of Computing
National College of Ireland

Supervisor: Hicham Rifai

National College of Ireland
Project Submission Sheet
School of Computing

Student Name:	Kajol Daiya
Student ID:	x19216831
Programme:	Data Analytics
Year:	2021
Module:	MSc Research Project
Supervisor:	Hicham Rifai
Submission Due Date:	31/01/2022
Project Title:	Instance Segmentation for Detecting Dental Caries in Panoramic X-rays using Detectron2
Word Count:	XXX
Page Count:	19

I hereby certify that the information contained in this (my submission) is information pertaining to research I conducted for this project. All information other than my own contribution will be fully referenced and listed in the relevant bibliography section at the rear of the project.

ALL internet material must be referenced in the bibliography section. Students are required to use the Referencing Standard specified in the report template. To use other author's written or electronic work is illegal (plagiarism) and may result in disciplinary action.

Signature:	Kajol Daiya
Date:	31st January 2022

PLEASE READ THE FOLLOWING INSTRUCTIONS AND CHECKLIST:

Attach a completed copy of this sheet to each project (including multiple copies).	☐
Attach a Moodle submission receipt of the online project submission, to each project (including multiple copies).	☐
You must ensure that you retain a HARD COPY of the project, both for your own reference and in case a project is lost or mislaid. It is not sufficient to keep a copy on computer.	☐

Assignments that are submitted to the Programme Coordinator office must be placed into the assignment box located outside the office.

Office Use Only	
Signature:	
Date:	
Penalty Applied (if applicable):	

Instance Segmentation for Detecting Dental Caries in Panoramic X-rays using Detectron2

Kajol Daiya
x19216831

Abstract

Dental disease is referred to as ”silent disease” since it does not cause pain until it has progressed to an advanced level. Dental disease is mostly avoided, but if not detected early, it can develop into periodontal infections and pus. For detecting oral illnesses, dentists rely solely on visual assessment using radiological images. Unfortunately, these radiographs have several drawbacks, including poor image quality, a low diagnosis rate, and a long processing time. This study sought to assist dentists by performing instance segmentation on panoramic Xrays for detecting five classes of caries. Faster RCNN R101-FPN and Faster RCNN X101-FPN pre-trained models are implemented and evaluated to examine the accuracy of the proposed dental caries detection model. The model's average precision @ IOU is 53.512 for segmentation and 66.18 for bbox, this determines how well the detecting bounding boxes match the ground truth bounding boxes. For the development of the dental care industry, the proposed system implemented cutting-edge computing algorithms and compared their results. The results of the experiments reveal that the Detectron2 model has proven to be accurate at recognizing five classes of dental caries on panoramic Xrays.

1 Introduction

Dental caries is a disease in which bacteria and plaque cause tissue injury on the tooth's enamel, which then spreads to the pulp and is by far the most common chronic disease worldwide. Though the number of major cavity lesions in people has decreased significantly, most people still have early lesions that need to be detected and treated. As per the recent dental health report, the majority of people have caries, and half of them ignore their dental caries, increasing the risk of other major dental diseases and orofacial suffering. Dental caries can be classed as secondary caries, proximal caries, rootpiece, caries involving pulp, and dentinal caries depending on the degree of the lesion. (Satcher 2017) According to the American Dental Association, even though no one dies from dental disease, dental issues can have a substantial impact on an individual's life, and apparently, a healthy mouth is a reflection of the overall health and well-being of an individual.
Dental x-rays images are a useful resource for dentists to accurately detect cysts, tumors, fractures, and other dental diseases that require more information than can be obtained by physically examining the patient. However, reading an x-ray is a difficult task that normally needs

years of training and expertise before a dental practitioner can provide accurate results. The analysis and processing of dental x-ray images are not only important for diagnosing but also for treating and investigating the nature of dental diseases, as well as identifying dental diseases in their early stages. Na'am et al. (2016) In the dental analogies, dental x-ray images capture many perspectives. These dental radiographs aid in the detection of embedding fragments, infection, fractured bones, lesions, and mandible issues. It is quite challenging for any dental practitioner to correctly detect dental caries with naked eyes or with dental diagnostic radiographs like Panoramic, Bitewing, or Periapical as they have some drawbacks. Kassebaum et al. (2015) When it comes to identifying dental caries, bitewing x-rays are the most prevalent, but they have some significant downsides, such as pain and disability during the bitewing x-ray and greater dosage of radiation. While panoramic radiographs use a lower radiation dosage, this image captures 's the entire mouth of an individual, including all of the patient's bones and jaw structure, making exact analysis of each tooth difficult. Panoramic radiography contains several flaws that result in poor analysis. According to different studies, dental diagnostic radiographs falsely classify 20% of atypical lesions as dental caries and are not very accurate. Another possible problem is that there are numerous types of x-rays, each with its own set of characteristics and challenges, making screening difficult.

1.1 Research Motivation and Background

We can argue dentistry is solely just toothaches, yet everyone has regular dental problems, and everyone has visited a dentist. In most European and Asian countries, the dental visits rate is 70-80 per year, and nearly every patient gets an OPG x-ray. Dental x-rays account for the majority of x-rays taken around the world. The dentist does the x-ray diagnosis manually, which is biased and time-consuming. On panoramic radiographs, three out of four caries lesions, early caries lesions, are often missed by dentists. Deep learning algorithms have distinct advantages that we may take advantage of if we adjust them to the problem case study. We have a lot of dental data to help us get more accurate in our estimations and to be able to work on untreated and early interdental caries, which are often missed by dentists since they are not professional radiologists and panoramic x-rays are difficult to interpret.
This dental caries diagnostic model will help dentists to acquire accurate caries detection which is missed with naked eyes on panoramic x-rays. In addition, this study presents a reliable object detection/ instance segmentation approach that may be used for various other dental disease detection/segmentation problems.

.

1.2 Research Question

Skin cancer, knee cartilage, diabetic retinopathy, pulmonary tuberculosis, and brain tumors are among the conditions for which deep learning CNN models are utilized for detection and classification. These models have demonstrated good precision, efficiency, and clinical potential in a variety of domains. However, there has been little research that uses deep CNN model architectures to examine detection and segmentation tasks in dentistry. For the development of the dental healthcare industry, this research experimented with a cutting-edge computing model Detectron2 for instance segmentation to accurately detect dental caries into 5 Classes: Dentinal Caries, Caries involving pulp, Rootpiece, Proximal Caries, and Secondary Caries.

Research Question: *"To what extend the state of art algorithm (Detectron2) can implement instance segmentation for precisely detecting dental caries in Panoramic x-rays?"*

1.3 Research Objectives and Contributions

In the Table 1, in order to investigate the research case study, the following objectives are set.

Table 1: Research Objective for Dental Caries Detection

Objectives	**Descriptions**	**Evaluation Method and Matrices**
1	Data gathering	
2	Data manual Annotation	
3	Data Pre processing	
4 4.1 4.2	Selecting Deep Learning algorithm Implementation of Detectron2 Comparison of Faster Mask RCNN models with Detectron2 default base model parameters	mAP, IoU, Accuracy
5	Results	

Major Contribution It's noteworthy that no previous research has looked into dental caries detection and classifying caries into 5 stages using Panoramic x-rays with deep learning model Detectron2. With this in mind, this research proposes a dental caries detection tool and evaluates the performance of Detectron2 which specializes in instance segmentation. This is accomplished by experimenting with the efficiency of Detectron2 for instance segmentation/object detection tasks on base models faster RCNN R101 FPN and faster RCNN X101 FPN. Thus this paper contributes; by experimenting with Facebook AI library Detectron2 for segmentation tasks for the development of the dental care industry.

The following is a breakdown of the research: Sections 2 and 3 review the literature on segmentation of teeth, caries detection using deep learning algorithms, Section 4 describes the implementation, evaluation, and results obtained from the deep learning models, Section 5 describes the discussion and Section 6 concludes the research work.

2 Critical Research on Dental Caries Detection and Teeth Segmentation (2012 - 2022)

Deep learning techniques improved the performance of automatic dental image analysis and dental caries detection. Due to the shortcomings of traditional methods on complex and difficult dental x-ray images, the accuracy of the deep learning techniques is quite remarkable on panoramic dental x-rays, Bitewing and RVG. Following that, this section will go over the studies on dental caries detection and segmentation of teeth in more detail from the last 10 years

2.1 Dental Image Processing and Segmentation Research

A number of recent research studies have introduced deep learning-based Computer-Aided Diagnosis (CAD) to detect caries on different dental x-ray radiographs. Jader et al. (2019) This work although did not perform teeth numbering task, were the first to examine the identification and segmentation of teeth on panoramic X-rays. They created the UFBA-UESC Dental Images Deep data set by modifying the UFBA-UESC Dental Images data set to incorporate information about tooth instances. The teeth of the binary masks that represented the full dental arch were manually separated. With 193 and 83 images, respectively, a Mask R-CNN with a ResNet-101 backbone was trained and tested. The finished network was then tested in binary pixel-wise way in the remaining 1224 images of the data set, yielding an F1 score of 88 percent. For dental image processing, technique CLAHE was used to improve the contrast of panoramic x-rays, which have low resolution and brightness. Bilateral filters were employed to keep the edges while also strengthening the picture resolution, according to Pandey et al. (2017). The drawback with this paper was that it did not cover segmentation of overlapping teeth or the use of segmentation methods to detect dental illness such as caries. Al-Sherif et al. (2012) suggested a method for teeth segmentation utilizing a Bitewing x-ray radiograph and an Energy-based algorithm, which achieved a higher accuracy rate than previous studies on teeth extraction and segmentation. The Gray Level Co-occurrence Matrix (GLCM) technique was used by Jusman et al. (2020) to analyze feature extraction performance of dental caries in x-ray images. This method is used to evaluate the pixels and quantization parameters of the GLCM, which is employed in an efficient dental caries classification system. Optimized segmentation methods by combining two phases, IC (Initial Contour) creation and intelligent level set segmentation, to get more accurate results. The segmentation approach was tested on 120 dental radiographs (X-rays) and found to be 90 percent accurate. Furthermore, the accuracy of dental caries identification is tested on a segmented image 155 by 98 percent. Prior to 2018, most academic research on automatic analysis of dental X-rays relied on automatic feature extractors. Due to the difficult nature of Panoramic x-ray images, the most of these works neglected these X-rays. In this research we use Panoramic x-rays to perform semantic and instance segmentation, the U-Nets suggested by cite Ronneberger is likely the most well-known FCN design for medical semantic segmentation. Hence, for semantic segmentation of dental panoramic radiographs, this study investigated FCNs based on U-Nets for semantic segmentation task.

2.2 Dental Caries Detection using Deep Learning Techniques Research

Convolutional neural networks (CNNs) have recently been used in deep learning-based approaches to achieve breakthrough outcomes above traditional methods. Modern CNN-based detection techniques can be divided into two categories which have two types of approaches: anchor-based and point-based. Zhou et al. (n.d.) To localize each object, anchor-based algorithms use exhaustive classifications on predefined anchor boxes and often use a non-maximum suppression strategy. Point-based object detection, on the other hand, tries to regress points to designate objects like the center point. For accurate object detection, multiple crucial points (e.g., the left-top and right-bottom corners) are synchronously regressed in addition to the central axis. According to recent studies, point-based procedures are more accurate and efficient than anchor-based methods in terms of effectiveness and precision Duan et al. (2019). Musri et al. (2021) set out to demonstrate how effective CNN methods are at detecting and diagnosing early dental caries on periapical x-rays.

Leopold et al. (2021) employed five convolutional neural networks to address the problem of dental caries detection: Resnet-152, Xception, AlexNet, VGG-16, and ResNext 101. One of the most important flaws in this suggested study was the manual image labeling techniques. The image registration technique, which was employed in neuroimaging, is one promising alternative offered by Leopold et al Chen et al. (2019) was successful in creating a neural network model based on fast R-CNN that autonomously annotates tooth in an x-ray with high accuracy scores that were comparable to those of a young dentist. This study found mix-matched annotation, as well as inaccurate and poor recognition of two partial teeth as complete teeth. A experimental work on caries detection in third molars is presented in Vinayahalingam et al. (2021) The researchers used a series of 100 Panoramic photos that had already been cropped to detect cavities in the third molars. The accuracy, sensitivity, and specificity of this automatic model were all 87 percent, 87 percent, and 86 percent, respectively. Using the MobileNet V2 model, AUC of 0.90 was calculated for detecting caries in third molars. The exploratory study had a flaw in that it contained cropped photographs of the third molars for detecting dental disease. Lee's research demonstrated a CNN model for dental caries diagnosis on bitewing radiographs using a U-shaped deep CNN (U-Net) as well as how the method can improve clinician efficiency. Moutselos et al. (2019) demonstrated that a deep learning model can classify dental occlusial caries using periodontal radiographs without requiring data pre-processing methods. For object mask, this study used deep neural networks Mask (R-CNN) extending Faster R-CNN. The researchers in this work used data augmentation to manage overfitting and transfer learning to improve prediction results.Deep learning-based object detection is growing rapidly, and there are numerous studies in the literature on the subject. The recent approaches can be divided into two groups: Region-Based Convolutional Neural Networks (R-CNNs) and (YOLO).Zhou et al. (n.d.) Because R-CNNs are trained in phases, it takes a long time to train them. Aside from training, the prediction stage takes a long time. To address these challenges, Girshick suggests a new model dubbed Fast R-CNN Girshick (2015). Instead of three independent modules, Fast R-CNN is trained as a single model. This architecture examines the photos and provides candidate regions, then extracts features from the candidates using a popular, pre-trained image classification model. Following that, the collected features are pooled in a Region of Interest (RoI) layer, which is followed by two fully linked layers. Finally, two more fully linked heads are available for bounding box regression and label classification.

2.3 Analytical Performance for Dental Caries Classification and Detection Research

Detection-based algorithms are more straightforward than Segmentation based algorithms, but they are more limited in scope because they only estimate a specific value for the image without

Table 2: Summary on Dental Caries Detection Techniques and Teeth Segmentation Literature Review

Authors	**Year**	**Algorithm Applied**	**Radiograph Type**	**Problem Case**	**Limitations**
Leopold, H. A., Singh, A., Sengupta, S. Lakshminarayanan	2021	Resnet, Xception, AlexNet, VGG16, ResNet	optical coherence tomography	Dental Caries Detection	Manual labeling
Chen, H., Zhang, K., Lyu, P., Li, H., Zhang, L., Wu, J	2019	Fast R-CNN	Periapical	Caries Detection	Flawed recognition of two half teeth as a whole tooth.
Musri, N., Christie, B., Ichwan, S. J. A.	2021	CNN	Periapical	Early Caries Detection	NA
Vinayahalingam, S., Kempers, S., Limon, L.,	2021	MobileNet V2	Panoramic	Caries Detection on the third molar	Included cropped panoramic x-rays
Park, K.J Kwak	2019	AlexNet GoogleNet	Panoramic	Teeth Segmentation	NA
Nadler, C	2019	Weighted K-NN	Panoramic	Dental Restoration; Segmentation	Relevantly small data set used.

indicating the disease's regional distribution. Whereas Segmentation algorithms, can provide a pixel-by-pixel identification of the given image. Pixel wise binary classification is commonly used in such models. This research done by Prajapati et al. for the classification of dental diseases was done on a short dataset, the transfer learning approach was used to achieve good accuracy in spite of data set size constrain. There were 251 x-rays in all, divided into three dental diseases: dental caries, Periapical Infection, and Periodontitis. VGG16 was utilized for image resizing and extraction of features. Transfer learning performed well in the classification task, however Convolutional neural networks did not due to the limited data set size.Haghanifar et al. (2020) designed a study on Panoramic radiographs for tooth extraction using a generic algorithm method. The ROI is used to detect and extract the jaws in order to extract a specific region from the data used by the author. There were 42 OPG photos in the dataset used in this study. For maxillary and mandibular, or upper and lower jaws, this mechanical teeth extraction model achieved an overall accuracy of 0.77.

Duong et al. (2021) recently developed an artificial intelligence integrated smartphone apps that gave patients an important indicator of their oral health. This study used mobile phone color photographs to recognize and classify and detect dental caries. A total of 620 images were collected and processed for this investigation. Researchers have previously employed a variety of machine learning methods to solve medical problems, but no previous dentistry research on caries has used mobile phone photographs as raw data for machine learning algorithms. Support Vector was the authors choice in this study because SVM classifiers are resistant to overfitting and determines their model size. The evaluation matrix for the SVM classifier is shown in Table 4: accuracy, sensitivity, and specificity.

Table 3: Matrix for evaluation Duong et al. (2021)

$Accuracy$	$Sensitivity$	$Specificity$
91.37%	87.1%	97.6%

Table 4: Performance based on accuracy score Prajapati et al. (2017)

$Algorithm$	$AccuracyScore$ (%)
CNN	0.7207
$TransferLearning$	0.8746
$TransferLearningwithfinetuning$	0.8746

3 Methodology and Design Workflow

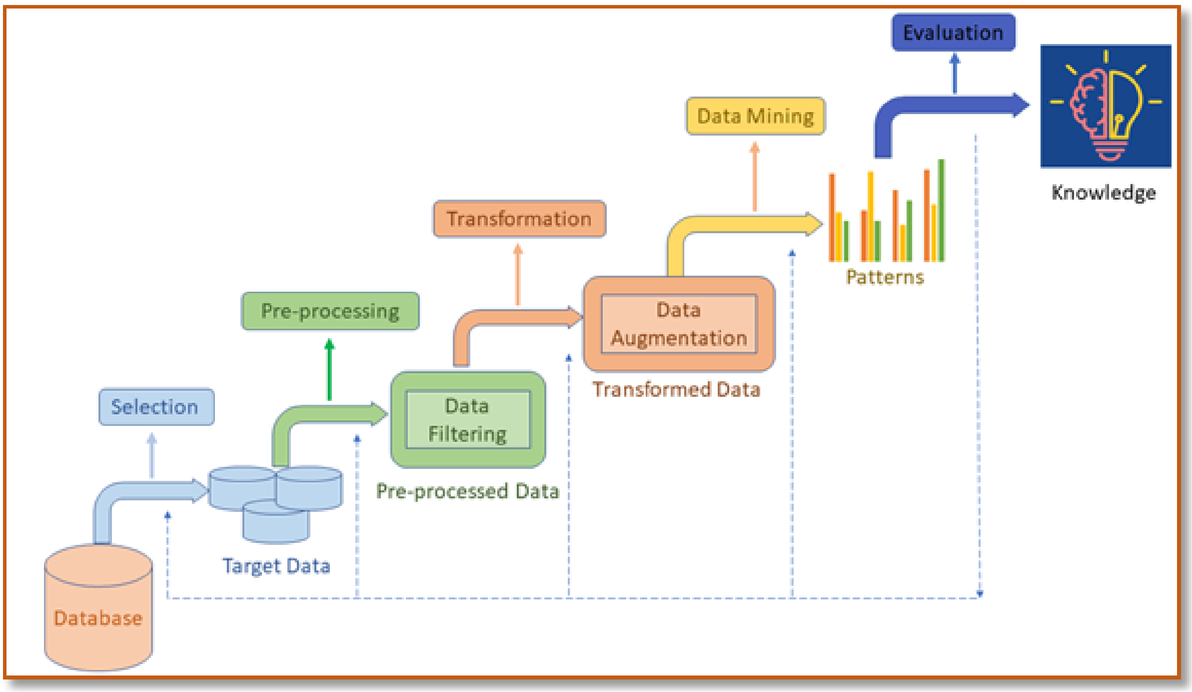


Figure 1: Dental Caries detection: KDD

To follow a structured data mining procedure, this research applies Knowledge Discovery in Databases (KDD) methodology starting with data collection, exploration of data to understand the data set. Proceeding with data pre-processing to improve the quality of panoramic x-ray images using pre-processing techniques like Gaussian Thresholding and CLAHE. Image resizing and other Data Augmentation are implied to transform the data in the Data Transformation stage. After that, splitting the training set into a test set allows us to quantitatively analyze the hyper-parameters for the model architectures. Beginning with the most often used deep learning architectures for dental caries detection and classification tasks, the strategies are used to enhance the base models by gaining knowledgeable insights, such as adjusting hyperparameters and training data augmentations.

3.1 Methodology

This study uses the knowledge discovery in database (KDD) methodology for segmentation and detection tasks. In the case of dental data, KDD approach may offer advantages over traditional statistical methods according to Gansky (2003) KDD is a critical method for identifying underlying trends and findings deep learning algorithms, according to various studies in Data Science. This methodology is a significant technique since it has the ability to combine numerous ways into one in order to develop an approach to determine sensible solutions.

Data Selection: The data set for this study came from a publicly accessible open-source platform. The dataset contains 116 panoramic x-rays with their relevant masks. The OPG X-ray covers the full region of the patient's mouth. The dental caries type is classed into 5 categories: Dentinal Caries, Proximal Caries, RootPiece, Caries involving pulp, and Secondary Caries, this is annotated manually in this dataset in coordination with 5 dental practitioners to attain accurate labels.

Data Pre-processing: Dental X-ray images tend to be quite noisy coming from various types of noise sources. To de-noise these x-ray images, denoising filters are used to highlight useful details in the x-ray and increase its image quality. Image thresholding and equalization are some of the tools that we have for image processing so image thresholding for segmentation tasks becomes a bit easier. With histogram equalization, we can stretch the histogram to span the entire range. Histogram Equalization considers the global contrast of the image, not just the local contrast. The result of Histogram equalization and Contrast Limiting Adaptive Histogram Equalizer was taken out. CLAHE does histogram equalization in small patches and it works very well and does contrast limiting.

1. CLAHE: A tool for enhancing image quality. To decrease noise, it works on pixel regions to identify the background and foreground separately.

2. Gaussian Filter: It distributes the data regularly between 0 and 1, which means it converts an image with a mean of 0 and a standard deviation of 1 using the formula below.

   $$g(a) = \frac{1}{\sigma\sqrt{2a}} e \frac{-a^2}{2\sigma^2}$$

   Here a denotes initial parameter, standard deviation is denoted by sigma, and variance is denoted by sigma square..

3. Thresholding: Every pixel value is compared to the target value in thresholding. It is taken as zero if the pixel value is less than the threshold, else it is assigned to a maximum value (generally 255). Thresholding is a widely used segmentation technique for distinguishing between foreground and background objects.

It can be seen in Figure 2, It is visible here that the CLAHE image provides a more detailed image compared to equalized image and in figure 3, binary threshold, binary inverted, TRUNC, TOZERO, and TOZERO inverted techniques are produced to perform image processing on the panoramic x-ray dataset.

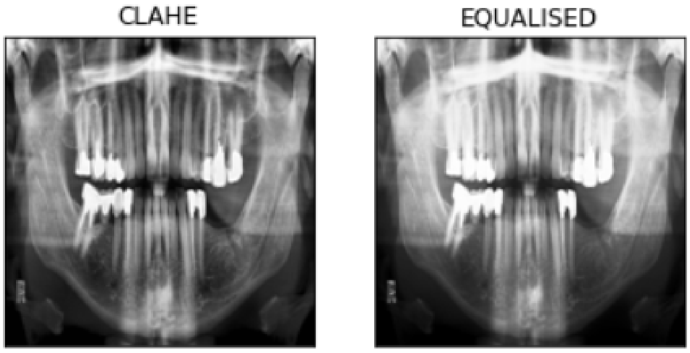


Figure 2: CLAHE and Histogram Equalization

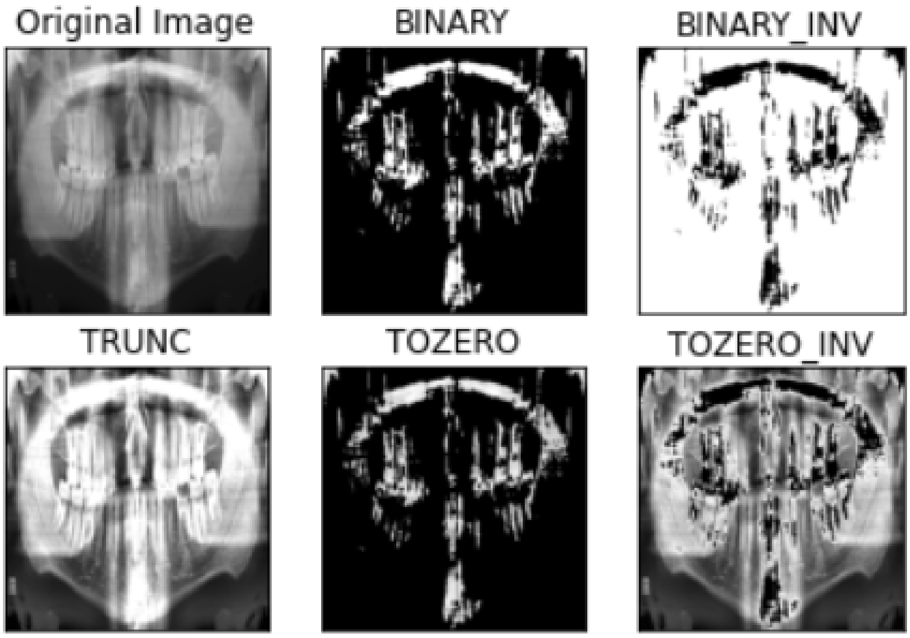


Figure 3: Thresholding

Data Transformation: Data augmentation is very useful to augment the data or increase the amount of training or validation data. Have chosen an image size of 256x 256 pixels, collect the images with LabelMe; an annotation tool built for annotating images online for research in computer vision. The dataset in this study contains only 116 images which are split into a train(70%), test(20%), validation(10%). The training dataset contains 93 relatively small images, also these images are distributed unevenly among 5 classes of dental caries. Therefore, image resizing, scaling, horizontal flipping, shear range, and zoom range are applied to address the small image dataset problem.

Feature Extraction: Feature Extraction is a method for reducing the number of features while preserving essential data and reducing duplicate data. This approach is important since it reduces computing time and improves deep learning skills. The mean, standard deviation, contrast, dissimilarity, and entropy were all collected from a panoramic image dataset in this study.

Data Mining: The Deep convolutional network Faster RCNN is mostly used by many researchers to detect dental caries, Faster RCNN network is used for object detection tasks. Hence, in this study Detectron2 to speed up our development process instead of creating a Faster R-CNN model from scratch. Detectron2 is a next-generation software system from Facebook AI Research that uses cutting-edge object detection algorithms. It's also usual to employ a feature extractor element of an architecture that's already been pre-trained on the image set, the feature extractor used here is ImageNet for the model to learn important features from the training image set. Many of these base models are available in the Detectron2 algorithm. R101-FPN1 and X101-FPN2 are two regularly used foundation models for Faster R-CNN. This study has chosen to investigate these two pre-trained models because, as compared to others, they had a higher Faster R-CNN box Average Precision (AP). On the pre-trained dataset (ImageNet). X101-FPN has a superior box AP on the ImageNet benchmark, but it takes longer to train/predict and may be overfitting in some circumstances. Considering this, this study also looks into R101-FPN.

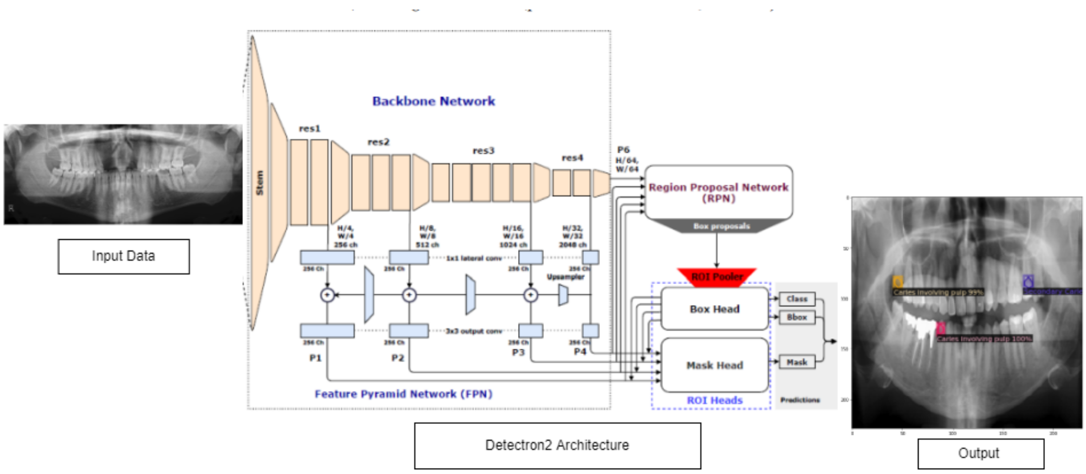


Figure 4: Detectron2 Architecture

Detectron2 is a renowned modular computer vision model library based on PyTorch. Wu et al. (2019) It is a two-stage network with three primary blocks: a Backbone Network, a Region Proposal Network (RPN), and an ROI head, as illustrated in extract extracted features from the source images. A stem block and four bottleneck blocks make up the ResNet model. The stem block, which has 77 convolution layers and a stride of 2, is employed. After that, the input image is downsampled twice using a max-pooling layer with a stride of 2. The stem block's output feature map is 64 H/4 W/4, where H and W denote the input image's height and width. The four bottleneck blocks are taken from the ImageNet base model ResNet model. The FPN is made up of the side and output convolution layers, as well as the four output characteristics maps from the ResNet bottleneck blocks (res1, res2, res3, and res4). 1x1

convolution layer is utilized for each side convolution layer. It converts 256 channel feature maps from output features from bottleneck blocks with varying channel numbers (256, 512, 1024, and 2048). From the res4 output, the FPN performs a forward process as shown in the Detectron2 Architecture (Figure 4) following that, a 3 3 output convolution layer is employed without modifying the channel numbers. P4 is the resultant feature map list. The res4 output is placed into the upsampler and lateral convolution is used to combine it with the res3 output. The generated feature map is also input into the output convolution and assigned the number P3. The technique is done twice, with the resultant feature maps labeled P2 and P1. Using a max-pooling layer with a stride = 2, the final P5 output feature map is simply a downsample of the res4 result. The ROI head block is made up of two distinct heads: a box head and a mask head. The ROI pooling procedure is used to feed the box recommendations into the box head. The class and bounding box estimation scores are the box head's ultimate outputs. The four output features maps from FPN, on the other hand, are sent into the mask head along with the box head's output. The output object's segmentation mask is mapped as a result of the prediction (Panoramic xray). Detectron2's final output image has three prediction maps for the object's class (object level identification), bounding box (location), and segmented mask (pixel level categorization).
Evaluation: The process of aggregating several input/output pairs is known as evaluation. In this study, two evaluators: COCOEvaluator and SemSegEvaluator are used to evaluate the evaluation matrix mAP, AP, and API for box detection, instance segmentation, and keypoint detection on the dental panoramic x-ray dataset. The evaluation matrix usually for segmentation tasks is chosen to be Precision and Recall.

- **Precision and Recall:** The capacity of a model to recognize only relevant things is referred to as precision. It is the proper percentage of true positives. Recall refers to a model's ability to locate all relevant cases (all ground-truth bounding boxes). Out of all available ground facts, it's the proportion of correct optimistic projections.To determine the precision and recall values, each detected bounding box must be classified using the IOU = 0.5 thresholds: True positive, False Positive, False Negative.

- **mAP:** The conventional precision metric used in image classification tasks is inapplicable for object detection/ instance segmentation tasks, unlike binary classification. This is why the mAP (Mean Average-Precision) algorithm is used to evaluate the results of the object detection/instance segmentation task in this study.

- **IoU:**Intersection over Union (IOU) is a measure of how much two bounding boxes overlap. It is used in computer vision to accurately identify an object.If the IOU is more than 0.5, the predicted bounding box is rated appropriately. This is only a human practice; you can set a different threshold, such as 0.6 or more, for reliable results.

Knowledge Representation: Finally, to verify the research aims, the evaluated results were visualized to better describe the data using Detectron2's class Visualizer and tensorboard to visualise loss value, accuracy, false positive, false negative results.

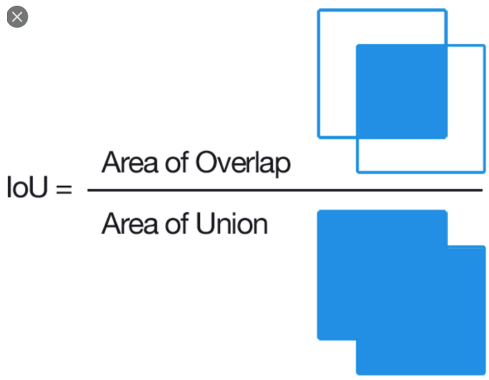


Figure 5: IoU formula

3.2 Overview of the Workflow of the proposed Dental Caries Diagnostic Tool

The model workflow is specified in figure 5, starting with the panoramic x-ray images data collection which was annotated by professional dentists practitioners with labelme annotation online tool, the data then were pre-processed for models; state of art model Detectron2 and Faster RCNN was implemented to get accurate object detection/ instance segmentation and later the model is trained and tested giving an accurate analysis of dental caries detection into five classes.

.

Figure 6: Overview of proposed model methodology

3.3 Conclusion

This study uses the methodology KDD and workflow design throughout the project to answer the research question and accurately detect the types of caries on panoramic Xrays, which has been shown to be effective. To achieve the goal, Google collaboratory notebook was used to implement the Detectron2 library for object detection/instance segmentation task, and other python libraries like Keras for preprocessing, matplotlib for data visualization, and tensorboard were used to visualize the output results.

4 Implementation, Evaluation and Result for Dental caries diagnostic tool

The model was built using Google Collaboratory, which has 11.62 GB of RAM and 65.50 GB of disk space. The data was loaded by mounting Google Drive, and Google Collaboratory includes runtime modifying choices. GPU Runtime is used to run both models. The labellme tool was used to annotate the x-ray images manually. The Detectron2, PyTorch, CUDA toolkit, Tensor Board, and additional python libraries like pycocotools were installed and downloaded, along with that python programming language is used to create models and produce COCO JSON files.

4.1 Implementation of the proposed model

The Detectron2 model is pre-trained on the COCO dataset, hence this study fine-tunes the detectron2 model on the panoramic x-ray image data set. The input data can be fed to deep learning models in a variety of forms, including Pascal VOC, YOLO format, COCO format, and others but only data in the form of COCO is accepted by Detectron2. As a result, retrieved annotations in the COCO format of the panoramic x-ray dataset using LabelMe, which consists of a JSON file. For object detection, we employ the Faster R-CNN model with the FPN backbone. The model zoo checkpoint is used to load the weights. In this case problem, the number of workers is set at 2. The batch size is set to 2, and the learning rate is set at 0.0025. Batch normalization has become one of the deep learning's accomplishments. By keeping the output distribution from one layer stable before forwarding to the next layer, it enables faster and more stable training. By avoiding vanishing gradients, this method also aids gradient descent. It subtracts the empirical averages over the batch divided by the observed standard deviations to normalize the output of the previous layer. The pixel means (cfg.MODEL.PIXEL MEAN) from all the images in the training set can be changed instead of the default values (generated from the ImageNet dataset) because Detectron2 recommends not changing the standard deviations. Though it enhance the performance this approach did not serve well. Another explanation could be that the baseline model was trained with standard means and standard deviations, so modifying them will affect the retrieved features from the baseline model. Augmentation during train time: Scale augmentation at train time improves results even more. A sample is scaled randomly between [500-1000] pixels during training and raises the number of iterations to 3500 (the learning rate is reduced by 10 at 2000 and 2400 iterations). Mask AP improves by 0.3 and box AP improves by 0.6 with train-time augmentation. We see an increase of 0.5 masks AP and 0.6 box AP by upgrading the 101-layer ResNeXt to its 152-layer counterpart. This demonstrates that a more detailed model can improve COCO outcomes. During test time augmentation, augmentations used is ResizeShortestEdge (short edge length=(800, 800), max size=1333, sample style='choice') which gives higher box and mask average precision score.

4.1.1 Experiments with Mask Faster RCNN with R101-FPN and Mask Faster R-CNN with X101-FPN

Even though it is faster to train a Faster R-CNN with R101-FPN as the baseline model, the accuracy is slightly good than with X101-FPN on the evaluation set at the test time prediction score thresholds of 0.65 and 0.71, respectively). When using RCNN with R101-FPN overall training speed is 2998 iterations in 1:08:06 that is 1.36 sec per iteration. And Overall training speed on 3498 iterations in 2:31:55 hours is (2.6060 s / iteration) for base model mask R-CNN X 101. Performing all of the experiments with these models takes time.

4.2 Evaluation Matrix Results

After training on the respective dataset, both models (Faster RCNN with R101-FPN and Faster RCNN with X101-FPN) had varying accuracy depending on the number of epochs and the training time. The total loss obtained after the training set for Faster R-CNN with R101-FPN and Faster R-CNN with X101-FPN is falling as seen in figure 7. If the total loss decreases it indicates that the training accuracy is high.

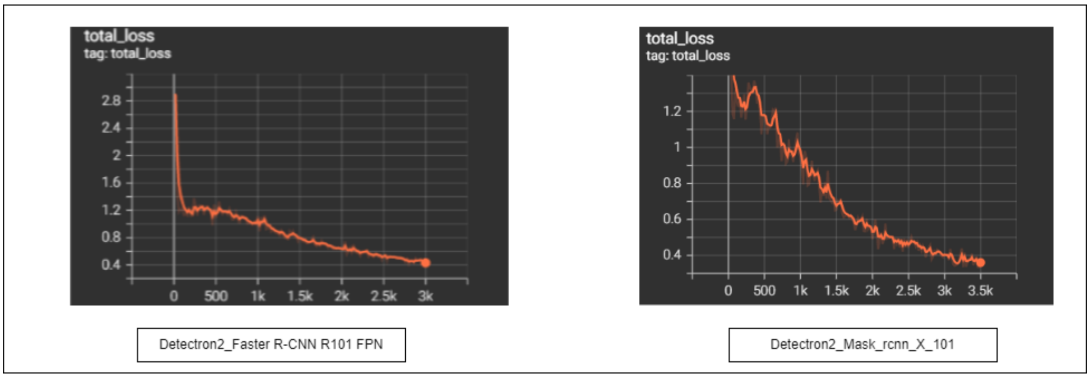


Figure 7: Total$_{loss}$

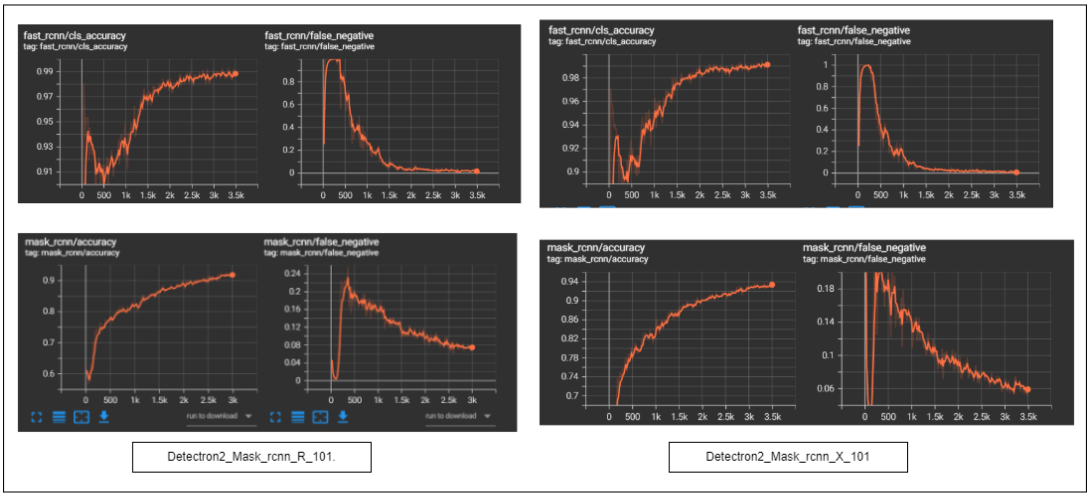


Figure 8: Accuracy and False negative Results

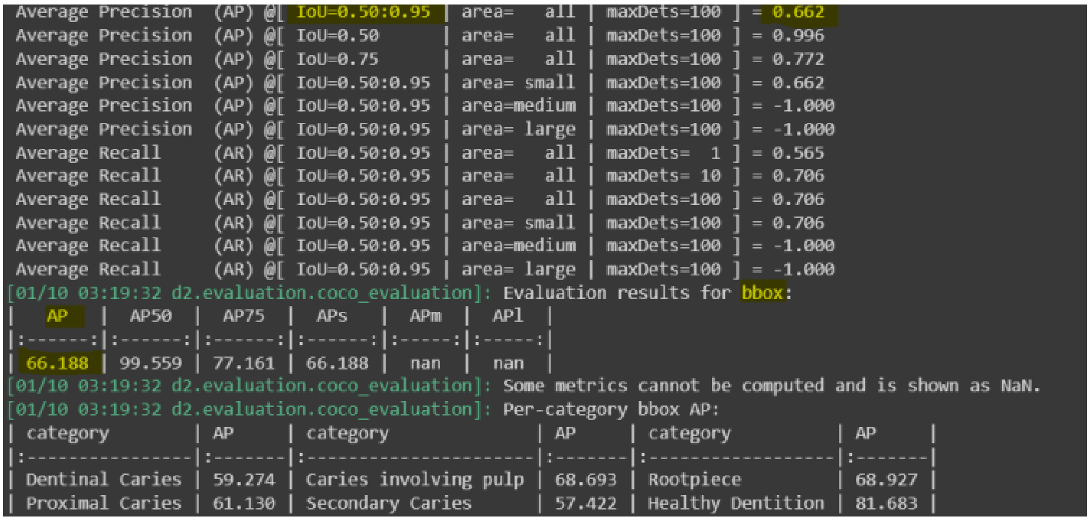
The prediction result for RCNN with R101-FPN and X101-FPN are relatively good and very similar to each other, the plots in Figure 8 show the result of false-negative that decreases with iterations which conclude that the training accuracy of the Detectron2 model is high. AP@IOU[.5:.05:.95] is a key assessment metric for object detection and instance segmentation The intersection-over-union (IOU) value between detected and ground truth boxes is determined to obtain precision and recall on the dataset. In Figure 9, the results of Average Precision @ IOU .5:.05:.95 is 53.512 for segmentation, and results of Average Precision @ IOU .5:.05:.95 is 66.18 for bbox. This metric (IOU) determines how well the detected bounding boxes match the ground truth bounding boxes. This is done by estimating the amount of overlap between the expected and ground-truth areas for each object type independently. The dental caries class Rootpiece has the highest AP of 68.927, as seen in Figure 10. For object detection, the COCO's AP@IOU[.5:.05:.95] is used as a benchmark according to various researches and experiments done for instance segmentation task. AP @.5 and AP @.75. : These two metrics are used to evaluate the precision x recall curve. For a more thorough analysis of the likeness of the ground truth and detection bounding boxes, the AP@.75 measure has been utilized. The segmentation and bbox results from AP@.75 are 51.005 and 71.161, respectively (Figure 9)

```
 Average Precision  (AP) @[ IoU=0.50:0.95 | area=   all | maxDets=100 ] = 0.535
 Average Precision  (AP) @[ IoU=0.50      | area=   all | maxDets=100 ] = 0.977
 Average Precision  (AP) @[ IoU=0.75      | area=   all | maxDets=100 ] = 0.510
 Average Precision  (AP) @[ IoU=0.50:0.95 | area= small | maxDets=100 ] = 0.535
 Average Precision  (AP) @[ IoU=0.50:0.95 | area=medium | maxDets=100 ] = -1.000
 Average Precision  (AP) @[ IoU=0.50:0.95 | area= large | maxDets=100 ] = -1.000
 Average Recall     (AR) @[ IoU=0.50:0.95 | area=   all | maxDets=  1 ] = 0.488
 Average Recall     (AR) @[ IoU=0.50:0.95 | area=   all | maxDets= 10 ] = 0.593
 Average Recall     (AR) @[ IoU=0.50:0.95 | area=   all | maxDets=100 ] = 0.593
 Average Recall     (AR) @[ IoU=0.50:0.95 | area= small | maxDets=100 ] = 0.593
 Average Recall     (AR) @[ IoU=0.50:0.95 | area=medium | maxDets=100 ] = -1.000
 Average Recall     (AR) @[ IoU=0.50:0.95 | area= large | maxDets=100 ] = -1.000
[01/10 03:19:32 d2.evaluation.coco_evaluation]: Evaluation results for segm:
|   AP   |  AP50  |  AP75  |  APs   |  APm  |  APl  |
|:------:|:------:|:------:|:------:|:-----:|:-----:|
| 53.512 | 97.707 | 51.005 | 53.512 |  nan  |  nan  |
[01/10 03:19:32 d2.evaluation.coco_evaluation]: Some metrics cannot be computed and is shown as NaN.
[01/10 03:19:32 d2.evaluation.coco_evaluation]: Per-category segm AP:
| category        | AP     | category             | AP     | category          | AP     |
|:----------------|:-------|:---------------------|:-------|:------------------|:-------|
| Dentinal Caries | 51.785 | Caries involving pulp | 43.117 | Rootpiece         | 54.281 |
| Proximal Caries | 47.771 | Secondary Caries     | 39.695 | Healthy Dentition | 84.422 |
```

Figure 9: Average Precision for segmentation

4.3 Visual Results

The visual output is shown in Figure 11, the output is obtained using Detectron2's class Visualizer. The model has performed instance segmentation using Panoramic X-ray accurately detecting 5 classes: Dentinal Caries, Proximal caries, Caries involving pulp, Rootpiece, and Secondary Caries. The results reported in this section answer the research question, and it can be concluded that this section fulfills objective 3 in table 1, namely the implementation of Detectron2 utilizing pre-trained models faster RCNN R101 FPN and mask faster RCNN X101 FPN models to detect dental caries into 5 classes. It also satisfies table 1's objective 4 of evaluating the models and displaying the findings to assess the model's performance.



```
Average Precision  (AP) @[ IoU=0.50:0.95 | area=   all | maxDets=100 ] = 0.662
Average Precision  (AP) @[ IoU=0.50      | area=   all | maxDets=100 ] = 0.996
Average Precision  (AP) @[ IoU=0.75      | area=   all | maxDets=100 ] = 0.772
Average Precision  (AP) @[ IoU=0.50:0.95 | area= small | maxDets=100 ] = 0.662
Average Precision  (AP) @[ IoU=0.50:0.95 | area=medium | maxDets=100 ] = -1.000
Average Precision  (AP) @[ IoU=0.50:0.95 | area= large | maxDets=100 ] = -1.000
Average Recall     (AR) @[ IoU=0.50:0.95 | area=   all | maxDets=  1 ] = 0.565
Average Recall     (AR) @[ IoU=0.50:0.95 | area=   all | maxDets= 10 ] = 0.706
Average Recall     (AR) @[ IoU=0.50:0.95 | area=   all | maxDets=100 ] = 0.706
Average Recall     (AR) @[ IoU=0.50:0.95 | area= small | maxDets=100 ] = 0.706
Average Recall     (AR) @[ IoU=0.50:0.95 | area=medium | maxDets=100 ] = -1.000
Average Recall     (AR) @[ IoU=0.50:0.95 | area= large | maxDets=100 ] = -1.000
[01/10 03:19:32 d2.evaluation.coco_evaluation]: Evaluation results for bbox:
|   AP   |  AP50  |  AP75  |  APs   |  APm  |  APl  |
|:------:|:------:|:------:|:------:|:-----:|:-----:|
| 66.188 | 99.559 | 77.161 | 66.188 |  nan  |  nan  |
[01/10 03:19:32 d2.evaluation.coco_evaluation]: Some metrics cannot be computed and is shown as NaN.
[01/10 03:19:32 d2.evaluation.coco_evaluation]: Per-category bbox AP:
| category        | AP     | category              | AP     | category          | AP     |
|:----------------|:-------|:----------------------|:-------|:------------------|:-------|
| Dentinal Caries | 59.274 | Caries involving pulp | 68.693 | Rootpiece         | 68.927 |
| Proximal Caries | 61.130 | Secondary Caries      | 57.422 | Healthy Dentition | 81.683 |
```

Figure 10: Average Precision for bbox

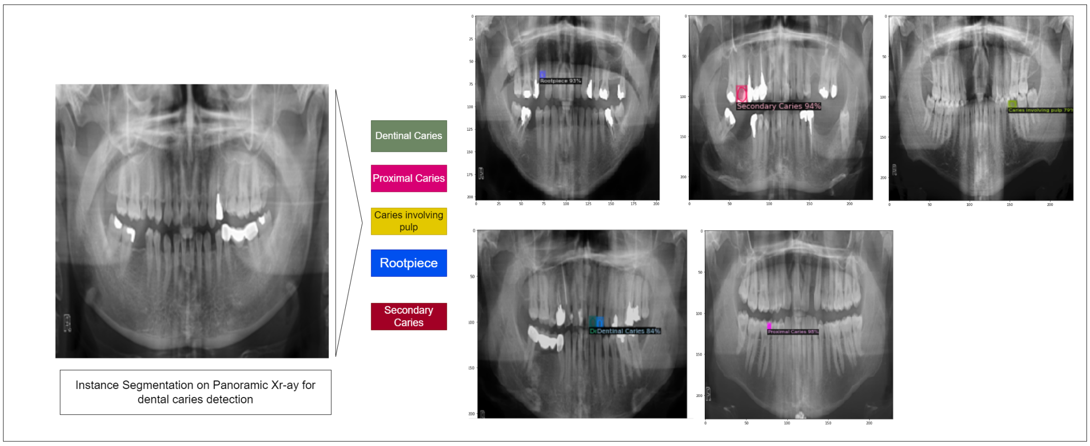


Figure 11: Detectron2 Instance Segmentation Visual Outcome

5 Discussion, Conclusion and Future Work

For many years, researchers have been attempting to segment teeth in dental X-ray images, relying primarily on unsupervised approaches. Recently in the year, 2020 a few research works used the supervised approach i.e.Mask RCNN and Faster RCNN to acquire results of semantic segmentation and instance segmentation on x-ray images for segmenting and detecting teeth. This research work followed a different direction, the approach was to address the dental clinical problem that is the detection of different stages/ classes of dental caries on panoramic x-ray. For multi-class detection, instance segmentation was performed to detect five different classes of dental caries, to the best of my knowledge, this is the first model that has detected classes of caries on panoramic x-ray experimenting with the state of art algorithm Detectron2. The panoramic x-ray dataset is used in this research to test Detectron's Faster RCNN implementation with various base models and parameters also other cutting edge approaches for object detection tasks that were examined like training and testing time augmentation. The results show that using Faster R-CNN with the X101-FPN base model and Detectron2's default setups produces reasonable prediction results of Average Precision which is 53.512 for segmentation and 66.188 for bbox, these results of Average Precision were not quite good as compared to other Detectron2 applications. The possibility for such results can be manual annotation and a small dataset size.

Deep learning requires a large amount of data and labeled data for segmentation tasks. Manual dental caries labeling takes a long time and is prone to errors. Errors may occur owing to varying difficulty to set bounding box precision or even a misunderstanding of caries types. In this study, five professional dental practitioners were in coordination for labeling 116 panoramic Xrays and classifying caries into Secondary Caries, Proximal Caries, Rootpiece, Caries involving pulp, and Dentinal Caries, yet there can be a possibility for error. Considering this, the results of instance segmentation are promising and reasonable, though not exceptional when compared to prior work on types of images of better quality. Two characteristics must be addressed to improve the results: (1) Increasing the number of Panoramic photos in the dataset, particularly the number of carious teeth. (2) Using more advanced parameters to tune Detectron2 to improve the model's performance. Dentists or radiologists provided precise annotations for caries regions will lead to correct and accurate segmentation of caries on Panoramic x-rays.

6 Acknowledgment

I would like to express my gratitude to my mentor Hicham Rifai, who has been incredibly supportive and encouraging right from the beginning of the research implementation phase. His timely assistance has greatly aided me in achieving the research's objectives. I would want to express my gratitude to the National College of Ireland in Dublin, which has been a rock throughout my master's degree, assisting me in learning new technologies through unique projects that have also shaped my technical background. Finally, I would want to express my gratitude to my family and friends, who have continuously encouraged me to work hard and boosted my confidence to confront any challenge in the task I have undertaken.

References

Al-Sherif, N., Guo, G. & Ammar, H. H. (2012), 'A new approach to teeth segmentation', *Proceedings - 2012 IEEE International Symposium on Multimedia, ISM 2012* (09), 145–148.

Chen, H., Zhang, K., Lyu, P., Li, H., Zhang, L., Wu, J. & Lee, C. H. (2019), 'A deep learning approach to automatic teeth detection and numbering based on object detection in dental periapical films', *Scientific Reports* **9**(1), 1–11.

Duan, K., Bai, S., Xie, L., Qi, H., Huang, Q. & Tian, Q. (2019), 'CenterNet: Keypoint triplets for object detection', *Proceedings of the IEEE International Conference on Computer Vision* **2019-October**(Iccv), 6568–6577.

Duong, D. L., Kabir, M. H. & Kuo, R. F. (2021), 'Automated caries detection with smartphone color photography using machine learning', *Health Informatics Journal* **27**(2).

Gansky, S. A. (2003), 'Dental data mining: potential pitfalls and practical issues.', *Advances in dental research* **17**, 109–114.

Girshick, R. (2015), 'Fast R-CNN', *Proceedings of the IEEE International Conference on Computer Vision* **2015 International Conference on Computer Vision, ICCV 2015**, 1440–1448.

Haghanifar, A., Majdabadi, M. M. & Ko, S.-B. (2020), 'PaXNet: Dental Caries Detection in Panoramic X-ray using Ensemble Transfer Learning and Capsule Classifier', pp. 1–14.
URL: *http://arxiv.org/abs/2012.13666*

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* pp. 400–407.

Jusman, Y., Tamarena, R. I., Puspita, S., Saleh, E. & Kanafiah, S. N. A. M. (2020), 'Analysis of Features Extraction Performance to Differentiate of Dental Caries Types Using Gray Level Co-occurrence Matrix Algorithm', *Proceedings - 10th IEEE International Conference on Control System, Computing and Engineering, ICCSCE 2020* (August), 148–152.

Kassebaum, N. J., Bernabé, E., Dahiya, M., Bhandari, B., Murray, C. J. & Marcenes, W. (2015), 'Global burden of untreated caries: A systematic review and metaregression', *Journal of Dental Research* **94**(5), 650–658.

Leopold, H. A., Singh, A., Sengupta, S. & Lakshminarayanan, V. (2021), 'Deep learning for ophthalmology using optical coherence tomography', *State of the Art in Neural Networks and their Applications* pp. 239–269.

Moutselos, K., Berdouses, E., Oulis, C. & Maglogiannis, I. (2019), 'Recognizing Occlusal Caries in Dental Intraoral Images Using Deep Learning', *Proceedings of the Annual International Conference of the IEEE Engineering in Medicine and Biology Society, EMBS* pp. 1617–1620.

Musri, N., Christie, B., Ichwan, S. J. A. & Cahyanto, A. (2021), 'Deep learning convolutional neural network algorithms for the early detection and diagnosis of dental caries on periapical radiographs: A systematic review', *Imaging Science in Dentistry* **51**, 1–6.

Na'am, J., Harlan, J., Madenda, S. & Wibowo, E. P. (2016), 'The algorithm of image edge detection on panoramic dental X-ray using multiple morphological gradient (mMG) method', *International Journal on Advanced Science, Engineering and Information Technology* **6**(6), 1012–1018.

Pandey, P., Bhan, A., Dutta, M. K. & Travieso, C. M. (2017), 'Automatic Image Processing Based Dental Image Analysis Using Automatic Gaussian Fitting Energy and Level Sets', *2017 International Work Conference on Bio-Inspired Intelligence: Intelligent Systems for Biodiversity Conservation, IWOBI 2017 - Proceedings* .

Prajapati, S. A., Nagaraj, R. & Mitra, S. (2017), 'Classification of dental diseases using CNN and transfer learning', *5th International Symposium on Computational and Business Intelligence, ISCBI 2017* pp. 70–74.

Satcher, D. (2017), 'Revisiting oral health in America: A report of the surgeon general', *American Journal of Public Health* **107**, S32–S33.

Vinayahalingam, S., Kempers, S., Limon, L., Deibel, D., Maal, T., Hanisch, M., Bergé, S. & Xi, T. (2021), 'Classification of caries in third molars on panoramic radiographs using deep learning', *Scientific Reports* **11**(1), 1–11.

Wu, Y., Kirillov, A., Massa, F., Lo, W.-Y. & Girshick, R. (2019), 'Detectron2', https://github.com/facebookresearch/detectron2.

Zhou, X., Austin, U. T., Wang, D., Berkeley, U. C. & Austin, U. T. (n.d.), '2019$_{C}VPR_{O}bjectasPoint'$.