**Assignment – 11**

1. What do REGION PROPOSALS entail?

Ans: Region proposals involve identifying potential objects or regions of interest within an image.

It's a crucial step in object detection tasks as it helps narrow down the search space for the presence of objects.

Various methods like selective search, EdgeBoxes, and region proposal networks (RPNs) can be used to generate region proposals.

1. What do you mean by NON-MAXIMUM SUPPRESSION?

Ans: NMS is a post-processing technique used in object detection to filter out redundant bounding boxes.

After obtaining multiple bounding boxes that potentially contain the same object, NMS removes overlapping boxes by retaining only the one with the highest confidence score.

It helps in reducing duplicate detections and improving the precision of object localization.

3. What exactly is mAP?

Ans: mAP is a metric commonly used to evaluate the performance of object detection algorithms.

It calculates the average precision across multiple classes or categories of objects in a dataset.

mAP considers both precision and recall, providing a comprehensive measure of the model's ability to detect objects accurately.

4. What is a frames per second (FPS)?

Ans: FPS is a measure of the speed or performance of a computer vision system, particularly in video processing tasks.

It represents the number of frames (images) processed or displayed by the system per second.

Higher FPS values indicate faster processing and smoother video playback.

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

Ans: IoU is a metric used to evaluate the accuracy of object localization in tasks like object detection and instance segmentation.

It measures the overlap between the predicted bounding box and the ground truth bounding box by calculating the ratio of the intersection area to the union area.

IoU values range from 0 to 1, with higher values indicating better alignment between predicted and ground truth boxes.

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

Ans: The precision-recall curve is a graphical representation that illustrates the trade-off between precision and recall at different decision thresholds.

Precision measures the proportion of true positive predictions among all positive predictions, while recall measures the proportion of true positive predictions among all actual positives.

The PR curve helps in evaluating the performance of a model across different thresholds and selecting an appropriate operating point based on the task requirements.

7. What is the term "selective search"?

Ans: Selective search is a region proposal algorithm used to identify potential object locations within an image.

It employs a hierarchical segmentation technique to group pixels into regions based on color, texture, and other low-level features.

Selective search generates a diverse set of region proposals that serve as input to object detection algorithms.

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

Ans: R-CNN (Region-based Convolutional Neural Network) comprises four main components:

Region Proposal: Identifying potential object regions using selective search or similar methods.

CNN Feature Extractor: Pretrained convolutional neural network (e.g., VGG, ResNet) used to extract features from proposed regions.

Region-wise Classifier: A classifier (often an SVM or softmax layer) trained to classify objects within proposed regions.

Bounding Box Regressor: A regressor trained to refine the bounding box coordinates of proposed regions.

9. What exactly is the Localization Module?

Ans: The localization module in object detection refers to the part of the network responsible for predicting the coordinates (x, y, width, height) of bounding boxes around detected objects.

It typically consists of regression layers trained to predict bounding box offsets or deltas from region proposals generated by the region proposal network (RPN) or other methods.

10. What are the R-CNN DISADVANTAGES?

Ans: Computational Complexity: R-CNN is computationally expensive during both training and inference due to the need to extract features for each region proposal separately.

Slow Inference Speed: Processing each region proposal individually leads to slow inference speed, limiting its applicability in real-time applications.

Lack of End-to-End Training: R-CNN involves a multi-stage pipeline with separate training stages for different components, making it less intuitive and harder to optimize compared to end-to-end approaches.

Memory Consumption: Storing and processing a large number of region proposals and extracted features require significant memory resources, especially for high-resolution images.