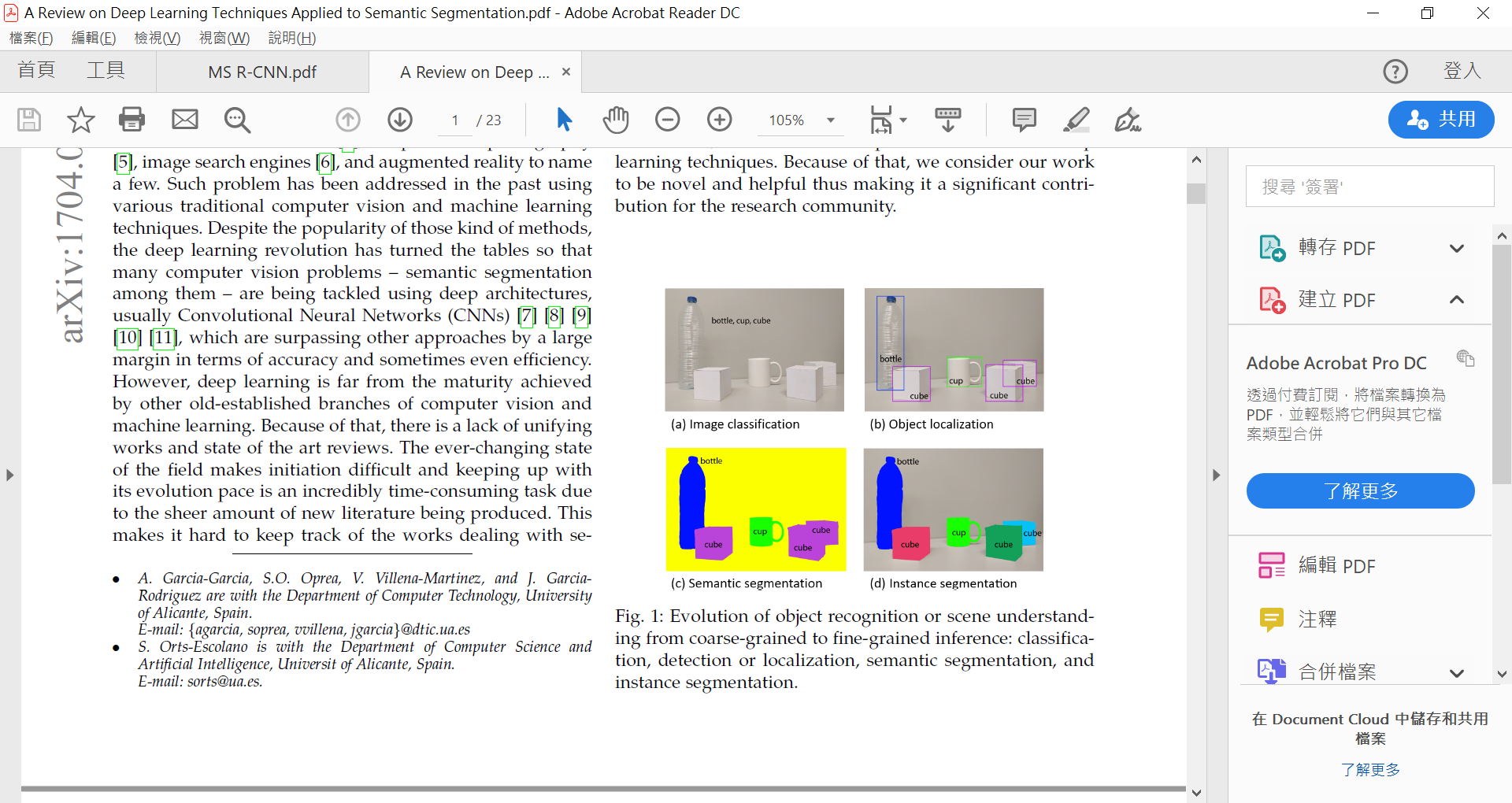
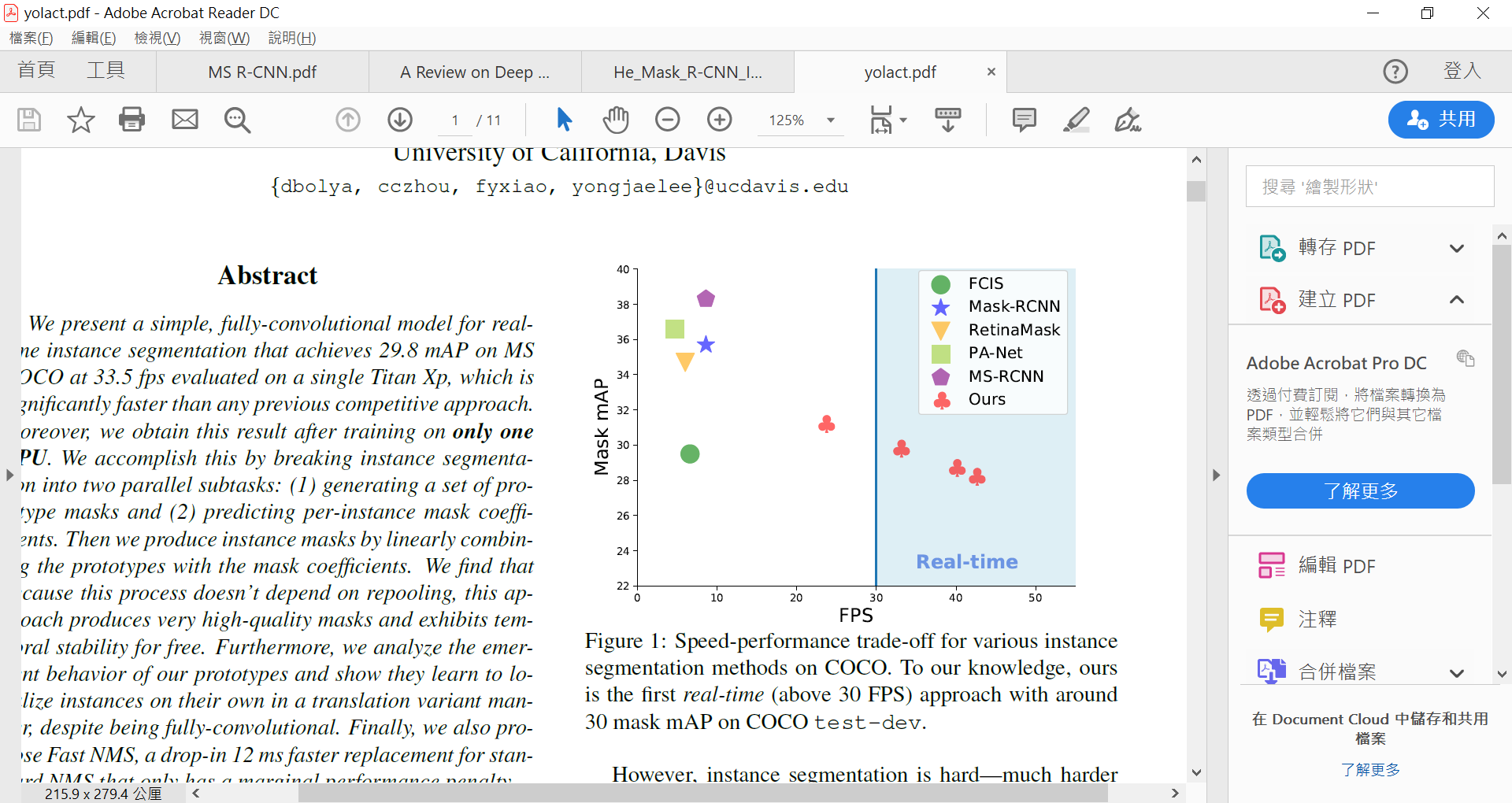
Now a day, deep learning has grown very fast in computer vision. We can roughly separate the problem in 4 stage. Image classification can only detect the class in the image. Object localization not only detect the class but also can locate each object in a bounding box. Semantic segmentation can separate the object by a mask, but it can’t distinguish different object in the same class. So the final goal is instance segmentation it can respectively get all the object in the image.



Yolo, R-CNN etc., these models are solving object localization problem.

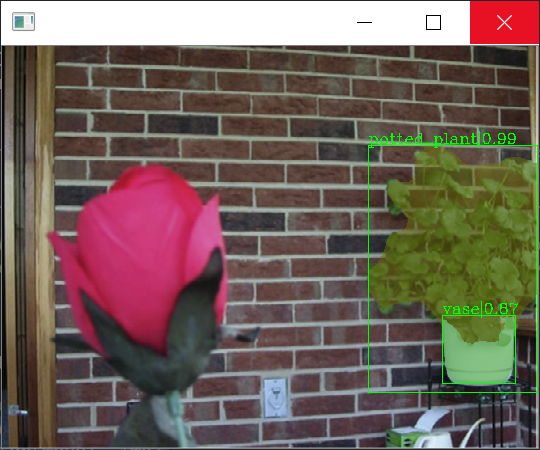
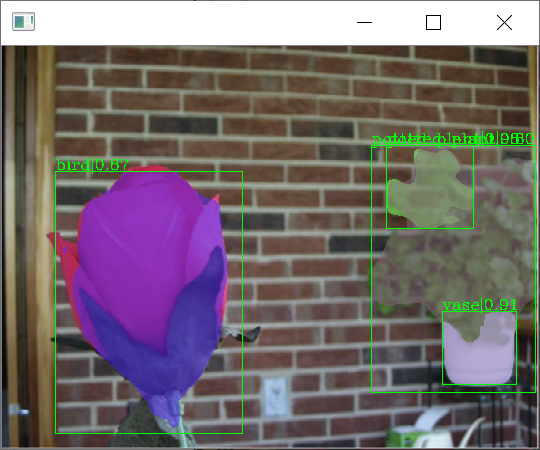
For solving instance segmentation problem, there are Mask-RCNN, MS-RCNN, Yolact etc. Mask-RCNN and MS-RCNN is derived from RCNN, and Yolact is derived from Yolo.



From the results on the above image we can see MS-RCNN has the greatest performance and we doesn’t need to deal the problem in real-time, so we choose MS-RCNN as our method.

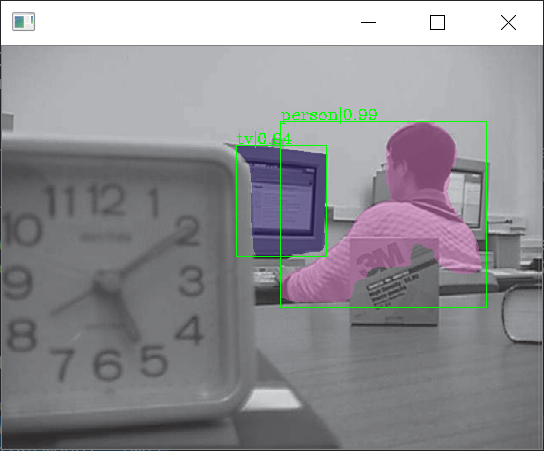
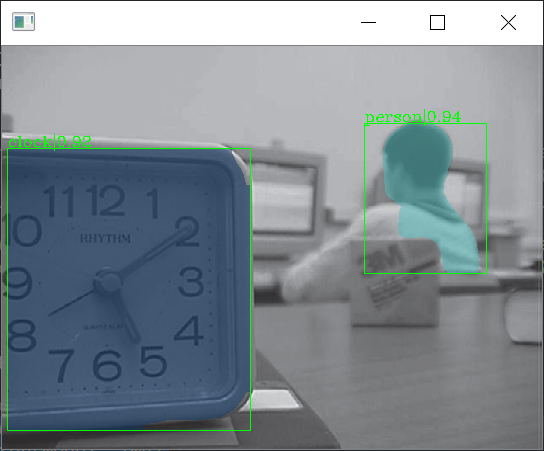
MS-RCNN is a new deep network derived from Mask-RCNN. MS-RCNN is a network which published in March 2019.

The result of Instance segmentation in MS-RCNN.



Problem 1. There has two images(near and far), so sometimes the same object will have two mask. How can I choose which mask to use?

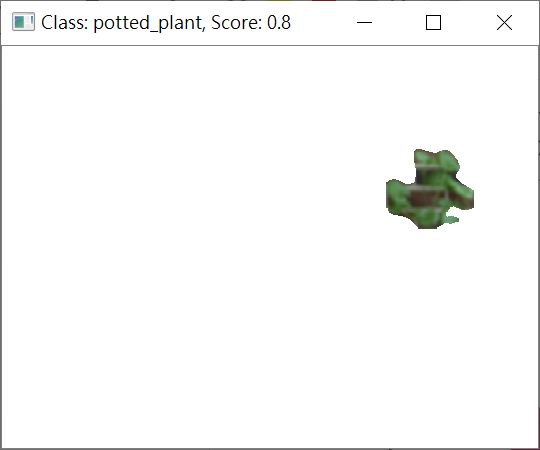
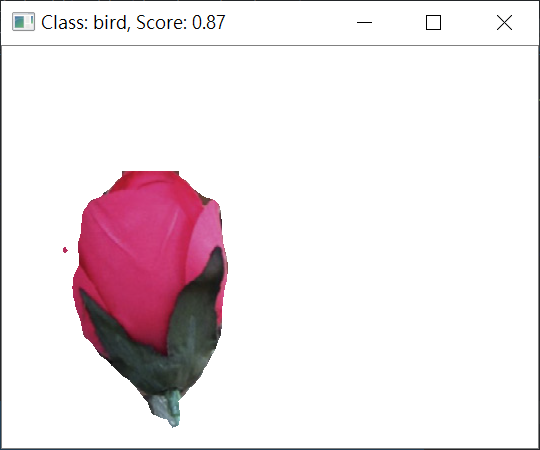
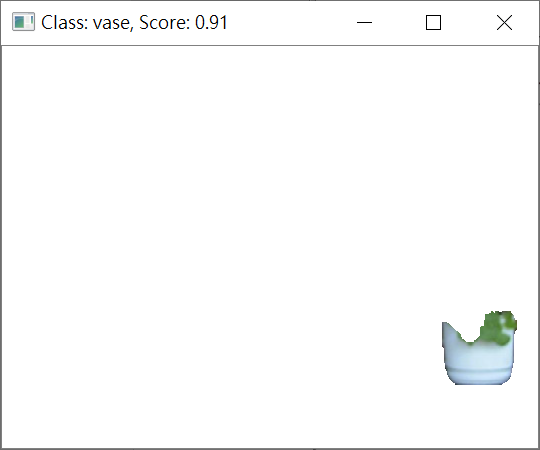
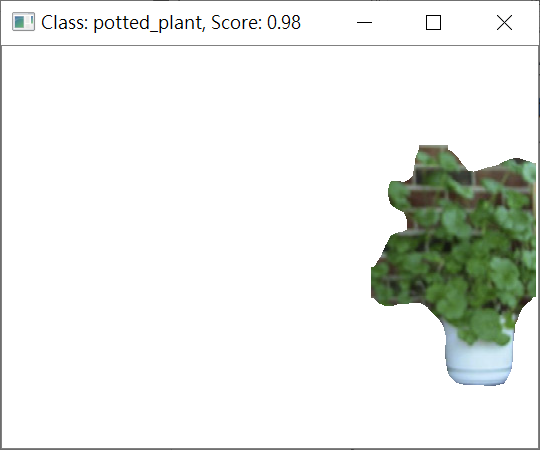
Near Far



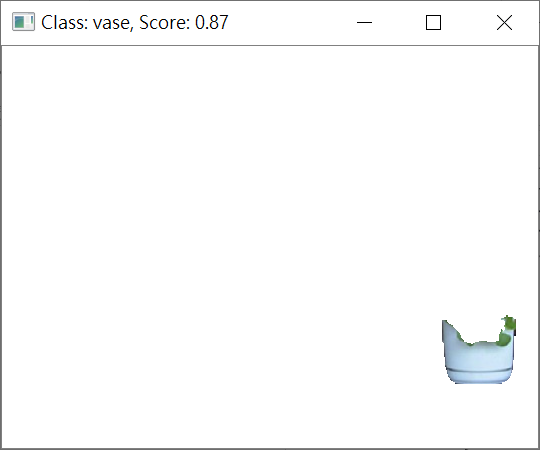
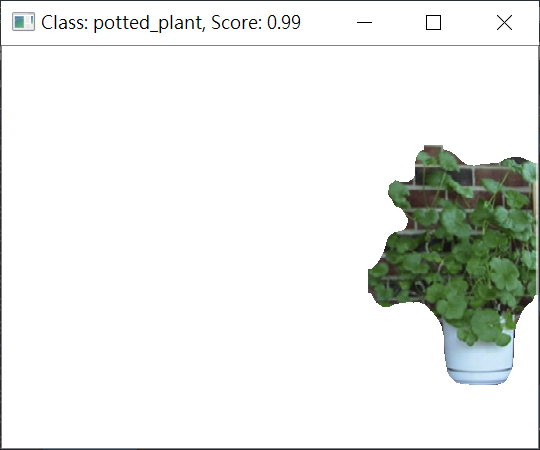
Near Far

The result down below is after deleting the other part in the image, only remain one object at a time.

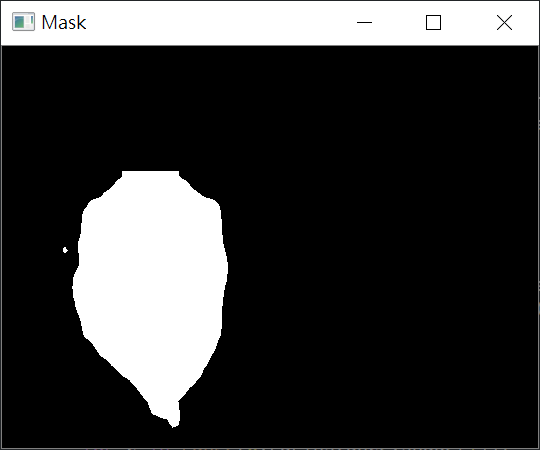
Near focus Instance segmentation:

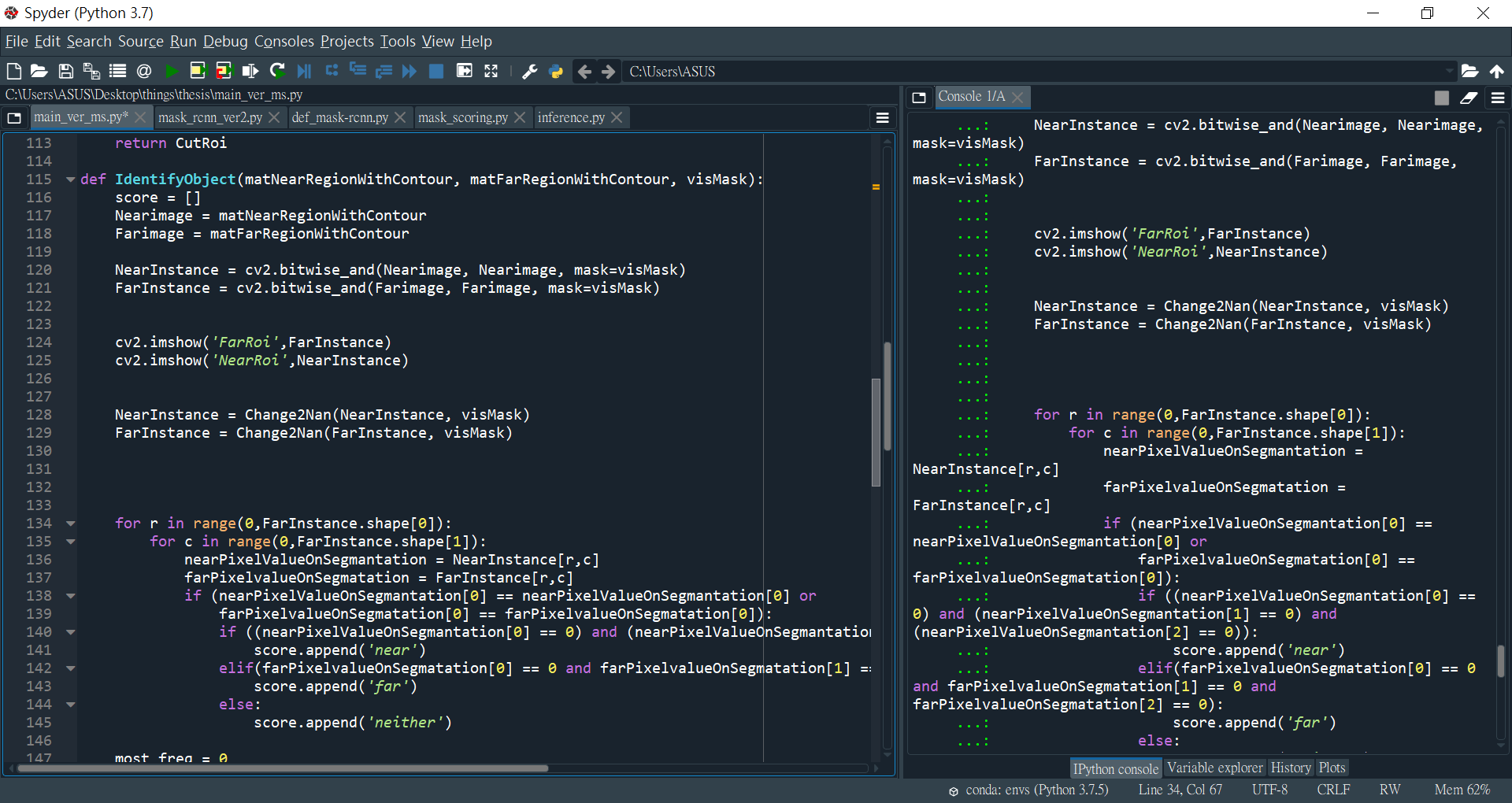
Far focus Instance segmentation:



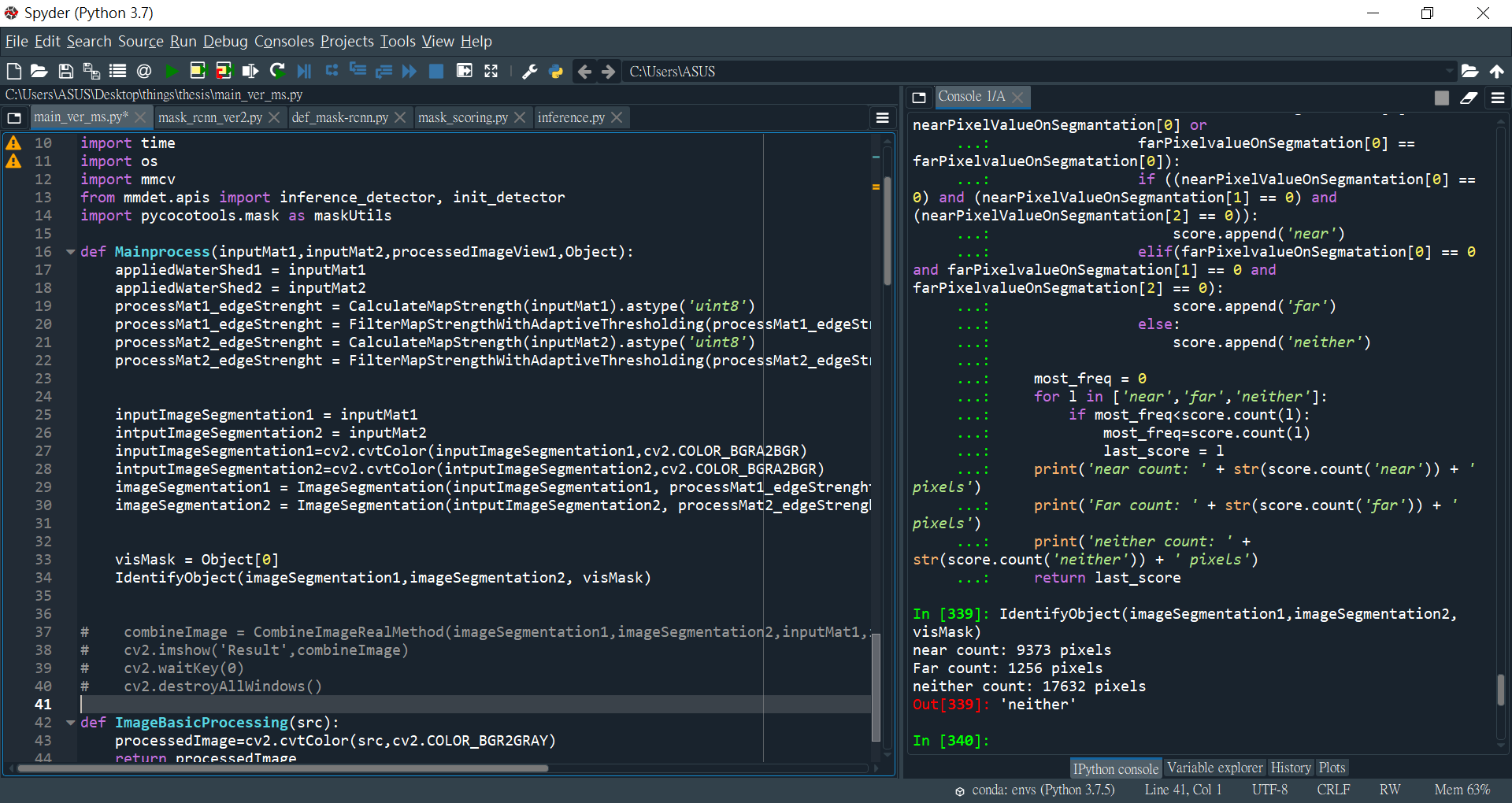
From the result of the MS-RCNN, we can get a Mask which can represent the exact location of the object.



Next step, we will calculate weather the object is in near focus, far focus or neither. Because of the mask, we can only calculate the object which is the white part of the mask. And then let each pixel has its own score.



The method I used here to detect focus of each pixel is by your code.



Problem 2. By this way, I get the result is on the left. This is calculating the flower which is the near focus object. But the score says the flower is neither near nor far focus, which isn’t what we supposed. Do we need to change another method to detect the focus? Or I done some mistakes in your code?