

ENGN-COMP-6528-Computer Vision-Midterm Exam

There are three question sessions
Q1-Q3.

Please name your submission as
COMP/ENGN6528_midsem_u1234567.docx
(or PDF as you prefer)

Q1.(26 points) Multiple Choices 26 Marks. There is at least one correct answer for each question. You can get the full mark only if all correct answers are selected. No mark will be granted for partially selected answers. (2 marks each)

1. (2 points) Which filter(s) could be used for removing salt-and-pepper noise?

- A. Gaussian Filter
- B. Bilateral Filter blurring
- C. Median Filter**
- D. Sobel Filter derivative

2. (2 points) Please select the linear filter(s).

- A. Gaussian filter
- B. Median filter
- C. Bilateral filter
- D. Sobel filter**

3. (2 points) Suppose a camera has a calibration matrix follows: $\begin{pmatrix} 1.5 & 0 & 0 \\ 0 & 1.5 & 0 \\ 0 & 0 & 1 \end{pmatrix}$

What does this tell you about the camera?

- A. This is not a valid camera matrix.
- B. The focal length is 1.5.**
- C. The camera has been rotated.
- D. The camera has been translated.
- E. None of the above.

4. (2 points) Which of the following are advantages of the HSV colour-space with respect to the RGB colour-space ?

- A. Brightness information is separated into a single channel.**
- B. It uses less memory to encode each pixel.
- C. Its fast to compute, so good for compression.
- D. The Hue and Saturation components of HSV are less sensitive to lighting variation.**
- E. It is a perceptually uniform space - distances in the colour space itself correspond to differences in human perception of colour.

5. (2 points) Which of the following is a (are) good property(ies) of an interest point?

- A. Strong 2D structure**
- B. Two eigenvalues of a local auto-correlation matrix (M) are both large.**
- C. Edges make good interest points.
- D. The derivatives are small.
- E. There are bright colours.

6. (2 points) Which of the following are not advantages of the SIFT feature over the Harris corner?

- A. Response is invariant to image rotation**
- B. Response is invariant to image scaling
- C. Response is invariant to image translation**

- D. Less computation required for computation of a SIFT feature.
- E. SIFT provides distinctive features for matching across images.
7. (2 points) Which of the following are properties of a Gaussian Filter?
- A. It is separable.
- B. Pixel weight is scaled according to the nearness of the pixel to the centre of the filter.
- C. The only parameter is the sigma of the Gaussian.
- D. Applying it always increase the overall image brightness.
- E. Its application sharpens edges in the image.
- F. Cascading the filter with size $\sigma = a$ followed by $\sigma = b$ has the same effect as applying a single filter with size $\sigma = (a + b)$
- $\sigma_{result} = \sqrt{\sigma_a^2 + \sigma_b^2}$
8. (2 points) Suppose you have a simple fully connected linear network, with bias, that takes a 32x32 greyscale image as input. The input is connected directly to a classification space with 10 output units. How many free parameters are there in the network?
- A. 10240
- B. 10250
- C. 10260
- D. 320
- E. 330
- F. 340
9. (2 points) Suppose we have a softmax function at the end of our network that has 4 outputs corresponding to 4 classes. Which of the following are possible outputs from the function?
- A. 0.1 1 0.3 0.7
- B. 0.3 0.4 0.2 0.1
- C. 1 0 1 0
- D. 0.4 0.1 0.4 0.1
- E. 0.5 0 1.5 0
- F. 0.5 0 -0.5 0
- G. 0.5 0 0.7 -0.2
- H. None of the above
10. (2 points) Which loss function is the best suited to classification problems? Note that y is the estimated value, y^* defines the ground truth. (Please select multiple answers if you consider them to be equally good).
- A. $\frac{1}{2}(y - y^*)^2$
- B. $y * \log(y) + (1 - y^*)\log(1 - y)$
- C. $|(y - y^*)|$
- D. None of these
11. (2 points) Suppose we have a 7x7 convolutional layer, followed by a ReLU, then sequentially another two 3x3 convolutional layers, each of them is followed by a ReLU. Note that the stride is equal to 1 for all convolutional layers. What is the effective receptive field of a node after the third convolutional layer?
- A. 7x7

- B. 9x9
- C. 10x10
- D. 3x3
- E. 11x11
- F. None of the above.

12. (2 points) Suppose a network has an input image of 240x240. A convolution layer of 5x5 with no padding, and a stride of two is applied. This is followed by a ReLU, and then a pooling layer of 2x2, stride 2. What is the size of the output of the pooling layer?

- A. 60x60
- B. 59x59
- C. 118x118
- D. 240x240
- E. 120x120
- F. None of the above.

$$\text{Output size} = \left\lfloor \frac{\text{Input size} - \text{Filter size}}{\text{Stride}} \right\rfloor + 1$$

13. (2 points) Which of the following does regularization in Neural Networks help with?

- A. keeping the weights small.
- B. preventing overfitting.
- C. making some of the weights zero.
- D. None of the above.

Q2. (34 points) Short Answer Questions. (No answer should be more than half a page.)

1. (3 points) Please provide one example each for a point operation and a neighborhood operation.
2. (5 points) Please describe the differences between Gaussian filter and Bilateral filter. What are the advantages and disadvantages of the bilateral filter?
3. (3 points) How does SIFT extract scale invariant features?
4. (3 points) What are the advantages of a separable filter?
5. (3 points) What are the advantages and disadvantages of forward warping and inverse warping for image transformation?
6. (5 points) Why is Harris corner detector rotation invariant?
7. (5 points) Please discuss whether bilateral filtering can remove salt and pepper noise. Please provide the reasoning.
8. (7 points) Compare the activation function shown in Fig. 1 with a Relu function. Discuss the differences and the behavior of this activation function when training a network.

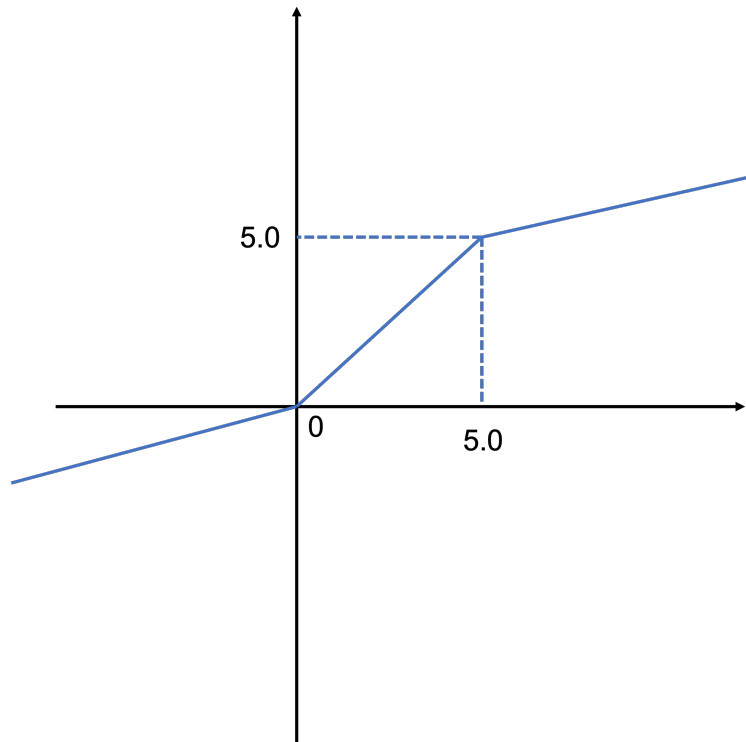


Figure 1: Activation Function Comparison

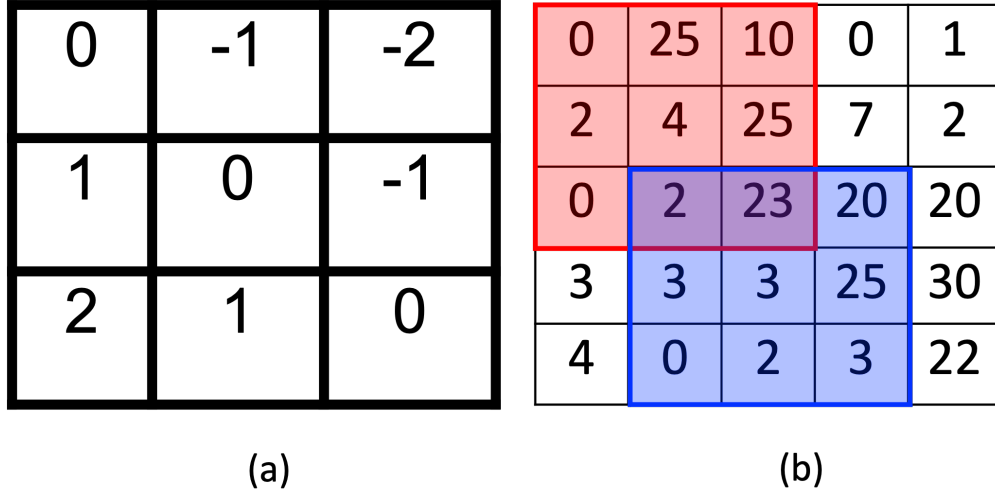


Figure 2: Image filter (a) and Image patch (b)

Q3. (40 points) Calculation questions. Please only provide the final answer.

1. (5 points) Given an input 5x5 image: $\begin{pmatrix} 230 & 22 & 9 & 4 & 7 \\ 210 & 31 & 20 & 12 & 2 \\ 109 & 2 & 4 & 3 & 2 \\ 33 & 3 & 2 & 4 & 5 \\ 21 & 4 & 5 & 6 & 3 \end{pmatrix},$

suppose a Deep CNN uses a filter (correlation) that it has learned: $\begin{bmatrix} 1.2 & 0.4 & 2 \\ 0.6 & 0.2 & 1 \\ 2 & 0.9 & 0.6 \end{bmatrix}$ to this image patch.

Then it is followed by a ReLU activation function. What will be the result for the centre pixel of this image?

2. (5 points) Suppose that Fig. 2 (b) is an image, and (a) is a filter.

Calculate the result if you were to apply the filter as a convolution operator for the red (upper left) and the blue square (bottom right) (Note: you are going to obtain one value for each region).

3. (15 points) Calculate the value of the blue patch in Fig. 3 using bilateral filtering. Assume the Domain

kernel is of size 5×5 , the standard deviation $\sigma_d = 3$, provided as $\begin{pmatrix} 0.0318 & 0.0375 & 0.0397 & 0.0375 & 0.0318 \\ 0.0375 & 0.0443 & 0.0469 & 0.0443 & 0.0375 \\ 0.0397 & 0.0469 & 0.0495 & 0.0469 & 0.0397 \\ 0.0375 & 0.0443 & 0.0469 & 0.0443 & 0.0375 \\ 0.0318 & 0.0375 & 0.0397 & 0.0375 & 0.0318 \end{pmatrix}$

the Range kernel is of size 5×5 and the standard deviation $\sigma_r = 50$. Please 1) provide the range filter associated to the pixel high-lighted in the Fig. 3, and 2) show the filtered value for the high-lighted pixel.

4. (15 points) Back propagation through the computational graph (see Fig. 4). The current value for $w_0 = 0.2, w_1 = 0.3, w_2 = 0.4, x_0 = 2, x_1 = 3$. p and q define the intermediate variables. l is the output of the computational graph. Please provide the gradient $\frac{\partial l}{\partial p}$ and $\frac{\partial l}{\partial q}$ based on the back-propagated gradient calculation.

128	128	100	100	103	50
150	100	120	30	53	54
150	112	127	40	35	20
132	125	112	43	20	10
133	130	100	30	10	20
140	130	120	20	10	20

Figure 3: Image Patch

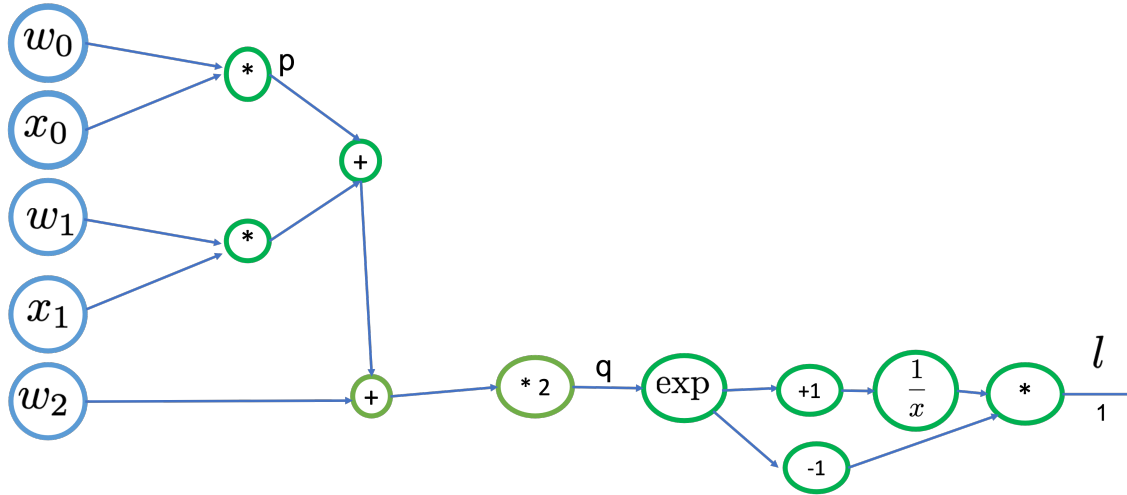


Figure 4: Computational Graph. $*$ is the multiplication operator, $+$ is addition operator, **exp** denotes the exponential operator, $+1$ defines the sum of the input with 1 and -1 defines the subtraction of 1 from input, $\frac{1}{x}$ defines the reciprocal operator.