Automated Segmentation of the Eustachian Tube – A Deep Learning Platform

Project Proposal
Team #3
March 1, 2022

Mentors: Dr. Francis Creighton, Prof. Mathias Unberath, Prof. Russell Taylor

Team Members: Ameen Amanian, Chanha Kim, Yuliang Xiao

Project Goal: To develop a deep learning platform for automated segmentation of the eustachian tube and integration into an image-registration surgical navigation system for performing eustachian tube dilation.

Clinical Significance

Eustachian tube dilation is a procedure approved for the surgical management of eustachian tube dysfunction (ETD) [1]. ETD results from impairment of middle ear ventilation and pressure regulation [2]. As a result, patients experience a range of symptoms from ear pain, pressure, cracking, to difficulty hearing which has a significant impact on patients' quality of life [1]. Given the proximity of the eustachian tube to certain critical structures such as the internal carotid artery, an automated system that assesses for anatomical variations would reduce potential risks associated with ETD [1]. Currently, existing registration-segmentation pipelines have varying accuracy and can be computationally expensive [3-4]. Furthermore, there is a lack of image-registration surgical navigation system utilizing automated segmentations of preoperative CT for this procedure. Therefore, we aim to assess the utility of and develop a deep learning pipeline to perform automated segmentation of the eustachian tube, define near-by critical structures, and establish the first pipeline that can be integrated into a surgical navigation system.

Technical Approach

The proposed pipeline (Figure 1) will utilize a ground truth dataset which will be co-registered as part of the preprocessing. The data will then serve as input to the deep learning pipeline for performing semantic segmentation. The proposed deep learning algorithms include nnUNet [5], VoxelMorph [6], and DeepReg [7]. The predicted segmentations will then be validated in comparison with the ground truth.

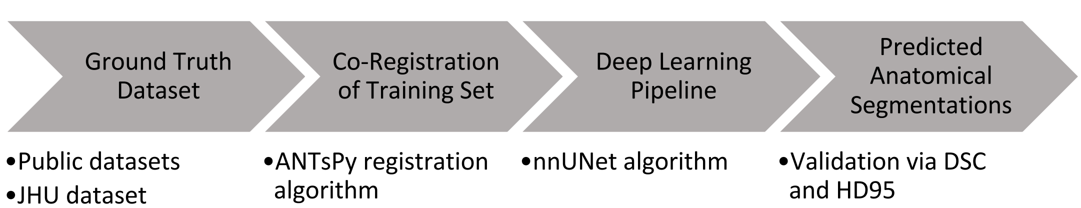


Figure 1 – Proposed Workflow

i. Data Preprocessing

In order to make the data compatible with the nnUNet, VoxelMorph and DeepReg frameworks, the raw images are required to be co-registered. We will be using ANTsPy, a Python library which includes blazing-fast IO, registration, segmentation, statistical learning, and visualization functionalities [8]. The workflow will include 1) Random selection of an image to be designated as a template, 2) Register the remaining images to the template, and 3) Apply the 'forward' deformation field to each image to ensure co-alignment within the dataset.

ii. nnU-Net Algorithm

For our project, we will be focusing on using nnU-net as the basis for our deep learning model for semantic segmentation of CT images as it has demonstrated an ability to handle dataset diversity found in the medical image segmentation domain (Figure 2) [5]. The workflow of nnU-net is as follows: nnU-net first uses its novel *heuristic rule* to determine the data-dependent hyperparameters, or *data fingerprints*, to automatically ingest the training data set. The *blueprint parameters* (e.g., loss function, and network architecture), *inferred parameters* (e.g., image resampling and batch size) along with the data fingerprint generate the *pipeline fingerprints*. The pipeline fingerprints then form *network training* for 2D, 3D, and 3D-Cascade U-Net using the hyperparameters determined so far. Using post-processing and ensembling strategies, nnU-Net uses the best configuration to produce the final prediction.

A motivation behind using nnU-net is its ability to handle a wide variety of target structures [5]. Unlike other deep learning models, nnU-net is an algorithm that is generalizable and has proven to surpass most existing approaches for data segmentation tasks. Furthermore, nnU-Net has a self-configuring ability in that it allows us to quickly train and use the model. Finally, the results can serve as a benchmark that can be improved upon if the training is not successful.

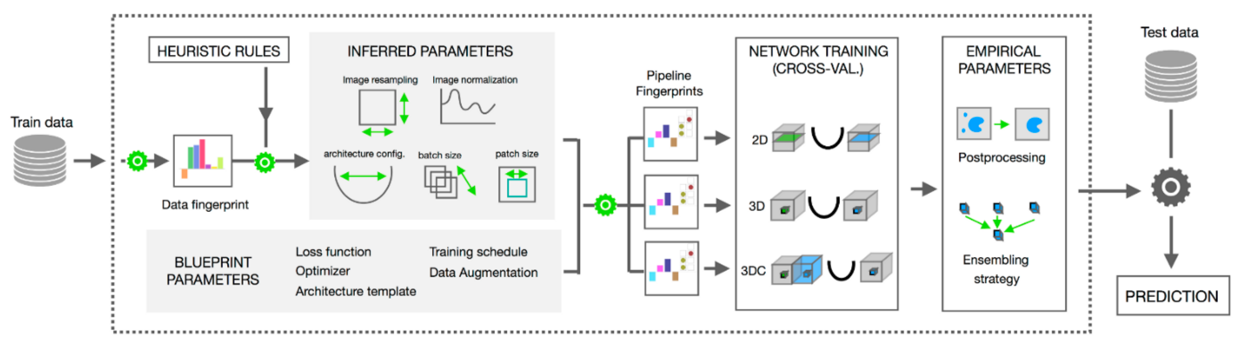


Figure 2 – nnU-Net Pipeline

The proposed deep learning workflow is shown in figure 3. First, we will generate a training and testing set and perform preprocessing on the training set (as discussed in previous section). The training set then serves as input into nnU-Net. Validation will be performed to see if the trained model generalizes well to test data (i.e., high Mean Surface Distance). Following nnU-Net validation, we can move on to building and training unsupervised learning models such as VoxelMorph and DeepReg.

If the training performance within nnU-Net requires improvement, we propose the following modifications to the nnU-net algorithm. First, we will incorporate CT-specific pre-processing which includes denoising, CT data interpolation with different splines, and finally windowing to increase the contrast across a region of interest. Second, we will perform manual adaptation of the loss function. nnU-net uses a dice loss (region-based loss function) or a cross-entropy loss (distribution-based loss function); however, we can cascade the dice and cross-entropy loss or provide weights to the background area of the label to soften the hard label used in loss functions. This can result in a regularization effect, increasing the robustness of the model and lowering the chances of overfitting [9]. Finally, we can extend or modify the heuristics used in nnU-net as suggested in the original nnU-net paper if the training fails, because the current heuristics may not be generalized enough to handle our domain-specific CT scans of the head.

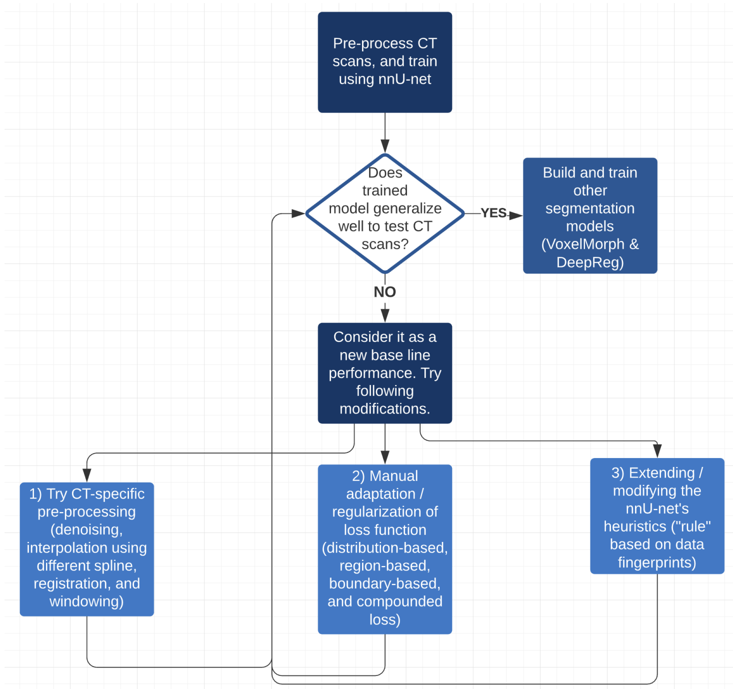


Figure 3 – Deep Learning Pipeline Workflow

iii. VoxelMorph Algorithm

VoxelMorph is a fast unsupervised-learning-based framework for deformable, pairwise medical image registration [6]. Compared to traditional registration methods, it treats registration as a function to map paired input images to a deformation field that make them aligned. Registration is formulated as an objective function and used in the convolution neural network to build the model that can optimize this function. In this algorithm, the first setting includes training the model to maximize standard image matching objective functions that are based on the image intensities. In the second setting, the auxiliary segmentations are leveraged in the training data, which increase accuracy when predicting on test datasets.

iv. nnUNet Model Validation

One measure used for model validation includes the dice similarity coefficient (DSC), a scoring system which measures volumetric overlap between two images [10]. However, as the eustachian tube is a very thin structure, the conventional use of DSC is not appropriate for our project. Thus, we will use the following metrics that capture the structure similarity.

Mean Surface Distance

The mean surface distance, d_{mean}, is the distance between the the surface (S) and the reference surface (S_{ref}) where $d(S,S_{ref})$ is the mean of distances between every surface voxel in S and the closest surface voxel in S_{ref}, while $d(S_{ref},S)$ is computed in a similar way.

$$d_{mean} = \frac{1}{2}\left[\bar{d}(S,S_{ref}) + \bar{d}(S_{ref},S)\right]$$

95% Hausdorff Distance (HD95)

The maximum Hausdorff distance (HD) is the maximum distance of a point in a set to the nearest point in the other set. More formally, the maximum Hausdorff distance from set X to set Y is a max-min function, defined as:

$$d_H(X,Y) = max\{d_{XY}, d_{YX}\} = max\left\{\max_{x\in X}\min_{y\in Y} d(x,y), \max_{y\in Y}\min_{x\in X} d(x,y)\right\}$$

95% HD is similar to the maximum HD; however, it is based on the calculation of the 95th percentile of the distances between boundary points in X and Y. The purpose for using 95% HD is to eliminate the impact of a very small subset of the outliers.

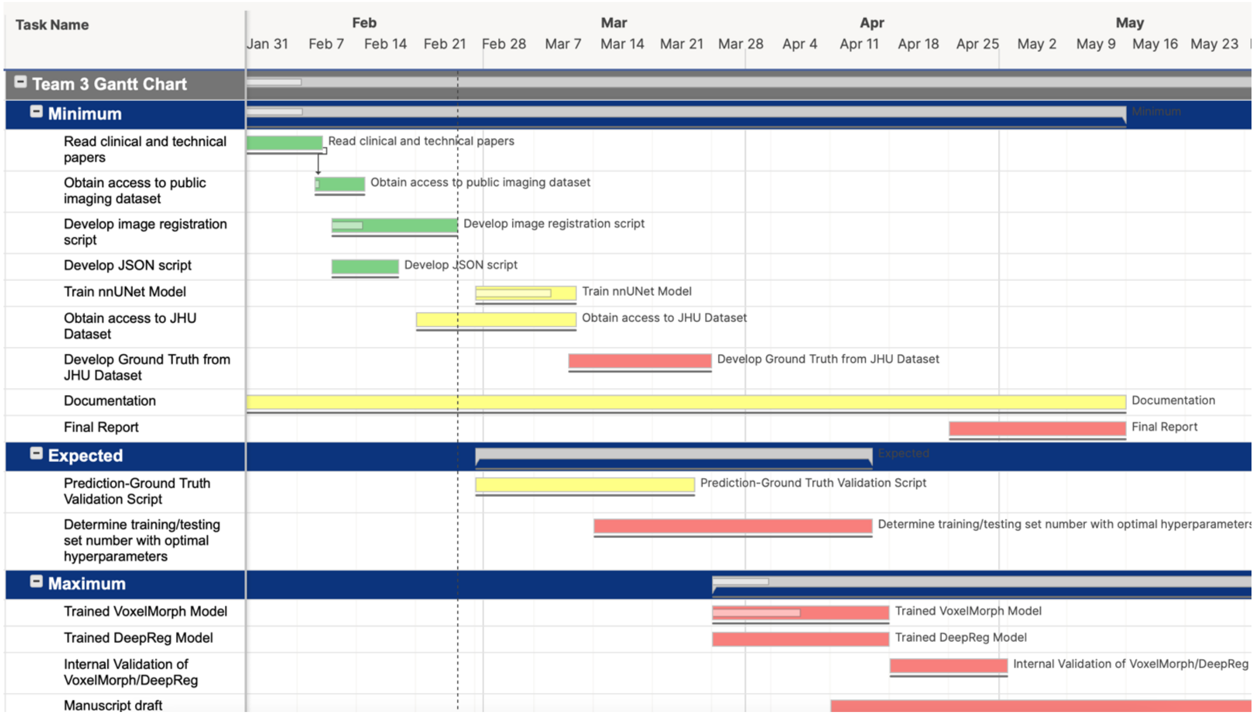
<u>List of deliverables</u>

Minimum	**Due Date**
CT co-registration script	February 25
Trained nnUNet model	March 11
Dataset containing ground truth segmentations	March 25
Documentation	May 13
Final report	May 13

Expected	**Due Date**
Validation script computing dice score and Hausdorff distance on predicted labels	March 25

Maximum	**Expected**
Trained VoxelMorph or DeepReg image registration models	April 15
Validation script comparing nnUNet model with VoxelMorph/DeepReg image registration models	April 30
Conference presentation and manuscript draft	July 1

Key Dates



Assigned Responsibilities

During the project, all team members will be involved in all aspects of the project. However, the tasks will be divided according to the following to meet the deliverables by the end of the semester:

Tasks	Assigned Member
Obtain publicly available images until IRB access granted	Ameen
Manual segmentation of the eustachian tube upon IRB approval	Ameen
Registration script via ANTs	Ameen and Yuliang
nnUNet model	Chanha, Yuliang, and Ameen
Validation scripts (DSC + HD95)	Chanha and Yuliang
Trained VoxelMorph Model	Yuliang
Trained DeepReg Model	Chanha
Ongoing documentation	Ameen, Chanha and Yuliang

Dependencies

Dependency	Solution	Alternative	Status	Date	Impact
Computation	Remote GPU access at Homewood	Google Colab MARCC	Successful access to remote GPU.	Feb 15	Will not be able to train neural network
Imaging Dataset	Pending access to IRB	Public datasets: New Mexico, NIH, Medical Decathlon	Using public datasets until IRB approved	March 15	Will not be able to segment ROI (ET)
Ground Truth	Annotations via 3D slicer	Segmentation of labels in public datasets	Performing segmentations on public dataset	March 25	Will not be able to train neural network

Management Plan

Meetings:
- Weekly meetings with mentors: Wednesdays 10:30AM-12:00PM
- Weekly meetings with LCSR Postdoc (Manish Sahu): Mondays 3:00-4:00PM
- Weekly team meetings: Fridays 3:00-4:00PM

File Management:
- GitHub: Source + Version Control
- Dataset: Microsoft Teams

Documentation:
- CourseWiki + Microsoft Teams

Reading List

1. Magro I, Pastel D, Hilton J, Miller M, Saunders J, Noonan K. Developmental Anatomy of the Eustachian Tube: Implications for Balloon Dilation. *Otolaryngol Head Neck Surg*. 2021;165(6):862-867. doi:10.1177/0194599821994817.
2. Froehlich MH, Le PT, Nguyen SA, McRackan TR, Rizk HG, Meyer TA. Eustachian Tube Balloon Dilation: A Systematic Review and Meta-analysis of Treatment Outcomes. *Otolaryngol Head Neck Surg*. 2020;163(5):870-882. doi:10.1177/0194599820924322.
3. Keschner D, Garg R, Loch R, Luk LJ. Repeat Eustachian Tube Balloon Dilation Outcomes in Adults With Chronic Eustachian Tube Dysfunction. *Otolaryngol Head Neck Surg*. August 2021. doi:10.1177/01945998211037975.
4. Isensee, F., Jaeger, P.F., Kohl, S.A.A. et al. nnU-Net: a self-configuring method for deep learning-based biomedical image segmentation. Nat Methods 18, 203–211 (2021). https://doi.org/10.1038/s41592-020-01008-z.
5. Balakrishnan G, Zhao A, Sabuncu MR, Guttag J, Dalca AV. VoxelMorph: A Learning Framework for Deformable Medical Image Registration. IEEE Trans Med Imaging. 2019 Feb 4. doi: 10.1109/TMI.2019.2897538. Epub ahead of print.
6. Fu Y, Brown NM, Saeed SU, et al., (2020). DeepReg: a deep learning toolkit for medical image registration. Journal of Open Source Software, 5(55), 2705. https: 1 //doi.org/10.21105/joss.02705.

References

1. Froehlich MH, Le PT, Nguyen SA, McRackan TR, Rizk HG, Meyer TA. Eustachian Tube Balloon Dilation: A Systematic Review and Meta-analysis of Treatment Outcomes. *Otolaryngol Head Neck Surg*. 2020;163(5):870-882. doi:10.1177/0194599820924322.
2. Magro I, Pastel D, Hilton J, Miller M, Saunders J, Noonan K. Developmental Anatomy of the Eustachian Tube: Implications for Balloon Dilation. *Otolaryngol Head Neck Surg*. 2021;165(6):862-867. doi:10.1177/0194599821994817.
3. Sinha A, Leonard S, Reiter A et al. Automated segmentation and statistical shape modeling of the paranasal sinuses to estimate natural variations. *Proc SPIE Int Soc Opt Eng*. 2016; 9784:97840D. doi: 10.1117/12.2217337.
4. Ding AS, Lu A, Li Z, et al. Automated Registration-Based Temporal Bone Computed Tomography Segmentation for Applications in Neurotologic Surgery. *Otolaryngol Head Neck Surg*. 2021; Online Ahead of Print. doi: 10.1177/01945998211044982.
5. Isensee F., Jaeger PF, Kohl SAA. et al. nnU-Net: a self-configuring method for deep learning-based biomedical image segmentation. *Nat Methods*. 2021;18, 203–211. doi.org/10.1038/s41592-020-01008-z.
6. Balakrishnan G, Zhao A, Sabuncu MR, Guttag J, Dalca AV. VoxelMorph: A Learning Framework for Deformable Medical Image Registration. *IEEE Trans Med Imaging*. 2019; Epub ahead of print. doi: 10.1109/TMI.2019.2897538.
7. Fu Y, Brown NM, Saeed SU, et al. DeepReg: a deep learning toolkit for medical image registration. *Journal of Open Source Software*. 2020; 5(55), 2705. doi.org/10.21105/joss.02705.
8. Avants, B. B., Tustison, N., & Song, G. (2009). Advanced normalization tools (ANTS). *Insight j*, *2*(365), 1-35.
9. Lu, W., Chaoli, W., Zhanquan, S., & Sheng, C. (2020). An Improved Dice Loss for Pneumothorax Segmentation by Mining the Information of Negative Areas. *IEEE Access PP(99):1-1*.
10. Zou KH, Warfield SK, Bharatha A, et al. Statistical validation of image segmentation quality based on a spatial overlap index. *Acad Radiol*. 2004;11(2):178-189. doi:10.1016/s1076-6332(03) 00671-8.
11. Dubuisson M-P, Jain AK. A modified Hausdorff distance for object matching. In: Proceedings of 12th International Confer- ence on Pattern Recognition. IEEE; 1994:566-568. doi:10.1109/ ICPR.1994.576361.