

02502 Image Analysis Exam Spring 2023

Technical University of Denmark

Written exam, May 17, 2023

Course name: Image Analysis

Course number: 02502

Number of Questions: 25

Aids allowed: All aids allowed.

Duration: 4 hours

Weighting: All questions are equally weighted

Notes: There are five possible answers to each question and a "do not know" option. A correct answer will be equivalent to 5 points. An incorrect answer will be equivalent to -1 points. Questions unanswered (equivalent to "do not know") will not produce points. The final grade is determined by the examiners.

Appendix: Remember to submit your code (Python files, Notebooks, PDF or similar) to the "mellemregninger" part of the exam.

Abdominal CT analysis

In collaboration with a hospital, you want to investigate how the kidneys and the livers are related to the overall health status of a person. You do that by analyzing the appearance of the kidneys and the livers as seen computed tomography (CT) scans of the human abdominal area.

To test the approach, you use it on slice of a CT scan. It is a DICOM file called 1-166.dcm. You have also been given expert annotations of the liver (LiverROI.png), the left kidney (KidneyRoi_l.png) and the right kidney (KidneyRoi_r.png).

Your analysis pipeline consists of the following steps:

1. Read the DICOM file and the expert annotations
2. Extract the pixel values of the liver, the left and the right kidney using the expert annotations
3. Compute the average Hounsfield unit value in the left and the right kidney (too see if there is a functional difference)
4. Compute the average and the standard deviation of the Hounsfield units in the liver
5. Compute a threshold for the liver, t_1 , that is the average liver Hounsfield unit minus the standard deviation
6. Compute a threshold for the liver, t_2 , that is the average liver Hounsfield unit plus the standard deviation
7. Create a binary image by setting all pixels that have a value that is between t_1 and t_2 to 1 and the rest to background.
8. Dilate the binary image with a disk shaped kernel with radius=3
9. Erode the binary image with a disk shaped kernel with radius=10
10. Dilate the binary image with a disk shaped kernel with radius=10
11. Extract all BLOBs in the binary image
12. Compute the area and the perimeter of all BLOBs
13. Remove all BLOBs with an area<1500 or an area>7000 or a perimeter<300
14. Compute the DICE score between the estimated liver and the ground truth liver (LiverROI.png)

Abdominal data (<https://designer.mcq.eksamen.dtu.dk/api/images/1e4b992b-fd0a-4c0f-bf8d-ed9eb35c7739>)

What is the average Hounsfield unit in the left and the right kidney?

Choose one answer

- ☐ 115, 119
- ☐ 78, 81
- ☒ 107, 108
- ☐ Do not know
- ☐ 126, 131
- ☐ 95, 98

What are the Hounsfield unit limits used for segmentation of the liver?

Choose one answer

- ☐ [43, 110]
- ☐ [67, 98]
- ☒ [51, 100]
- ☐ [75, 89]
- ☐ [32, 78]
- ☐ Do not know

What is the DICE score between the found liver estimate and the ground truth liver?

Choose one answer

- ☐ 0.56
- ☒ 0.78
- ☐ 0.87
- ☐ 0.81
- ☐ 0.67
- ☐ Do not know

Forensic glass analysis

Sometimes glass fragments are found at crime scenes and it would be beneficial to identify the type of glass (tableware, window, vehicles, lamps). Therefore, a chemical analysis of different types of glass have been performed and it is the goal to do a statistical analysis of these components. The analysis includes: the refractive index (RI), the amount of Sodium (Na), Magnesium (Mg), Aluminum (Al), Silicon (Si), Potassium (K), Calcium (Ca), Barium (Ba) and Iron (Fe).

To remove redundancy, we start by performing a Principal Component Analysis (PCA) of the data. The measurements are scaled before computing the covariance matrix and the PCA by:

- 1) Subtract the mean from the data
- 2) Compute the minimum and maximum value of each measurement (for example the minimum and maximum Calcium content).
- 3) Divide each measurement by the difference between the maximum and the minimum value of that type of measurement (Calcium content is divided by the maximum-minimum calcium content)

The data for the question glass_data.txt

(<https://designer.mcq.eksamen.dtu.dk/api/images/57354828-bab7-48d2-be2b-ec6ac516de60>)

When the PCA has been computed, we compute how much each component explain of the total variation. How much of the total variation is explained by the first three components?

Choose one answer

- ☐ 62%
- ☐ 56%
- ☐ 83%
- ☒ 76%
- ☐ Do not know
- ☐ 69%

The covariance matrix is computed before doing the PCA. What is the value of the diagonal matrix at position (0, 0) ?

Choose one answer

- ☐ 0.412
- ☐ Do not know
- ☒ 0.018
- ☐ -0.031
- ☐ 1.231
- ☐ -0.102

What is the first value of the Sodium (Na) measurements after subtraction of the mean and scaling?

Choose one answer

- ☒ 0.03
- ☐ -0.01
- ☐ Do not know
- ☐ 0.23
- ☐ 0.005
- ☐ -0.73

After the PCA analysis, the data is projected onto the principal components. The absolute value is then computed off all the projected values and the maximum absolute value is found. What is this value?

Choose one answer

- ☒ 1.12
- ☐ 0.19
- ☐ Do not know
- ☐ 2.98
- ☐ 4.43
- ☐ 2.34

Advanced 3D image registration

You are using SimpleITK to register one image to another image.

First you try to apply the following steps to the moving image:

1. Apply a 30 degrees roll.
2. Translate 10 units in the X direction.
3. Apply a 10 degrees yaw.

You combine these transformations into an affine transformation matrix.
What is this matrix?

Choose one answer

☐
$$\begin{pmatrix} 0.87 & -0.5 & 0 & 10 \\ 0.42 & 0.73 & -0.71 & 0 \\ 0.42 & 0.73 & 0.71 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

☐
$$\begin{pmatrix} 0.53 & -0.81 & 0.25 & 0 \\ 0.81 & 0.40 & -0.43 & 0 \\ 0.25 & 0.43 & 0.87 & 10 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

☒
$$\begin{pmatrix} 0.85 & -0.17 & 0.49 & 9.84 \\ 0.15 & 0.98 & 0.09 & 1.74 \\ -0.5 & 0 & 0.87 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

☐ Do not know

☐
$$\begin{pmatrix} 0.87 & -0.5 & 0 & 10 \\ 0.35 & 0.61 & -0.71 & 0 \\ 0.35 & 0.61 & -0.71 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

☐
$$\begin{pmatrix} 0.50 & -0.87 & 0 & 8.67 \\ 0.87 & 0.50 & 0 & 5 \\ 0 & 0 & 1 & 10 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

Product names on boxes

A customer has asked you to help them to design a system that can recognize product brands on cardboard boxes. Your first system does:

1. Converts the input image from RGB to HSV
2. Creates a new image that only contains the H component of the HSV image
3. Creates a binary image by setting all pixels with an H value of $0.3 < H < 0.7$ to 1 and the rest of the pixels to 0
4. Performs a morphological dilation with a disk shaped structuring element with radius=8 on the binary image.

cardbox photo (<https://designer.mcq.eksamen.dtu.dk/api/images/20cf5ac7-f4cb-4e07-ba8e-c023971c82b9>)

What is the number of foreground pixels (value = 1) in the resulting image?

Choose one answer

- ☐ Between 30000 and 40000
- ☐ Between 60000 and 70000
- ☒ Between 50000 and 60000
- ☐ Do not know
- ☐ Between 40000 and 50000
- ☐ Between 70000 and 80000

Shoe comparison

A sports shoes company wants a new system to compare existing designs with their new models. They have provided you with two photos of a new (shoe_2) and an old shoe (shoe_1). Your initial idea is to:

1. Do a landmark based registration of shoe_1.png (the source) to shoe_2.png (the destination) using a similarity transform.
2. Extract the found scale of the transform (using for example **tform.scale**).
3. Compare the values of the alignment error, F (sum of squared distances), before and after the registration.
4. Compare the value of the blue component of the colors values of the aligned images.

The manually placed landmarks on shoe_1.png are ([40, 320], [425, 120], [740, 330]) and on shoe_2.png ([80, 320], [380, 155], [670, 300]).

Shoe photos (<https://designer.mcq.eksamen.dtu.dk/api/images/da19129b-d107-442c-99b6-764846c90285>)

After the source photo has been transformed (using `tform.inverse`) both the transformed image and the destination image are converted to bytes using `img_as_ubyte`. Finally, the blue component of both images are extracted at position (200, 200). What is the absolute difference between these values?

Choose one answer

- ☐ 16
- ☐ 5
- ☐ 12
- ☐ 21
- ☒ 26
- ☐ Do not know

What is the scale of the similarity transform after it has been estimated?

Choose one answer

- ☒ 0.83
- ☐ 0.78
- ☐ 0.92
- ☐ 0.67
- ☐ 0.57
- ☐ Do not know

What is the change of the landmark alignment error, F , when comparing it before and after the transformation?

Choose one answer

- ☐ 15672
- ☐ 12131
- ☐ 14511
- ☒ 10277
- ☐ 8523
- ☐ Do not know

Video change detection

You have developed a new video change detection system. To test if it works, you have acquired two images. One is the background (background.png) and the second is a new frame (new_frame.png). They are both RGB images of size (600, 450).

Your algorithm works like this:

1. Converts both images to gray scale using **color.rgb2gray**. Both images are now floating point images where the pixel values are between 0 and 1.
2. Updates the background image by $\text{new_background} = \alpha * \text{background} + (1 - \alpha) * \text{new_frame}$. Alpha = 0.90.
3. Computes the absolute difference image between the new frame and the new background.
4. Computes how many pixels in the difference image that have a value above 0.1. These are the *changed pixels*.

The photos used for change detection

(<https://designer.mcq.eksamen.dtu.dk/api/images/7b0fb781-bbd9-4960-90c3-03bc49216c20>)

How many pixels are detected as changed?

Choose one answer

- ☐ Between 20000 and 25000
- ☐ Between 5000 and 10000
- ☐ Between 10000 and 15000
- ☒ Between 25000 and 30000
- ☐ Between 15000 and 20000
- ☐ Do not know

What is the average value of the estimated new background image in the pixel region [150:200, 150:200] ?

Choose one answer

- ☐ 0.50
- ☐ 0.20
- ☒ 0.60
- ☐ 0.30
- ☐ 0.40
- ☐ Do not know

Character recognition

You would like to sort all your books and magazines by their title and you start by making a system that can recognize letters.

Photo of book covers (<https://designer.mcq.eksamen.dtu.dk/api/images/92056198-6823-4803-9f42-6383935cfa71>)

To try to find the red letters, we first extract the R, G, B color channels from the image. Secondly we create a new binary image from the RGB image by setting all pixel with $R > 100$ and $G < 100$ and $B < 100$ to 1 and the remaining pixels to 0.

The binary image is eroded using a disk shaped structuring element with radius=3. How many foreground pixels are there in the eroded image?

Choose one answer

- ☐ Between 40000 and 50000
- ☐ Between 0 and 10000
- ☐ Between 10000 and 20000
- ☐ Between 20000 and 30000
- ☒ Between 30000 and 40000
- ☐ Do not know

You would like to pre-process the image before the analysis and you try with the following approach:

1. Convert the input photo from RGB to gray scale
2. Apply a median filter to the gray scale image with a square footprint of size 8

What is the value at the pixel at (100, 100) in the resulting image?

Choose one answer

- ☐ Do not know
- ☐ 0.81
- ☐ 0.67
- ☒ 0.45
- ☐ 0.23
- ☐ 0.17

To try to find the letters, we perform the following operation:

1. Compute a binary image where the pixels that has an RGB value with $R > 100$ and $G < 100$ and $B < 100$ are set to 1 and the rest of the pixels to 0
2. Erode the binary image using a disk-shaped structuring element with radius=3
3. Compute all the BLOBs in the image
4. Compute the area and perimeter of all found BLOBs
5. Remove all BLOBs with an area < 1000 or an area > 4000 or a perimeter < 300

What letters are left in the photo after these operations?

Choose one answer

- ☐ TE
- ☐ LU
- ☒ AB
- ☐ Do not know
- ☐ AS
- ☐ SOL

Video surveillance system

You are developing a new surveillance system that consists of a camera connected to a computer using a USB-2 connection. The images are RGB (8-bits per channel) images with a size of 1600 x 800 pixels. When you tested your system, you did some timings of your system. You found out that your image analysis algorithm takes 230 milliseconds to process one frame. You also found out that the camera can successfully send 6.25 images per second to the computer (this is an average value measured over a minute).

What is the overall system framerate of your system (the maximum frame-per-second)?

Choose one answer

- ☐ 7.45
- ☐ 6.25
- ☐ Do not know
- ☐ 5.47
- ☒ 4.30
- ☐ 2.71

How much data can the USB-2 connection transfer from the camera to the computer on the system you tested?

Choose one answer

- ☐ 22 megabytes per second
- ☐ 20 megabytes per second
- ☐ 18 megabytes per second
- ☐ 16 megabytes per second
- ☐ Do not know
- ☒ 24 megabytes per second

PizzAI

A pizza startup, PizzAI, wants to use machine learning to match customer taste to the optimal pizza. They have provided you with ten training photos of their current selection of pizzas. You start by using image based principal component analysis to make a statistical model of the visual appearance of the pizzas in the training set.

You start by computing the average pizza and then you use PCA from `sklearn.decomposition` to compute 5 principal components of the 10 images.

PizzAI photos (<https://designer.mcq.eksamen.dtu.dk/api/images/98a53a89-57d1-4a6e-a8f1-4213432afebf>)

One of your friends is an experimental eater and wants to taste the pizza that is visually as far away from the average pizza as possible. You compare all pizzas with the average pizza and select the one which has the largest sum of squared differences compared to the average pizza. Which pizza do you serve for your friend?

Choose one answer

- ☒ BigSausage
- ☐ CucumberParty
- ☐ BewareOfOnions
- ☐ Do not know
- ☐ TheBush
- ☐ PaleOne

The company has asked you to compute a measure of the variation on their menu. After doing the PCA, you compute how much the first principal component explains of the total variation. It is:

Choose one answer

- ☒ 23%
- ☐ 38%
- ☐ 42%
- ☐ 57%
- ☐ 68%
- ☐ Do not know

The company has asked you to define their signature pizzas. The ones that are most varied. After computing the PCA, you project all pizzas on to the PCA space. You find the two pizzas that are the furthest away on the first principal axes. One in the positive and one in the negative direction. What pizzas do you suggest to be signature pizzas?

Choose one answer

- ☐ TheBush and WhiteSnail
- ☐ Do not know
- ☒ FindTheOlives and BigSausage
- ☐ PaleOne and SnowAndGrass
- ☐ Leafy and GreenHam
- ☐ CucumberParty and WhiteSnail

An international student at DTU, is missing his favorite pizza. The student has asked his family to send him a photo (super_pizza.png) of this amazing pizza. Which pizza on the PizzAI menu looks most similar to this pizza. You find the solution by projecting the photo of the wanted pizza on to PCA space and finding the closest menu pizza in PCA space. Which pizza is that?

Choose one answer

- ☐ PaleOne
- ☐ FindTheOlives
- ☒ GreenHam
- ☐ TheBush
- ☐ BewareOfOnions
- ☐ Do not know

Concert light design

One of your friends is a concert light designer. She thinks it could be great to have an automatic system to get feedback on the overall light distribution on the stage. It would also be nice if the system can rotate the image to adjust for the camera position. Your first attempt do the following:

1. Rotates the image 11 degrees with a rotation center of (40, 40).
2. Transform the image from RGB to gray scale.
3. Computes an automatic threshold using Otsu's method.
4. Computes the percentage of foreground pixels.

You have been given a single photo to test your algorithm.

Concert lights (<https://designer.mcq.eksamen.dtu.dk/api/images/a5a37299-1ea6-47a2-b0ba-61c8eeff54f5>)

What is the percentage of foreground pixels in the final image+

Choose one answer

- ☐ 57
- ☐ 37
- ☒ 26
- ☐ 49
- ☐ 19
- ☐ Do not know

What is the threshold found by Otsu's method?

Choose one answer

- ☐ 0.17
- ☐ 0.29
- ☐ Do not know
- ☐ 0.73
- ☒ 0.42
- ☐ 0.57