Appendix 1.

**Mask R-CNN model architecture.** Mask R-CNN is an instance segmentation algorithm, which is often used for various tasks such as object detection and image segmentation. The Mask R-CNN architecture combined the Faster- R-CNN algorithm and Fully Convolutional Network (FCN). FCN was used to generate the corresponding MASK branch. Mask R-CNN consists of two stages. The first stage was to generate the regions of proposals which mapped from the input images to the feature maps. The second stage was to classify the proposals and predicted the bounding boxes and the masks.

**Detectron2.** Detectron2 is a library provided by Facebook AI Research. This library provides state-of-the-art detection and segmentation algorithms.  
[https://github.com/facebookresearch/detectron2](https://github.com/facebookresearch/detectron2)

**Input.** Cut out the slice (480x480 pixels) from the original OCT image (1008x596 pixels). This slice was used for training and predicting of the Mask R-CNN.

**Backbone.** The purposed method used ResNet-50 as the backbone from the model zoo of detectron2.  

**Training detail.** The purposed method trained the Mask R-CNN from detectron2. Training parameters learning rate = 0.003 (cfg.SOLVER.BASE_LR = 0.003) and iterations = 85000 (cfg.SOLVER.MAX_ITER = 85000).