Mandibular Canal Segmentation

**Introduction:**

The human mandible, also known as the lower jaw, is a complex anatomical structure and it is the only movable bone in the facial area. This bone facilitating the function of facial expression and speech. More important structures are the two mandibular canals which are located underneath the teeth. The anatomy of canals constitutes the two openings which are called mandibular foramen and mental foramen as shown in figure 1 Mandibular is responsible for motor and sensory innervations to muscleand teeth respectively.

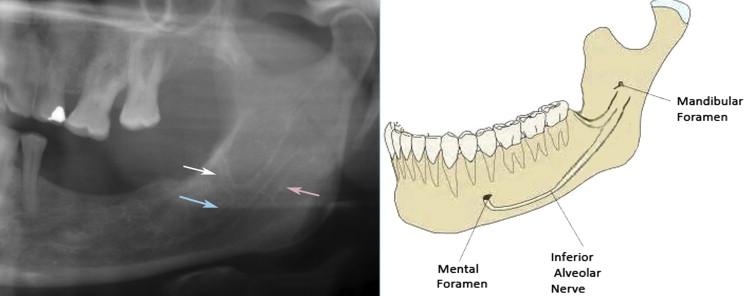


Fig 1: Anatomy of Mandibular canal

For 3D diagnostic and operation planning, cone beam computed tomography (CBCT) is widely used in dentomaxillofacial radiology. CBCT allows accurate imaging of hard tissue with less radiation dosage and it is more cost-effective and easily available. Mandibular canal location is of prime importance in dental implantology as before surgical operation the location, place, and size of implant should be determined. And the popular approach is to segmentation of canal across the cross-sectional slices using 3D imaging software. The labeling is very tedious and time- consuming thus there is a need for automated methods to do that cumbersome task.

**Related methods:**

There are several methods available that are semiautomated and very few are based on deep learning-based. Below is the table which shows the techniques and their corresponding accuracy.

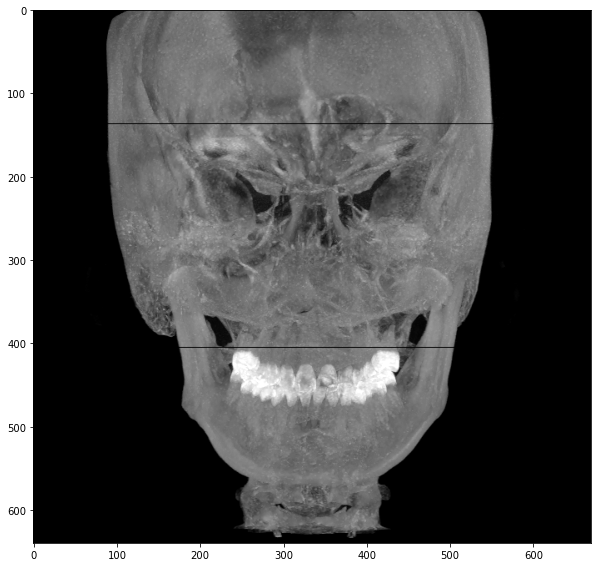
| **Paper Title** | **Modality/Volumes** | **Technique** | **Dice Score** |
| --- | --- | --- | --- |
| Deep Learning Method for Mandibular Canal Segmentation in Dental Cone Beam Computed Tomography Volumes [1]. | CBCT/ 637 | Deep learning/ Unet | 0.57 Left Canal  0.58 Right Canal |
| Automatic Mandibular Canal Detection using a deep learning convolutional neural network [2]. | CBCT/102 | Deep learning /Segnet, Unet | 0.49,0.577 |
| Automatic Segmentation of mandibular canal in cone beam CT images using conditional statistical shape model and fast marching [3]. | CBCT/ 120 | Image processing/ Conditional Statical Shape model | 0.9138 |

# Data Analysis

Data Analysis of the raw scans and their annotations showed us that the data was divided into three different groups of scans with varying contrast. The following figures shows MIP (coronal) of a sample from each group at same window level suitable for a group of scans:



For different windows suitable to each separately.



These are Type 1,2 and 3 scans respectively.

# **Preprocessing**

**Objective:** to pre-process the raw data and extract meaningful information prior to training of a Deep Neural network. This includes localization of the raw data to include only the mandibular region making it easy for a DNN to pay attention to only the region where the mandibular canal can be present. The other part of pre-processing is applying windowing to the raw data that gives the best contrast enabling the DNN to easily learn the location of the canal.

## Localization

### Experiment 1:

In the first experiment we started by applying image processing techniques to detect facial landmarks such as nose,ears and chin to crop the raw scan. Detecting the nose and chin gives us the cropping locations for depth of the mandibular region, whereas the ears gives us the cropping locations for the width of the mandibular region.

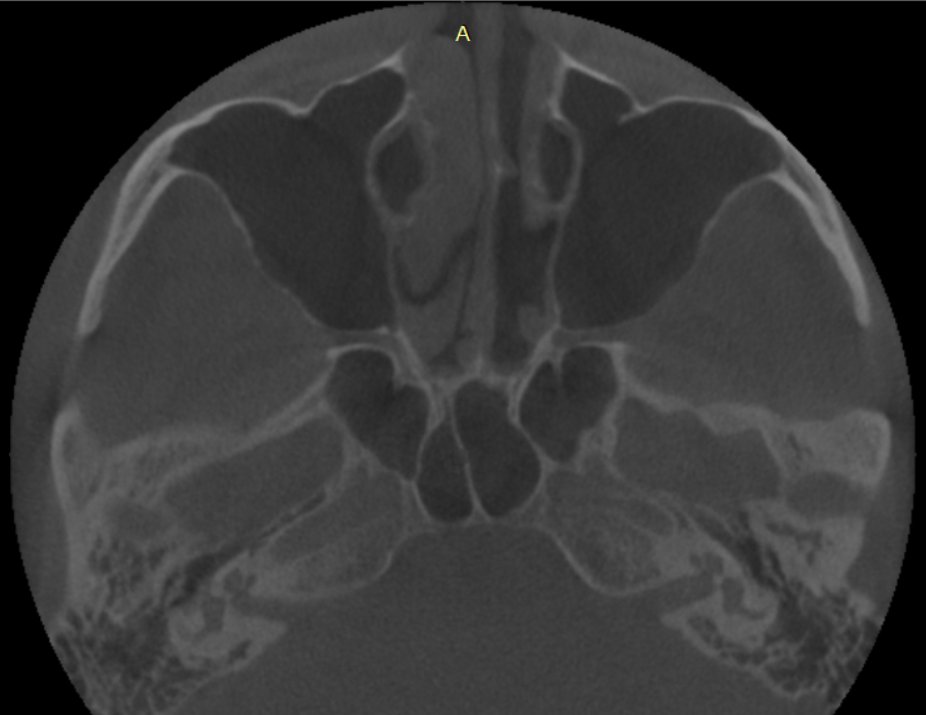
**Detection of Nose:** A very visible feature of the nose is that it is generally the part of the head thats the most protruding and located in the top two thirds of the head. This enabled us to extract the nose easily for scans that had visible noses, but didn’t work well for scans that didn’t have clearly defined facial boundaries.

Fig 2: a) sample scan with clear facial boundary. b) sample scan without clear facial boundary.

**Detecting ears:** Similar situation arised when detecting ears. As Fig 2 shows the ears are also not easily detectable.

### Experiment 2:

In this experiment we forewent detecting facial landmarks and instead focused on features within the head to localize the mandibular region. This included teeth, chin, dental region detection.

**Detecting teeth:** For this purpose, we performed segmentation of teeth via basic thresholding first but the results weren’t up to par so we trained a DNN to aid in teeth detection (more on this particular experiment is present in the next section about Segmentation via Deep learning) and the following shows a sample of result we obtained:

The starting of the teeth along depth was noted for the upper bound of the mandibular region.

Fig 3: Teeth segmentation mask

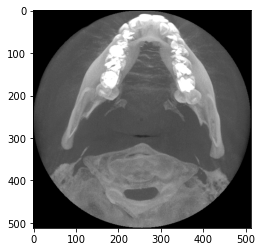
**Dental Region detection:**  For this we found the Maximum Intensity projection of the teeth region along the depth as shown in fig 4. Thresholding is then applied to get the exact dental mask (shown in fig 5) which ten gives us the upper and lower bounds for the rows and columns of the mandibular region.



Fig 4: MIP of teeth region along depth Fig 5: Thresholding to get dental mask

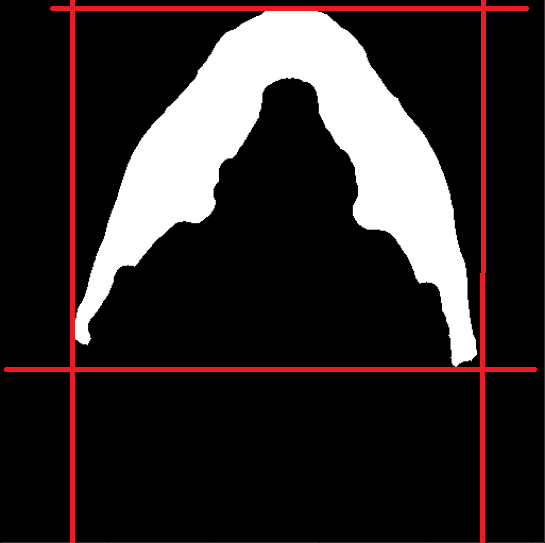


Fig 6: Locations for upper and lower bounds of mandibular region.

**Chin detection**: The last location of the lower bound for the depth of the mandibular region which is found by detection of chin. For this we first find the Maximum intensity projection of the head as coronal view as shown in Fig 7. The sharp change total head area per row at the end of the jaw and start of the chin enables us to easily detect the lower bound for depth of the mandibular region.

Fig 7: Coronal MIP of the head

Localization was then performed using the above algorithm for all the scans and their annotations. To check performance the total canal area (in number of pixels) in raw annotations was compared with the localized annotations and we got the following results:

Data: **997** scans

Scans with 0 error : **93.78%** (935 scans)

Scans with less than 0.1% error: **97.4%**

Scans with less than 5% error: **98.6%**

## Dynamic Windowing

While choosing the optimum window levels, same window levels for all three types of scans didn’t give us very good results and we had to come up with an algorithm which would dynamically check the histogram of intensities of the scan and perform windowing to it that would give us the best contrast for the mandibular canal.

For this purpose we chose a slice of the scan,along depth, that passes through the teeth to get all the varying intensity within that slice. This was the slice that corresponded to the centre column. We plotted a histogram of intensities of this slice with 10 bins total. Then we checked the highest intensity as well as the frequency of the pixels corresponding to that intensity level.

**Type 2:** If the highest intensity of the slice crosses 3200, the scan was classified as type 2.

**Type 3:** If the highest intensity of the slice was less than 3200 and the pixels with intensity greater than 2000 had a frequency greater than 400, then the scan was classified as Type 3.

**Type 1:** If none of the pixels with frequency greater than 400 lied beyond the intensity level of 2000 were classified as type 1 scan.

Then windowing can be applied with regards to the purpose of windowing, canal segmentation or teeth segmentation.

The following windows were chosen:

|  | **Canal Segmentation** | | **Teeth Segmentation** | |
| --- | --- | --- | --- | --- |
| **Minimum** | **Maximum** | **Minimum** | **Maximum** |
| **Type 1** | -250 | 1500 | 250 | 1500 |
| **Type 2** | -1000 | 5000 | 500 | 4050 |
| **Type 3** | -250 | 1500 | 500 | 4050 |

# Deep learning-based experiments

## Mandibular Canal Segmentation

We have used following models in the experimentation process. They are the most commonly used 3D segmentation models in literature.

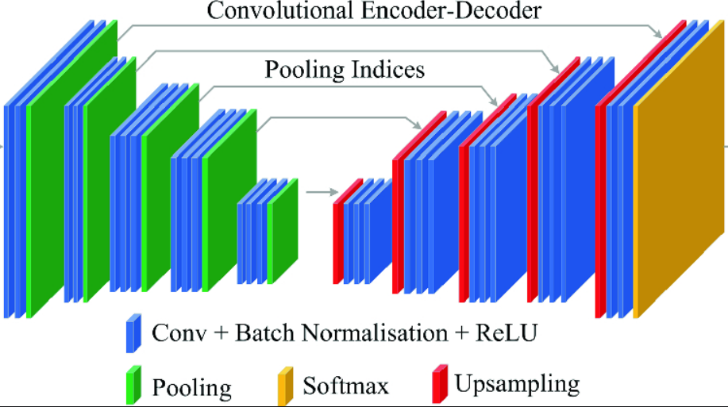


Fig 2: Architecture of SegNet [4]

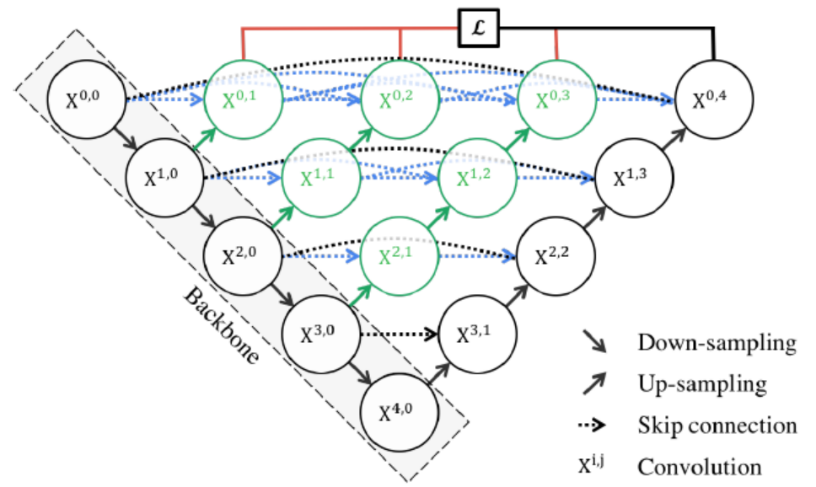


Fig 3: Architecture of NestedUnet or Unet++ [5]

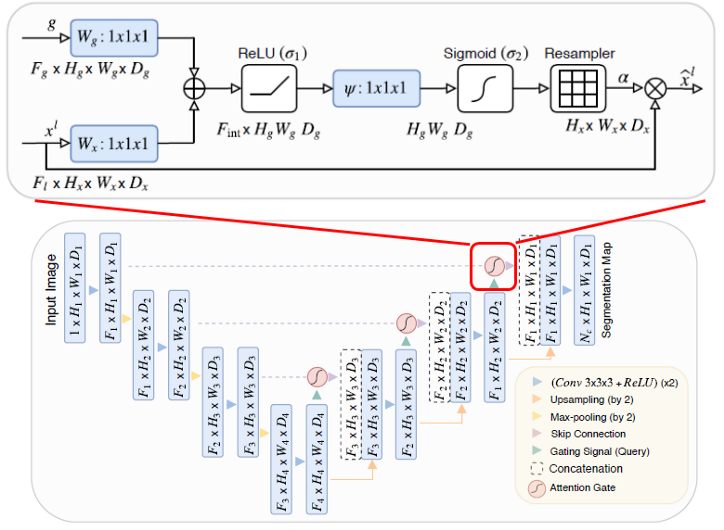


Figure 4: Architecture of Attentionunet [6]

**Dataset:**

Dataset is CBCT 3D volumes collected from different hospitals.   
  
More information can be added **(Miss Rabeea)**

| Dataset Dimension | Voxel Spacing | Number of Volumes |
| --- | --- | --- |
| 512x512x512 | -- | -- |

**Experiment 1:**

In this experiment we tried to evaluate the model which takes input image for left or right canal only.

**Details:**

| **Total Samples** | **Training Samples** | **Validation Samples** | **Volume Size** | **Learning Rate** | **Batch Size** | **Model** | **Dice Score** |
| --- | --- | --- | --- | --- | --- | --- | --- |
| 140 | 110 | 30 | 128x128x128 | 0.0001 | 4 | **Unet++** | **63.97** |
| SegNet | 62.34 |

Table 2: Details about experiment no. 1.

**Result:**

As we can see from the table the dice is low and so we need to design a different approach for preprocessing of volumes.

**Experiment 2:**

In this experiment we increase our dataset and change our preprocessing and crop the both canals using landmarks and then test the images on different models.

**Details:**

| **Total Samples** | **Training Samples** | **Validation Samples** | **Volume Size** | **Learning Rate** | **Batch Size** | **Model** | **Dice Score** |
| --- | --- | --- | --- | --- | --- | --- | --- |
| 700 | 500 | 200 | 128x128x128 | 0.0001 | 4 | Unet++ | 72.35 |
| SegNet | 72.14 |
| **Attention Unet** | **73.1** |

Table 2: Details about experiment no. 2.

**Result:** We have gained state of the art accuracy but we run more experiment for localization of canal

**Experiment 3:**

In this experiment we done preprocessing for localization of canal and made rectangular mask around that.

**Details:**

| **Total Samples** | **Training Samples** | **Validation Samples** | **Volume Size** | **Learning Rate** | **Batch Size** | **Model** | **Dice Score** |
| --- | --- | --- | --- | --- | --- | --- | --- |
| 1003 | 700 | 200 | 256x256x256 | 0.0001 | 4 | Unet++ | - |
| SegNet | - |
| Attention Unet | 89.1 |

Table 3: Details about experiment no. 3.

**Result:** (Continued)

## Teeth Segmentation

**Introduction:**

For orthodontic diagnosis and treatment planning, digital dentistry has been developing and playing an important role in the past decade. 3D teeth segmentation and acquisition is one of the key elements in digital dentistry. For acquiring complete 3D teeth models there are two mainstream technologies, i.e, Intraoral scanning and cone-beam computed tomography (CBCT). With intraoral scanning, one can obtain surface geometry but it does not provide information about tooth roots whereas CBCT provides more crucial and comprehensive for overall 3D volumetric information and all oral tissues. CBCT has been widely used in dentistry for oral surgery and digital orthodontics. Originally, 3D tooth instance segmentation and identification have the main application in digital orthodontics but in this experiment, we will use teeth segmentation for obtaining dental arch which leads to obtaining a panoramic view of tooths.

We performed several experiments to obtain the optimal model and then tune the hyperparameters to get the desired results. Previously, many methods have been proposed for teeth segmentation using model-driven approaches [1-4] but the main deficiency is that they required hand-crafted features to obtain segmentation results. In contrast in our experiments, we used data-driven-based approaches to get 3D end-to-end volumetric teeth segmentation. The details about experiments and data are given below.

**Data :**

We used 760 CBCT volumes to perform deep learning-based teeth segmentation and we divide the data into training, validation, and testing with 430, 100, and 230 volumes respectively. The data visualization along with its mask is given below.

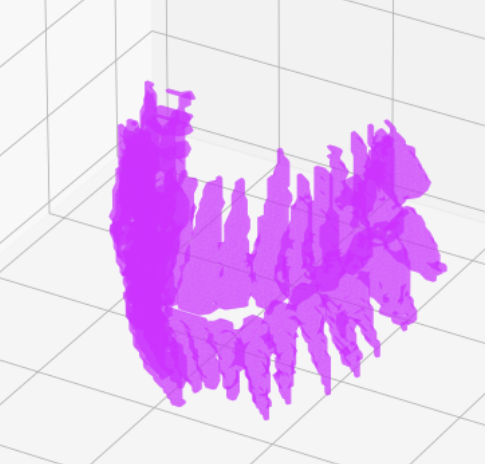
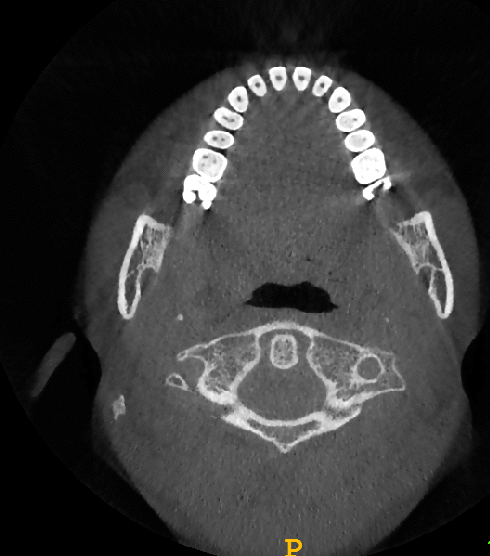


Figure 1: Visualization of Image and mask

**Model:**

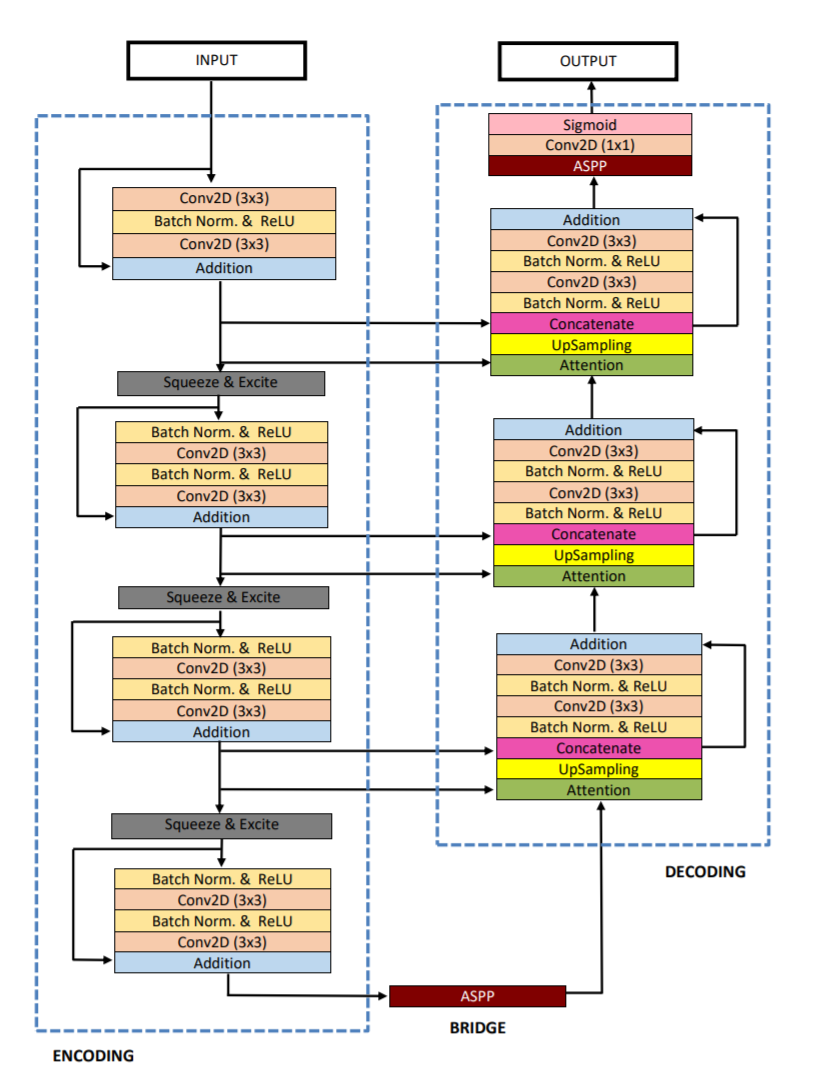
The model for training that has been used is ResUNet++ [5] which is encoder-decoder architecture inspired by U-Net [6] and ResUNet [7]. The model uses residual blocks, squeeze and excitation blocks, Atrous Spatial Pyramid Pooling (ASPP), and attention blocks. The model diagram is as follows

Figure 2: Model block Diagram

**Experiments and Results:**

The CBCT original volume size is quite large we resize the data into different sizes and train the model on ResUNet++ and evaluated on Dice similarity coefficient (DSC) metric and results are given below.

| Experiment Number | Model | Patch size | Dice score (original size) |
| --- | --- | --- | --- |
| 1 | ResUNet++ | 128x128x128 | 74.67 |
| 2 | ResUNet++ | 144x176x176 | 79.58 |
| 3 | ResUNet++ | **192x192x192** | **81.34** |

**References:**

[1] H Akhoondali, RA Zoroofi, and G Shirani. Rapid automatic segmentation and visualization of teeth in ct-scan data

[2] Sandro Barone, Alessandro Paoli, and ARMANDO VIVIANO Razionale. Ct segmentation of dental shapes by anatomy-driven reformation imaging and b-spline modelling. International journal for numerical methods in biomedical engineering, 32(6):e02747, 2016.

[3] Dong Xu Ji, Sim Heng Ong, and Kelvin Weng Chiong Foong. A level-set based approach for anterior teeth segmentation in cone beam computed tomography images. Computers in biology and medicine, 50:116–128, 2014.

[4]Yuru Pei, Xingsheng Ai, Hongbin Zha, Tianmin Xu, and Gengyu Ma. 3d exemplar-based random walks for tooth segmentation from cone-beam computed tomography images. Medical physics, 43(9):5040–5050, 2016.

[5] https://arxiv.org/pdf/1911.07067.pdf

[6] O. Ronneberger, P. Fischer, and T. Brox, “U-net: Convolutional networks for biomedical image segmentation,” in Proceedings of International Conference on Medical image computing and computer-assisted intervention. Springer, 2015, pp. 234–241

[7] Z. Zhang, Q. Liu, and Y. Wang, “Road extraction by deep residual unet,” IEEE Geoscience and Remote Sensing Letters, vol. 15, no. 5, pp. 749–753, 2018

**References:**

[1] Jaskari, J., Sahlsten, J., Järnstedt, J., Mehtonen, H., Karhu, K., Sundqvist, O., Hietanen, A., Varjonen, V., Mattila, V. and Kaski, K., 2020. Deep Learning Method for Mandibular Canal Segmentation in Dental Cone Beam Computed Tomography Volumes. *Scientific reports*, *10*(1), pp.1-8.

[2] Kwak, G.H., Kwak, E.J., Song, J.M., Park, H.R., Jung, Y.H., Cho, B.H., Hui, P. and Hwang, J.J., 2020. Automatic mandibular canal detection using a deep convolutional neural network. *Scientific reports*, *10*(1), pp.1-8.

[3] Abdolali, F., Zoroofi, R.A., Abdolali, M., Yokota, F., Otake, Y. and Sato, Y., 2017. Automatic segmentation of mandibular canal in cone beam CT images using conditional statistical shape model and fast marching. *International journal of computer assisted radiology and surgery*, *12*(4), pp.581-593.

[4] Badrinarayanan, V., Kendall, A. and Cipolla, R., 2017. Segnet: A deep convolutional encoder-decoder architecture for image segmentation. *IEEE transactions on pattern analysis and machine intelligence*, *39*(12), pp.2481-2495.

[5] Çiçek, Ö., Abdulkadir, A., Lienkamp, S.S., Brox, T. and Ronneberger, O., 2016, October. 3D U-Net: learning dense volumetric segmentation from sparse annotation. In *International conference on medical image computing and computer-assisted intervention* (pp. 424-432). Springer, Cham.

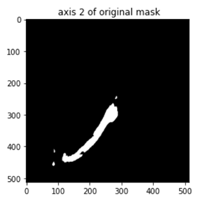
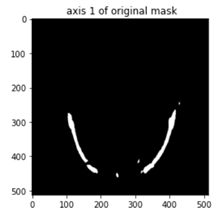
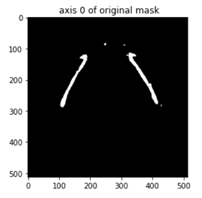
[6] Oktay, O., Schlemper, J., Folgoc, L.L., Lee, M., Heinrich, M., Misawa, K., Mori, K., McDonagh, S., Hammerla, N.Y., Kainz, B. and Glocker, B., 2018. Attention u-net: Learning where to look for the pancreas. *arXiv preprint arXiv:1804.03999*.

## 

# Presentation and Post-processing of results

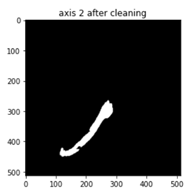
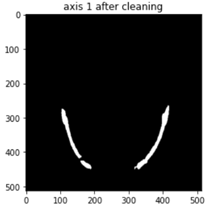
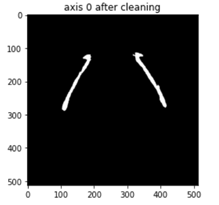
## Post-Processing of Mandibular Canal

The original mask is shown below:



### CLEANING:

For cleaning the mask, morphological operation of removing small objects was used. Below are the results:



## 

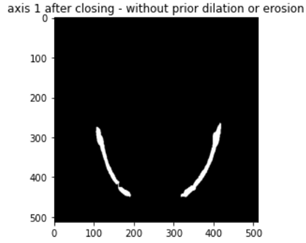
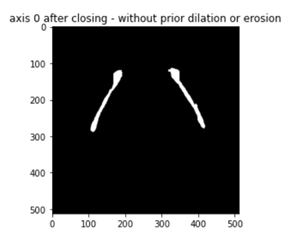
### CLOSING:

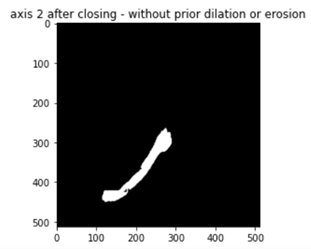
To fill the gaps two morphological techniques were experimented:

1. Direct closing
2. Dilation followed by erosion

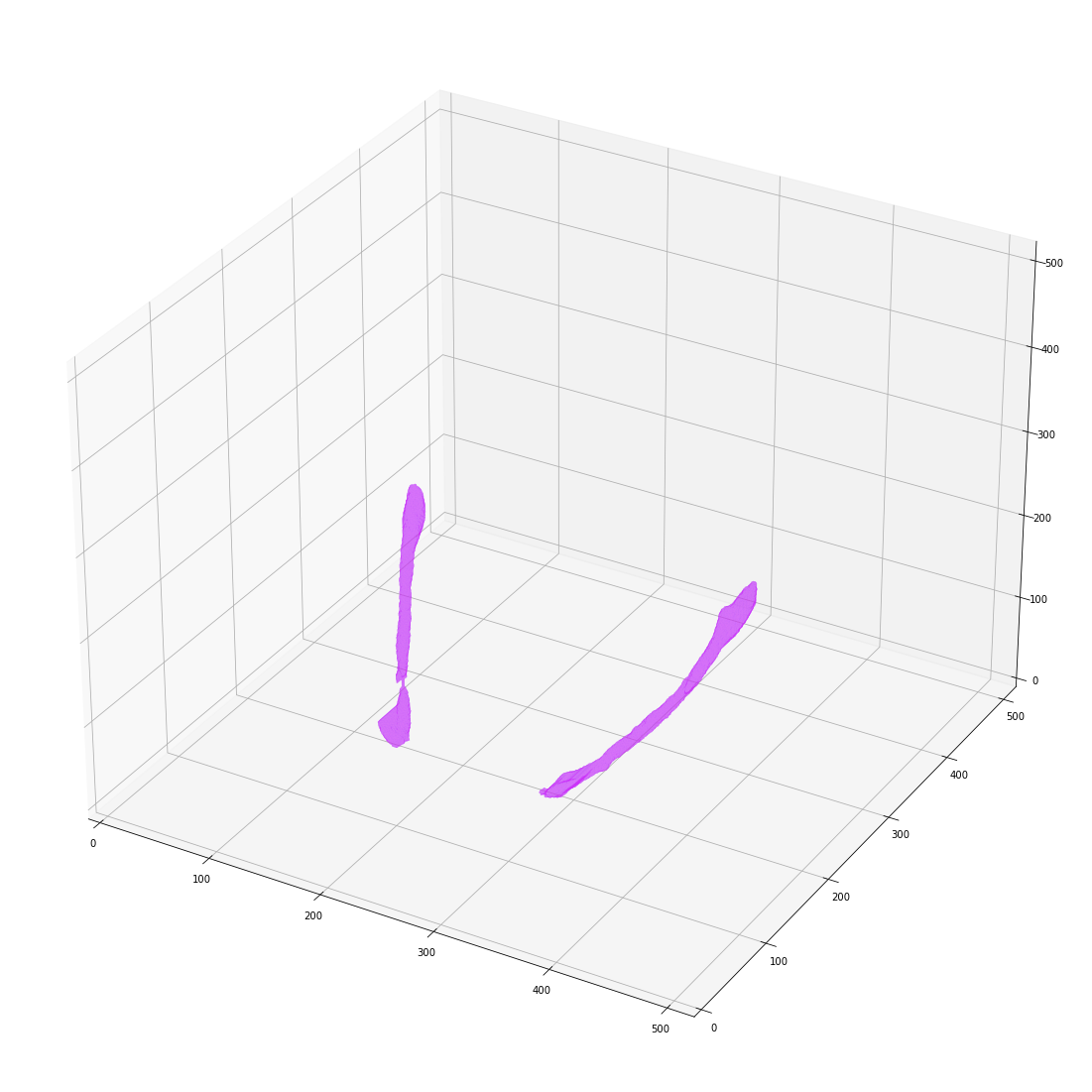
### Direct Closing:

The first approach was using the function of closing directly. A kernel of size 120x120x120 was used.





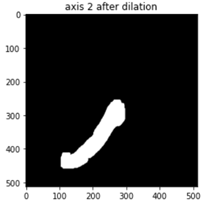
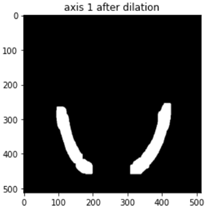
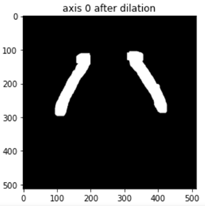
For better visualization, 3D plot was also plotted as shown below:

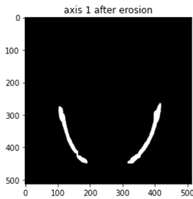
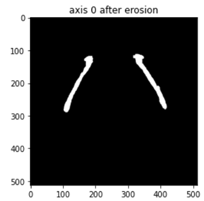


### Dilation followed by Erosion:

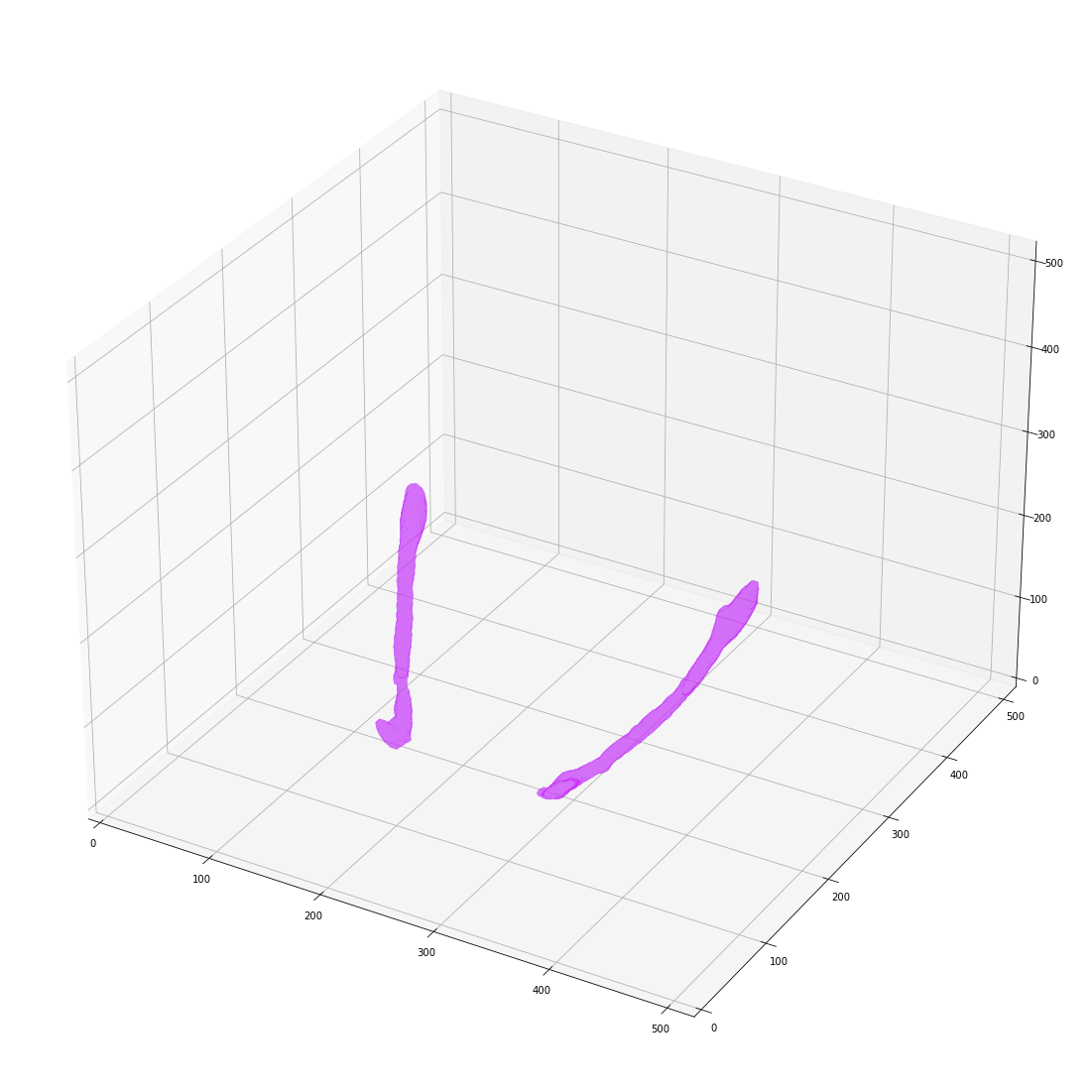
The previous approach didn’t fill the gaps as expected, therefore this approach was implemented.

For dilation a kernel of size 20x18x18 was used whereas for erosion it was 20x15x15. These were set using the hit and trial method, selected where there were best results.





3D plot after this process was:



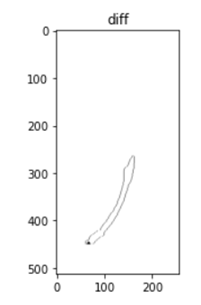
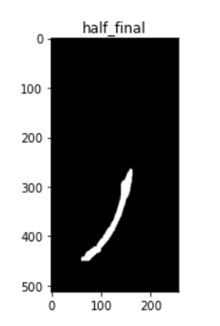
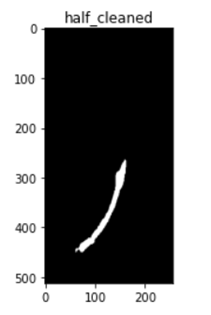
### Findings:

Although the gap was filled nicely in the second approach, it came with certain expenses.

Due to dilation, there were extra pixels included as a part of the Mandibular Canal.

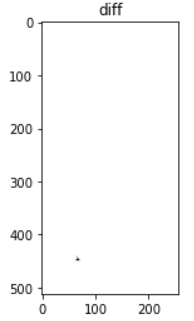
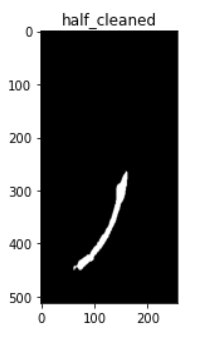
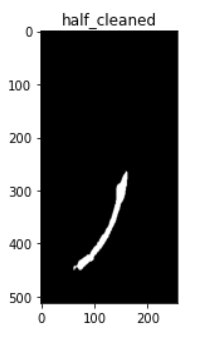
**Percentage error** was





Where as, for direct cleaning **percentage error** was





Direct closing has very less error but as seen in the 3D plot it didn't fill the gap as expected where for the other technique, the percentage error was really large. So, some other technique has to be found to cater to this problem.

## Presentation of Results

**Objective:** The objective is to present the mandibular canal with reference to the head through multiple projections. Currently we are providing the following presentation formats:

* Panorama View
* Para-sagittal view
* Maximum Intensity Projection
  + Coronal
  + Sagittal
* Axial view of the mandibular region

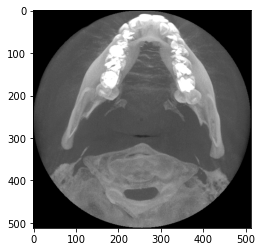
For both Panorama and Para-sagittal view the first step is to develop an algorithm that automatically draws a dental arch that passes through the teeth.

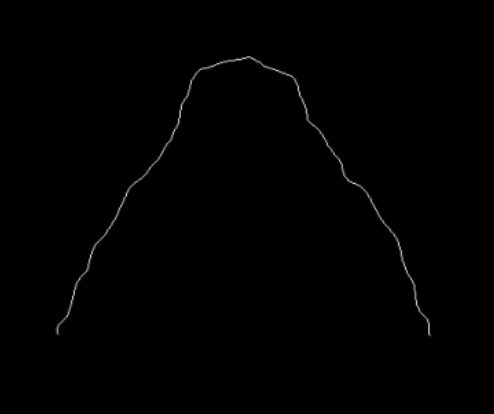
### Panorama and Para-sagittal Views

#### Experiment 1:

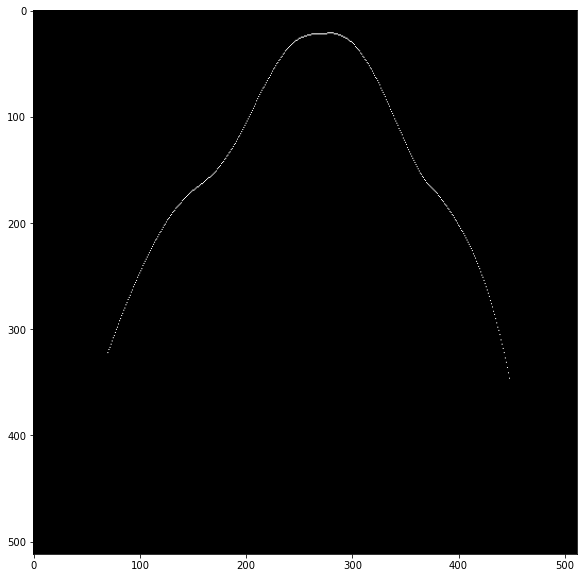
**Dental Arch generation**

In the first experiment we generated a dental arch that would be used for generation of both the views.

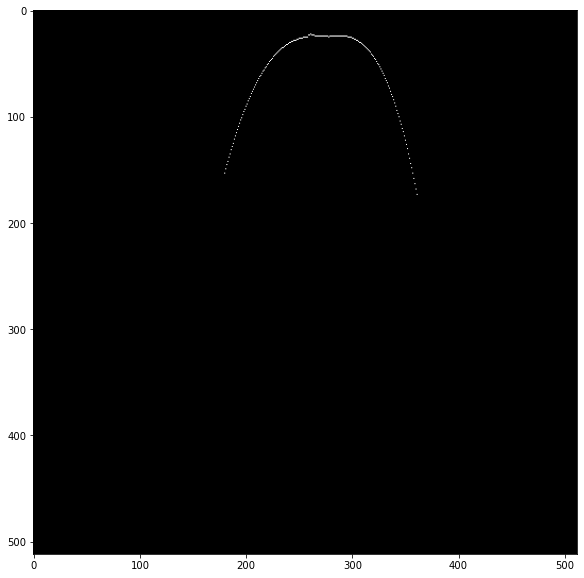
For this we performed the following steps:

1. Performed windowing based on the scan type:
   * Type 1: 0 to 2500
   * Type 2: -500 to 4500
   * Type 3: -500 to 3500
2. From teeth segmentation, we noted the region of teeth and took a Maximum Intensity projection along depth to extract the dental region.
3. Performed simple thresholding that would generate a mask for only the dental area. To refine this mask we perform the morphological operations of hole filling, closing and removing small masks.
4. However sometimes the bottom region also gets extracted but we cater to that by selecting the region that's both the largest and is on top.
5. Now we find the skeleton of this region using skimage library’s skeletonize function. There are several instances where the skeleton generated is not clean and has several branches extending from it. This was catered to by designing a function to detect all the branches of the skeleton (using ) and then recursively remove all the small branches until finally only a single skeleton remains.





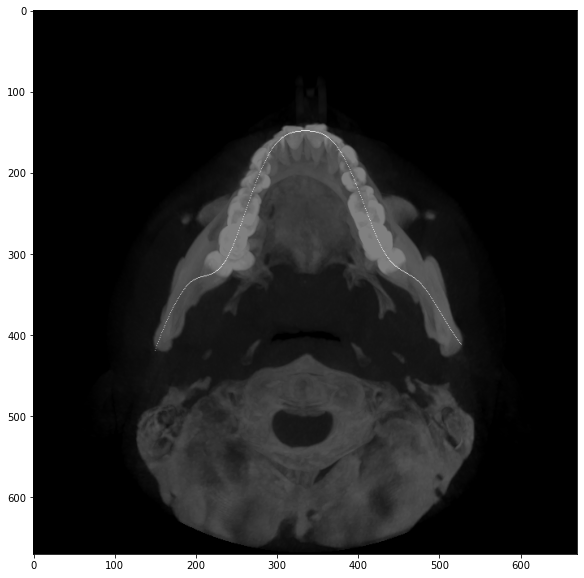
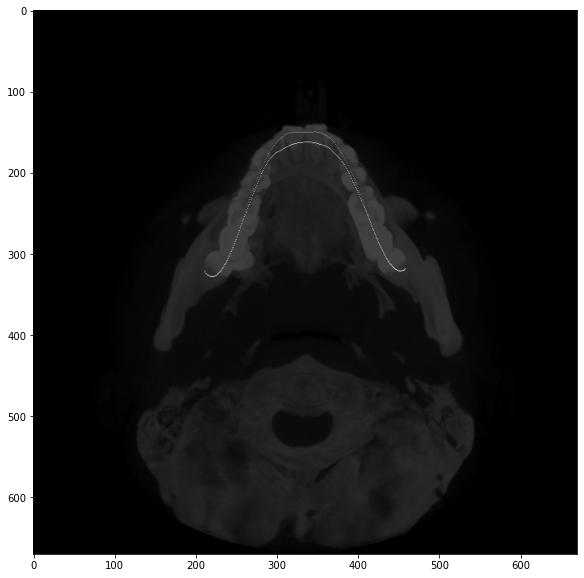
1. The result is then smoothed through Savgol Filter with window size of 85 and 3rd order interpolation.
2. The result does not necessarily pass through all the teeth so we use the teeth mask previously generated to find the path of the dental arch along the teeth.
3. We find the MIP of the binary mask of the teeth and perform steps 3 to 6 on this mask as well.



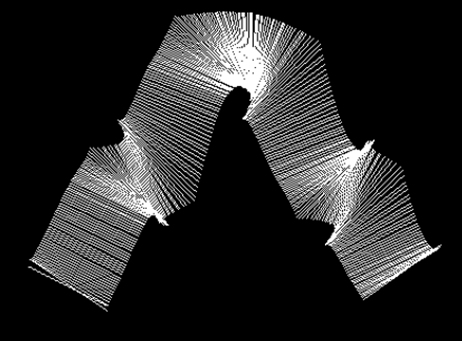




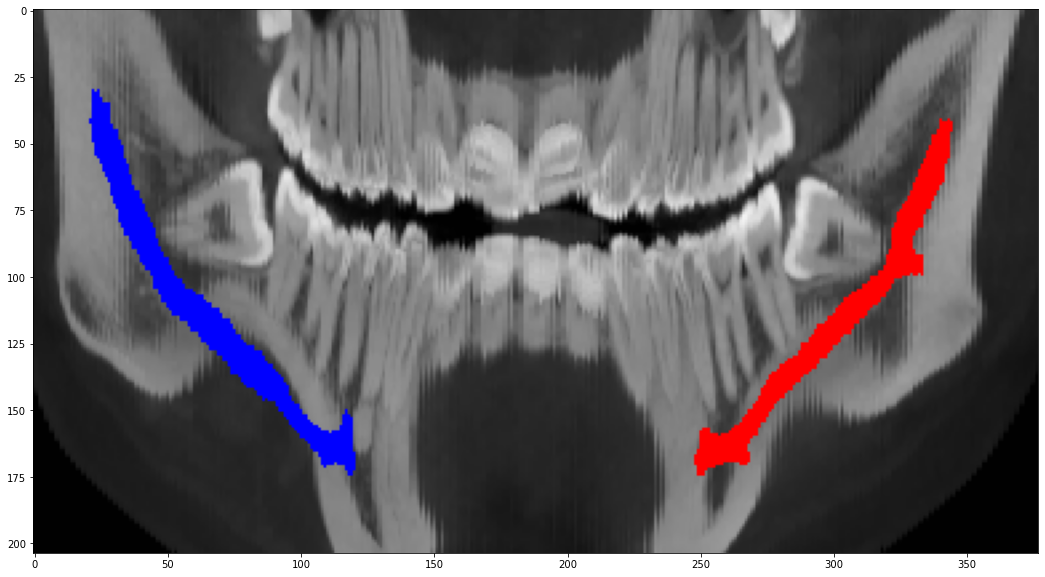
1. This sometimes doesn't pass through the milk teeth and just passes through their roots. In order to cater to that we developed a function that will calculate the offset of this curve. Th amount to offset the curve by is found through the middle position between the top of the teeth binary mask and the top of the teeth arch. After finding the offset arch we only keep the offset arch’s top most part and interpolate it with original teeth arch’s side parts.





We then find the perpendicular lines passing through each point of the dental arch by choosing a point and finding its gradient to the next 25th point on the arch (This smoothens the angle change of the perpendiculars)

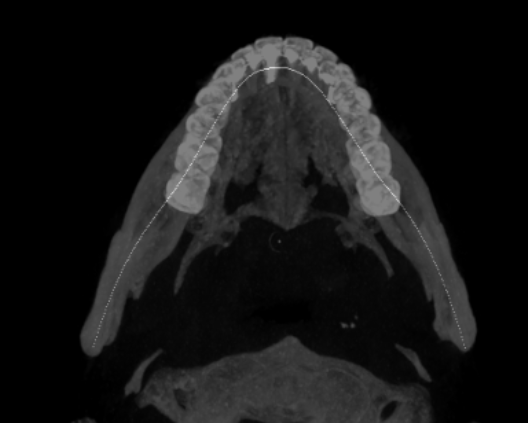
**Panorama:** To generate the Panorama view we performed the following steps:

1. Each perpendicular cuts through the cross-section of the dental region in axial view. The entire depth of a perpendicular line is a slice that starts from the top of the head and ends at the bottom. Taking the Maximum Intensity projection of the original scan values, of each perpendicular slice along the perpendicular gives us a 1D array.
2. We also find the MIP of the canal mask generated in a similar fashion which results in a 1D array.
3. We started from the right most perpendicular and then stacked each consecutive perpendicular’s MIP as we followed the dental arch to the left. This gives us a 2D array where the rows represent the entire depth of the original scan and the columns are equal to the number of points on the dental arch.
4. The same is done for the canal mask, and then both are cropped in their depths by noting the depth where the canal mask starts in the mask panorama.
5. Once we have the Panoramic views for both, the scan and the canal mask, we superimpose the canal mask onto the scan and label each canal with a different color. Blue represents the right canal and red represents the left canal.

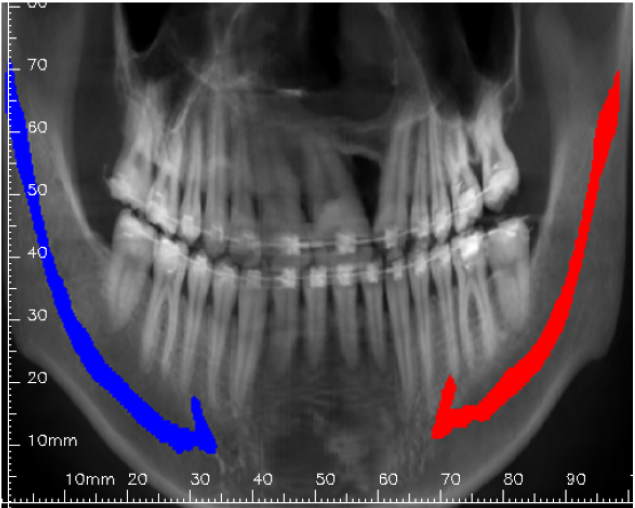


**Para-sagittal View:** For Para-sagittal view the perpendicular we drew previously give us the para-sagittal slices at each point on the dental arch. Therefore for each perpendicular, we extract its entire depth sice on the original scan as well as the canal mask. The canal mask gives us the position of the canal on that particular slice and a patch of 100 by 100 is cropped around the canal center on both the original para-sagittal slice as well as the canal masks’.

#### Experiment 2:

We realised that the panorama results were not up to par with standard reports so we performed another experiment where we generated a simpler and separate dental arch for the panorama view. In this after performing upto steps 1 to 5, we 

1. Selected 15 control points on the dental arch where three of the points are the topmost and both ending points of the arch. The rest are extracted at equal intervals.
2. Performed linear interpolation between these points.
3. Smoothed the results using Savgol filter with window size 109 and order 3.
4. We then generated perpendiculars and followed all the steps required to generate panorama view with gaslight change. Instead of taking MIP of each perpendicular slice, we took its Average Intensity Projection and got the following result which was on par with standard reports:



#### 

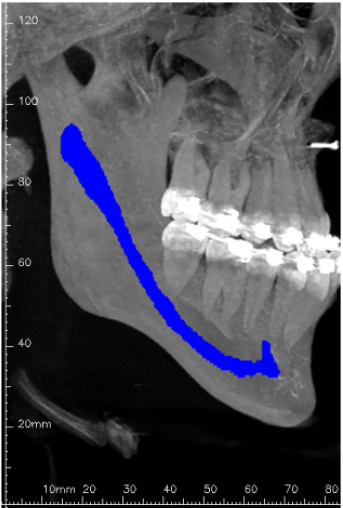
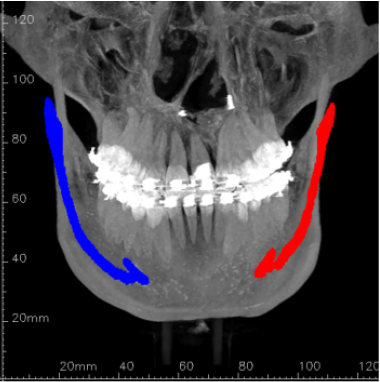
#### Experiment 3:

In this experiment we added teeth contour and canal contour to the para-sagittal slices.

We also added scales to each view by using the pixel spacing and slice thickness to get the accurate physical distance on the scale.

Each picture also labels the minimum distance between canal and tooth in that para-sagittal slice.

#### Experiment 4:

We also presented the results by taking MIPs of the entire head (original scan) through front and sagittal views. Same was done for the canal mask and the result superimposed on the MIP of the head. Blue represents the right canal and red represents the left canal.

[Link to folder containing dicom results](https://drive.google.com/drive/folders/1JtqzpWecCnTKClm57mj-pugTx0Y2ZSAQ?usp=sharing)

# Improvement of AI Module

Previously the canal segmentation model was trained on data manually localized (+-10 pixels from each dimension) and showed a decrement in accuracy when tested on scans localized through our image processing algorithm.

## Retraining on data localized from Localizer (Image Processing)

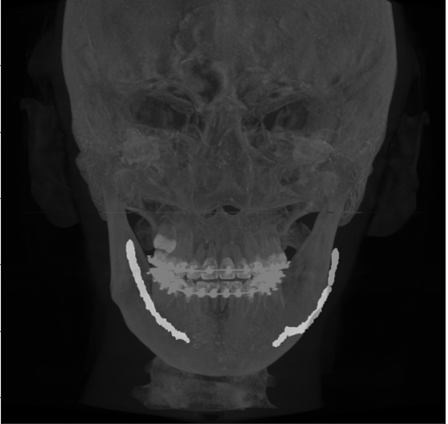
We trained two separate models on scans localized from our localizer and compared the results with our previous training.

| **Model** | **Dice score** |
| --- | --- |
| 3D Resplusplus (previous) | 59% |
| 3D UNET | 65% |
| 3D Resplusplus | 67.5% |

## Improvement of Localizer

To improve the accuracy further we saw the need to improve our localizer. For this we forewent Image processing, and trained a 3D Resplusplus model on full scans for canal segmentation. From the predicted naive canal masks, we extracted the required 3D patch (+-20 pixels in every dimension from the start of the first canal to the end of the last canal). This patch was our localized region of the scan. The localizer gave a 90.1 % IOU score.

## Retraining on data localized from Localizer (DNN)

We retrained the 3D Resplusplus model on the localized area and found a lot of faulty raw annotations or scans like:

After performing a detailed analysis on scans with less than 0.5 dice, we extracted out all the faulty scans from the data. The link to the visual analysis of each scan being removed is attached:

[Mandibular canal Segmentation faulty scans.pptx](https://docs.google.com/presentation/d/16VoeCG66KlavUC5aly35JAnEdBw1Q9Bw/edit?usp=sharing&ouid=116962362872131887109&rtpof=true&sd=true)

After retraining without faulty scans and testing we achieved **71.35%** accuracy for the left canal and **71.13%** accuracy for the right canal.

## Improvement through 2D segmentation

For further improvement we tried the technique of segmenting 2D para-sagittal slices of the localized region for a single canal. After training this we took joint prediction from 3D and 2D networks, and post-processed the resulting mask and achieved an accuracy of 69% for individual canal, which was lower than 3D predictions alone.

#### Current accuracy: 71.35 left canal & 71.13 right canal