Pattern Recognition – Final Exam

Model Correction

Master 2 – Artificial Intelligence & Data Science

Part 1: Multiple Choice Questions (7.5 points)

Each correct answer is worth **0.75 points**.

Q	**Answer**	**Justification**
1	B	R-CNN introduced selective search for region proposals.
2	B	Fast R-CNN shares convolutional features using RoI Pooling.
3	C	Faster R-CNN replaces selective search with an RPN.
4	C	All variants rely on a convolutional backbone.
5	B	One-stage detectors predict boxes and classes jointly.
6	B	SSD uses multi-layer feature maps for scale variation.
7	B	YOLO predicts an object if its center lies in the grid cell.
8	B	YOLO struggles with small object detection.
9	B	SSD default boxes grow in size with depth.
10	C	NMS removes redundant overlapping boxes.

Score: 7.5 / 7.5

Part 2: Exercises (12.5 points)

Exercise 1: Evolution of the R-CNN Family

(a) Computational Savings (5 points)

Let the time required for a single CNN forward pass be denoted by T.

- **R-CNN**: Each of the 2000 region proposals is processed independently.

$$\text{Total time} = 2000 \times T$$

- **Fast R-CNN**: The entire image is processed only once.

$$\text{Total time} = 1 \times T$$

Computational savings:

$$2000T - T = 1999T$$

Conclusion: Fast R-CNN saves **1999 CNN forward passes** compared to R-CNN.

(b) Motivation for Replacing Selective Search with RPN (5 points)

Selective search was replaced by the Region Proposal Network (RPN) due to several limitations:

- It is computationally expensive and slow.
- It is based on hand-crafted heuristics.
- It is not learnable and cannot be trained end-to-end.
- It does not benefit from GPU acceleration.

The RPN addresses these issues by:

- Being fully learnable.
- Sharing convolutional features with the detector.
- Running efficiently on GPUs.
- Enabling end-to-end training of the detection pipeline.

Conclusion: RPN improves both **speed** and **proposal quality**.

(c) Feature Sharing via RoI Pooling (5 points)

In Fast R-CNN, the input image is processed once by a convolutional network to produce a feature map. Each region proposal is then projected onto this shared feature map.

RoI Pooling:

- Extracts a fixed-size feature map for each proposal.
- Allows all proposals to share the same convolutional features.

In contrast, R-CNN processes each region independently, preventing feature sharing and resulting in high computational cost.

Key idea: RoI Pooling enables efficient feature reuse across regions.

Exercise 2: Non-Maximum Suppression (NMS)

(a) Purpose of NMS (5 points)

In object detection, multiple bounding boxes often correspond to the same object. Non-Maximum Suppression (NMS) is used to:

- Remove redundant overlapping detections.
- Retain only the most confident bounding box for each object.

This improves detection clarity and reduces false positives.

(b) NMS Algorithm (5 points)

Given bounding boxes $B = \{b_1, \ldots, b_n\}$ with scores $S = \{s_1, \ldots, s_n\}$ and IoU threshold T:

1. Sort all bounding boxes by decreasing confidence score.
2. Select the box with the highest score and add it to the final set.
3. Remove this box from the list.
4. Compute IoU between this box and all remaining boxes.
5. Remove boxes whose IoU is greater than or equal to T.
6. Repeat until no boxes remain.

(c) Effect of the IoU Threshold (5 points)

If two boxes have $\text{IoU} = 0.6$ and the threshold is $T = 0.5$:

- Since $0.6 > 0.5$, the box with the higher confidence score is kept.
- The other box is suppressed.

- Increasing T results in more boxes being kept.
- Decreasing T results in more aggressive suppression.

Final Result

Total Score: 20 / 20