Automated Segmentation of the Eustachian Tube – A Deep Learning Platform

Ameen Amanian[1,2], Yuliang Xiao[2], Zhiwei Gong[2], Deepa Galaiya[1], Russell Taylor[2], Mathias Unberath[2], Manish Sahu[2], and Francis Creighton[1,2]

[1]*Department of Otolaryngology – Head and Neck Surgery, Johns Hopkins University School of Medicine, Baltimore, Maryland;*
[2]*Laboratory for Computational Sensing and Robotics, Johns Hopkins University Whiting School of Engineering, Baltimore, Maryland*

Clinical Problem

- Eustachian tube dysfunction (ETD) results from impairment in middle ear ventilation and pressure regulation and has a significant impact on patients' quality of life.
- There are a lack of validated clinical tools for diagnosis of eustachian tube dysfunction. Therefore, clinicians have turned toward imaging to better understand the ET anatomy.
- Current works on morphological analysis of the ET for diagnostic purposes rely on manual annotation of computed tomography (CT) imaging. However, manual segmentation is cumbersome and difficult to translate into the clinical domain.

Proposed Solution

- **Develop a deep learning pipeline** to perform automated segmentation of the eustachian tube.
- **We hypothesize** that ET and nearby critical structures can be automatically segmented on CT imaging using a deep learning approach.

The Solution

Ground Truth Development

- Eustachian tube (ET), Internal carotid artery (ICA), and Torus Tubarius (TT) were manually annotated using an open source software, 3D slicer.
- Dataset: 22 training data, 9 test data with dimension 512x512xN (N is the number of slices) and 0.4 mm slice thickness.

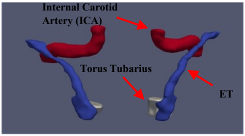

Figure 1 – Ground Truth Annotations

nnUNet Implementation

- No-New U-Net framework with 100 epochs evaluated by dice + focal loss function was made to learn the empirical parameters.
- The Dice Similarity Coefficient (DSC) and Average Hausdorff Distance (AHD) was

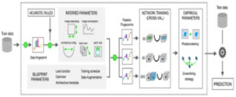

Figure 2 – nnUNet Network Architecture

Results and Discussions

22 CT scans were used for training the segmentation network. Predictions were assessed on 9 test CT volumes.

The framework showed increased accuracy along the bony and pharyngeal openings of the ET whereas there was decreased accuracy along the mid-cartilaginous segment as assessed via a heat map (Figure 3).

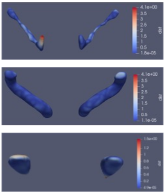

Figure 3 – Heatmap between ground truth and prediction (sequence: ET (top), ICA (Mid), Torus (bottom))

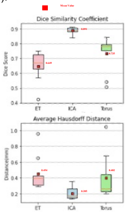

Figure 4 – Performance metrics trend

The network was limited in its performance when predicting the middle part of eustachian tube which is a segment without a clear boundary. However, this segment has not been shown to be associated with ETD .

Conclusion

We have developed the first deep learning framework which performs automated segmentation of the eustachian tube and surrounding anatomical structures with promising results.

Overall, we learned that the eustachian tube is a complex anatomical structure and its path is difficult to manually segment on CT images.

Given that MRI images have better soft tissue detail, next steps will entail developing our ground truth on MRI images and using registration-label propagation techniques, we will transfer the labels onto CT images to enhance the reliability of ground truth segmentations.

Statement of Impact

This pipeline serves as a tool for studying large datasets within the clinical domain in an efficient manner. Given its automated nature, it shows promise in being integrated into the current clinical workflow for the diagnosis and treatment of patients with ETD.