



JK Lakshmipat University, Jaipur
Institute of Engineering and Technology
End Term Examination, November, 2024
B.C.A. Semester V, 2024-25

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 they can be achieved?
 - (b) How is mean filtering different from median filtering? Where are they used?
 - (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 $\hat{y}$.

$$\hat{y} = \begin{bmatrix} 10.05 \\ 0.87 \\ -3.26 \end{bmatrix}$$

- (d) How many threshold activation perceptrons will we need in an MLP to model an octagonal decision region (a decision region bounded by a eight-sided polygon) over a 2D input space? Explain.

PART II: Attempt all questions.

2. What are region proposals? How are they useful during object detection? (3 - COs 3,4)
3. For a trained object detection model, how do we find true positive, false positive and false negative predictions? What is mAP metric and what are the steps to compute it? (3 - COs 3,4)
4. A CNN model generates the following probability values for the binary classification task, where 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. What are the two assumptions that are necessary for calculation of 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)

$$I = \begin{bmatrix} 0 & 0 & 0 & 200 & 200 & 200 \\ 0 & 0 & 0 & 200 & 200 & 200 \\ 0 & 0 & 0 & 200 & 200 & 200 \\ 200 & 200 & 200 & 0 & 0 & 0 \\ 200 & 200 & 200 & 0 & 0 & 0 \\ 200 & 200 & 200 & 0 & 0 & 0 \end{bmatrix}$$

- (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. Given a subset of ImageNet dataset with 50 categories (all grayscale images) and with uniform image sizes as 64×64 . A CNN architecture is constructed with the following requirements:
- R1 The network takes the 64×64 sized images as input (grayscale images) and produces a output vector of size 50 (for 50 classes).
- R2 The network contains following layers in order:
- CONV1 (64 3×3 filters, stride = 1, padding = 1)
 - MAXPOOL1 (2×2 filter, stride = 2)
 - CONV2 (128 3×3 filters, stride = 1, padding = 1)
 - MAXPOOL2 (2×2 filter, stride = 2)
 - CONV3 (128 3×3 filters, stride = 1, padding = 1)
 - MAXPOOL3 (2×2 filter, stride = 2)
 - FC1 (128 output size)
 - FC2 (50 output size)

For the above network, show the input sizes of tensors, output activation map dimensions, and number of parameters for each layer (include bias parameters). Also, calculate the total number of parameters for the network. (6 - COs 2,3,4)