Exam

02506
Spring 2023

General information

- The exam consists of 20 questions. All questions have equal weight: the correct answer gives 1 point, and the incorrect or missing answer gives 0 points. You only need to submit the answers to the questions. You should not upload any notes or calculations.
- There is one correct answer for each question. Some of the numeric results have been rounded, and might slightly deviate from your result. This should not prevent you from being able to pick the correct answer.
- Each page contains one question. If there are illustrations and images, those refer to the question on that page.
- The notation in the questions is the same as in the course note.
- Data and images are provided for some questions. The question text states which data or image should be used.
- To load arrays saved as text files you can use `numpy.loadtxt` in Python and `dlmread` in Matlab.

Relevant links

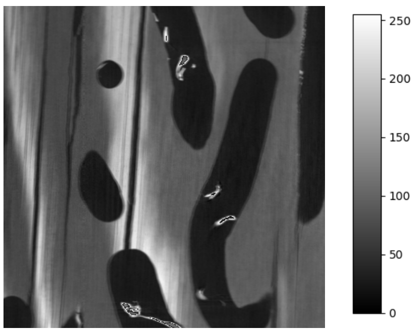
- Course note
  `http://www2.imm.dtu.dk/courses/02506/LECTURE_NOTES_2023.pdf`
- Resources are available for a subset of questions in the form of data, images, and code. Filenames typeset in `typewriter font` indicate that you can find files in the materials folder
  `http://www2.imm.dtu.dk/courses/02506/EXAM2023/DATA_2023.zip`

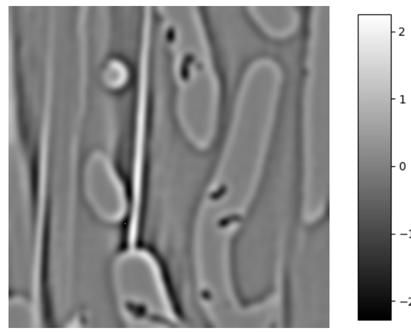
Question 1

Gaussian filtering The image on the left shows an original image and the image on the right shows the result after computing the Laplacian of the Gaussian

$$\nabla^2 L = L_{xx} + L_{yy}.$$

The images are given in the files `wood.png` and `wood_LoG.tif`.





Which standard deviation σ is used to compute the Laplacian of Gaussian image?

(a) $\sigma = 0.0$
(b) $\sigma = 1.2$
(c) $\sigma = 2.0$
(d) $\sigma = 2.4$
(e) $\sigma = 3.3$
(f) $\sigma = 4.8$
(g) $\sigma = 5.5$
(h) $\sigma = 6.4$
(i) $\sigma = 8.3$
(j) $\sigma = 13.4$
(k) $\sigma = 22.8$
(l) $\sigma = 30.3$

Question 2

Mean bending Consider a closed curve represented by N points and available as `zealand.txt`.

The bending at a curve point $\mathbf{p}_i$ is defined as

$$b_i = ||\mathbf{p}_{i-1} + \mathbf{p}_{i+1} - 2\mathbf{p}_i||$$

where the first and the last point need to be appropriately handled. The mean curve bending is defined as

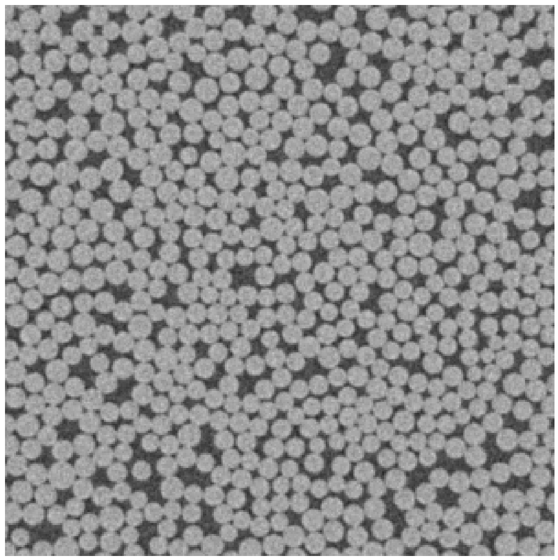
$$M = \frac{1}{N}\sum_{i=1}^{N} b_i \,.$$

What is the mean curve bending of this curve?

(a) -2.4
(b) -8.9
(c) 2.8
(d) 5.2
(e) 11.2
(f) 45.8
(g) 126.6
(h) 197.5
(i) 395.0
(j) 205368.1

Question 3

Blob detection The image below shows a slice of a CT scan of composite fibers. The image file is called `fiber.png`. You should detect blobs at the scale $t = 9$ by finding the extrema (8 nearest neighbors) of the Laplacian of the Gaussian image at that single scale.



How many dark and bright blobs do you detect?

(a) Dark: 10, bright: 17
(b) Dark: 17, bright: 10
(c) Dark: 108, bright: 102
(d) Dark: 102, bright: 108
(e) Dark: 123, bright: 324
(f) Dark: 324, bright: 123
(g) Dark: 576, bright: 312
(h) Dark: 312, bright: 576
(i) Dark: 672, bright: 633
(j) Dark: 633, bright: 672
(k) Dark: 1230, bright: 32
(l) Dark: 32, bright: 1230
(m) Dark: 1526, bright: 1461
(n) Dark: 1461, bright: 1526
(o) Dark: 1768, bright: 1836
(p) Dark: 1836, bright: 1768

Question 4

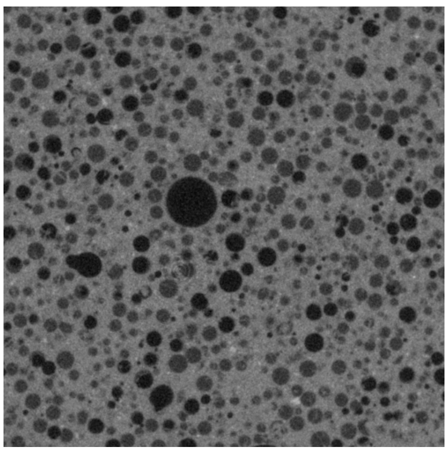
Blob size The image to the left shows a slice from a CT scan of a concrete sample and on the right, you see the same image with blobs detected. The original image is in the file `cement.png` detected blobs are listed in the text file `cement.txt`. The first column is the blob response β, the second column is the y-coordinate (row), the third column is the x-coordinate (column), and the fourth column is the scale t at which the blob was detected. An estimate of the blob radius r (in pixels) is given by

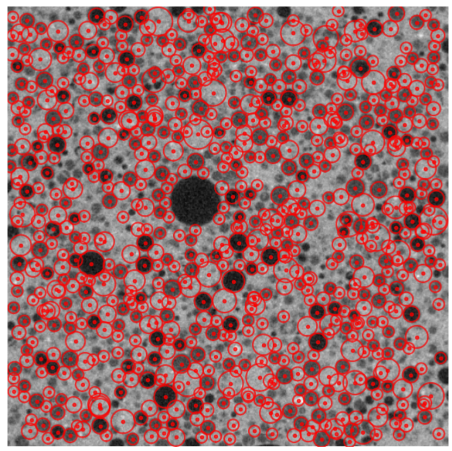
$$r = \sqrt{2t}\ .$$

We are interested in blobs with a with a sufficiently large response

$$|\beta| > 10\ .$$

Furthermore, we are only interested in dark blobs.





What is the average diameter (in pixels) of the detected high-response dark blobs?

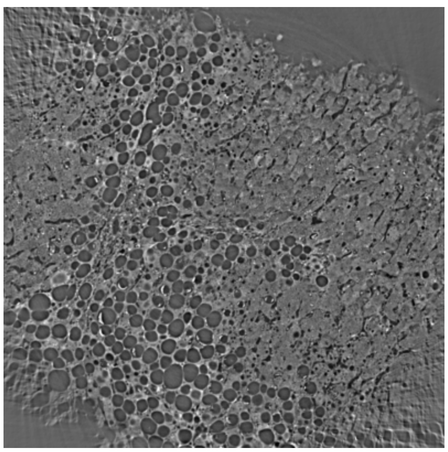
(a) 3.5
(b) 5.7
(c) 7.4
(d) 7.7
(e) 7.9
(f) 10.5
(g) 14.9
(h) 15.4
(i) 15.7
(j) 16.8
(k) 19.4
(l) 21.4
(m) 22.6
(n) 25.4
(o) 31.4
(p) 115.4

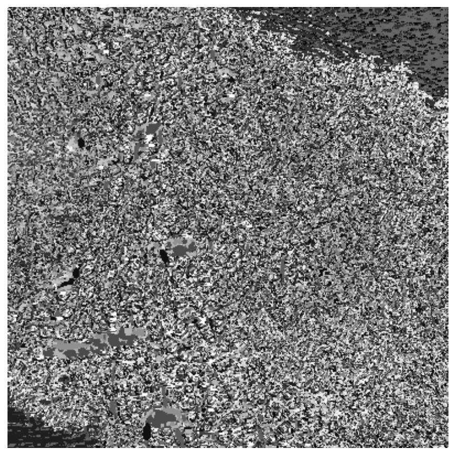
Question 5

Feature-based segmentation The image below to the left shows a slice of a CT scan of liver tissue, which should be segmented using feature-based segmentation as done in the Lecture note Chapter 3.

A dictionary with 256 dictionary elements has been trained. Based on the dictionary, the assignment image to the right has been obtained. The assignment image is called `liver_assignment.png`.

Furthermore, a dictionary of label-wise probabilities has been obtained, and these probabilities are given in the file `dictionary_probabilities.txt`. There are three columns in the file corresponding to the three labels, such that column 1 is the probability of label 1, column 2 is the probability of label 2, etc. Each row corresponds to a label index, such that the first row corresponds to pixels assigned to index 0 in the assignment image, the second row to index 1, etc. Each pixel should be assigned a label according to the largest label-wise probability.





What is the proportion of the image that will be labeled with label 3?

(a) 0
(b) 0.089
(c) 0.215
(d) 0.300
(e) 0.376
(f) 0.408
(g) 0.504
(h) 0.623
(i) 0.896
(j) 0.923
(k) 1

Question 6

Gaussian features This question is based on the approach for computing Gaussian features as described in the Lecture note Chapter 3.

Which statement about Gaussian features is correct?

(a) Gaussian features can be computed densely over the image, such that each pixel is represented by a feature vector. This is the only feature type that has this property, and therefore the only feature that can be used for feature-based segmentation.

(b) Gaussian features are densely computed features obtained at multiple scales. To ensure accurate matching of the features, it is essential that the features are computed with increasing scale, such that the part at the lowest scale is concatenated with a part at a higher scale, and so forth.

(c) Gaussian features are dense features that are inspired by Gaussian scale space. They are computed by first detecting what type of feature is in a pixel at a given scale. These include edges, blobs, ridges, saddle points, and flat areas. Then the feature with the maximum response at the scale is chosen. Finally, features at multiple scales are combined into one feature vector.

(d) For feature-based segmentation, Gaussian features are computed as dense features that characterize the texture in a neighborhood around a feature. By computing the features using several σ parameters going from small to larger, we make sure to capture both fine-scale textures and coarser textures.

Question 7

Point correspondance Two sets of points $\mathbf{p}$ and $\mathbf{q}$ have been matched and their coordinates are given in the files `points_p.txt` and `points_q.txt`. We have the transformation

$$\hat{\mathbf{q}} = s\,\mathbf{R}\,\mathbf{p} + \mathbf{t}$$

where s is the scale, t is the translation, and the rotation matrix

$$\mathbf{R} = \begin{bmatrix} \cos\theta, & -\sin\theta \\ \sin\theta, & \cos\theta \end{bmatrix}.$$

Here, θ is the rotation angle.

What are the parameters, scale s, translation t, and rotation angle θ in degrees of the transformation that minimizes $\sum_{i=1}^{n}(\hat{\mathbf{q}}_i - \mathbf{q}_i)^2$?

(a) $s = 1.195, \quad \mathbf{t} = [-3.72, 12.17]^T, \quad \theta = 111.32°$
(b) $s = 0.84, \quad \mathbf{t} = [17, -4]^T, \quad \theta = 48°$
(c) $s = 1.2, \quad \mathbf{t} = [-4, 12]^T, \quad \theta = 112°$
(d) $s = 0.676, \quad \mathbf{t} = [-37.14, 4.37]^T, \quad \theta = 301.59°$
(e) $s = 0.837, \quad \mathbf{t} = [17.13, -3.72]^T, \quad \theta = 48.41°$
(f) $s = 1.48, \quad \mathbf{t} = [37.14, -4.37]^T, \quad \theta = 248.68°$
(g) $s = 1.52, \quad \mathbf{t} = [-17.14, 14.37]^T, \quad \theta = 189.12°$
(h) $s = 0.655, \quad \mathbf{t} = [50.46, 15.84]^T, \quad \theta = 17.18°$
(i) $s = 0.235, \quad \mathbf{t} = [12.14, -14.37]^T, \quad \theta = 84.31°$
(j) $s = 1.09, \quad \mathbf{t} = [-98.14, 114.37]^T, \quad \theta = 307.81°$

Question 8

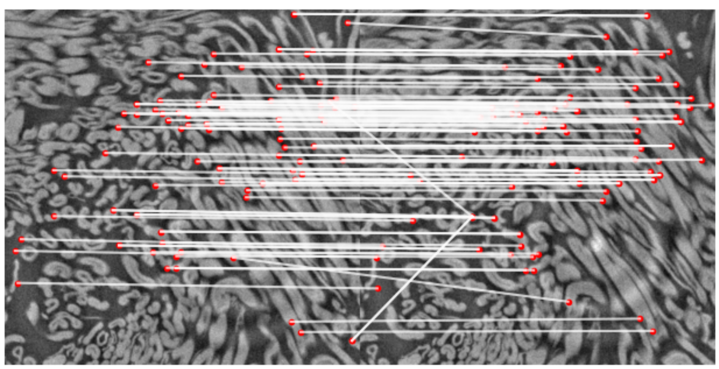
SIFT matching Below two images are matched using SIFT. The detected SIFT are given in the files `sift_a.txt` and `sift_b.txt`. Each row contains a SIFT feature, with the first two columns being the y (row) and x (column) coordinates respectively, and the rest being an L_2 normalized 128-dimensional descriptor, δ. Your task is to match each descriptor in δ_a to all descriptors in δ_b using the Euclidean distance d. A match to a descriptor in δ_a is accepted if the distance d_{closest} to the closest descriptor in δ_b divided by the second closest distance $d_{\text{second closest}}$ in δ_b is below 0.8:

$$c = \frac{d_{\text{closest}}}{d_{\text{second closest}}} ,$$

where

$$\text{Match} = \begin{cases} \text{True} & \text{if} \quad c < 0.8 \\ \text{False} & \text{else} \end{cases}$$

We can assume that there is only a translation between the two images and ignore the outliers.



What is the length of the translation?

(a) 0
(b) 3.03
(c) 3.26
(d) 4.62
(e) 5.47
(f) 9.60
(g) 10.63
(h) 11.32
(i) 12.13
(j) 52.27
(k) 92.09

Question 9

MRF prior Which statement is correct?

(a) When using Markov Random Fields to segment an image, the prior energy is expressed as one-clique potentials. Typically, it is measured as the sum of the squared difference between the pixel intensities and the mean intensity of the label of a pixel.

(b) In Markov Random Fields, the prior energy is expressed in terms of the local potentials of two-cliques. The prior typically models the smoothness of the solution, and a high weight to the prior will give a smoother MRF solution.

(c) The prior energy of an MRF is measured from the obtained configuration, i.e. the realized labeling of the image when an optimal solution is obtained. The prior energy is measured as the sum of all one- and two-clique potentials for the obtained configuration.

(d) The prior energy is a measure of the fit of an image to an optimal configuration of an MRF. Due to the spatial smoothness obtained when using an MRF, pixels will be labeled differently than the mean intensity they are closest to. The prior measures this by summing the difference between the intensities of the pixels and the mean of the label.

Question 10

Energy change Consider an MRF segmentation as in the Lecture note with

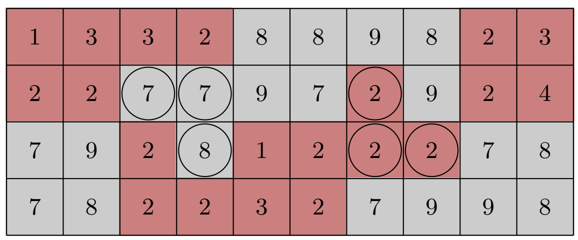
$$\mu_{\text{red}} = 2\,, \quad \mu_{\text{gray}} = 8\,,$$

and first-order neighborhood.

The smoothness parameter in the vertical direction (used, for example, between pixels with values 1, 2, 7, and 7, in the first column) was set to a larger value than the smoothness in the horizontal direction

$$\beta_{\text{vertical}} = 16, \quad \beta_{\text{horizontal}} = 12\,.$$

The configuration C_{old} is indicated with the red and the gray color in the illustration. The configuration C_{new} is obtained by flipping the label of the six pixels which are circled in the illustration.



What is the change in the posterior energy $E(C_{\text{new}}) - E(C_{\text{old}})$?

(a) -420
(b) -176
(c) -17
(d) 0
(e) 1
(f) 4
(g) 16
(h) 20
(i) 24
(j) 32
(k) 368
(l) 420

Question 11

External force A curve is evolved to segment the object in the image `flower.png`. The curve is available as `snake.txt` where every line of the file contains (r, c) coordinates of the curve points. The first curve point is indicated with a small circle.

What is the scalar component of the external force f_{ext} on the first curve point?

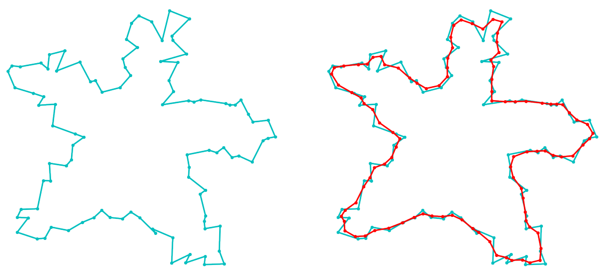
(a) -7999.4
(b) -6846.6
(c) -3799.8
(d) -2728.8
(e) -1727.6
(f) -1474.2
(g) 3799.8
(h) 5188.8
(i) 7133.4
(j) 11011.4

Question 12

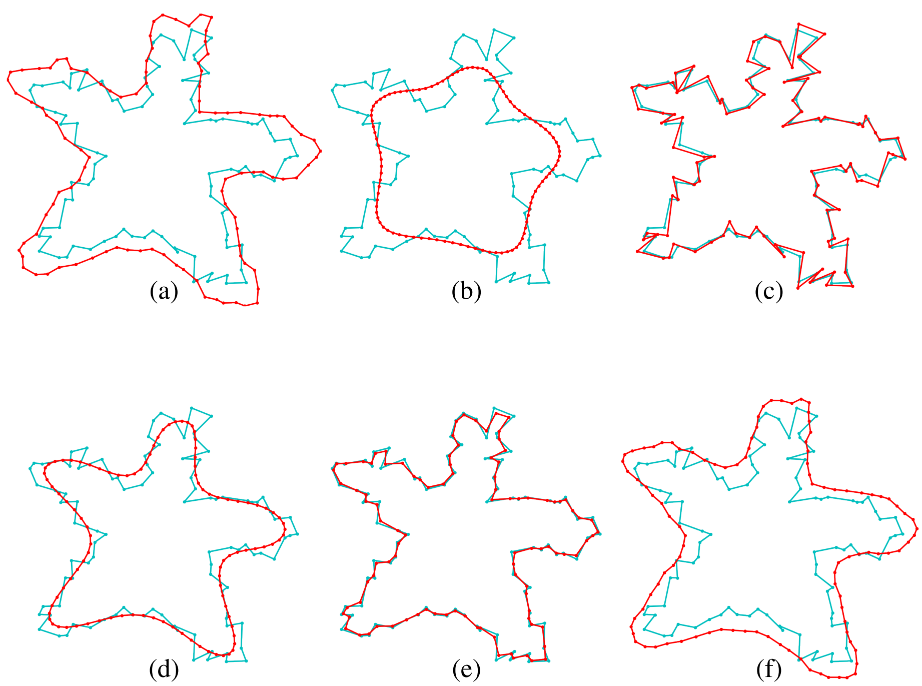
Beta effect A curve is smoothed using a backward Euler scheme as in the Lecture note Chapter 6. When the regularization parameters are set to

$$\alpha = 0.2,\ \beta = 0.2$$

you get the smoothing effect shown on the right.



Which result do you expect to obtain get when you keep α unchanged and set $\beta = 20$?

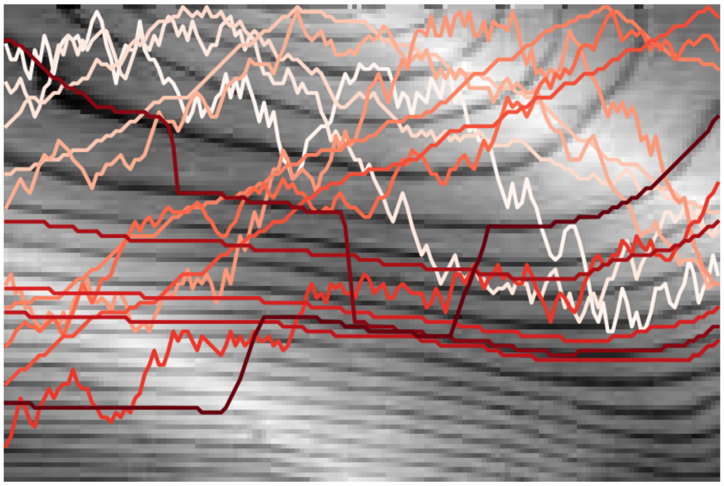


(a) (b) (c)

(d) (e) (f)

Question 13

Surface cost Consider an image `lines.png` and a set of 15 surfaces saved as `lines.txt`, visualized below.



Every line in the text file contains one surface, and the coordinates given are 0-indexed row coordinates. Surfaces are also 0-indexed so the first line contains a surface with the index 0.

We are interested in finding the surface with the lowest on-surface cost $c_{\text{on}} = I$ where I is the image. Furthermore, we are only interested in surfaces that fulfill a smoothness constraint $\Delta_x = 3$.

What is the index of the surface with the lowest on-surface cost that fulfills the smoothness constraint?

(a) 0
(b) 1
(c) 2
(d) 3
(e) 4
(f) 5
(g) 6
(h) 7
(i) 8
(j) 9
(k) 10
(l) 11
(m) 12
(n) 13
(o) 14

Question 14

Surface detection Consider an image I and a result of the optimal surface detection using an on-surface cost c_{on}.

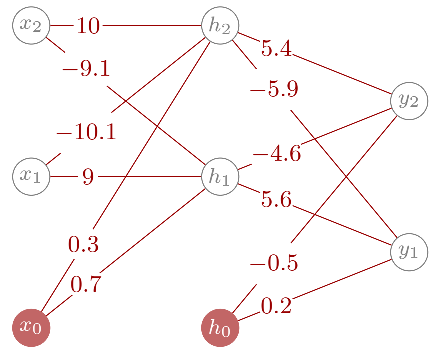
6	9	6	5	2	3	9
7	8	3	7	**4**	6	1
2	8	**5**	**6**	9	**5**	**4**
6	**7**	8	1	4	6	9
1	8	8	4	6	2	6
5	2	3	7	3	6	8

Which settings (constraint and cost) were used to obtain this result?

(a) $\Delta_x = 3$ and $c_{\text{on}} = I$
(b) $\Delta_x = 0$ and $c_{\text{on}} = 10 - I$
(c) $\Delta_x = 1$ and $c_{\text{on}} = |I - 5|$
(d) $\Delta_x = 2$ and $c_{\text{on}} = |I - 5|$
(e) $\Delta_x = 1$ and $c_{\text{on}} = |I - 6|$

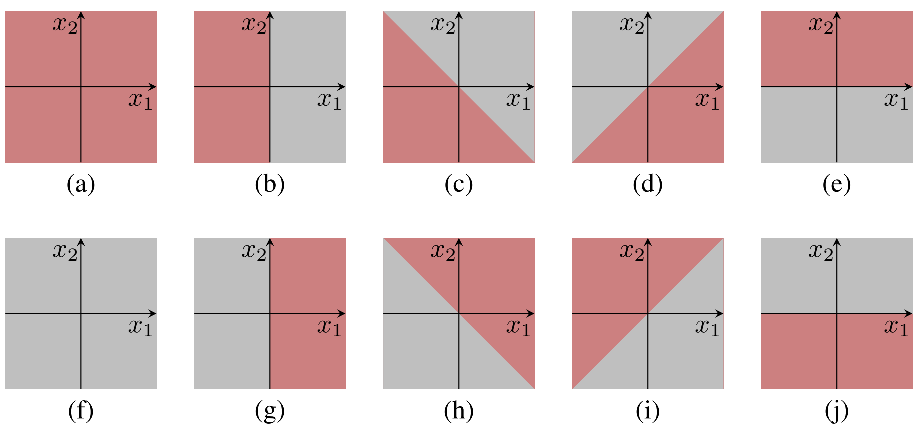
Question 15

Trained MLP Consider the classification network



with ReLU activation in the hidden layer and softmax in the last layer. Inputs are given as pairs (x_1, x_2) from $[-1, 1] \times [-1, 1]$. Each input point is assigned to either a class 1 (drawn below in gray) or a class 2 (drawn in red).

Which of the following patterns has the network been trained on?



Question 16

Partial derivatives Consider a 5-class classification network with softmax activation in the last layer and a cross entropy loss.

For a certain input, the output values are

$$\mathbf{y} = [\ 0.05, \quad 0.2, \quad 0.4, \quad 0.1, \quad 0.25\ ]$$

while we know that the target is

$$\mathbf{t} = [\ 0, \quad 1, \quad 0, \quad 0, \quad 0\ ].$$

What are the values of the partial derivatives

$$\frac{\partial L}{\partial y_i}$$

for this specific input?

(a) $[\ -0.95, \quad 0.2, \quad -0.6, \quad -0.9, \quad -0.75\ ]$
(b) $[\ -0.05, \quad 0.8, \quad -0.4, \quad -0.1, \quad -0.25\ ]$
(c) $[\ 0, \quad -5, \quad 0, \quad 0, \quad 0\ ]$
(d) $[\ 0, \quad -0.2, \quad 0, \quad 0, \quad 0\ ]$
(e) $[\ 0, \quad 0.2, \quad 0, \quad 0, \quad 0\ ]$
(f) $[\ 0, \quad 1, \quad 0, \quad 0, \quad 0\ ]$
(g) $[\ 0.05, \quad -0.8, \quad 0.4, \quad 0.1, \quad 0.25\ ]$
(h) $[\ 0.05, \quad 0, \quad 0.4, \quad 0.1, \quad 0.25\ ]$
(i) $[\ 0.1, \quad -0.6, \quad 0.8, \quad 0.2, \quad 0.5\ ]$
(j) $[\ 20, \quad 5, \quad 2.5, \quad 10, \quad 4\ ]$

Question 17

Momentum Which statement is correct?

(a) Momentum is a method that is used for accelerating the training by accumulating an exponentially decaying moving average of gradients. This is done by keeping the gradients from the previous step, then updating the gradient using a linear combination of current gradients and the kept gradients, and finally using the updated gradients in the backpropagation step.

(b) Momentum is a data augmentation method that is used for obtaining larger gradients and thereby faster convergence of neural networks. The data augmentation involves random flips, rotations, and nonlinear deformations.

(c) Momentum is used for obtaining an unbiased estimate of the gradient by taking the average gradient over a minibatch of images that are obtained by randomly permuting the training set.

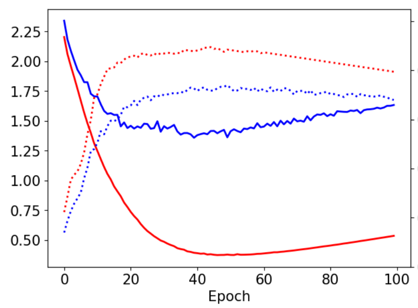
(d) Momentum is a parameter initialization strategy that ensures that a neural network will converge to an acceptable solution in an acceptable amount of time. This is done by randomly initializing the weights in the range

$$W_{i,j} \sim U\left(-\sqrt{\frac{6}{m+n}}, \sqrt{\frac{6}{m+n}}\right).$$

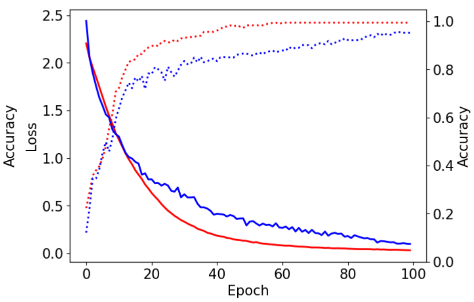
Question 18

Classification training You are training a classification network on images of hand-written digits. You have divided the available images into a training and a validation set. You plot the learning curves using the legend shown below.

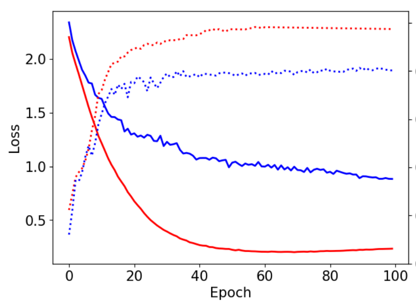
What do you expect to see?



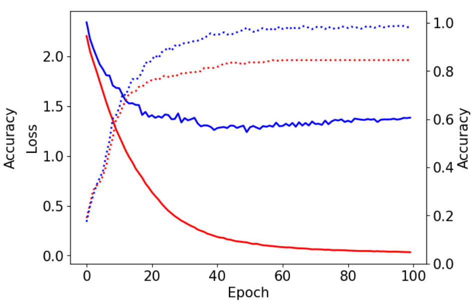
(a)



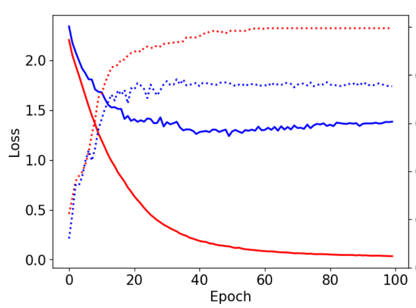
(b)



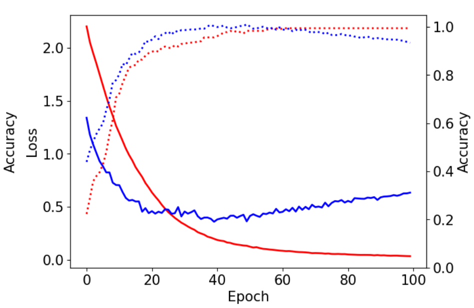
(c)



(d)



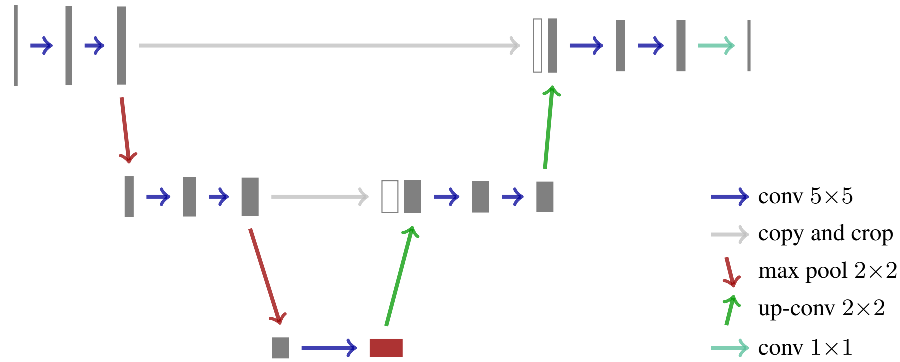
(e)



(f)

Question 19

Network bottleneck A neural network similar to a U-net is created for segmenting images. The network is outlined in the sketch.



The neural network takes in an image of size 268×352 and processes it without any padding. The bottleneck hidden layer, at the bottom of the sketch, is marked in red.

What is the spatial size of the bottleneck hidden layer?

(a) 41×60
(b) 45×66
(c) 53×74
(d) 54×76
(e) 57×78
(f) 63×82
(g) 67×88
(h) 122×164
(i) 134×176
(j) 204×288
(k) 260×344
(l) 268×352

Question 20

Data augmentation Which statement about data augmentation is correct?

(a) Data augmentation is a technique for dealing with class imbalance by which certain samples, which are considered difficult to classify, are sampled more often during training than other samples in order to force the network to learn these samples better and to prevent overfitting of the easy samples.

(b) Data augmentation is a method by which the data is sorted and organized into minibatches in a way that ensures that each minibatch contains samples that are the most distinct from each other. This ensures that the network converges optimally by seeing the widest range of samples possible during each mini-batch update.

(c) Data augmentation is a method for synthetically generating new samples by making small modifications to existing samples in the dataset in order to increase the variability and size of the dataset and reduce overfitting.

(d) Data augmentation is a technique by which the samples in the validation set are selected in a manner that more accurately models the overall data distribution. In doing so, we can obtain more accurate estimates of generalization performance when evaluating on the validation dataset.