



JK Lakshmipat University, Jaipur
Institute of Engineering and Technology
Computer Vision (CS1228)
End Term, B.Tech., VII Semester

Date: December 4, 2023

Duration: 3 hours

MM: 30

Course Instructor: Dr. Arpan Gupta

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

PART I: Attempt all questions.

1. Short-answer questions / MCQs: (2x5 - COs 1,2,3,4)

- (a) Which of the following method(s) is/are used for creation of a scale invariant feature descriptor? Justify your answer(s).
1. Gaussian Filtering
 2. Pyramids
 3. Non-Maximum Suppression
 4. Average Precision
- (b) IoU metric (or its variant) can be used for which of the following task(s). Justify your answer(s).
1. Detection of defects in a machine vision system.
 2. Classification of MNIST dataset images for 10 categories of handwritten digits.
 3. Localizing the actions in untrimmed video sequences.
 4. Compressing the videos into MP4 format.
- (c) Median Filtering operation is computationally better than mean filtering operation (True/False). State with reason.
- (d) 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 $\hat{y}$.

$$\hat{y} = \begin{bmatrix} 2.33 \\ -1.46 \\ 0.56 \end{bmatrix}$$

- (e) 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?

PART II: Attempt all questions.

2. What is Histogram Equalization. How is it different from Contrast Stretching?

(3 - COs 1,2,4)

3. What do you mean by “bias-variance trade-off” with respect to overfitting and underfitting of a CNN model. Mention two ways in which to reduce overfitting the CNN model on image data.

(3 - COs 3,4)

4. What is Orientation Normalization in SIFT descriptor? How does it ensure rotation invariance? Illustrate with the help of a diagram. (3 - COs 1,2,4)
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 1,4)

PART III: Attempt all questions.

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

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

- (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 ($64 \ 3 \times 3$ filters, stride = 1, padding = 1)
- MAXPOOL1 (2×2 filter, stride = 2)
- CONV2 ($128 \ 3 \times 3$ filters, stride = 1, padding = 1)
- MAXPOOL2 (2×2 filter, stride = 2)
- CONV3 ($128 \ 3 \times 3$ 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. (4 - COs 3,4,5)