554SM – COMPUTER VISION AND PATTERN RECOGNITION

Written Examination

January 28, 2019

Name: .....

Student Number: .....

Instructions:

- Answer the multiple-choice questions (for each question, *only one choice is correct*).
- Answer the essay question.
- Fill in the answers to the multiple-choice questions on the answer sheet (last page).

1. Which of the following statements about instance recognition from local features is **wrong**?

   (a) it is best suitable for planar objects, or deformations explained by affine transformation
   (b) it is usually accompanied by a geometric consistency check
   (c) its performance drops dramatically in the presence of even small occlusions
   (d) none of the above

2. The normalized Laplacian of Gaussian filter

   (a) is useful in calculating the direction of an edge
   (b) is invariant to contrast
   (c) can be used for detecting blobs in scale-space

3. The PCA-SIFT descriptor

   (a) builds a histogram of gradient directions for a window centered in the keypoint
   (b) does not build any histogram
   (c) builds a histogram of gradient directions for 16 sub-windows of a window centered in the keypoint

4. In convolutional neural networks, the fully connected layers are typically

   (a) the last layers
   (b) the first layer
   (c) interspersed across the whole network

5. The following filter

$$\begin{bmatrix} 1 & -1 & 2 \\ 2 & -2 & -4 \\ 1 & -1 & 2 \end{bmatrix}$$

   (a) is not separable
   (b) is separable

6. A morphological operation using a 3×3 structuring element containing all 1s has been applied to the image on the left to get the image on the right. What morphological operation has been applied?

0	0	0	0	0	0	0
0	0	1	1	1	0	0
0	1	1	1	1	1	0
0	1	1	1	1	1	0
0	1	1	1	0	0	0
0	0	1	0	0	0	0
0	0	0	0	0	0	0

original

0	0	0	0	0	0	0
0	0	0	0	0	0	0
0	0	0	1	0	0	0
0	0	1	0	0	0	0
0	0	0	0	0	0	0
0	0	0	0	0	0	0
0	0	0	0	0	0	0

transformed

(a) closing
(b) erosion
(c) opening
(d) dilation

7. The Viola-Jones detector is fast

(a) because it rejects the majority of the non-face windows with few computations
(b) because it exploits the GPU
(c) because it employs boosting

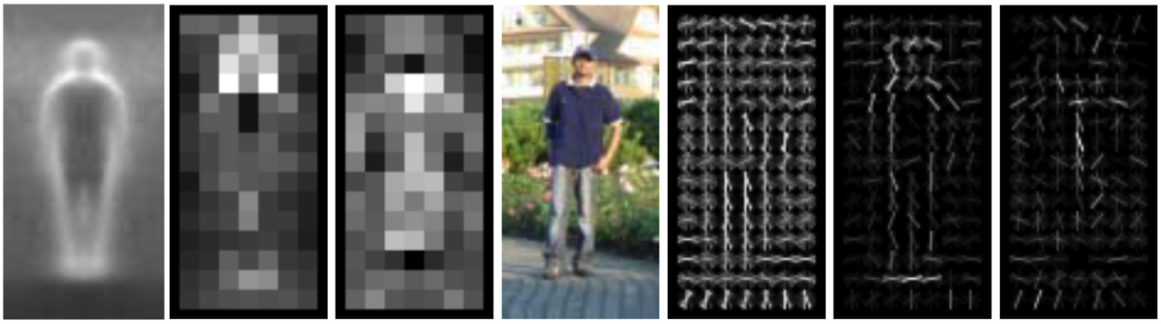
8. When training an n-class SVM based on the Directed Acyclic Graph, how many binary classifiers we need to train?

(a) $n - 1$
(b) $n(n-1)/2$
(c) n
(d) $n(n+1)/2$

9. The median filter

(a) is linear
(b) is preferred to the Gaussian filter when the noise is Gaussian
(c) is robust to outliers

10. Consider the following figure, taken from the original Histogram of Oriented Gradients paper.

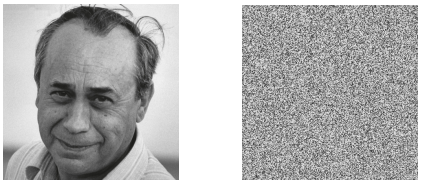


(1) (2) (3) (4) (5) (6) (7)

Which of the following captions refers to image (6)?

(a) HoG descriptor
(b) HoG descriptor weighted by positive SVM weights
(c) HoG descriptor weighted by negative SVM weights
(d) average gradient image over training examples
(e) none of the above

11. Suppose you have a training set of N labeled face/non-face images $\mathcal{S} = \{I_1, \ldots, I_N\}$ of the same size, say 256×256. You train a convolutional neural network of a given structure to distinguish between face and non-face and you get some performance on the test set $\mathcal{T} = \{J_1, \ldots, J_M\}$. Let define a permutation operator $\phi(\cdot)$ that permutes the pixels of 256×256 images (an example is shown below).



I $\phi(I)$

Now you train a convolutional neural network (having the same structure) on the training set $\tilde{\mathcal{S}}_1 = \{\phi(I_1), \ldots, \phi(I_N)\}$. What can be said about the performance obtained on the test set $\tilde{\mathcal{T}} = \{\phi(J_1), \ldots, \phi(J_M)\}$?

(a) in general, nothing can be said
(b) it is likely to be the same
(c) it is likely to be worst
(d) it is likely to be better

12. Let #SV be the number of support vectors of an SVM. Which of the following statements is **wrong**?

(a) the computational effort required for evaluating an instance depends on #SV
(b) #SV depends on the regularization parameter C
(c) the generalization capability of the SVM is an increasing function of #SV
(d) none of the above

13. As a consequence of the depth-speed ambiguity

(a) the solution of the structure from motion problem is defined up to a scale factor
(b) the faster the objects, the more uncertain their location
(c) close object moving quickly cannot be distinguished by distant object moving slowly

14. Consider the hysteresis thresholding and let θ_H and θ_L be the high and low thresholds, respectively. Which of the following statements is **wrong**?

    (a) pixels whose value is below θ_H are rejected if they are connected to pixels whose value is below θ_L
    (b) pixels whose value is below θ_L are certainly rejected
    (c) pixels whose value is above θ_H are certainly accepted
    (d) pixels whose value is above θ_L are accepted if they are connected to pixels whose value is above θ_H

15. The decision function of a soft margin SVM

    (a) depends on the unbounded support vectors but not on the bounded support vectors
    (b) does not depend on the support vectors
    (c) depends on the bounded support vectors but not on the unbounded support vectors
    (d) depends on both the unbounded and bounded support vectors

Essay question: describe briefly the main characteristics of the Viola-Jones approach to object detection.

Answer (do not exceed the frame below):

Answer Sheet

Question#	Answer
1	
2	
3	
4	
5	
6	
7	
8	
9	
10	
11	
12	
13	
14	
15	

The space below is reserved to the instructor

multiple choice (0 to 15) _____

essay question (0 to 5) _____

oral discussion (0 to 5) _____

project (0 to 5) _____

total _____