Semantic Segmentation of Extraocular Muscles on Computed Tomography Images using Convolutional Neural Networks

THE UNIVERSITY OF CHICAGO
UChicago Medicine

Ramkumar Shanker[1], Anthony Moliterno[1], Ellen Yu[1], Daniel Ginat[2], Shree Bharadwaj[1]

1. Graham School MScA, The University of Chicago 2. Department of Radiology, The University of Chicago

Background

- Thyroid eye disease (TED) is an autoimmune disorder characterized by enlargement of the extraocular muscles (EOM)
- The frequently affected EOMs in TED are Superior rectus (SR), Lateral rectus (LR), Medial rectus (MR) and Inferior rectus (IR) (shown in Fig 1)
- Computed Tomography (CT) is a preferred non-invasive method to facilitate TED diagnosis. Figure 2 is the demo-CT of enlarged LR/ MR in TED patient compared with normal EOMs
- The limitations of current EOMs measurement are:
  - Candidate CT slide selection is subjective to the examiner
  - 2D diameter of EOM may not always be the reliable indicator of muscle enlargement across all slices
  - making measurements manually is time-consuming

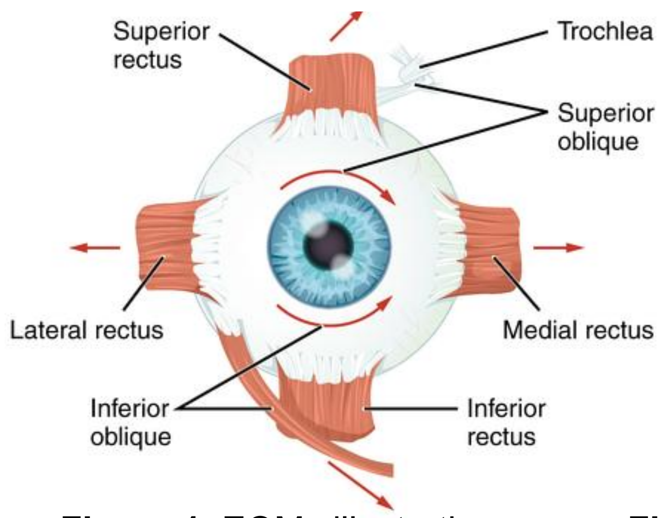


Figure 1. EOMs illustration

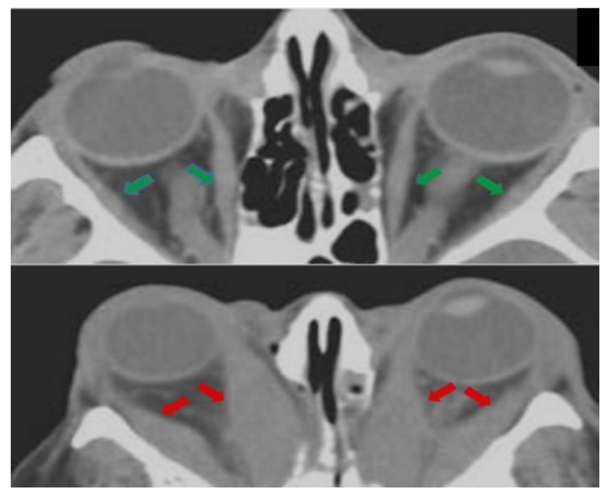


Figure 2. Comparison of normal sized EOM (green arrows) with enlarged EOMs in TED patient (red arrows).

Research Purpose and Objective

- Establish an automated semantic segmentation process to identify each major EOMs (SR, LR, MR, IR) and measure their 3D volumes which will enhance the overall accuracy compared with 2D measurement
- Develop a model to predict TED based on 3D volume measurement of EOMs and other TED related factors to improve the consistency between different examiners

Datasets

Source	Dataset Title	Subjects	Study Desc	Notes
Radiology department, the University of Chicago	Orbital CTs of patients with/ without visual enlargement of EOMS	104	Orbit	71 patients with at least one enlarged EOM, 29 patients with normol EOM
The Cancer Imaging Archive	Head-and-neck squamous cell carcinoma patients with CT taken during pre-treatment, mid-treatment and post-treatment	31 (93 scans)	Head-Neck	HNSCC[1]
AIDOC	Orbit CTs	30	Orbit	-

- Dataset Partition: Train (86%), Validation (7%), Test (7%).

Methods

- **Hypothesis**: EOM segmentation using 3D predictions will account for muscle contours across slices than that using stacked 2D predictions
- **Experiments on EOM prediction make performance comparisons:**
  - **2D coronal model** – using Coronal view slices and 2D conv kernels
  - **2D ensemble model**:
    1. Step 1: Predict only LR and MR using only axial slices
    2. Step 2: Predict only SR and IR using only sagittal slices
    3. Step 3: Predict all LR, MR, SR and IR using only coronal slices
    4. Step 4: Ensemble predictions (AND condition) of Step 1, 2 and Step 3
  - **3D model** – using 3D patches drawn from the overall scan
- CT Image pre-processing and manual segmentation using 3D-Slicer (Fig 3) with multi-class labels (Medial-1, Lateral-2, Superior-3, Inferior-4)

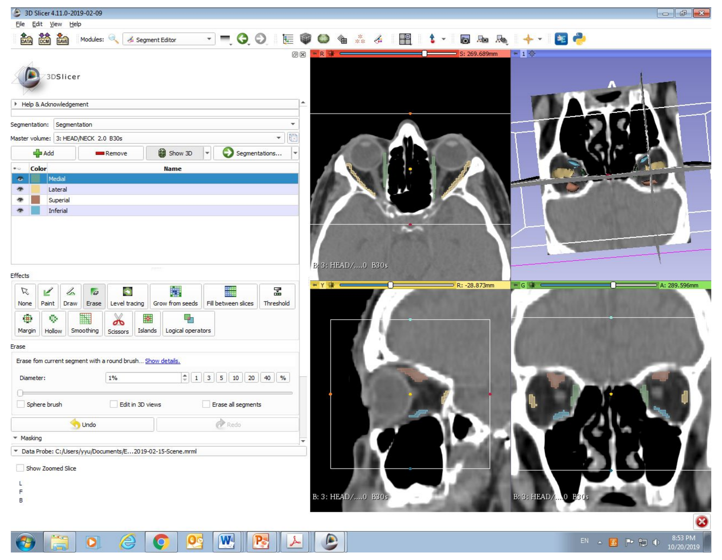


Figure 3. Illustration of 3D-Slicer tool for manual Segmentation.

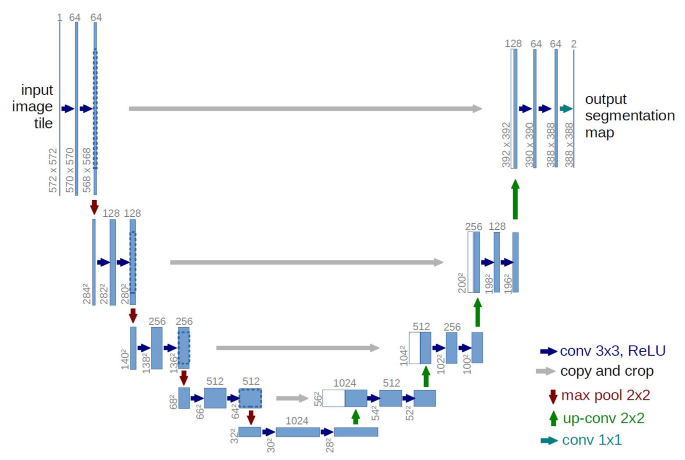


Figure 4. U- Net architecture [2].

Results

- Training results using optimal settings for network hyperparameters
- Training on “mscagpu” on midway2 (~3 minutes per epoch)

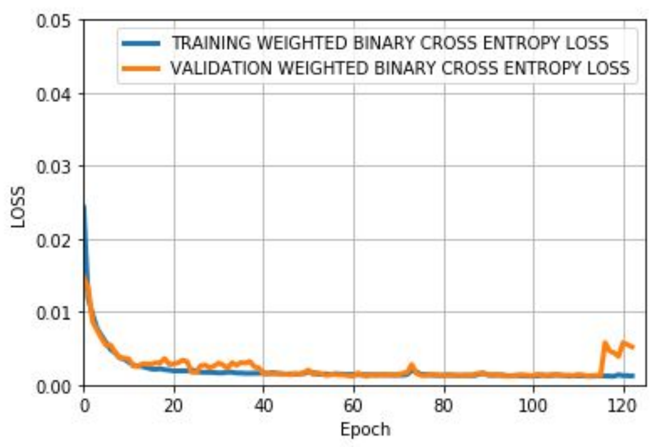


Figure 5. Loss curves for 2D Unet with Weighted Binary Cross Entropy

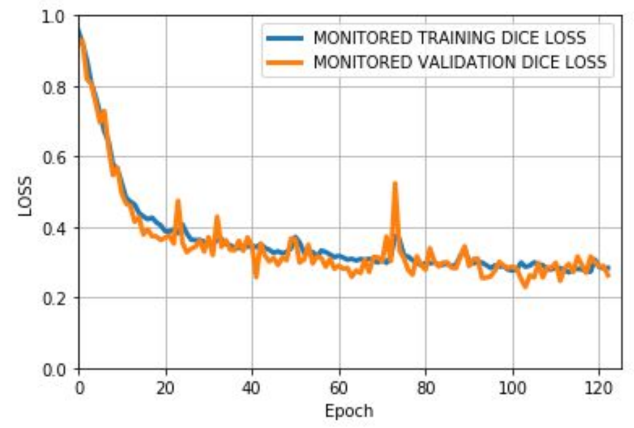


Figure 6. Monitored Loss curves for 2D Unet with Dice Coefficient Loss

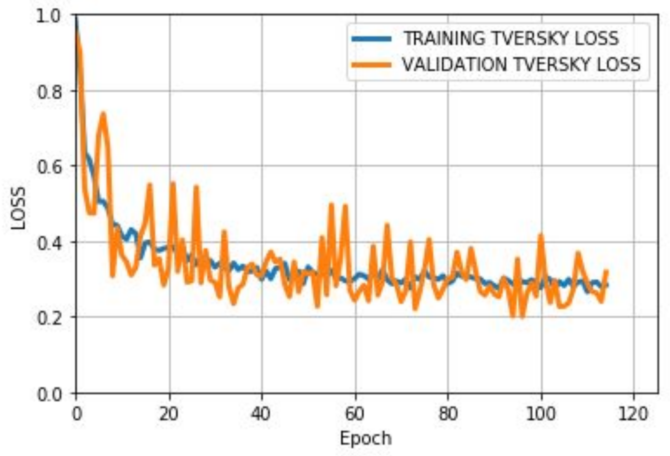


Figure 7. Loss curves for 3D Unet with Focal Tversky Loss[4]

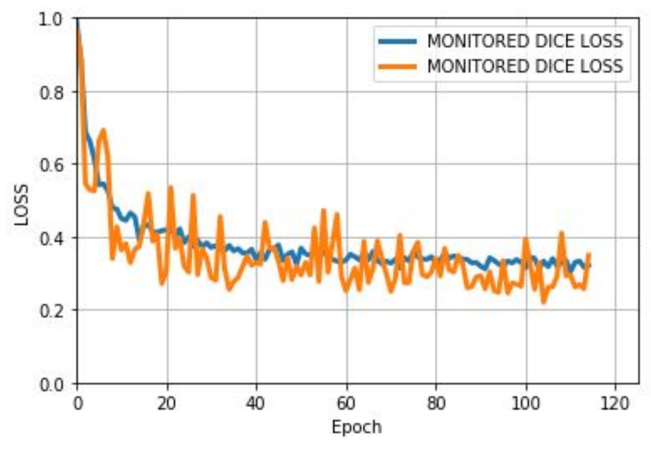
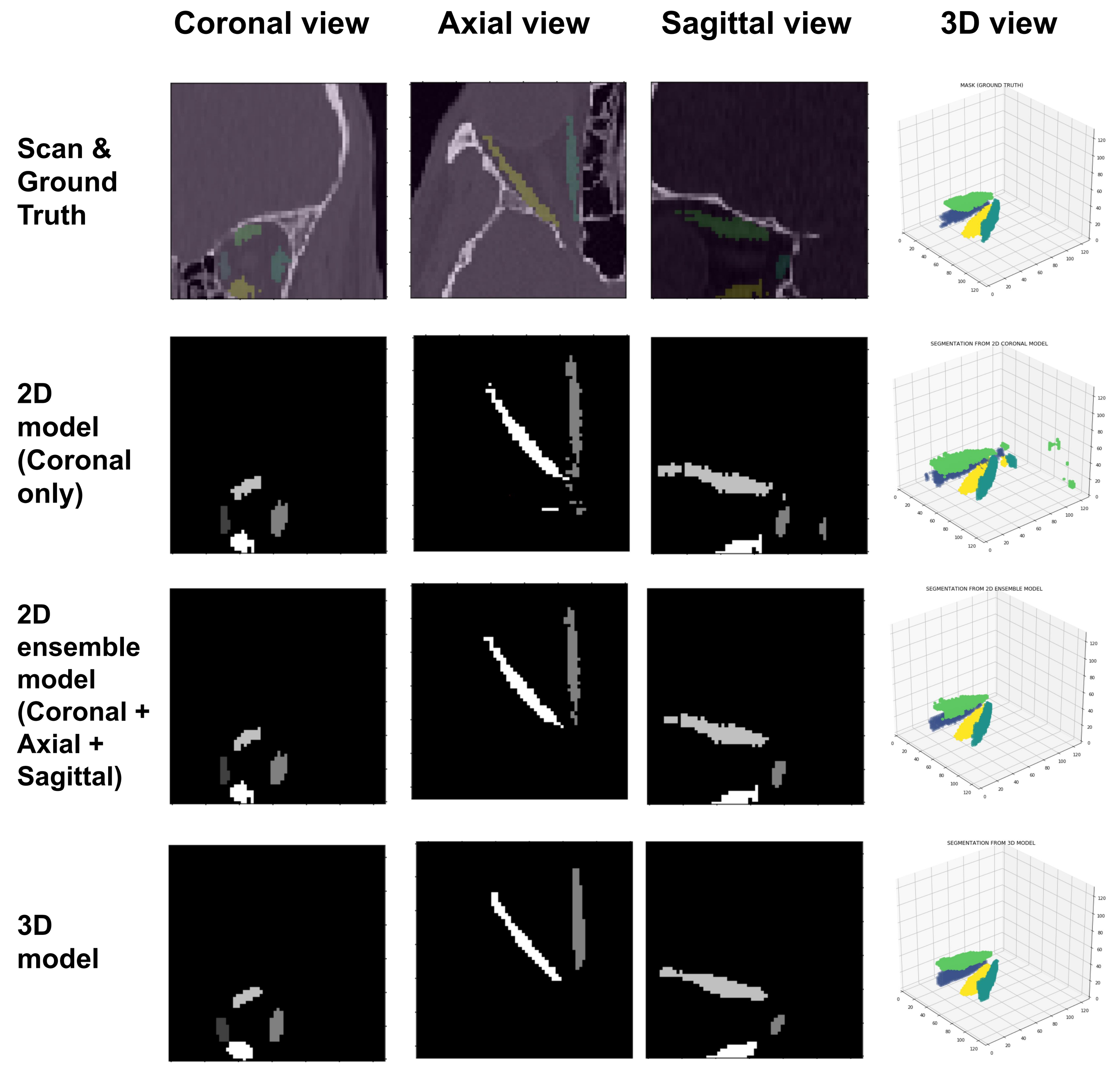


Figure 8. Monitored Loss curves for 3D Unet with Dice Coefficient Loss

Qualitative Performance



Discussion

- During inference, trained network takes ~3 seconds to segment one 64x64x64 image
- To enable CNNs to learn in 3D, we **resample the DICOM pixel arrays** using slice thickness, into real-life “voxels” of 1mm x 1mm x 1mm size
- Stacking 2D coronal segmentations into 3D produces irregular shapes when looked at from Sagittal and Axial views, but **3D segmentations accurately capture slice-to-slice contours**
- However, **ensemble segmentations from 3 moderately accurate 2D models capture 3D contours almost as accurately as a 3D model but with a fraction of training cost**

References

1. Tatiana Bejarano, Mariluz De Ornelas Couto, Ivaylo B. Mihaylov. **Head-and-neck squamous cell carcinoma patients with CT taken during pre-treatment, mid-treatment, and post-treatment Dataset.** The Cancer Imaging Archive 2018
2. Olaf Ronneberger, Philipp Fischer, and Thomas Brox **U-Net: Convolutional Networks for Biomedical Image Segmentation** Computer Science Department and BIOSS Centre for Biological Signalling Studies,University of Freiburg, Germany
3. Patrice Y. Simard, Dave Steinkraus, John C. Platt **Best Practices for Convolutional Neural Networks Applied to Visual Document Analysis** ICDAR 2003
4. Nabila Abraham, Naimul Mefraz Khan **A Novel Focal Tversky loss function with improved Attention U-Net for lesion segmentation IEEE ISBI 2019**