Sample Questions : Section 1 & 2.

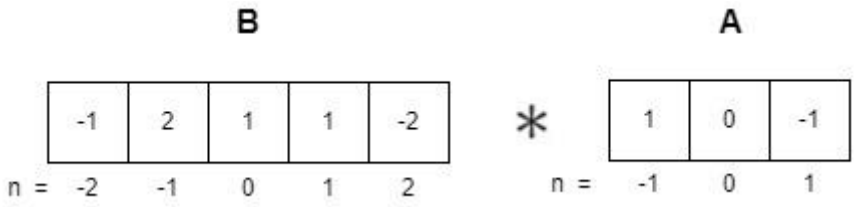
Dr. Tony Scanlan Machine Vision & Image Processing

Question 1:

(i) What are the two key stages in producing a digital image?
(ii) What is the function of the shutter in a digital camera?
(iii) How can a colour image be produced using a single image sensor?
(iv) What colour space do most display devices use?
(v) What physical principle is used in image sensors to convert incident light to electrons?

Question 2:

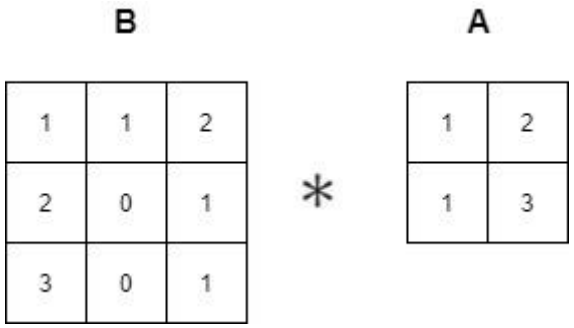
Given the 1D convolution equation $(x * h)[n] = \sum_{k=-\infty}^{\infty} x[k]h[n-k]$ Find the result of convolving the 1D kernel A with the sequence B. (Show how the delayed and reversed kernel is used to determine the convolution result over the range values of $n = [-3,3]$ in your answer).



(Note that you can check your answer using this online calculator.)

Question 3:

Find the result of convolving the 2x2 kernel **A** with the 3x3 image **B**. (Assuming no-padding this is also called "Valid Padding")



Question 4:

Find the result of convolution of the 3x3 kernel **A** with the 5x5 image **B** at the position shown in the red square only.

B

1	1	1	2	2
1	1	1	2	2
2	2	0	1	1
3	3	0	1	1
3	3	0	1	1

$*$

A

1	-2	1
2	4	2
1	-2	1

Question 5:

If we have an image of size 16 and a kernel of size 3 with unit strides and no zero padding, what size will the output image be? (Use the relationships in the Guide to Convolutional Arithmetic to calculate the answer)

Question 6:

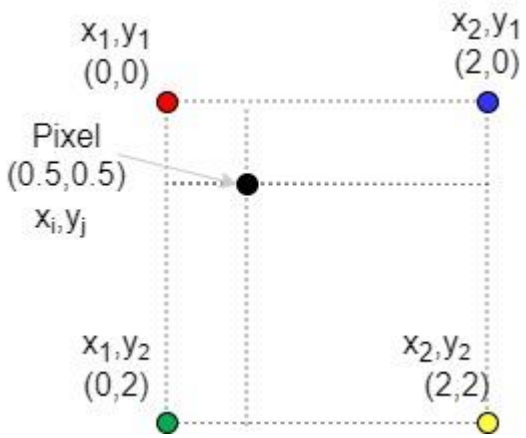
If we have an image of size 24 and a kernel of size 5 with unit strides how much padding is required to make the output size the same as the input? Also show the output size is the same as the input with calculation. (Use the relationships in the Guide to Convolutional Arithmetic to calculate the answer)

Question 7:

I. According to the Nyquist/Shannon sampling theorem, what is the maximum signal frequency that can be sampled in a discrete time signal processing system.
II. What is meant by "Aliasing" in sampling? How is Aliasing mitigated in a typical sampling system?
III. Why do image frequencies arise in a sampled signal processing system?
IV. What is the function of an interpolation filter in signal reconstruction? How can we achieve perfect reconstruction of a signal?
V. What can happen to a sampled signal if it is reconstructed with a function such as linear interpolation and then re-sampled?
VI. How is frequency measured in a digital Image?
VII. What is the effect of aliasing in a digital image?
VIII. Why would we expect to see jagged lines in images that are reduced in size using nearest neighbour interpolation.
IX. What is the Gibb's phenomenon and why is it observed when a Sinc Interpolation filter is used to re-size digital images?

Question 8:

A pixel is to be bilinearly interpolated from the 4 nearest pixels (Intensities of Pixels: I_{red} = 10, I_{blue}= 5, I_{Green}= 12, I_{yellow} = 8). Calculate the interpolated pixel value.



Question 9:

Perform Separable Convolution between the Image A and Kernel B. Assume that we want to keep the input and output image the same size (use one ring of zero padding to achieve this).

A

3	1	2
2	4	1
3	0	2

$*$

B

1	2	1
0	0	0
-1	-2	-1

Question 10:

I. What do we mean when we say a 2D is kernel is Separable?
II. Give a definition of the superposition principle as it applies to image processing (in words)?
III. What do we mean if a linear filter displays Shift Invariance?
IV. What is the commutative property of convolution?
V. What is the associative property of convolution?

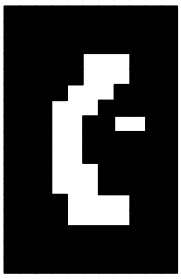
Question 11:

Determine the result of applying a 3 x 3 median filter to the image shown.

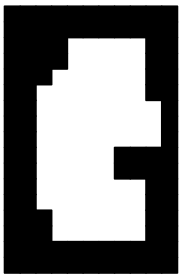
11	10	8	7	15
8	9	6	8	12
7	2	5	9	14
11	10	8	13	11
6	12	19	20	18

Question 12:

Which morphology operation has been applied to go from binary image (a) to binary image (b).



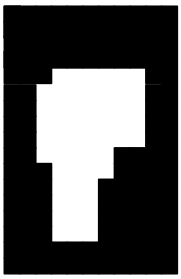
(a)



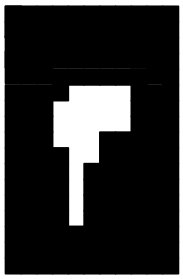
(b)

Question 13:

Which morphology operation has been applied to go from binary image (a) to binary image (b).



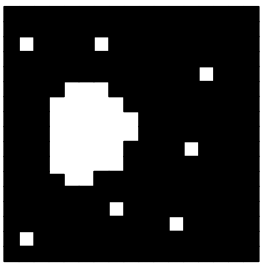
(a)



(b)

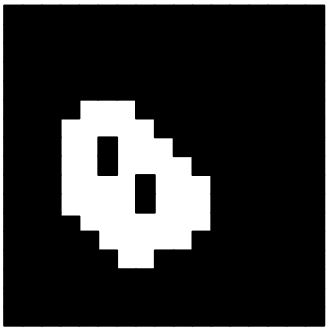
Question 14:

Which morphology operation would we apply to the binary image below in order to remove the noise pixels yet keep the size of the foreground object the same.



Question 15:

Which morphology operation would we apply to the binary image below in order to remove the small holes in the white region of the image yet keep the size of the foreground object the same.



Question 16:

(i) What is the general principle used to find edges in images?
(ii) What will happen to a noisy image if an edge detector is applied?
(iii) What direction of edges will an X-direction kernel recover?
(iv) Names 3 types of edge detector?
(v) What is the principle difference between the Laplacian kernel and the other kernels?
(vi) In the Gaussian Derivative kernel, what is the effect of the σ parameter?
(vii) What is the characteristic effect seen in an image that has been filtered by the Laplacian kernel?

Question 17:

B

1	1	1	2
1	1	2	2
1	2	2	4
2	2	4	4

X

-1	0	1
-2	0	2
-1	0	1

Y

1	2	1
0	0	0
-1	-2	-1

$$G = \sqrt{G_x^2 + G_y^2} \qquad \theta = tan^{-1}\left(\frac{G_y}{G_x}\right)$$

Part (a) Given the x-direction Sobel kernel **X** and the y-direction Sobel kernel **Y** and the image **B**, find the gradient images **Gx** and **Gy** by convolving the **X** and **Y** kernels with **B**. (In this question assume the kernels have already been flipped)

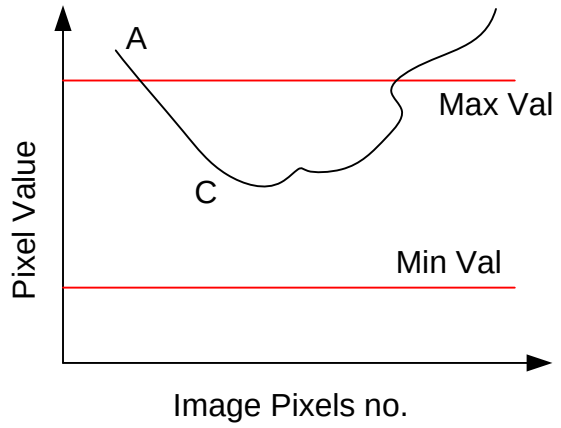
Part(b) Now that you have determined **Gx** and **Gy** find the magnitude **G** and angle **θ** of the combined gradient matrices using the equations provided.

Question 18:

List the principle steps (There are 4) in the canny edge detection algorithm.

Question 19:

In the Canny edge detection algorithm Hysterersis Thresholding is applied. As shown in the image below the pixel from an edge and the threshold values (max val & min val) are plotted. Pixel C has a value between the thresholds and is connected to pixel A. Will pixel C be retained by the algorithm?



Question 20:

(i) In the linear Hough Transform lines in the image space are transformed to points in parameter space. How are the lines corresponding to edges in the image identified?

(ii) What is the fundamental problem with the representation of lines in parameters space as a slope m and intercept b, as used in the linear hough transform?

Question 21:

Choose the correct statement from the following:

(i) The Harris corner detector operates by finding intensity changes moving in a small neighbourhood of the pixel.

(ii) The harris corner algorithm uses gradient information along the x-axis only.

(iii) The harris corner algorithm uses frequency occurrence of pixel information.

Question 22:

In the Harris Corner Detection Algorithm the second moment matrix is formed from approximate derivatives of the image convolved with a Gaussian window. Given the second moment matrix $M = \begin{bmatrix} A & C \\ C & B \end{bmatrix} = \begin{bmatrix} 10 & 1 \\ 1 & 10 \end{bmatrix}$ calculate the corner response function R and determine if the current point is an edge, corner or flat region (allow $\alpha = 0.04$).

Question 23:

Select the correct statements from the list below:

(i) In the Harris corner algorithm an Edge will have one high and one low curvature.

(ii) In the Harris corner algorithm all edges will have all high curvatures.

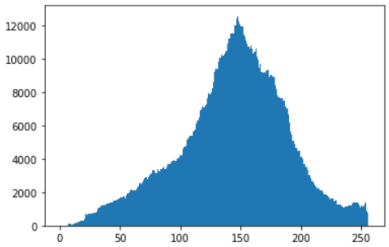
(iii) In the Harris corner algorithm all edges will have all minimal curvatures.

(iv) In the Harris corner algorithm all corners will have all high curvatures.

Question 24:

In image (a) we would like to segment the foreground image of the man from the background, given the image histogram (b) explain why this will not be possible.

(a)



(b)

Question 25:

(a) In the watershed algorithm without markers, where does the flooding start from?
(b) What happens when the water from two sources meet?
(c) What is the common problem encountered if we use the watershed algorithm without markers?
(d) Why do we apply the distance transform in the watershed algorithm with markers.

Question 26:

In the clustering technique which can be used to segment an image, which algorithm is commonly used to cluster the pixels based on their colour?

Sample Questions : Section 3 & 4.

Dr. Tony Scanlan Machine Vision & Image Processing

Question 1:

(i) List some qualities (four) of a good feature detector.

(ii) In feature matching of images, descriptor vectors from pairs of similar images must be compared. How is a match between two descriptor vectors detected?

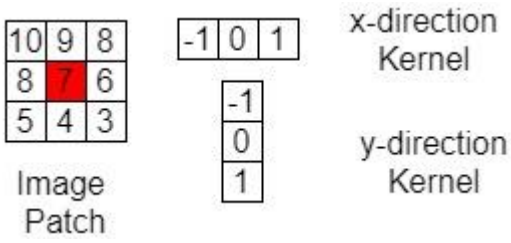
(iii) If the Laplacian of Gaussian "blob" feature detector is to be made scale invariant, how will this be achieved?

(iv) What are the 4 key steps in the SIFT transform.

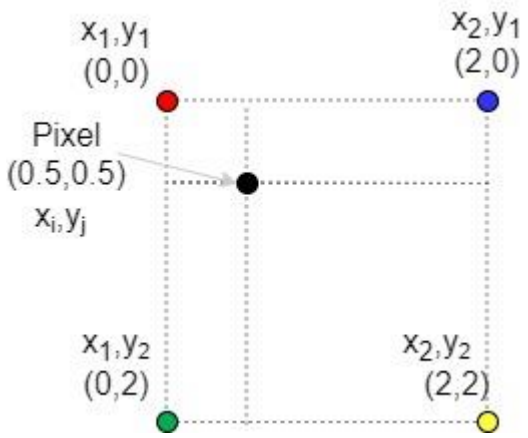
(v) What is the reason behind determining the characteristic orientation in the SIFT keypoint descriptor?

Question 2:

(i) Calculate the magnitude and gradient of the red pixel in the image patch below, use the simple 1D kernels to obtain x and y components of the gradient.



(ii) A pixel with a calculated magnitude of μ =10 is to be divided between the 4 nearest histograms (centres as shown in the diagram). Calculate the contribution to the histogram with centre (0,0).



Question 3:

(i) How would we define supervised learning?

(ii) What are the three key components in a supervised learning system?
(iii) What is the difference between Classification and Regression?

Question 4:

(i) If a classifier model has a high bias and low variance, what type of fitting behaviour are we likely to see if the model is trained on the data?
(ii) If a classifier model has a low bias and high variance, what type of fitting behaviour are we likely to see if the model is trained on the data?
(iii) List some ways to minimise overfitting during training?
(iv) What does the term Generalisation mean for a classifier?

Question 5:

(i) If we build an urban speed limit (50Km/hr) sign detection system and we find that out of 200 detections it correctly recognises 85 (50Km/hr) signs, it incorrectly identified 12 signs (i.e. 100km/hr, 80km/hr as 50km/hr) and it failed to identify 9 (50Km/hr) signs, calculate the Accuracy Precision and Recall of the system.
(ii) If a classifier had high precision but low recall, what would we expect to observe with regards to the number of False Positives and False Negatives
(iii) If a classifier had low recall but high Precision, what would we expect to observe with regards to the number of False Positives and False Negatives

Question 6:

(i) We want to use a Naïve Bayes Classifier to classify types of fruit (Banana, Orange or Other) based on a series of features as shown in the table below. If a fruit is picked at random and displays the features of being "Long", "Sweet" and "Yellow" what is the most likely class? (determine the probabilities to justify your answer)

Type/Feature	**Long**	**Not Long**	**Sweet**	**Not sweet**	**Yellow**	**Not Yellow**	**Total**
Banana	400	100	350	150	450	50	500
Orange	0	300	150	150	300	0	300
Other	100	100	150	50	50	150	200
Total	500	500	650	350	800	200	**1000**

(ii) We want to use a Naïve Bayes Classifier to classify types of dogs (Labrador or Chihuahua) based on a series of features as shown in the table below. If a dog is picked at random and displays the features of being "small", "Not black coat" and "Not short hair" what is the most likely class? (determine the probabilities to justify your answer)

Type/Feature	**Small**	**Not Small**	**Black Coat**	**Not Black Coat**	**Short hair**	**Not short hair**	**Total**
Labrador	50	250	120	180	250	50	300
Chihuahua	650	50	150	550	200	500	700
Total	700	300	270	730	450	550	**1000**

Question 7:

What image recognition challenges will we encounter when trying to detect “Deckchairs” in the above image?

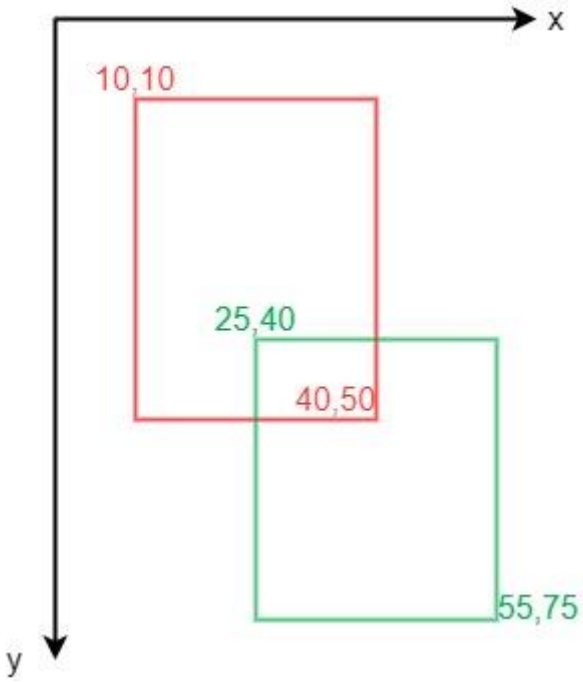
Question 8:

(i) Why is a sliding window often necessary for object detection?

(ii) What additional problem with a fixed detector window size does using an image pyramid solve?

(iii) What is the main drawback with using a sliding window and image pyramid approach.

(iv) Calculate the Intersection over Union of the two bounding boxes in the image below.



Question 9:

I. What is the disadvantage of using simple correlation based similarity for template matching?

II. Why is normalised Cross Correlation computationally expensive?

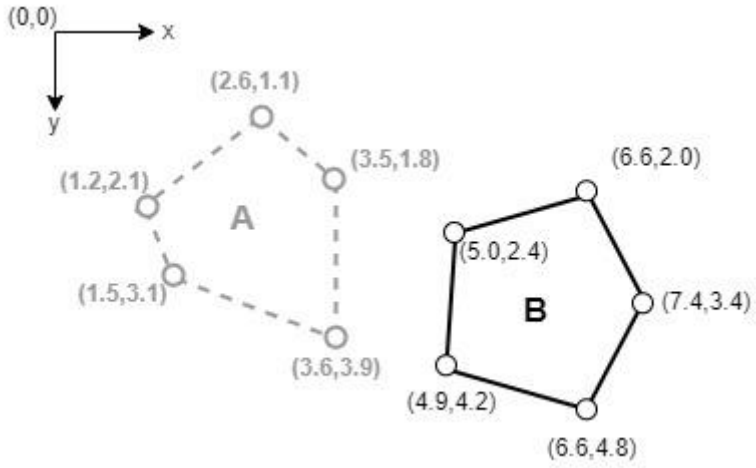
III. Why is the shortest distance a poor metric to use for template matching with a binary image (& binary template)?

Question 10:

(i) Determine the directed Hausdorff distance $h(A,B) = \max_{a\in A} \min_{b\in B} \|a-b\|$ between the point sets A (grey dashed line) and B (shown below) shown below.

(ii) Determine the directed Hausdorff distance $h(B,A) = \max_{b\in B} \min_{a\in A} \|a-b\|$ between the point sets A (grey dashed line) and B (solid line) shown below.

(iii) Determine the Hausdorff distance $H(A,B) = \max(h(A,B), h(B,A))$

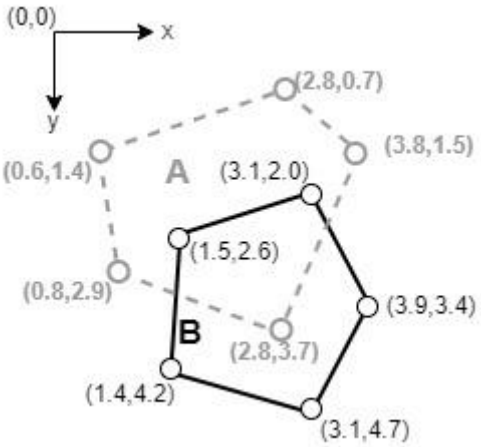


Question 11:

(i) Determine the directed Hausdorff distance $h(A,B) = \max_{a\in A} \min_{b\in B} \|a-b\|$ between the point sets A (grey dashed line) and B (solid line) shown below.

(ii) Determine the directed Hausdorff distance $h(B,A) = \max_{b\in B} \min_{a\in A} \|a-b\|$ between the point sets A (grey dashed line) and B (solid line) shown below.

(iii) Determine the Hausdorff distance $H(A,B) = \max(h(A,B), h(B,A))$



Question 12:

(i) List 3 key enhancements or approaches within the Viola Jones Algorithm that enable fast detection

(ii) Obtain the integral image of the image patch (a), and find the sum of the pixels under the shaded region using values from the integral image. (For the sum calculation write your answer in the form A-B-C+D = _)

1	0	1	1	2
0	2	1	2	0
2	1	1	0	1
1	1	2	1	2
1	1	2	0	2

(a)

(iii) Obtain the integral image of the image patch (b), and find the sum of the pixels under the shaded region using values from the integral image. (For the sum calculation write your answer in the form A-B-C+D = _)

1	2	0	1	2
1	1	1	2	0
0	2	1	0	1
1	1	0	1	2
1	1	2	1	3

(b)

Question 13:

(i) In the Histogram of Orientated Gradients representation, what image feature is detected.

(ii) Given the gradient information (Magnitude and Angle) for a 4 x 4 pixel cell and assuming that the magnitude is to be interpolated between the two nearest neighbouring histogram bins. (a) Determine which bin centres will have some magnitude assigned by the pixel in the red square (b) and calculate the magnitudes that will be assigned to each of the bins.

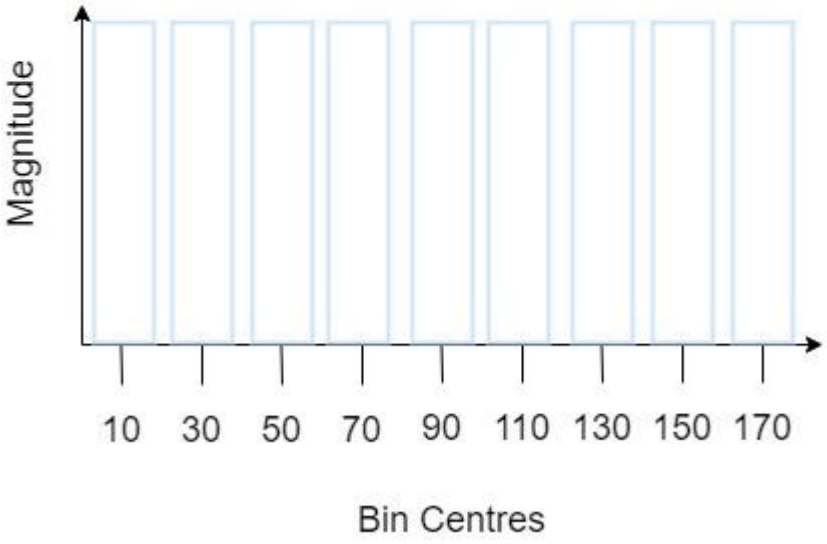
Magnitude

35	20	8	2
30	10	5	1
25	12	6	2
28	8	4	1

Angle (Degrees)

45	49	55	60
44	48	54	61
44	50	54	59
39	53	55	63

(a) 4x4 Cell Magnitude and Angle



(b) Empty Histogram of Gradients

(iii) If a Histograms of Gradient based descriptor is comprised of: cells of 4 x 4 pixels, blocks of 2 x 2 cells, histograms with 9 orientations in each cell and 6 x 10 overlapping blocks at the top level; What is the length of the Descriptor vector?

Sample Questions : Section 5 & 6.

Dr. Tony Scanlan Machine Vision & image Processing

Question 1:

(i) In principle component analysis if we multiply the original data by all the Eigenvectors. Will the resulting data be covariant not covariant?

(ii) In principle component analysis if the original data is multiplied by a subset of the eigenvectors, what will happen to the data?

(iii) Which of the following statements are true about Principle Components Analysis.

a. The principal components are eignvectors of the centered data matrix.
b. The principal components are eigenvectors the sample covariance matrix.

Question 2:

Obtain the eigenvalues and eigenvectors of the 2 x 2 matrix $A = \begin{pmatrix} 0 & 1 \\ -2 & -3 \end{pmatrix}$ (show your working out)

Question 3:

Obtain the eigenvalues and eigenvectors of the 2 x 2 matrix $A = \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix}$ (show your working out)

Question 4:

(i) Consider the following data matrix $\boldsymbol{X} = \begin{bmatrix} 4 & 2 & 5 & 1 \\ 1 & 3 & 4 & 0 \end{bmatrix}$ compute the principle components for $\boldsymbol{X}$. (Matrix columns correspond to 2 dimensional data points, Note that you can use python or other computer programme to determine eigenvalues/vectors in this question)

(ii) Which eigenvector would we choose if we wanted to reduced the dimensionality of the data to 1 dimension?

(iii) What are the values of the data along this reduced dimension?

Question 5:

What advantage has the Fisherface method over the Eigenfaces method of facial recognition.

Question 6:

(i) Answer if the statements below are true or false

a. The bag of words algorithm matches similar documents by comparing word frequency
b. The bag of features representation can be used for image classification
c. The bag of features uses an orderless representation

(ii) Which popular feature(or keypoint) descriptor do we use in the bag of features to initially represent the image features.

(iii) Briefly explain (i.e give main points) how the Bag of features representation is generated for an image.

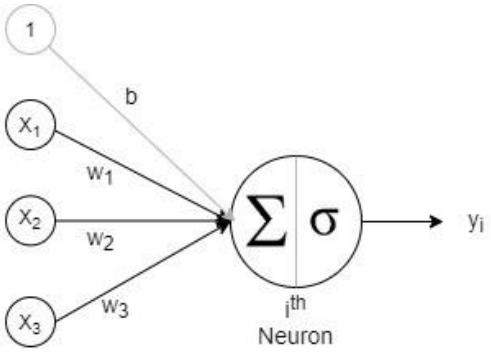
(iv) Briefly explain how images can be retrieved with Bag of features and an inverted file index.

(v) In the bag of words we want to determine the Tf-idf (Term Frequency-inverse document frequency) weighted word frequency t_i to a word i that appears $n_{id} = 20$ times in a document, given that $n_d = 2500$ is the total number of words in the document, n_i =1520 is the number of times the word appears in the database and $N = 750{,}000$ is the number of documents in the database.

(vi) In the bag of words we want to determine the Tf-idf (Term Frequency-inverse document frequency) weighted word frequency t_i to a word i that appears $n_{id} = 310$ times in a document, given that $n_d = 3500$ is the total number of words in the document, n_i =63,524 is the number of times the word appears in the database and $N = 750{,}000$ is the number of documents in the database.

Question 7:

(i) For the simple neuron as shown in the image, with inputs $\boldsymbol{x} = [0.9, 0.7, 0.2]$, weights $\boldsymbol{w} = [0.5, 0.4, -0.2]$ and bias $b = 0.3$ calculate the output y_i of the neuron (assuming a linear activation function).



(ii) If the input to a rectified linear unit is z =-0.5 what is the output?

(iii) If the input to a sigmoid activation unit is $z = -0.1$, what is the output given the sigmoid function $\sigma(z) = \frac{1}{1+e^{-z}}$

(iv) Calculate the output probabilities for the softmax activation function with inputs $z_i = [2, 1.5, 2.2, 0.9]$ given the equation for the softmax activation $\sigma(z_i) = \frac{e^{z_i}}{\sum_{j=1}^{K} e^{z_j}}$.

(v) When training a neural network what happens if the learning rate is too high?

Question 8:

(i) What is a key difference between Deep learning based systems and earlier approaches to building machine vision systems?

(ii) What combination of three technologies helped drive the revolution in deep learning?

(iii) List some disadvantages of Deep Learning?