1. What do REGION PROPOSALS entail?

Ans. Region proposal algorithms seek to replace the traditional image pyramid and sliding window approach. These algorithms: Accept an input image. Over-segment it by applying a superpixel clustering algorithm.

2. What do you mean by NON-MAXIMUM SUPPRESSION? (NMS)

Ans. Non Maximum Suppression (NMS) is a technique used in numerous computer vision tasks. It is a class of algorithms to select one entity (e.g., bounding boxes) out of many overlapping entities. We can choose the selection criteria to arrive at the desired results.

3. What exactly is mAP?

Ans. The feature maps of a CNN capture the result of applying the filters to an input image. I.e at each layer, the feature map is the output of that layer. The reason for visualising a feature map for a specific input image is to try to gain some understanding of what features our CNN detects.

4. What is a frames per second (FPS)?

Ans. When CNN series develops to Faster Region with CNN (R-CNN), the Mean Average Precision (mAP) has reached 76.4, whereas, the Frame Per Second (FPS) of Faster R-CNN remains 5 to 18 which is far slower than the real-time effect.

5. What is an IOU (INTERSECTION OVER UNION)?

Ans. Intersection over union (IoU) is known to be a good metric for measuring overlap between two bounding boxes or masks. Image is created by Oleksii Sheremet with Microsoft Visio. If the prediction is completely correct, IoU = 1. The lower the IoU, the worse the prediction result.

6. Describe the PRECISION-RECALL CURVE (PR CURVE)

Ans. A precision-recall curve is a plot of the precision (y-axis) and the recall (x-axis) for different thresholds, much like the ROC curve. A no-skill classifier is one that cannot discriminate between the classes and would predict a random class or a constant class in all cases.

7. What is the term "selective search"?

Ans. Selective Search is a region proposal algorithm for object detection tasks. It starts by over-segmenting the image based on intensity of the pixels using a graph-based segmentation method by Felzenszwalb and Huttenlocher.

8. Describe the R-CNN model's four components.

Ans. R-CNN is a two-stage detection algorithm. The first stage identifies a subset of regions in an image that might contain an object. The second stage classifies the object in each region. Applications for R-CNN object detectors include: Autonomous driving.

9. What exactly is the Localization Module?

Ans. What is CNN localization?

Image localization is a spin-off of regular CNN vision algorithms. These algorithms predict classes with discrete numbers. In object localization, the algorithm predicts a set of 4 continuous numbers, namely, x coordinate, y coordinate, height, and width, to draw a bounding box around an object of interest.

10. What are the R-CNN DISADVANTAGES?

Ans. The R-CNN model has some drawbacks: It is a multi-stage model, where each stage is an independent component. Thus, it cannot be trained end-to-end. It caches the extracted features from the pre-trained CNN on the disk to later train the SVMs.