INDIAN SCHOOL AL WADI AL KABIR

Class: X	Department: Computer Science
WORKSHEET	ARTIFICIAL INTELLIGENCE (417) Part B - Unit 5 - COMPUTER VISION

SECTION A: Objective Type Questions (1 Mark Each)

*Q1. Which of the following processes is categorized as an **Image Processing** operation rather than a **Computer Vision** task?*

- A) Detecting facial features to log student attendance in a smart classroom
- B) Identifying traffic signs and pedestrians for an autonomous vehicle navigation system
- C) Adjusting the brightness, contrast, and scaling of a digital photograph
- D) Delineating individual coral reefs from aerial drone imagery using segmentation

*Q2. In an 8-bit **Grayscale image**, what visual representation corresponds to a pixel numerical value of **0**?*

- A) Maximum brightness (Pure White)
- B) Total absence of light and color (Pure Black)
- C) Mid-tone gray balance
- D) Full saturation of the primary Green channel

Q3. An AI-powered camera in an automated retail warehouse scans a shelf, identifies three distinct cereal boxes, and draws individual rectangular bounding boxes around each box while assigning category labels. Which Computer Vision task is being executed?

- A) Image Classification
- B) Classification + Localization
- C) Object Detection
- D) Instance Segmentation

*Q4. Why are **corners** considered superior image features compared to flat surfaces or straight edges in computer vision applications?*

- A) Corners require fewer computational bits to store in memory.
- B) Moving a corner patch in any direction produces a distinct, unique change in visual appearance.
- C) Corners contain higher RGB channel intensity values than flat regions.
- D) Corners remain completely unaffected by image convolution operators.

Q5. The mathematical operation in image processing that performs an element-wise multiplication between an input image array and a small sliding filter matrix called a kernel, followed by summing the results, is known as _____.

*Q6. What is the primary function of the **Rectified Linear Unit (ReLU)** layer in a Convolutional Neural Network (CNN)?*

- A) It flattens a multi-dimensional feature matrix into a one-dimensional vector.
- B) It replaces all negative values in the feature map with zero, introducing non-linearity.
- C) It downsamples the image dimensions by calculating spatial pixel averages.
- D) It aligns kernel centers with image edges using zero-padding.

Q7. users to point their camera at foreign text on street signs and translates the words into their native language instantly by overlaying translated text on the camera view. Name the two core technologies combined in this computer vision application.

*Q8. In a Convolutional Neural Network, a **Max Pooling** operation with a 2 \times 2 filter is applied to the following pixel block: \begin{pmatrix} 14 & 22 \\ 9 & 18 \end{pmatrix} Which value will be selected and returned by the Max Pooling operation for this block?*

- A) 14
- B) 15.75
- C) 18
- D) 22

Q9. Assertion-Reasoning Question

- **Assertion (A):** Object Detection is computationally more complex than standard Image Classification.
- **Reasoning (R):** Object Detection requires the model to both identify object categories and accurately localize their spatial extent using bounding boxes across single or multiple instances.
- **Options:**
    - A) Both A and R are true, and R is the correct explanation of A.
    - B) Both A and R are true, but R is not the correct explanation of A.
    - C) A is true, but R is false.
    - D) A is false, but R is true.

*Q10. How many total numerical values are required to store an **RGB color image** with a resolution of 1000 \times 800 pixels in computer memory?*

- A) 800,000 values
- B) 1,600,000 values
- C) 2,400,000 values
- D) 4,800,000 values

SECTION B: Short Answer Type Questions (2 Marks Each)

Q11. Differentiate between **Computer Vision** and **Image Processing** based on their core objectives and provide one standard example for each.

Q12. Explain how an **RGB image** is represented and stored in computer memory, and contrast it with a **Grayscale image** in terms of channels and pixel storage.

Q13.Differentiate between **Object Detection** and **Instance Segmentation** in Computer Vision, highlighting how each task delineates detected objects in an image.

Q14. Agricultural drones equipped with cameras are increasingly deployed for crop monitoring. Identify two specific parameters that computer vision algorithms analyze from aerial crop photographs, and explain how this automation improves farming efficiency compared to traditional manual inspection.

Q15.Define the term Resolution.

Detailed Answer Key & Official Marking Scheme

Section A: Objective Type Answers & Explanations

1. **Correct Answer: C (Adjusting the brightness, contrast, and scaling of a digital photograph)**
    - *Explanation:* Image processing focuses on manipulating or enhancing raw images (subset), whereas Computer Vision extracts high-level semantic meaning to make automated predictions (superset).
2. **Correct Answer: B (Total absence of light and color / Pure Black)**
    - *Explanation:* In an 8-bit pixel system (0–255), 0 represents total absence of light/color (black), while 255 represents complete presence (pure white).
3. **Correct Answer: C (Object Detection)**
    - *Explanation:* Object detection identifies multiple instances of real-world objects and delineates their positions using rectangular bounding boxes.
4. **Correct Answer: B (Moving a corner patch in any direction produces a distinct, unique change in visual appearance)**
    - *Explanation:* Flat regions look identical everywhere, and edges look identical along their length. Moving a corner in any direction produces a distinct visual shift, making corners optimal anchor features.
5. **Correct Answer: Convolution (or Convolution Operator)**
    - *Explanation:* Convolution is the fundamental operation where a filter matrix slides over image pixel arrays to perform element-wise multiplication and summation.
6. **Correct Answer: B (It replaces all negative values in the feature map with zero, introducing non-linearity)**
    - *Explanation:* The Rectified Linear Unit function applies \text{Output} = \max(0, \text{Input}), eliminating negative activation values and sharpening edge transitions for feature extraction.
7. **Correct Answer: Optical Character Recognition (OCR) and Augmented Reality (AR)**
    - *Explanation:* OCR reads and extracts text characters from visual frame pixels, while AR overlays the translated text seamlessly onto the camera preview display.
8. **Correct Answer: D (22)**

- *Explanation:* Max Pooling evaluates all elements in the 2 \times 2 grid \{14, 22, 9, 18\} and selects the maximum scalar value, which is 22.

9. **Correct Answer: A (Both A and R are true, and R is the correct explanation of A)**
    - *Explanation:* Object Detection goes beyond single-label classification by identifying multiple object boundaries and locations, adding spatial search complexity.
10. **Correct Answer: C (2,400,000 values)**
    - *Explanation:* Total pixels = 1000 \times 800 = 800,000. Since an RGB image contains 3 color channels (Red, Green, Blue), total values = 800,000 \times 3 = 2,400,000.

Section B: Model Answers & Evaluation Guidelines

Q11. Computer Vision vs. Image Processing

- **Model Answer:**
    - **Computer Vision:** Focuses on extracting meaningful information and high-level understanding from visual data to enable automated decision-making and predictions. *Example:* Facial recognition attendance or autonomous vehicle obstacle detection.
    - **Image Processing:** Focuses on manipulating, transforming, or enhancing raw input images to improve visual quality or prepare them for further analysis. *Example:* Image rescaling, noise reduction, or brightness adjustment.
- **Marking Breakdown:**
    - **1 Mark:** Correctly contrasting the core objectives (understanding/prediction vs. image manipulation/enhancement).
    - **1 Mark:** Providing valid, accurate examples for both domains (1/2 mark each).

Q12. RGB vs. Grayscale Image Representation

- **Model Answer:**
    - **Grayscale Image:** Represented by a single 2D array (1 plane/channel). Each pixel is stored as a single 8-bit byte value ranging from 0 (black) to 255 (white). Image size equals \text{Height} \times \text{Width}.
    - **RGB Image:** Represented by three distinct 2D color planes/channels (Red, Green, Blue). Each pixel contains a set of three independent numerical values (0–255), representing the intensity of each primary color. The channels combine to render full color.
- **Marking Breakdown:**
    - **1 Mark:** Explaining Grayscale representation (single channel, 0–255 pixel value).
    - **1 Mark:** Explaining RGB representation (3 channels—Red, Green, Blue, three values per pixel).

Q13. Object Detection vs. Instance Segmentation

- **Model Answer:**
    - **Object Detection:** Identifies and categorizes real-world objects in an image and locates them by drawing rectangular **bounding boxes** around each detected object.
    - **Instance Segmentation:** Goes beyond bounding boxes by detecting objects, classifying them, and labeling **every individual pixel** belonging to each object instance, creating precise outline boundaries (segments).

- **Marking Breakdown:**
    - **1 Mark:** Defining Object Detection with reference to bounding boxes.
    - **1 Mark:** Defining Instance Segmentation with reference to pixel-by-pixel outline labeling.

Q14. Application Case Study: Agricultural Drones

- **Model Answer:**
    - **Analyzed Parameters (Any two):** Leaf pixel color variations (detecting nutrient deficiencies or chlorophyll loss), crop canopy coverage, pest infestation spots, and soil moisture patterns.
    - **Efficiency Improvement:** Automated drone vision rapidly scans vast farmlands in minutes with consistent digital precision, eliminating time-consuming, labor-intensive, and error-prone manual human field walks. It enables early detection of crop stress before visible damage spreads across fields.
- **Marking Breakdown:**
    - **1 Mark:** Identifying two valid crop/plant parameters analyzed by computer vision (1/2 mark each).
    - **1 Mark:** Explaining the gain in efficiency over traditional manual inspection methods.

Q15. Resolution Definition:

Resolution refers to the total number of pixels contained within a digital image. It measures the level of detail and visual clarity—the higher the pixel count, the more closely the digital representation matches the original image