



JK LakshmiPat University, Jaipur
Institute of Engineering and Technology
End Term Examination, April, 2024
B.C.A. Semester VI, 2023-24

Roll No. : _____

CS1234: APPLIED COMPUTER VISION

Duration: 2.5 hours

MM: 30

Instructions: There are total 7 questions in 3 parts and all the questions are compulsory. Please make appropriate assumptions where required. **Scientific Calculators are allowed.**

PART I: Attempt all questions.

1. Short-answer questions: (2x4 - COs 1,2,3,4)

- (a) What do you mean by rotation and scale invariance? Briefly explain how we can achieve it?
- (b) Median Filtering operation is computationally better than mean filtering operation (True/False). State with reason.
- (c) An output layer of an MLP generates the following scores for a test image in a dataset with 3 classes. What is the output relative probabilities for the respective classes, if softmax is applied to $\hat{y}$.

$$\hat{y} = \begin{bmatrix} 5.05 \\ 0.87 \\ -2.26 \end{bmatrix}$$

- (d) What is Histogram Equalization. How is it different from Contrast Stretching?

PART II: Attempt all questions.

- 2. What are activation functions? Mention their advantages. Which activation functions are prone to vanishing gradient problem and why? (3 - COs 1,3)
- 3. Briefly explain the generic steps for object detection that need to be carried out for images. Which evaluation metric is used for determining whether a detection is correct or not? (3 - COs 2,4)
- 4. A CNN model generates the following probability values for the binary classification task, where $\mathbf{Y}$ are the ground truth labels denoting whether an object of interest is present in an image or not. (3 - COs 2,3)

Y	1	1	1	0	0	1	0	1	0	0
$h_{\theta}(x)$	0.14	0.83	0.42	0.71	0.31	0.26	0.75	0.92	0.55	0.64

Find the precision, recall, and F1 scores for the above model assuming threshold probability value of 0.7.

- 5. Define optical flow and mention some applications where optical flow is used. Why is brightness constancy assumption required for determining the optical flow? (3 - COs 1,3)

PART III: Attempt all questions.

6. Given the following image patch with graylevel values as shown below: (4 - COs 1,2,3)

$$I = \begin{bmatrix} 0 & 0 & 0 & 100 & 100 & 100 \\ 0 & 0 & 0 & 100 & 100 & 100 \\ 0 & 0 & 0 & 100 & 100 & 100 \\ 100 & 100 & 100 & 0 & 0 & 0 \\ 100 & 100 & 100 & 0 & 0 & 0 \\ 100 & 100 & 100 & 0 & 0 & 0 \end{bmatrix}$$

- (a) Apply filtering on I for each of the following masks. Do not use padding.

$$A = \begin{bmatrix} 1 & 0 & -1 \\ 2 & 0 & -2 \\ 1 & 0 & -1 \end{bmatrix} \quad B = \begin{bmatrix} 1 & 2 & 1 \\ 0 & 0 & 0 \\ 1 & -2 & -1 \end{bmatrix}$$

- (b) Which mask (out of A and B) is used to detect vertical edges and which one is used to detect horizontal edges? Mention the name of this filtering operation when these specific mask values are used.
- (c) Let a filter of size $F \times F$ be convolved with image patch I of size $N \times N$ ($N > F$), with a stride of S , then derive the formula for obtaining the output size of the patch (ignore padding).
7. Design a CNN architecture that takes an input image size of 128×128 and classifies the image into 5 object categories. Consider the following constraints for the layers and learnable parameters. Make appropriate assumptions if needed: (6 - COs 2,3,4)
- Create the layers such that the total number of learnable parameters are less than 2,000,000.
 - Take convolution kernels of sizes 3×3 , with padding, pooling layers of 2×2 with stride 2, appropriate activation layers, and FC layers at the end.

Mention the output activation map dimensions after each layer and the number of learnable parameters for each.