Automatic Tooth Segementation and Classification in Dental Panoramic X-Ray images

A report of Mini Project on Image Processing (EC386)
Submitted by

Abhinav Pavithran	17EC202
Abhishek Ghadge	17EC203
Nadeem Roshan	17EC226

Under the guidance of
Dr. Shyamlal

in partial fulfilment of the requirements for the award of the degree of
BACHELOR OF TECHNOLOGY

DEPARTMENT OF ELECTRONICS AND COMMUNICATION ENGINEERING
NATIONAL INSTITUTE OF TECHNOLOGY KARNATAKA
SURATHKAL, MANGALORE - 575025

Contents

1 Abstract

In dentistry, radiological examinations help specialists by showing structure of the tooth bones with the goal of screening embedded teeth, bone abnormalities, cysts, tumors, infections, fractures, problems in the temporomandibular regions, just to cite a few. Sometimes, relying solely in the specialist's opin- ion can bring differences in the diagnoses, which can ultimately hinder the treatment. Although tools for complete automatic diagnosis are not yet expected, image pattern recognition has evolved towards decision support, mainly starting with the detection of teeth and their components in X-ray images. Tooth detection has been object of research during at least the last two decades, mainly relying in threshold and region-based methods. Following a different direction, this paper proposes to explore a deep learning method for instance segmentation of the teeth. It is noteworthy that this image type is the most challenging one to isolate teeth, since it shows other parts of patient's body (e.g., chin, spine and jaws). We propose a segmentation system based on mask region- based convolutional neural network to accomplish an instance segmentation. Performance was thoroughly assessed from a 1500 challenging image data set, with high variation and containing 10 categories of different types of buccal image.

2 Introduction

From images obtained by X-rays, dentists can analyze the entire dental structure, planning (if necessary) patient's treatment. Indeed X-ray images are a tool that is used in dental medicine to check the state of the teeth, gums, jaws and bone structure of a mouth, allowing diagnosis of buccal problems. In dentistry, X-rays are divided into two categories: Intraoral, a radiographic technique performed with the film positioned in the buccal cavity (the X-ray image is obtained inside the mouth), and extraoral, in which the patient is positioned between the radiographic film and the X-ray source (the X- ray image is obtained outside the patient's mouth). In these two categories, there are three types of dental X-rays that are used most often in dental examinations: Extraoral panoramic radiography - also called panoramic X-ray or orthopanto- mography; intraoral bitewing radiography - or bitewing X- ray; and periapical intraoral radiography or only periapical X-rays. Figure 1 illustrates examples of these X-ray image types. Particularly, panoramic X-ray is a useful exam to complement the clinical examination in the diagnosis of dental diseases (caries or endodontic diseases). This type of examination al- lows the visualization of dental and buccal irregularities, such as: Teeth included, bone abnormalities, cysts, tumors, cancers, infections, post-accident fractures, temporomandibular joint disorders that cause pain in the ear, face, neck and head region. Commonly, dentists request

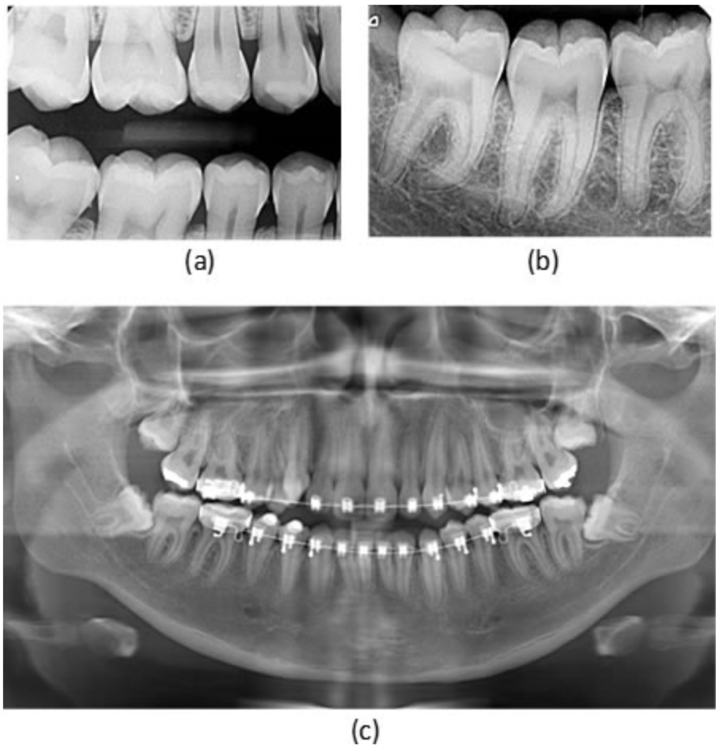

Figure 1: Types of X-ray images: (a) Bitewing X-ray; (b) Periapical X-ray; (c) Panoramic X-ray.

panoramic view of the mouth as a preoperative examination of the teeth, and bone surgeries of the temporomandibular region [1], [2].

3 Literature Survey

Panoramic radiography, also called panoramic x-ray, is a two-dimensional (2-D) dental x-ray examination that captures the entire mouth in a single image, including the teeth, upper and lower jaws, surrounding structures and tissues. (radio-graph) is a noninvasive medical test that helps physicians diagnose and treat medical conditions.Because your mouth is curved, the panoramic x-ray can sometimes create a slightly blurry image where accurate measurements of your teeth and jaw are not possible. If your dentist or surgeon needs more information, a computed tomography (CT) scan or magnetic resonance imaging (MRI) may be ordered. This limitation can be changed with development and application of our algorithms.

Instance Segmentation, which seeks to obtain both class and instance labels for each pixel in the input image, is a challenging task in computer vision. State-of-the-art algorithms often employ two separate stages, the first one generating object proposals and the second one recognizing and refining the boundaries. Further, proposals are usually based on detectors such as faster R-CNN which search for boxes in the entire image exhaustively.

In this paper [2], they propose a novel algorithm that directly utilizes a fully convolutional network (FCN) to predict instance labels. Specifically, we propose a variational relaxation of instance segmentation as minimizing an optimization functional for a piecewise-constant segmentation problem, which can be used to train an FCN end-to-end.

In this paper [3], A comprehensive research on all available papers on the platfrom of : IEEE Xplore , ScienceDirect , Google Scholar and Scopus is done. The research found that all papers use Region-based: The goal of the region-based method is to divide an image into regions, based on discontinuities in pixel intensity levels.
Threshold-based. The rationale of the intensity threshold application in image segmentation starts from the choice of a threshold value. Pixels whose values exceed the threshold are placed into a region, while pixels with values below the threshold are placed into an adjacent region.
Cluster-based. Clustering is a method used to make automatic grouping of data according to a certain degree of similarity between the data. The criterion of similarity depends on the problem to be solved
Boundary-based. Boundary-based methods are used to search for discontinuities (point and edge detection) in the gray levels of the image. Thirty-four percent (34%) of the relevant papers used boundary-based segmentation methods.

4 Backgound(Description of Dataset)

Dataset comprises of 1500 challenging panoramix x-ray images, with high variation and containing 10 categories of different types of buccal image.

Number	Category	Images	Average no. 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 with all the teeth, containing teeth with restoration and without dental appliance	220	32
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	29
10	Images missing teeth, containing teeth without restoration and without dental appliance	115	28

5 Methodology

We used Mask R-CNN Model for the task of instance segmentation of panoramic x-ray radiograph images. There are two stages of Mask RCNN. First, it generates proposals about the regions where there might be an object based on the input image. Second, it predicts the class of the object, refines the bounding box and generates a mask in pixel level of the object based on the first stage proposal. Both stages are connected to the backbone structure.
Mask R-CNN architecture is depicted in Fig. 3. As an extension of the Faster R-CNN , Mask R-CNN includes a branch of convolutional networks to accomplish the instance segmentation task. After extracting features from ResNet101, these features compose a feature pyramid network (FPN), where ultimately anchors are defined and regions of interest (RoIs) are extracted. These two stages (FPN + anchors) form the region proposal network (RPN). After that, RoIs are aligned to have the same size.

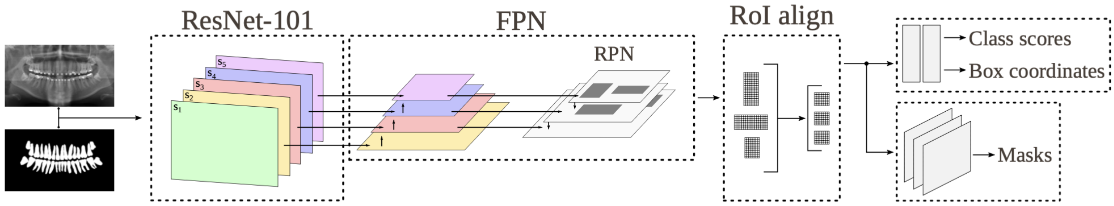

Figure 2: Training process of the segmentation system. From left to right: X-ray images and annotation masks as inputs, ResNet101 backbone with 5-stage feature extractor (from S1 to S5), where the output of each ResNet stage, but S1, forms a layer in the feature pyramid network (FPN); anchors are determined over FPN, and regions of interest (RoI) are computed (defining the region proposal network (RPN)) and, finally, aligned (RoI aligned). Outputs are the class scores and box coordinates, given by full connected network, and masks, given by a fully convolutional network.

At the end, each fixed-size feature is: i) Classified as tooth or background (class scores); ii) localized by regressing the bounding box coordinates; and iii) per-pixel segmented by the fully convo- lutional network (FCN) in each detected tooth bounding box (masks).
On our data set, only 193 buccal images were annotated. This amount of annotations was not sufficient to train Mask R-CNN from scratch due to the number of free parameters in the deep learning network. To cope with the lack of annotated data, pre-trained weights were taken from MSCOCO data set, which has 80 annotated objects for instance segmentation task. We only used the pre-trained weights in the backbone (ResNet 101) of the Mask-RCNN network. Just the weights of the top layers (RPN and so forth) were initialized with our data set.

5.1 Data Pre-processing

193 images were used for training (6987 teeth), while 83 images (3040 teeth) were used as a validation set to fine-tuning the deep network. The process of splitting the teeth in each image was done so in order to train the Mask R-CNN with more samples (now, objects are the teeth, rather than an entire dental arch) than the originally gathered data set. This strategy demonstrated to be effective, even considering that 1224 test images was used as the data set with entire dental arch annotation. After training the Mask R-CNN with 6987 tooth images (from 193 images), and fine-tuning the network parameters with the 3040 tooth images (from 83 images), 1224 dental arch images were used to evaluate the Mask R-CNN (by using original annotation of the mouth).

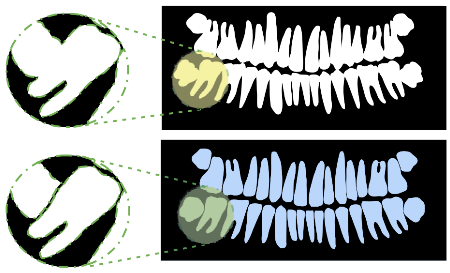

Figure 3: Process of separating the teeth in the data set

6 Model Framework

The framework for the described experiment is shown in fig.4

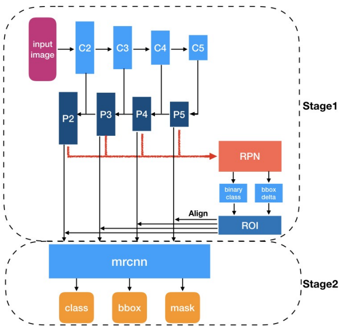

Figure 4: MRCNN framework

7 Results and Discussions

The model was trained and tested on various panoramic x-ray images and the results for 5 images from the test set are shown below in fig. 5
Our Model is able to segment the image near perfectly, correctly identifying the regions of the image which are teeth almost in all cases observed
The model was tested for VOC-Style mAP @ IoU=0.5 and achieved a score of 0.9159

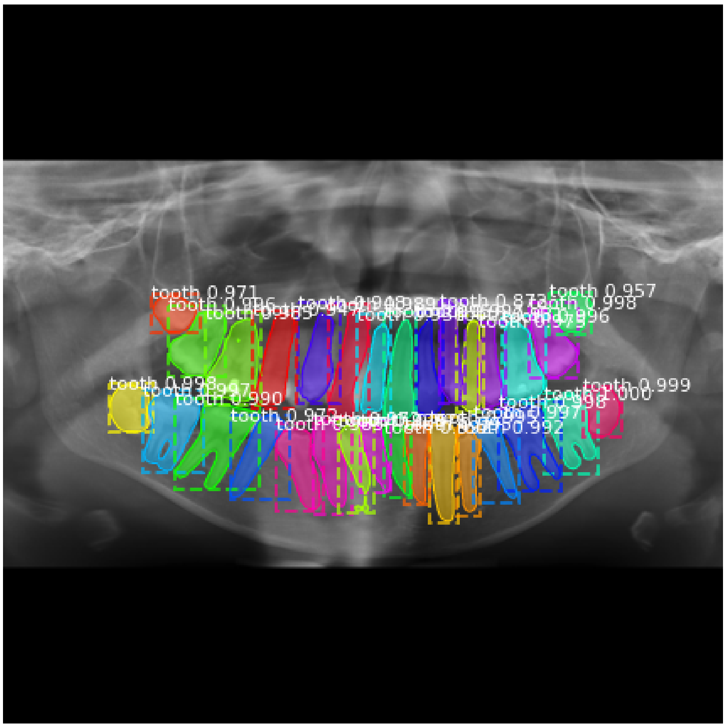

Figure 5: Results on single image with confidence scores

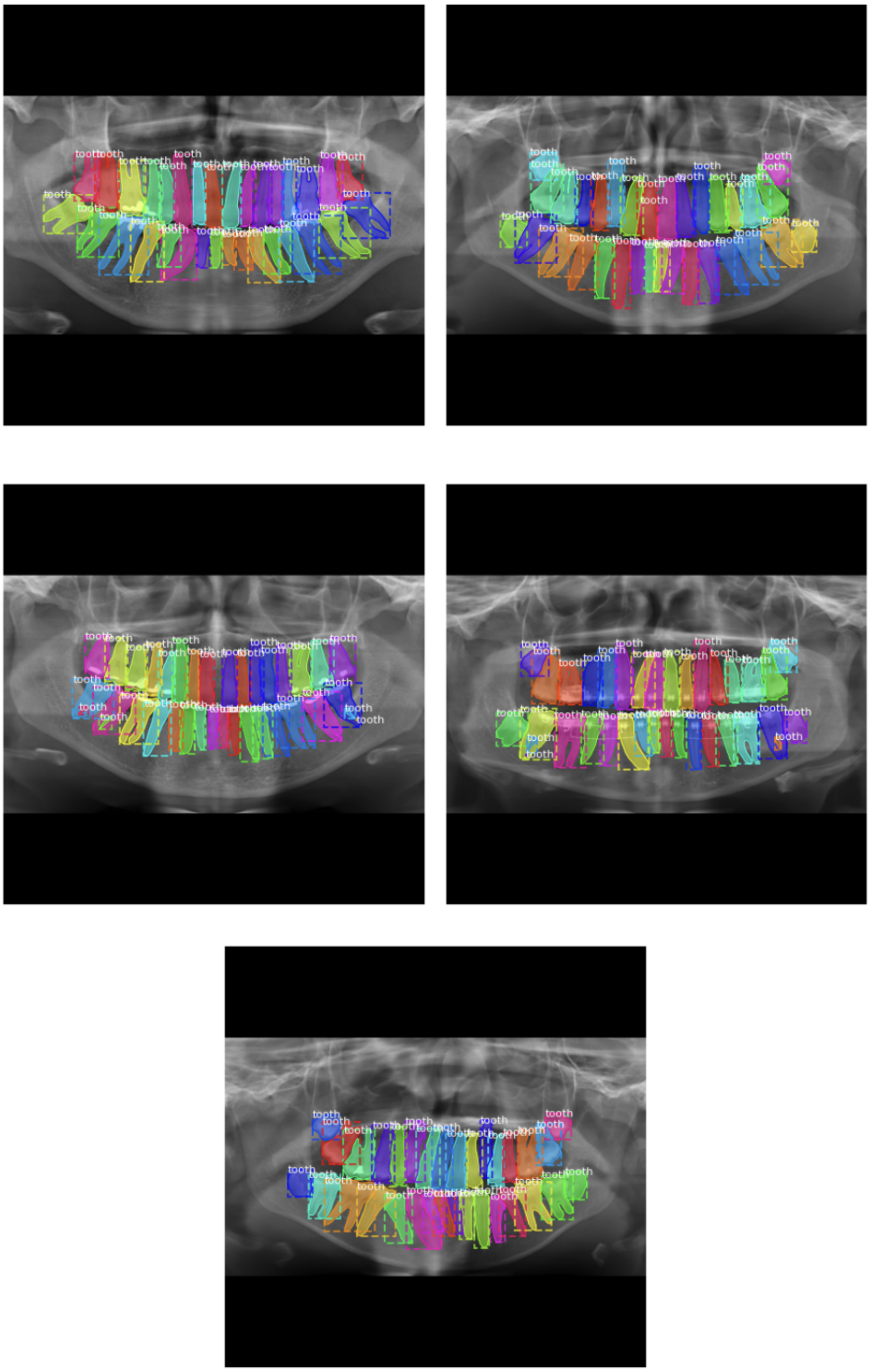

Figure 6: Results on 5 images from the test set

8 Conclusion

Segmenting teeth in dental X-ray images has been pursuit for many years, mainly relying in unsupervised methods. Al- though many approaches were proposed and tested, successful results were still far from being reached.
Segmenting tooth in buccal images are mandatory for more complex tasks in decision support systems. This is the first step to detect not only teeth and their constituent parts, but also artifacts (e.g., prosthesis), tooth problems, and even missing teeth. Considering that our proposed deep learning system demonstrated promising results on a challenging data set, future work resides on the instance segmentation of each component part of the mouth and teeth, as well as detection of missing teeth, all these with the goal of automatically generating medical reports.

9 Future works

The future works include the following:

- Further classifying each tooth by type (eg: incisor, molar etc.)
- Creating an end-to-end model that is able not only to detect and correctly classify teeth but also generate a diagnosis for any abnormalities

10 References

[1]G. Jader, J. Fontineli, M. Ruiz, K. Abdalla, M. Pithon and L. Oliveira, "Deep Instance Segmentation of Teeth in Panoramic X-Ray Images," 2018 31st SIBGRAPI Conference on Graphics, Patterns and Images (SIBGRAPI), 2018, pp. 400-407, doi: 10.1109/SIBGRAPI.2018.00058.

[2]Deep Variational Instance Segmentation J. Yuan, C. Chen, F. Li Published 2020,Computer Science,ArXiv,Corpus ID: 220686370

[3]Automatic segmenting teeth in X-ray images: Trends, a novel data set, benchmarking and future perspectives Gil Silva 1 , Luciano Oliveira 2 Ivision Lab, Federal University of Bahia, Brazil Matheus Pithon 3 Southeast State University of Bahia, Brazil

[4]https://www.radiologyinfo.org/en/info/panoramic-xray: :text=Panoramic%20radiography %2C%20also%20called%20panoramic,to%20that%20of%20a%20horseshoe.