Deep Segmentation for Medical Volume Rendering

[ECS 289H] Final Project Report

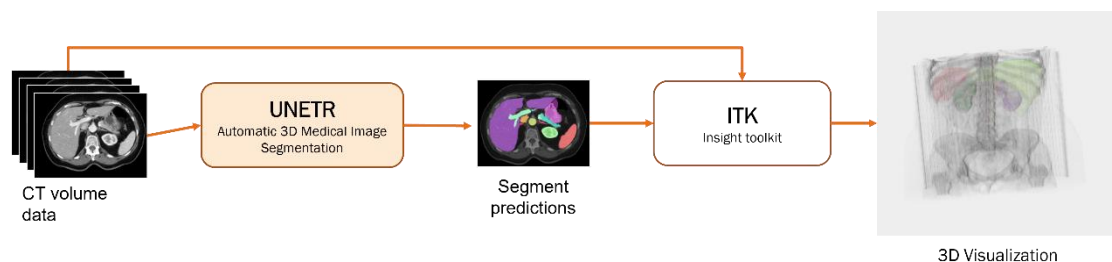Li-wen Lin 920253827

Mar. 13, 2022

Motivation and Objectives

In the fall quarter of 2021, I participated in Prof. Ma's Advance Visualization course and learned about volume rendering techniques and its applications. During volume rendering, it is often difficult for users, especially novice users, to tune the transfer function in order to visualize the portion of interest. This problem may result from low contrast intensities of different segments (e.g., organs in human body CT scan) thus making the segments hard to separate by tweaking the opacity of different intensity values. To tackle this problem, I want to adopt deep learning method to assist the separation of organs inside human body.

Background and Related Work

Convolutional Neural Networks (CNN) have shown prominence for past medical image segmentation applications [4, 5]. However, the locality of convolutional layers in CNNs limits its capability of learning long-range social dependencies. Recently, multiple methods were proposed that explore the possibility of using transformer-based models for the task of 2D image segmentation [6, 7], since transformer encoders have great capability of modeling long-range dependencies and capturing global context. UNETR [1] is a U-shaped encoder-decoder network using a transformer encoder. This model was proposed by Hatamizadeh et al. in 2021 and is tailored for 3D segmentation, which is a good fit for dealing with volumetric data.

Methodology

CT volume data → UNETR (Automatic 3D Medical Image Segmentation) → Segment predictions → ITK (Insight toolkit) → 3D Visualization

1. In the UNET architecture, a 3D input volume is divided into a sequence of uniform non-

overlapping patches and projected into an embedding space using a linear layer.

2. There is an official implementation of UNETR. I set up the environment using Anaconda virtual environments. The main required packages are PyTorch and MONAI [11].

3. The abdominal volume data is fed into the network for inference of organ segment predictions. Inference is done by using sliding window with overlapping portions between the neighboring patches. I tested the same data with overlapping portions set to 0.0 and 0.5 to compare the prediction results and inference time.

4. For visualization, I originally planned on using INVIWO [2] to utilize its visualization pipeline and built-in transfer function tool. However, I realized that it would be redundant to visualize the segmentation result on a third-party software when there are plenty of ways to visualize it in a python environment. In the end, I choose to use ITK (Insight Toolkit) [10] for volume visualization with segmentation results.

Libraries

1. UNETR is implemented with PyTorch and MONAI (AI toolkit for healthcare imaging) [11].

2. ITK is a library developed for image registration and segmentation routines originally developed for the Visible Human Project. It provides useful interactive visualization widgets suitable for this project [8].

3. BTCV dataset is from the BTCV Challenge. It consists of 30 subjects with abdominal CT scans where 13 organs were annotated by medical professionals. Each CT scan consists of 80 to 225 slices with 512*512 pixels and slice thickness ranging from 1 to 6 mm.
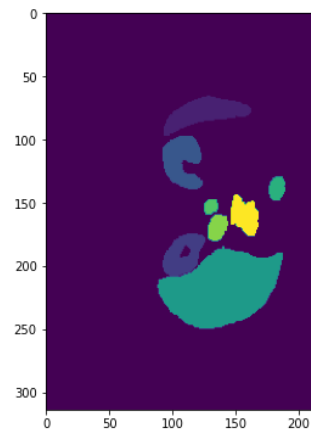
Result and Evaluation

- Calculate **dice score** to show how similar presdiction results are similar to ground truth.

$$\text{Dice score} = \frac{2 \cdot \text{number of true positives}}{2 \cdot \text{number of true positives} + \text{number of false positives} + \text{number of false negatives}}$$

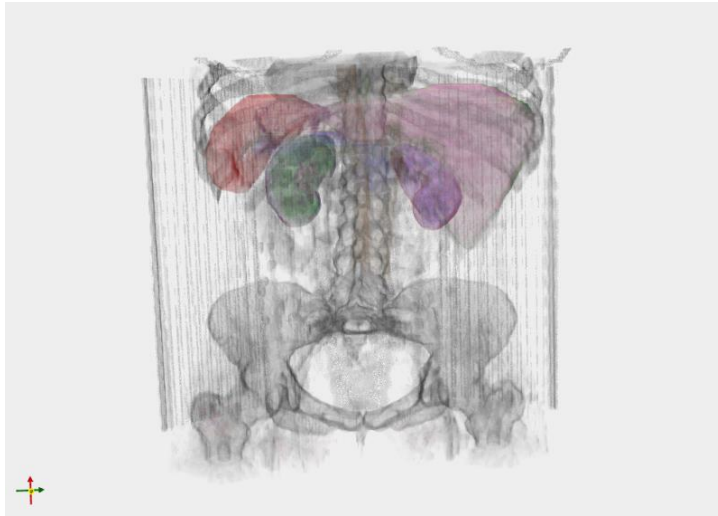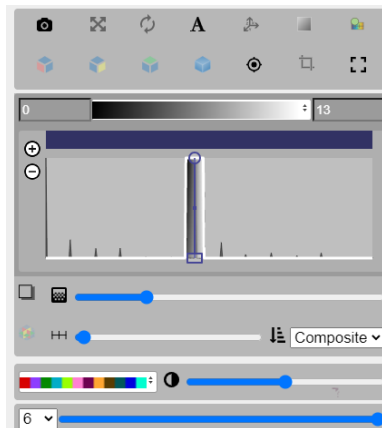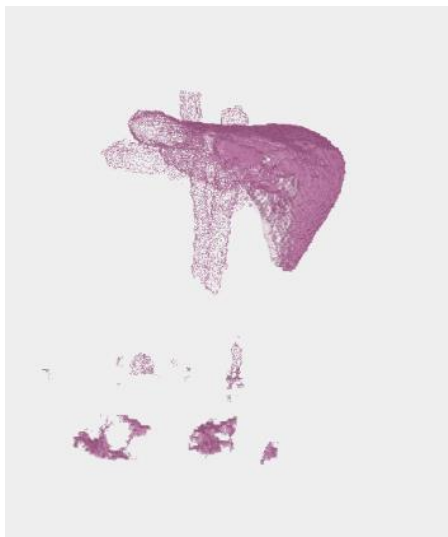● Prediction results:
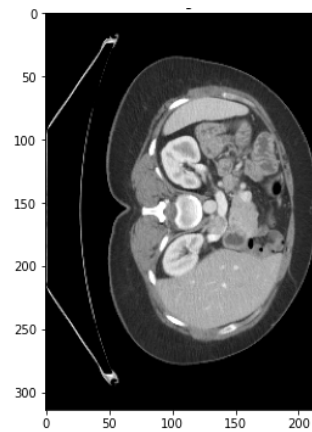
| CT slice | Ground Truth |

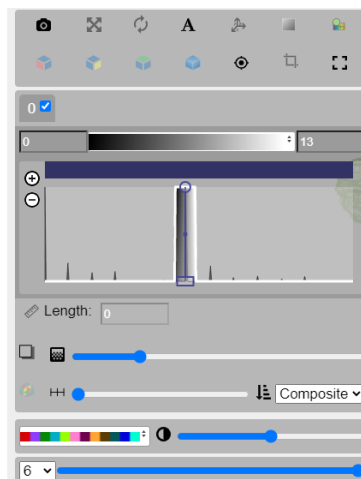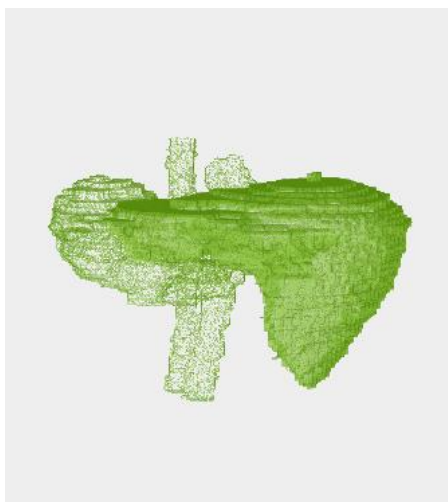|	Overlapping = 0.0	Overlapping = 0.5
Dice score	0.7136022788046353	0.7549031257410542
Inference time	103.0689 seconds	257.4765 seconds

- **3D visualization result:**

Visualize input volume and label map using segment result

Visualize of segment result only (liver)

Visualize of ground truth (liver)

Discussion and future works

From the rendering results, we can see that organs can be shown clearly seen with the help of deep segmentation results mapped to the volume.
The second visualization of segmented liver shows that there are still parts being labeled falsely. I suspect that this results partially from the ground truth label that was used for training, as we can see the attempt of visualizing liver also shows other organs.

This project can be improved by more accurate annotations of training data, or preprocessing the input volume such as denoising and edge detection may result in better separation of organs.

References

[1] Hatamizadeh et al., "UNETR: Transformers for 3D Medical Image Segmentation," 2022 IEEE/CVF Winter Conference on Applications of Computer Vision (WACV), 2022, pp. 1748-1758, doi: 10.1109/WACV51458.2022.00181.

[2] D. Jönsson et al., "Inviwo — A Visualization System with Usage Abstraction Levels," in IEEE Transactions on Visualization and Computer Graphics, vol. 26, no. 11, pp. 3241-3254, 1 Nov. 2020, doi: 10.1109/TVCG.2019.2920639.

[3] D. Engel and T. Ropinski, "Deep Volumetric Ambient Occlusion," in IEEE Transactions on Visualization and Computer Graphics, vol. 27, no. 2, pp. 1268-1278, Feb. 2021, doi: 10.1109/TVCG.2020.3030344.

[4] Qi Dou, Hao Chen, Yueming Jin, Lequan Yu, Jing Qin, and Pheng-Ann Heng. 3d deeply supervised network for automatic liver segmentation from ct volumes. In International conference on medical image computing and computer-assisted intervention, pages 149–157. Springer, 2016.

[5] Qikui Zhu, Bo Du, Baris Turkbey, Peter L Choyke, and Pingkun Yan. Deeply-supervised cnn for prostate segmentation. In 2017 international joint conference on neural networks (IJCNN), pages 178–184. IEEE, 2017.

[6] Sixiao Zheng, Jiachen Lu, Hengshuang Zhao, Xiatian Zhu, Zekun Luo, Yabiao Wang, Yanwei Fu, Jianfeng Feng, Tao Xiang, Philip HS Torr, et al. Rethinking semantic segmentation from a sequence-to-sequence perspective with transformers. arXiv preprint arXiv:2012.15840, 2020.

[7] Mathilde Caron, Hugo Touvron, Ishan Misra, Herv´e J´egou, Julien Mairal, Piotr Bojanowski, and Armand Joulin. Emerging properties in self-supervised vision transformers. arXiv preprint arXiv:2104.14294, 2021.

[8] Matthew Brett, "Analysis of Functional Magnetic Resonance Imaging in Python", 2007

[9] https://www.synapse.org/#!Synapse:syn3193805/wiki/89480

[10]https://github.com/InsightSoftwareConsortium/itkwidgets

[11]https://github.com/Project-MONAI