|  |  |  |  |
| --- | --- | --- | --- |
| Question | Q1 | Q2 | Q3 |
| Marks | 6 | 6 | 8 |
| Total | 20 | | |

**Full Name (as in course registration): Kar Chaudhuri Anirban**

**Date: 07/02/2021**

**Course:** **Vision Systems**

**Question 1.** For each question, select the **single** most appropriate answer, and **provide your justification** (no longer than two sentences).

a). The LBP code studied in our course is robust to

1. Image flip (e.g., flip the image horizontally or vertically)
2. Image rotation (e.g., rotate the image clockwise or counter-clockwise)
3. Global lighting change (e.g., light is turned on or turned off in the classroom)
4. None of above

Answer: C. This method involves thresholding followed by frequency distribution of pixel values and averaging that removes noise problem, depicted by unusually high or low pixel values that have low counts.

b). Suppose you are applying a sliding-window-based pre-trained HoG object classifier on the test image to perform object detection, increasing the stride (the distance the sliding window moves at each step) of the sliding window will tend to

1. Increase object detection accuracy, increase computational cost (run time).
2. Increase object detection accuracy, decrease computational cost (run time).
3. Decrease object detection accuracy, increase computational cost (run time).
4. Decrease object detection accuracy, decrease computational cost (run time).

Answer: D.  Larger stride will result in a smaller number of feature vectors classification per image

& decreases accuracy. Computation cost is lower as there are lesser iterative cycles for computation.

Q**uestion 2.** An image with a resolution of pixels and its intensity values are illustrated below. For the center pixel (highlighted in black color in the image), apply the local binary pattern method to calculate its local binary code. Show your calculations.

|  |  |  |
| --- | --- | --- |
| 5 | 8 | 6 |
| 11 | 7 | 9 |
| 3 | 10 | 2 |

For each pixel, compare the pixel to each of its (eight) neighbours and use label 1 if the center pixel's value is greater than the neighbour; otherwise, use label 0.

|  |  |  |
| --- | --- | --- |
| 0 | 1 | 0 |
| 1 | 7 | 1 |
| 0 | 1 | 0 |

X

|  |  |  |
| --- | --- | --- |
| 1 | 2 | 4 |
| 128 | 7 | 8 |
| 64 | 32 | 16 |

= 0\*1 + 1\*2 + 0\*4 + 1\*8 + 0\*16 + 1\*32 + 0\*64 + 1\*128

= 0 + 2 + 0 + 8 + 0 + 32 + 0 + 128

= **170**

**Question 3.** Suppose you have a large collection of photos from your holiday trips, including (1) photos of single person face (either yourself or other person), (2) photos with multiple faces, and (3) photos without people face. These photos are stored in your personal laptop, as illustrated in the right Figure. **Propose a computer vision system that uses what you have learned in this class to find the subset of the images that contains your face alone from this large collection of photos.** In your answer, you need to describe what computer vision methods need to be applied to achieve this objective. You could draw a flow chart or a system architecture to justify your answers.



HOG classifier, based on edges, will detect presence of multiple human faces, single or zero human faces in an image

Generate positive and negative pairs of photos and then train a Siamese Neural Network

Viola Jones algorithm for human face recognition

Siamese Network performance evaluated based on counts of false positives & negatives

Trained Siamese Neural Network will classify positive (same class) & negative pairs (different class) based on contrastive loss

When training a Siamese Network with a Contrastive loss, it will take **two inputs data** to compare **at each time step**. These two input data could either be similar or dissimilar. This is modelled by a binary class variable Y whose values are: 0 if dissimilar & 1 if similar. This loss function is known as contrastive loss.