
Organ Detection on Medical Image using YOLO and UNet

— Yudong Sun, Anjia Wang —
Apr. 29th, 2019

Content

1. Introduction
2. YOLO
3. Unet
4. Results
5. Conclusion

Introduction

- Traditionally a large amount of CT Medical Images are examined by medical experts manually to recognize certain organs.
- Nowadays, deep learning is able to automatically recognize the organs on CT images, which saves significant time.
- Deep learning approaches:
 - You Only Look Once(Yolo)
 - U-Net: Convolutional Networks for Biomedical Image Segmentation(U-Net)

First Part – You Only Look Once (YOLO)

- A single neural network predicts bounding boxes and class probabilities directly from full images in one evaluation.
- Optimized end-to-end directly on detection performance.

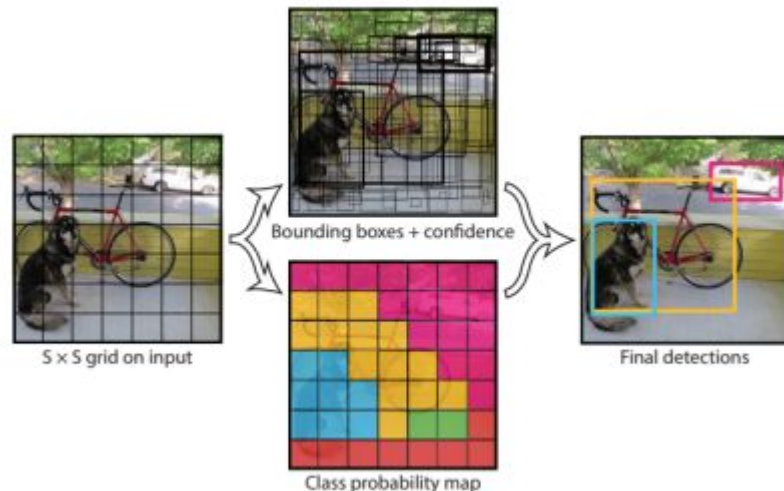
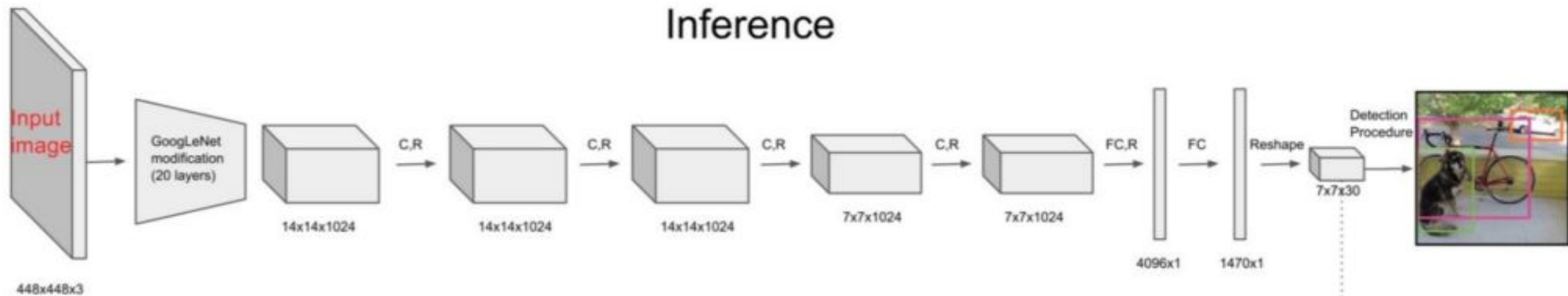


Figure 2: The Model. Our system models detection as a regression problem. It divides the image into an $S \times S$ grid and for each grid cell predicts B bounding boxes, confidence for those boxes, and C class probabilities. These predictions are encoded as an $S \times S \times (B * 5 + C)$ tensor.

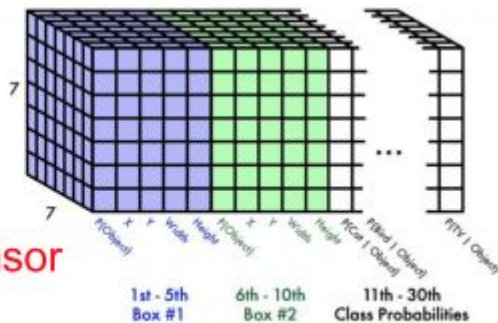
You Only Look Once (YOLO)



Each cell predicts:

- For each bounding box:
 - 4 coordinates (x, y, w, h)
 - 1 confidence value
- Some number of class probabilities

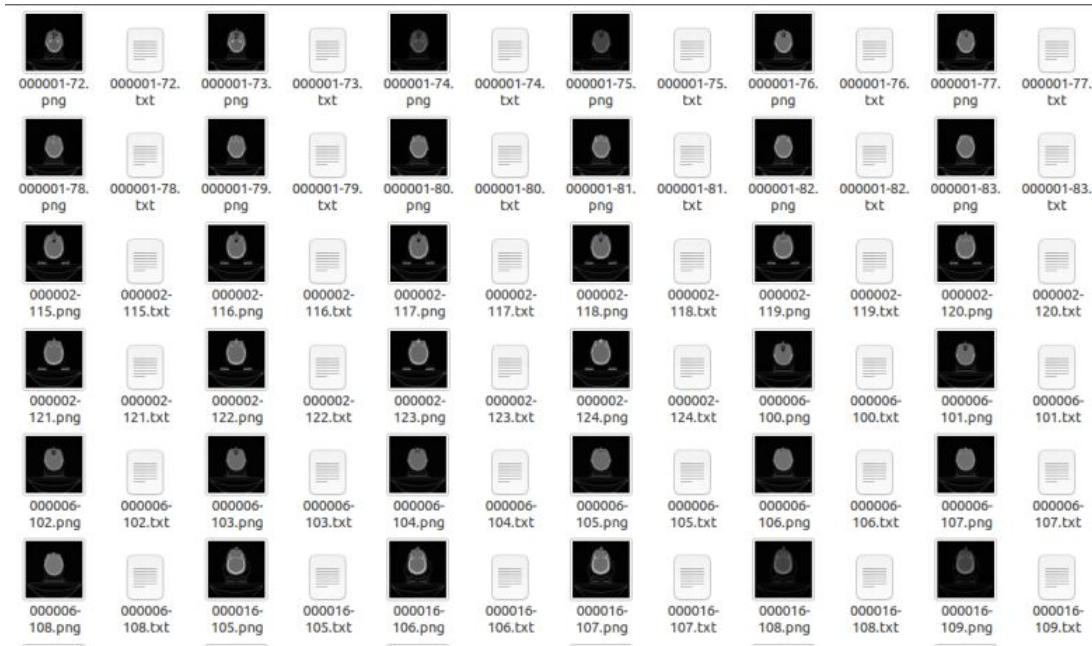
$S * S * (B * 5 + C)$ tensor



Data Preprocessing

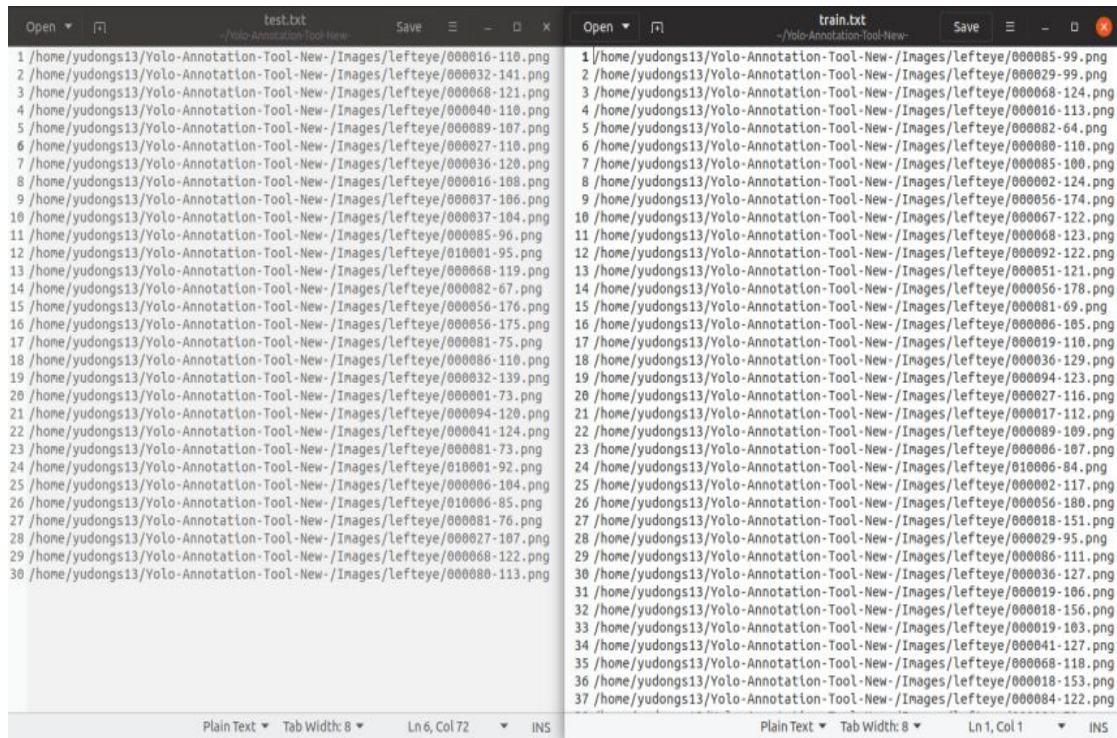
- Main thing for creating text file of images:
 - Create .txt-file for each .jpg-image-file - in the same directory and with the same name, but with .txt-extension
 - For img1.jpg you should create img1.txt containing:
<object-class> <x> <y>
<width> <height>. <x> <y> - are center of rectangle (are not top-left corner)

```
1 0.716797 0.395833 0.216406 0.147222
```



Yolo Training

- We tell YOLOv2 what images form our actual training set, and what will serve as test set: the test.txt and train.txt files. We use 305 images to run the model.
- To run this model and the other models in this project, we used the Tesla K80 GPU, kepler architecture, Intel Xeon.

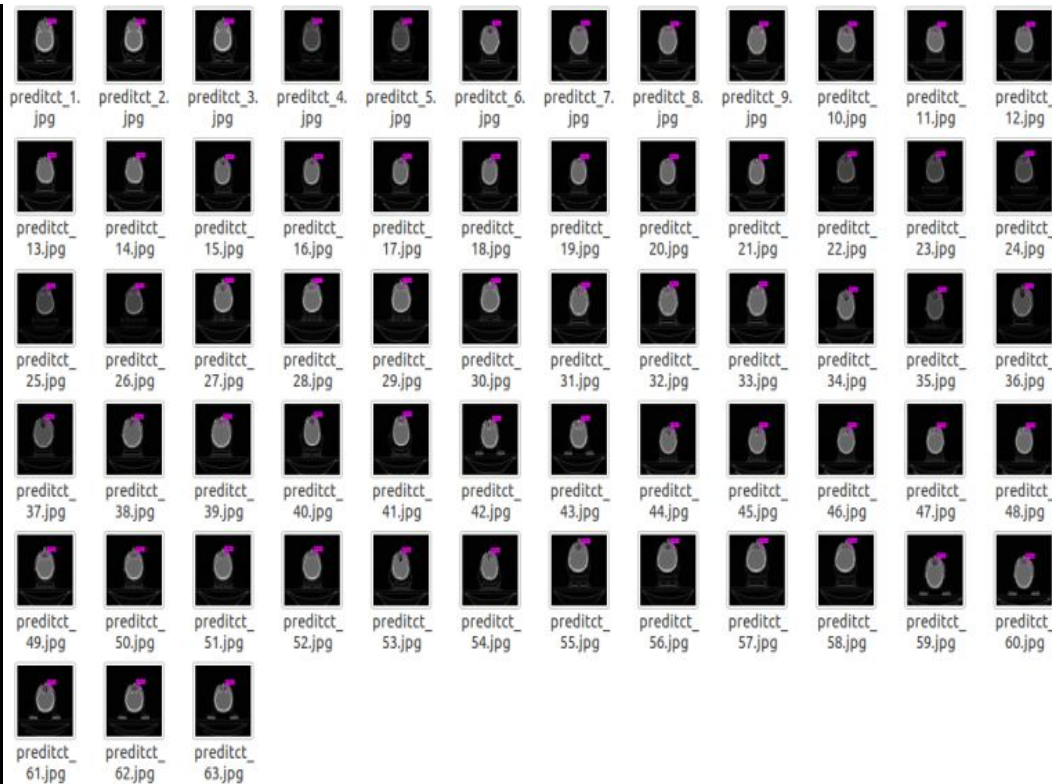
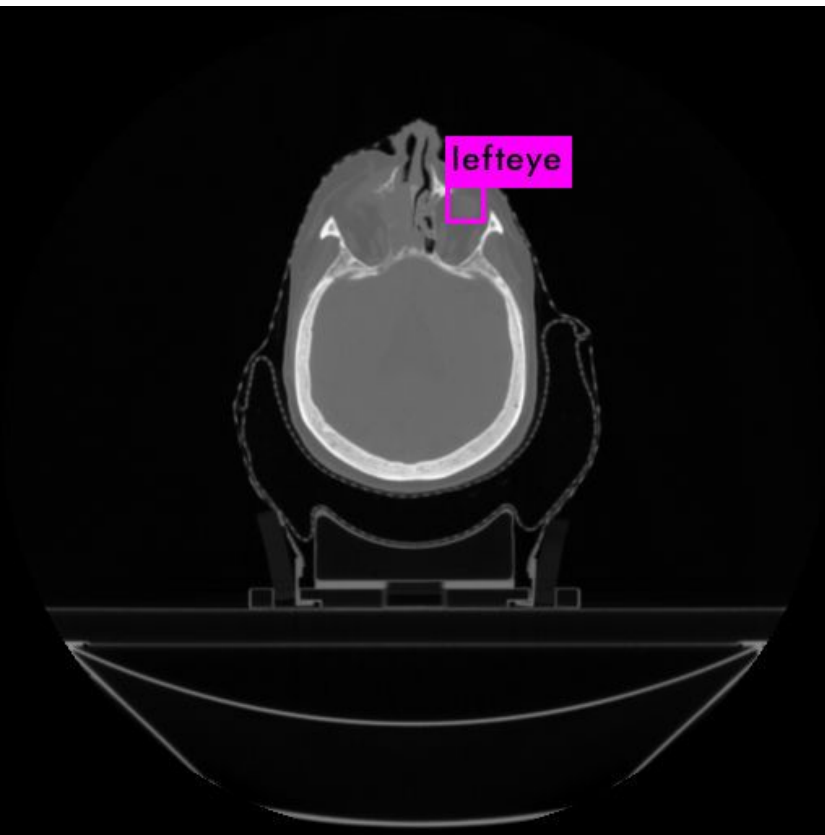


The image shows two side-by-side text editors. The left editor, titled 'test.txt', contains a list of 30 image file paths, each starting with '/home/yudongs13/Yolo-Annotation-Tool-New/Images/lefteye/' followed by a unique ID. The right editor, titled 'train.txt', contains a list of 37 image file paths, each starting with '/home/yudongs13/Yolo-Annotation-Tool-New/Images/lefteye/' followed by a unique ID. Both editors have a status bar at the bottom indicating 'Plain Text', 'Tab Width: 8', 'Ln 6, Col 72', and 'INS'.

```
test.txt
1 /home/yudongs13/Yolo-Annotation-Tool-New/Images/lefteye/000016-110.png
2 /home/yudongs13/Yolo-Annotation-Tool-New/Images/lefteye/000032-141.png
3 /home/yudongs13/Yolo-Annotation-Tool-New/Images/lefteye/000068-121.png
4 /home/yudongs13/Yolo-Annotation-Tool-New/Images/lefteye/000040-110.png
5 /home/yudongs13/Yolo-Annotation-Tool-New/Images/lefteye/000089-107.png
6 /home/yudongs13/Yolo-Annotation-Tool-New/Images/lefteye/000027-110.png
7 /home/yudongs13/Yolo-Annotation-Tool-New/Images/lefteye/000036-120.png
8 /home/yudongs13/Yolo-Annotation-Tool-New/Images/lefteye/000016-108.png
9 /home/yudongs13/Yolo-Annotation-Tool-New/Images/lefteye/000037-106.png
10 /home/yudongs13/Yolo-Annotation-Tool-New/Images/lefteye/000037-104.png
11 /home/yudongs13/Yolo-Annotation-Tool-New/Images/lefteye/000085-96.png
12 /home/yudongs13/Yolo-Annotation-Tool-New/Images/lefteye/010001-95.png
13 /home/yudongs13/Yolo-Annotation-Tool-New/Images/lefteye/000068-119.png
14 /home/yudongs13/Yolo-Annotation-Tool-New/Images/lefteye/000082-67.png
15 /home/yudongs13/Yolo-Annotation-Tool-New/Images/lefteye/000056-176.png
16 /home/yudongs13/Yolo-Annotation-Tool-New/Images/lefteye/000056-175.png
17 /home/yudongs13/Yolo-Annotation-Tool-New/Images/lefteye/000081-75.png
18 /home/yudongs13/Yolo-Annotation-Tool-New/Images/lefteye/000086-110.png
19 /home/yudongs13/Yolo-Annotation-Tool-New/Images/lefteye/000032-139.png
20 /home/yudongs13/Yolo-Annotation-Tool-New/Images/lefteye/000001-73.png
21 /home/yudongs13/Yolo-Annotation-Tool-New/Images/lefteye/000094-120.png
22 /home/yudongs13/Yolo-Annotation-Tool-New/Images/lefteye/000041-124.png
23 /home/yudongs13/Yolo-Annotation-Tool-New/Images/lefteye/000081-73.png
24 /home/yudongs13/Yolo-Annotation-Tool-New/Images/lefteye/010001-92.png
25 /home/yudongs13/Yolo-Annotation-Tool-New/Images/lefteye/000006-104.png
26 /home/yudongs13/Yolo-Annotation-Tool-New/Images/lefteye/010006-85.png
27 /home/yudongs13/Yolo-Annotation-Tool-New/Images/lefteye/000081-76.png
28 /home/yudongs13/Yolo-Annotation-Tool-New/Images/lefteye/000027-107.png
29 /home/yudongs13/Yolo-Annotation-Tool-New/Images/lefteye/000068-122.png
30 /home/yudongs13/Yolo-Annotation-Tool-New/Images/lefteye/000080-113.png

train.txt
1 /home/yudongs13/Yolo-Annotation-Tool-New/Images/lefteye/000085-99.png
2 /home/yudongs13/Yolo-Annotation-Tool-New/Images/lefteye/000029-99.png
3 /home/yudongs13/Yolo-Annotation-Tool-New/Images/lefteye/000068-124.png
4 /home/yudongs13/Yolo-Annotation-Tool-New/Images/lefteye/000016-113.png
5 /home/yudongs13/Yolo-Annotation-Tool-New/Images/lefteye/000082-64.png
6 /home/yudongs13/Yolo-Annotation-Tool-New/Images/lefteye/000080-110.png
7 /home/yudongs13/Yolo-Annotation-Tool-New/Images/lefteye/000085-100.png
8 /home/yudongs13/Yolo-Annotation-Tool-New/Images/lefteye/000082-124.png
9 /home/yudongs13/Yolo-Annotation-Tool-New/Images/lefteye/000056-174.png
10 /home/yudongs13/Yolo-Annotation-Tool-New/Images/lefteye/000067-122.png
11 /home/yudongs13/Yolo-Annotation-Tool-New/Images/lefteye/000068-123.png
12 /home/yudongs13/Yolo-Annotation-Tool-New/Images/lefteye/000092-122.png
13 /home/yudongs13/Yolo-Annotation-Tool-New/Images/lefteye/000051-121.png
14 /home/yudongs13/Yolo-Annotation-Tool-New/Images/lefteye/000056-178.png
15 /home/yudongs13/Yolo-Annotation-Tool-New/Images/lefteye/000081-69.png
16 /home/yudongs13/Yolo-Annotation-Tool-New/Images/lefteye/000086-105.png
17 /home/yudongs13/Yolo-Annotation-Tool-New/Images/lefteye/000019-118.png
18 /home/yudongs13/Yolo-Annotation-Tool-New/Images/lefteye/000036-129.png
19 /home/yudongs13/Yolo-Annotation-Tool-New/Images/lefteye/000094-123.png
20 /home/yudongs13/Yolo-Annotation-Tool-New/Images/lefteye/000027-116.png
21 /home/yudongs13/Yolo-Annotation-Tool-New/Images/lefteye/000017-112.png
22 /home/yudongs13/Yolo-Annotation-Tool-New/Images/lefteye/000089-109.png
23 /home/yudongs13/Yolo-Annotation-Tool-New/Images/lefteye/000086-107.png
24 /home/yudongs13/Yolo-Annotation-Tool-New/Images/lefteye/010006-84.png
25 /home/yudongs13/Yolo-Annotation-Tool-New/Images/lefteye/000082-117.png
26 /home/yudongs13/Yolo-Annotation-Tool-New/Images/lefteye/000056-180.png
27 /home/yudongs13/Yolo-Annotation-Tool-New/Images/lefteye/000018-151.png
28 /home/yudongs13/Yolo-Annotation-Tool-New/Images/lefteye/000029-95.png
29 /home/yudongs13/Yolo-Annotation-Tool-New/Images/lefteye/000086-111.png
30 /home/yudongs13/Yolo-Annotation-Tool-New/Images/lefteye/000036-127.png
31 /home/yudongs13/Yolo-Annotation-Tool-New/Images/lefteye/000019-106.png
32 /home/yudongs13/Yolo-Annotation-Tool-New/Images/lefteye/000018-156.png
33 /home/yudongs13/Yolo-Annotation-Tool-New/Images/lefteye/000019-103.png
34 /home/yudongs13/Yolo-Annotation-Tool-New/Images/lefteye/000041-127.png
35 /home/yudongs13/Yolo-Annotation-Tool-New/Images/lefteye/000068-118.png
36 /home/yudongs13/Yolo-Annotation-Tool-New/Images/lefteye/000018-153.png
37 /home/yudongs13/Yolo-Annotation-Tool-New/Images/lefteye/000084-122.png
```

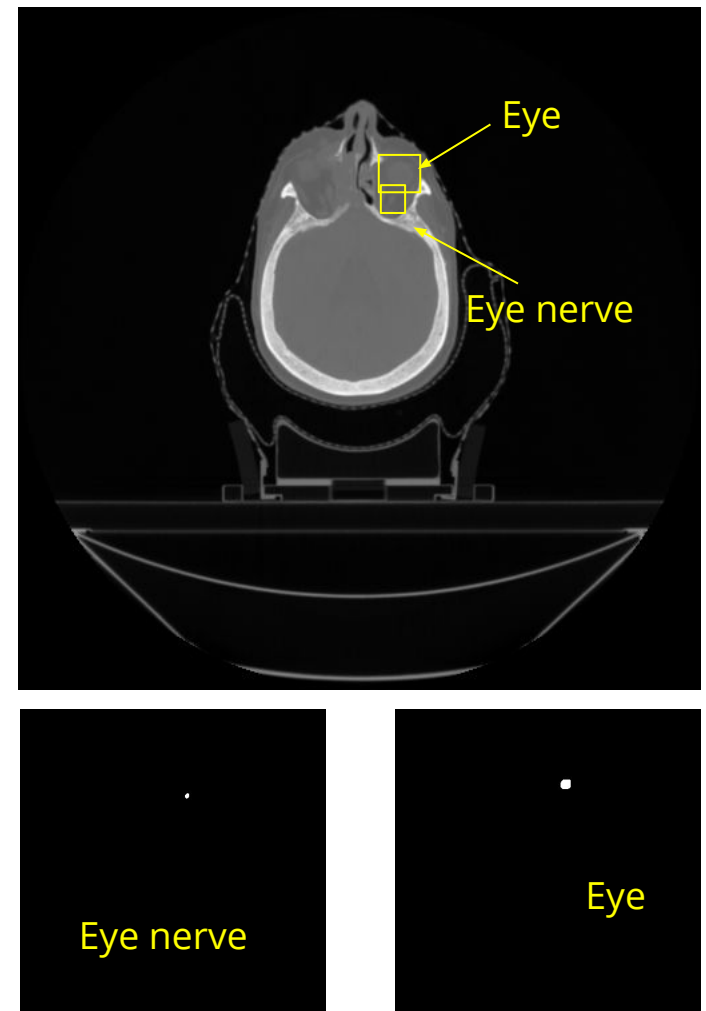

YOLO test results



UNet

UNet is convolutional neural network developed at the Computer Science Department of the University of Freiburg, Germany in 2015.

UNet framework is implemented using Tensorflow and widely used for medical image segmentation.



UNet Results

To recognize eye nerve, MD Anderson cancer center obtained 66% accuracy using around 10 thousands of CT images, two Nvidia Geforce GTX 1080 Ti graphic cards

Due to resources limit including hardware, dataset and timeframe, no positive results are obtained from us.

	Dataset	GPU	Accuracy
MD Anderson	~10,000	Nvidia GTX 1080Ti * 2 (Pascal)	~66%
2-man crew	~60	Nvidia Tesla K80 (Kepler)	~0%

A black square with a single white dot in the upper right quadrant, representing the ground truth for eye nerve recognition.

Ground truth

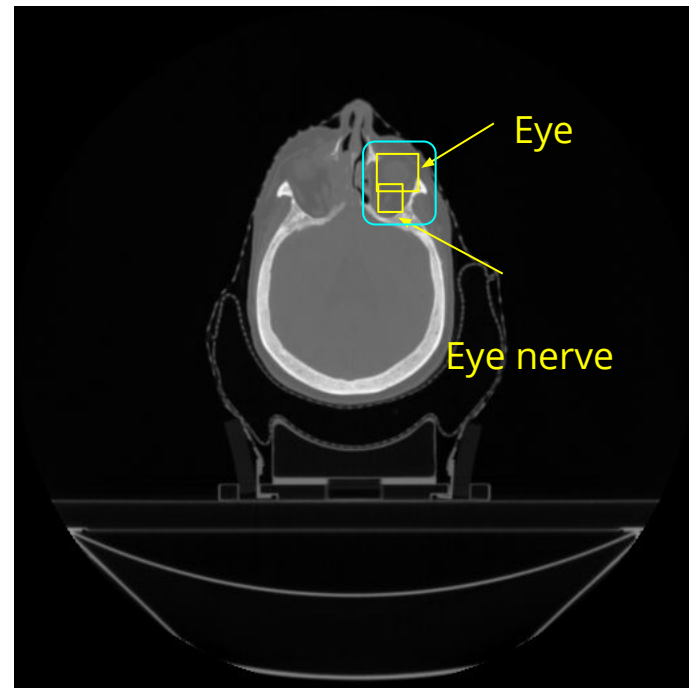
A black square with no visible features, representing the prediction for eye nerve recognition.

Prediction

YOLO and UNet

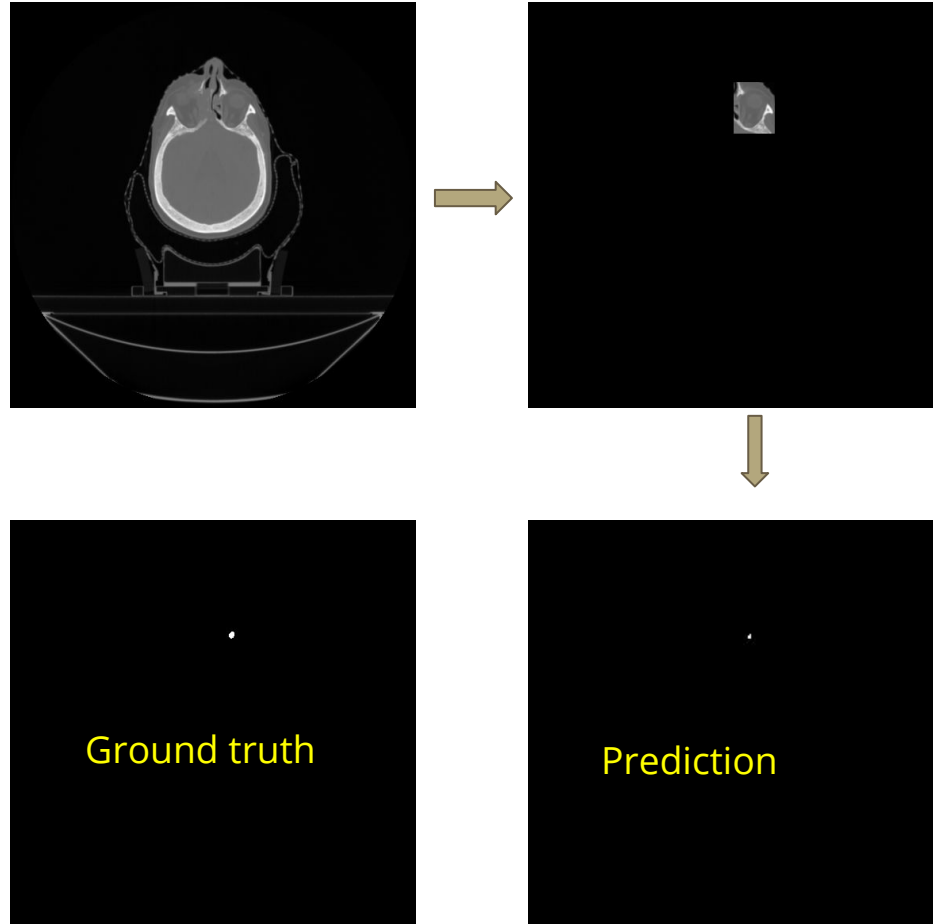
Due lacking of dataset, the model can't be trained with enough input. The aimed object can't be distinguished from background.

Thus, YOLO is used to detect a much smaller area, which is not as hard as finding nerves but actually contains nerves.



Results

1. YOLO detects left eye (86%).
2. Keep the image around the bounding box of left eye, the rest is overwritten with black.
3. Preprocess all the images for training and testing.
4. Train the UNet model as usual.



Conclusion

1. Without enough dataset, the performance of sole UNet is unacceptable.
2. Combination of YOLO+UNet works much better than sole UNet.
3. Next step: Preprocess more CT images up to 2,000 for model training.

*Thank you
and any
question
please?*