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.

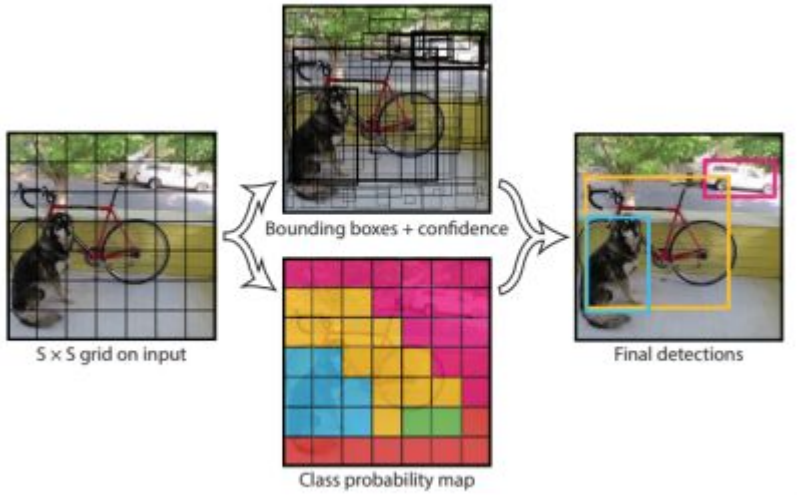
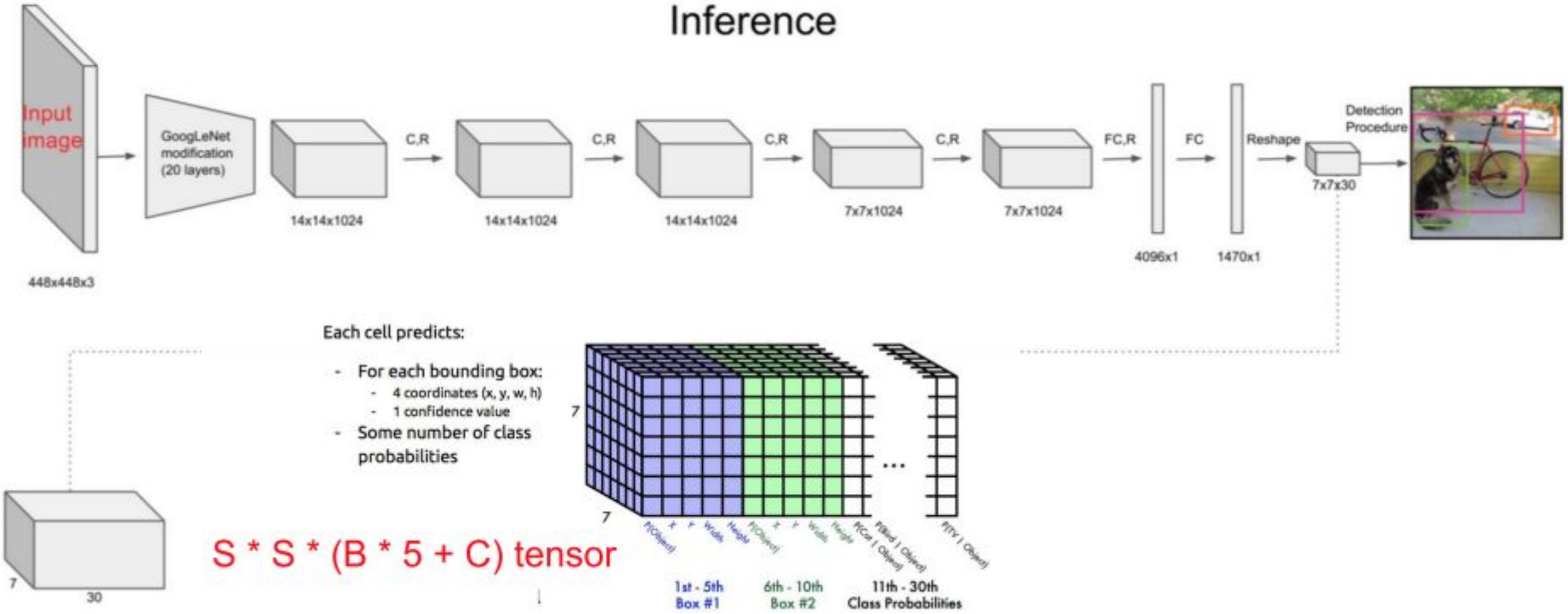


Figure 2: The Model. Our system models detection as a regression problem. It divides the image into an $S \times S$ grid and for each grid cell predicts B bounding boxes, confidence for those boxes, and C class probabilities. These predictions are encoded as an $S \times S \times (B * 5 + C)$ tensor.

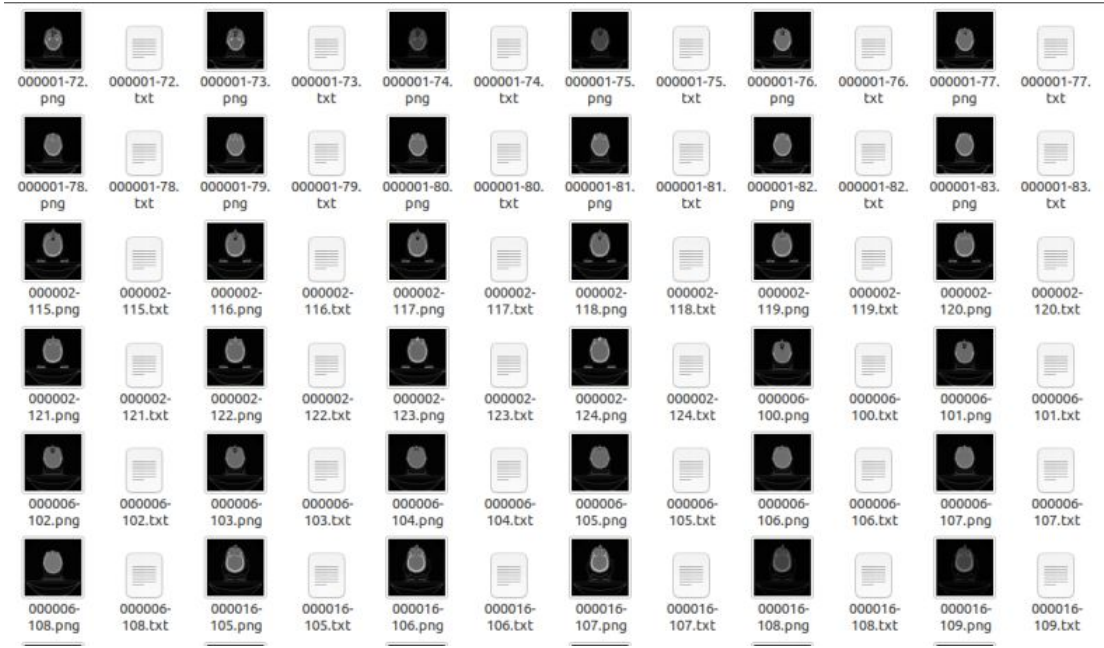
Redmon, J., Divvala, S., Girshick, R., & Farhadi, A. (2016). You only look once: Unified, real-time object detection. In *Proceedings of the IEEE conference on computer vision and pattern recognition* (pp. 779-788). https://arxiv.org/pdf/1506.02640.pdf
Lecture 8: Spatial Localization and Detection: http://cs231n.stanford.edu/slides/2016/winter1516_lecture8.pdf

You Only Look Once (YOLO)



Data Preprocessing

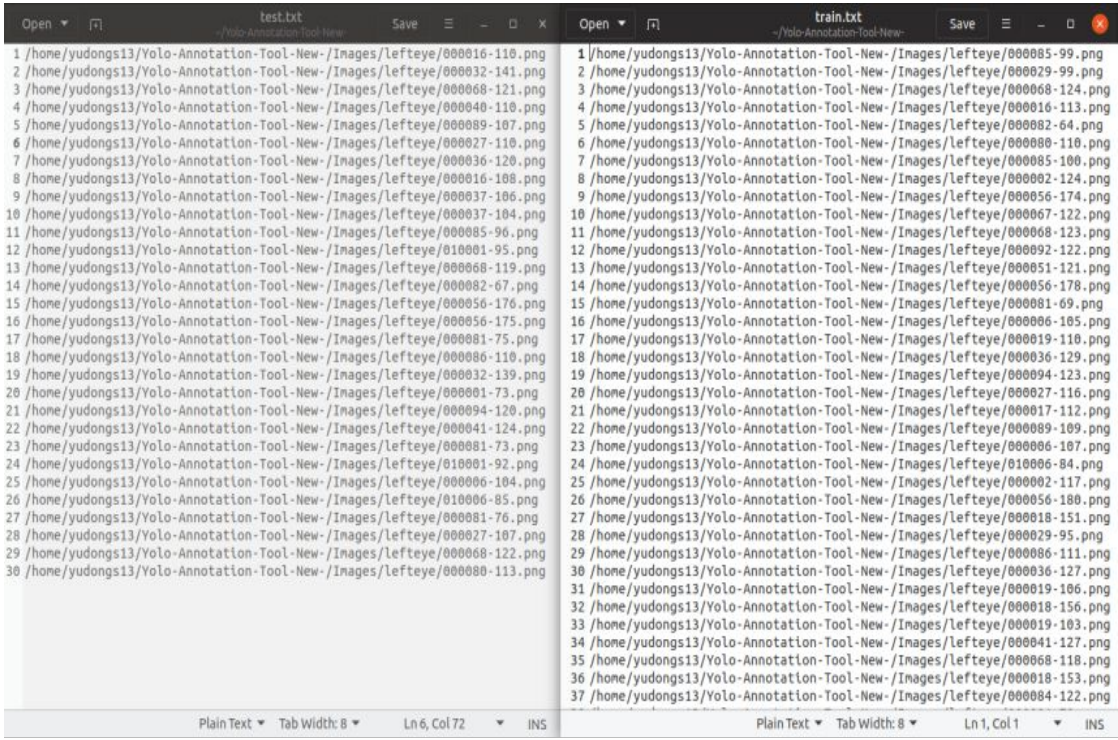
- Main thing for creating text file of images:
  - Create .txt-file for each .jpg-image-file - in the same directory and with the same name, but with .txt-extension
  - For img1.jpg you should create img1.txt containing: <object-class> <x> <y> <width> <height>. <x> <y> - are center of rectangle (are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



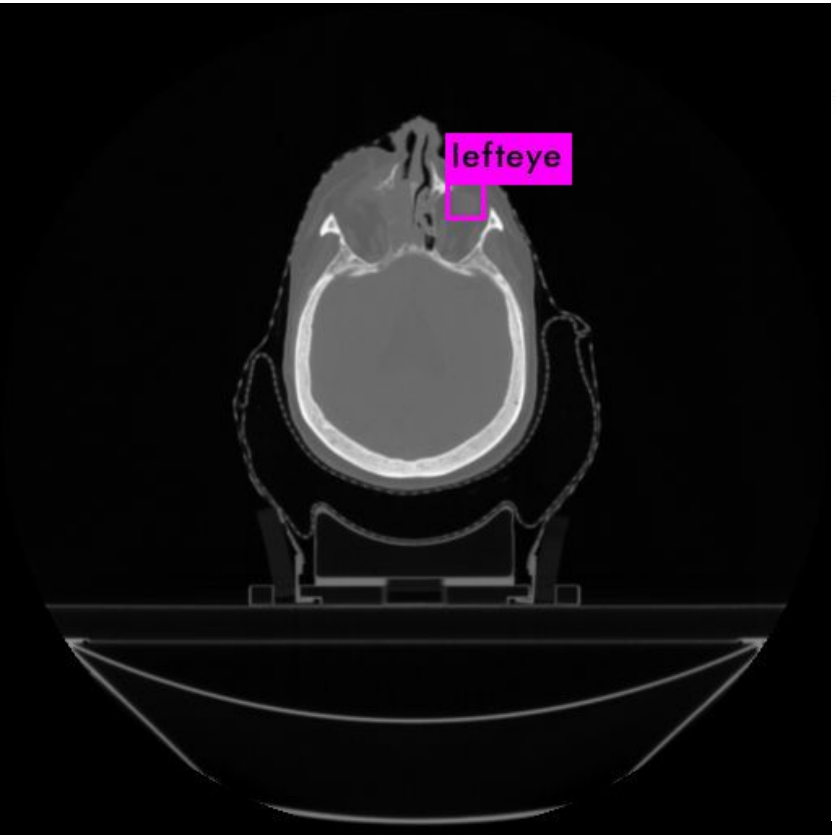

```
1 0.716797 0.395833 0.216406 0.147222
```

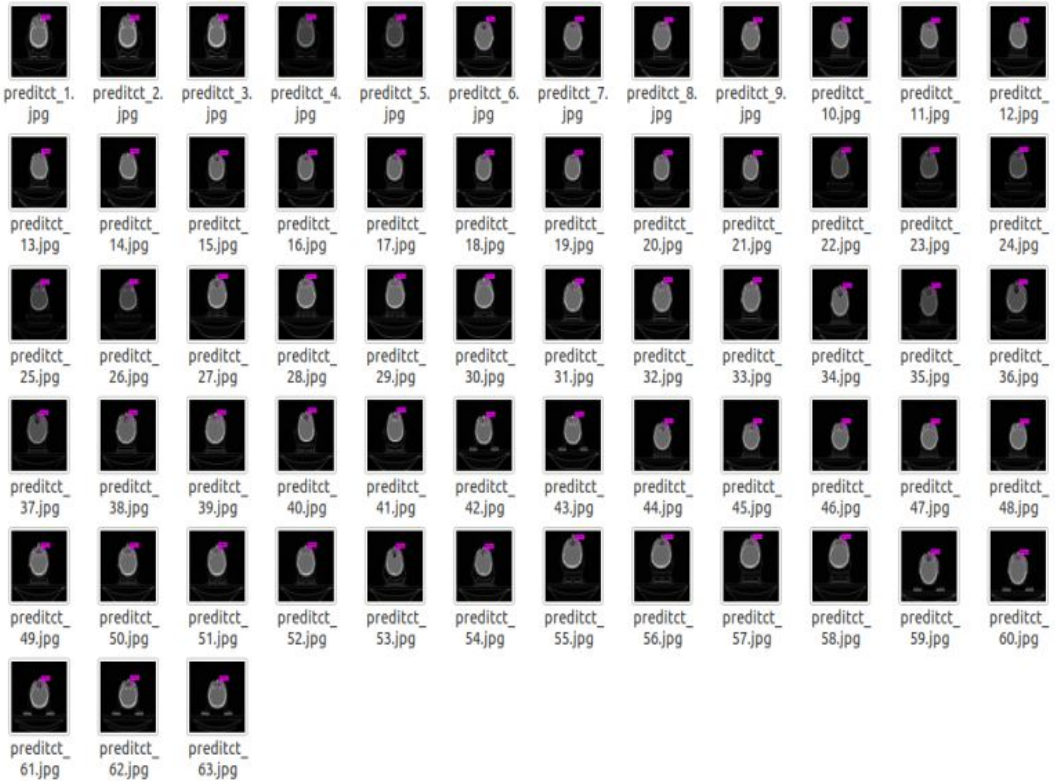
Yolo Training

- 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.



YOLO test results

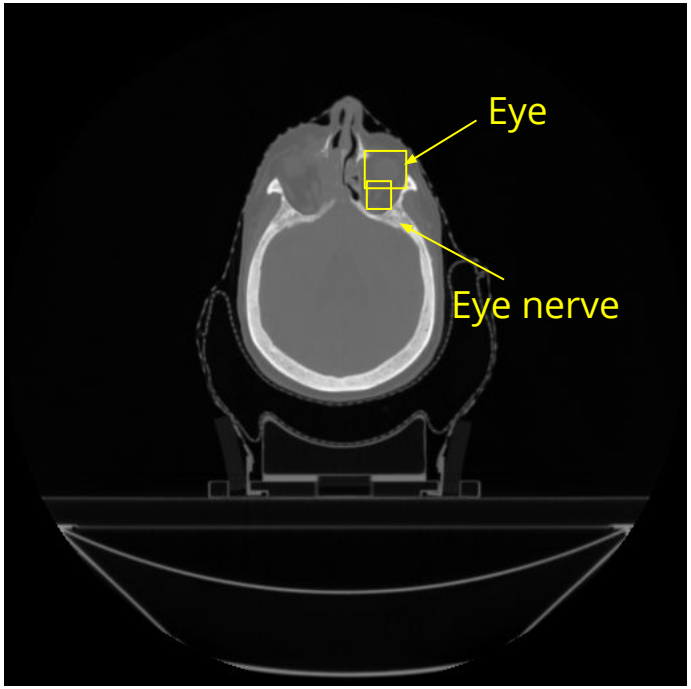


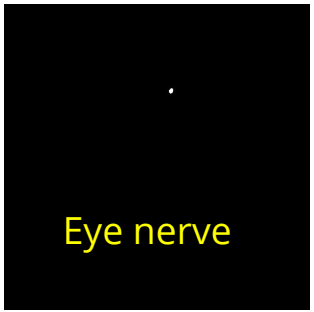


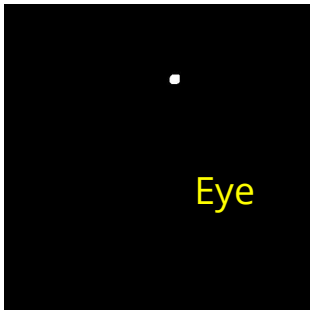
UNet

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.





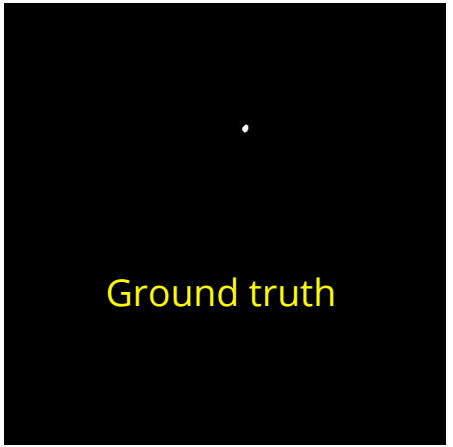


UNet Results

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

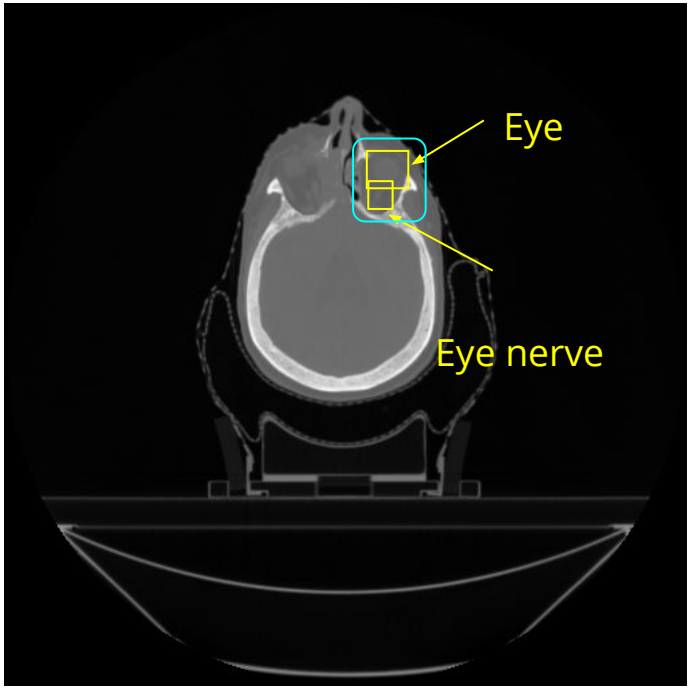
Due to resources limit including hardware, dataset and timeframe, no positive results are obtained from us.

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



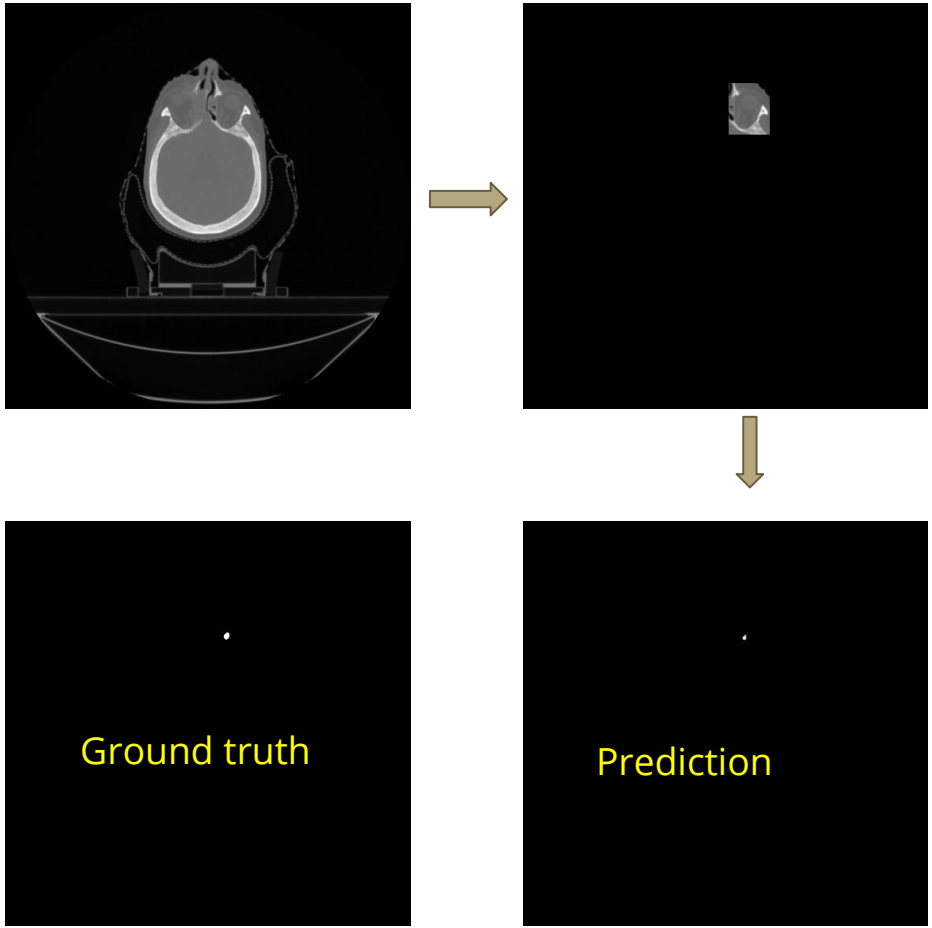
YOLO and UNet

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?