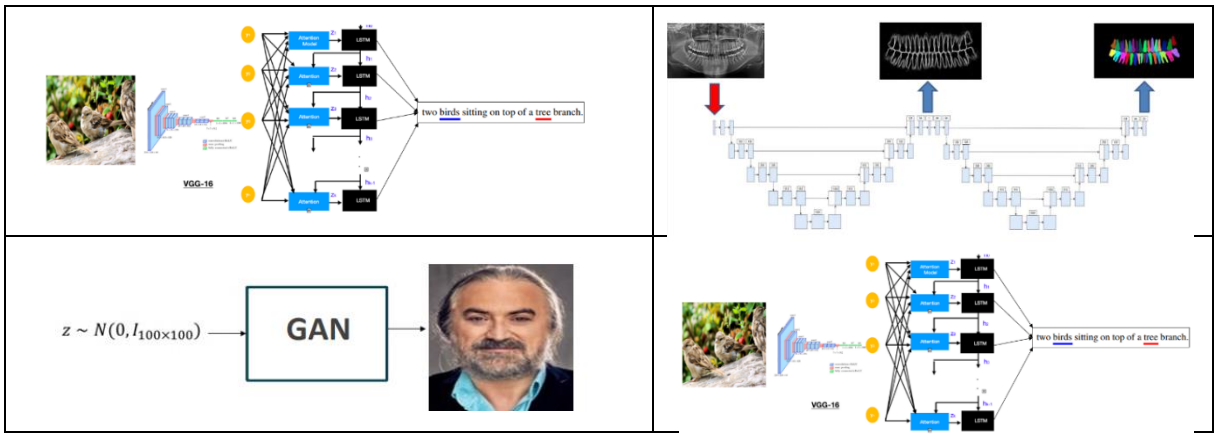
Q1-) Describe the computer vision tasks from the given images.



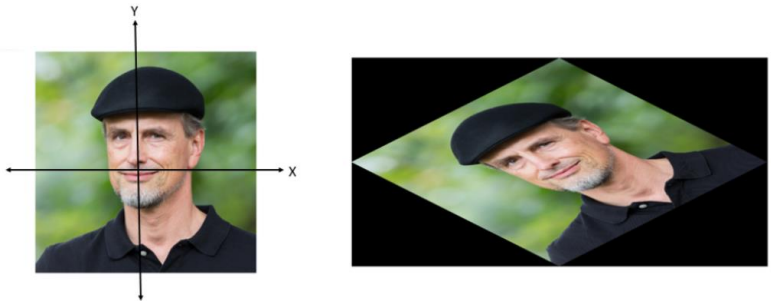
Answer:

Captioning	Segmentation
Generation	Classification

Q2-) For the given image of Jürgen Schmidhuber, match the Hue, Saturation and Value channels.

Answer: Value, Saturation, Hue.

Q3-) For the image on the left, X-Y axis and rotation center, what is the correct transformation matrix to obtain the second image?



Answer: The image is first scaled in X dimension by two. Then, it is rotated by 45 degrees counterclockwise. Thus,

$$\begin{bmatrix} \cos\left(\frac{\pi}{2}\right) & -\sin\left(\frac{\pi}{2}\right) & 0 \\ \sin\left(\frac{\pi}{2}\right) & \cos\left(\frac{\pi}{2}\right) & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 2 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 2\cos\left(\frac{\pi}{2}\right) & -\sin\left(\frac{\pi}{2}\right) & 0 \\ 2\sin\left(\frac{\pi}{2}\right) & \cos\left(\frac{\pi}{2}\right) & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

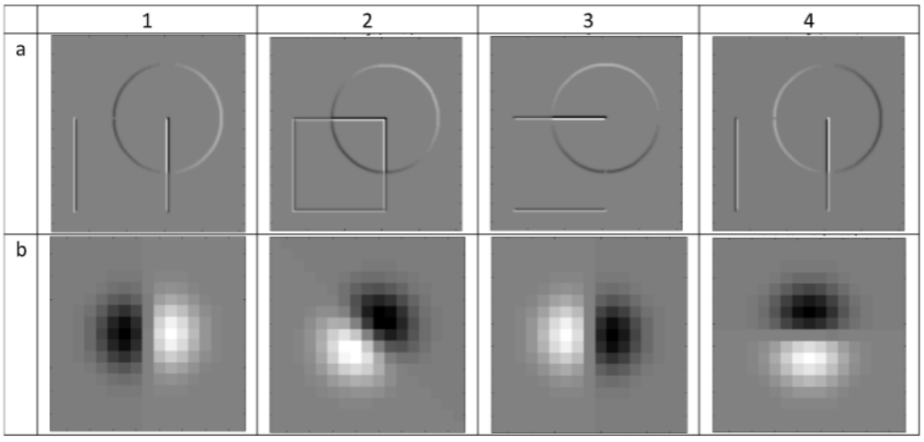
However, other answers containing scaling and a 45 degrees of rotation are also acceptable.

Q4-) What is the effect of convolving the following 5x5 filter to an image?

$$\begin{bmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

Answer: The image is shifted to the left by two pixels. However all left/right shift answers are accepted.

Q5-1) For the example image, some image filters and their responses (cross correlation) are given below in changed order. Match each filter with its response.



Answer: a1-b3, a2-b2, a3-b4, a4-b1

Q5-2) For a patient, a brain slice is obtained by an MR Scanner. An automatic tumor segmentation algorithm is tried to be employed to detect the tumor, however our algorithm did not output meaningful results because of the noises given in the image. Take a closer look at the picture. What kind of noise is this?

Answer: Salt & Pepper noise.

Q5-3) What kind of filter should we use to reduce the noise?

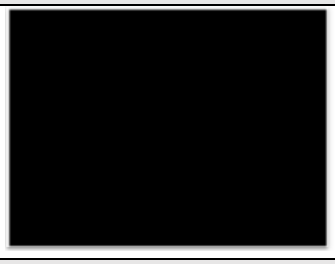
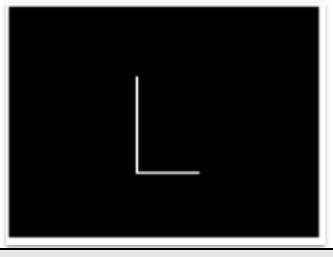
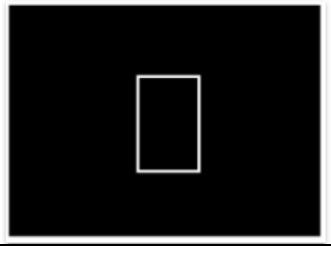
Answer: Median filter.

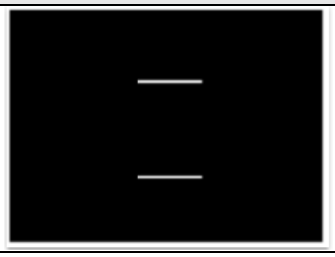
Q6-1) Find the result of the following operation.

$$* \begin{bmatrix} -1 & -2 & -1 \\ 0 & 0 & 0 \\ 1 & 2 & 1 \end{bmatrix} = I_1$$

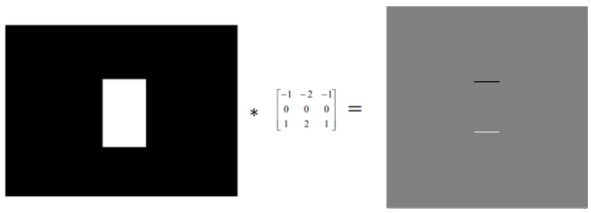
$$* \begin{bmatrix} -1 & 0 & 1 \\ -2 & 0 & 2 \\ -1 & 0 & 1 \end{bmatrix} = I_2$$

$abs(I_1) + abs(I_2) = ?$

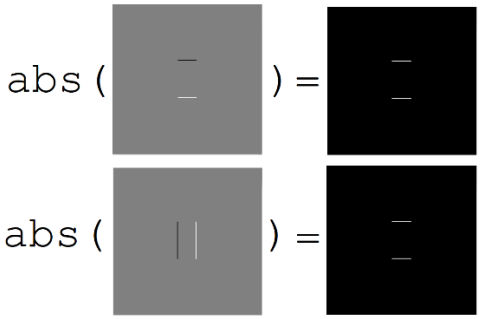
I	II
	
III	**IV**
	

V	VI
	

Answer: The process is given below.

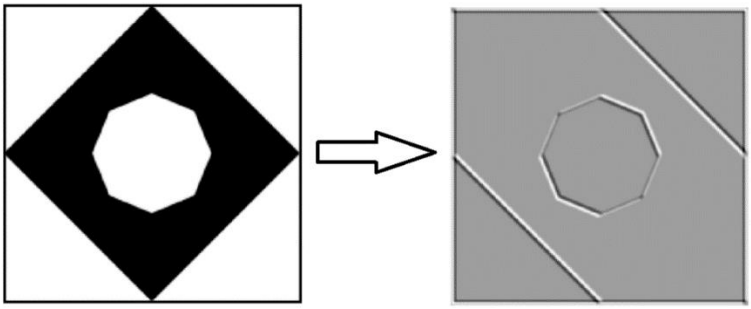


Here the black lines represent negative values and White lines represent positive values. Then finding the abstract images, we can obtain double lines.

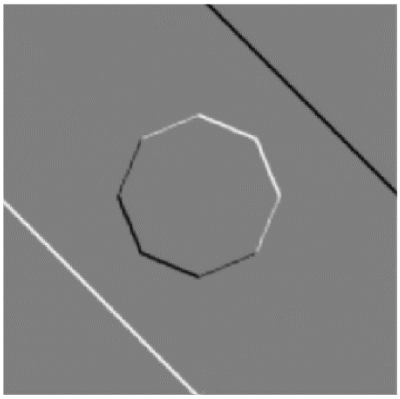


Summing these twow images, we can obtain III.

Q6-2) Using the Robinson Compass masks, how could we obtain the second picture from the first picture?



Answer: We should use the SW mask twice. After the first iteration, the result is as given below. As in 6.1, White line is positive and black line is negative.



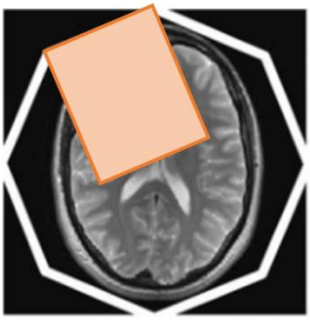
In the second iteration, according to negative-positive and positive-negative crossings, new lines will appear.

Also, using the NE filter twice outputs the same result.

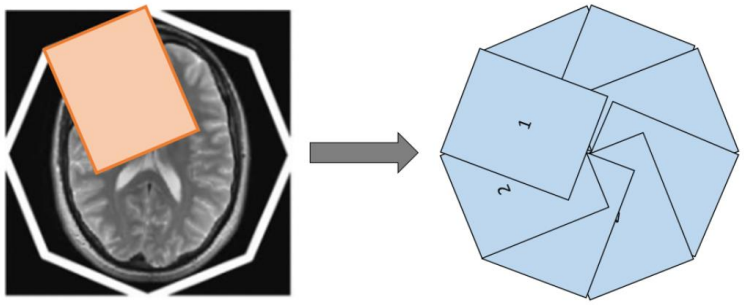
Q6.3) Suppose that, we are using a normalized image filter (sum of the kernel values is 1) on an image. The image has float pixel values which are in the interval (0,1). What is the interval for output pixel values?

Answer: The question has incomplete information of having only positive values. If so the answer is (0,1). However, all answers are accepted.

Q7.1) In many MR Scanners, it desired to scan the brain with multiple scanners placed around the patient's head (Parallel MR Imaging) for efficiency. Each of these scanners capture a local area, but there are partial intersections between these areas. An example of 8 scanners and an area scanned by one of the scanners is given in the figure. Imagine a case, where we have 8 output images from each scanner and do not know which scanner they belong to.

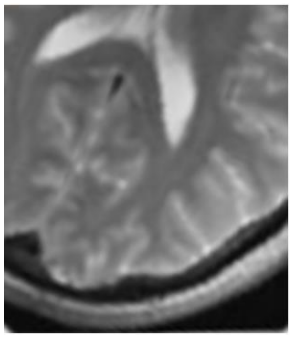


How to find the images from the neighbor scanners to obtain the full scan? (e.g. Plane 1 and 2)



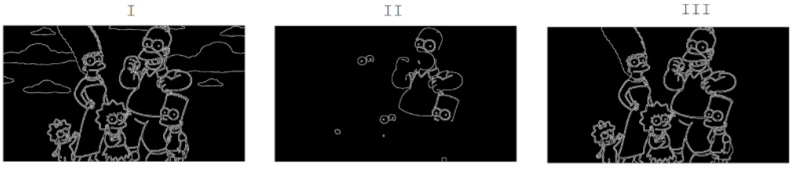
Answer: Template matching

Q7.2) Suppose that after obtaining the full image, we lost the information of where one of the planes is placed. How can we find the exact place of the following plane?



Answer: Template matching

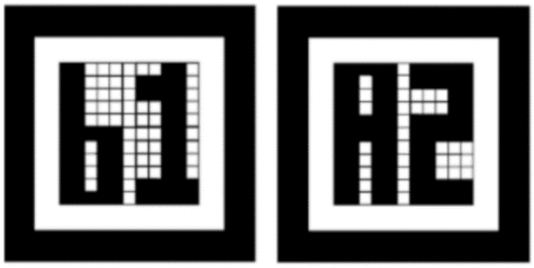
Q7.3) For the given image, Canny edge detection algorithm is used with different threshold values. Compare the threshold values.



Answer:

Highest threshold: II
Mid-Level threshold: III
Lowest threshold: I

Q8) Suppose that, for a computer vision conference, the organizers are planning to develop an informative mobile application. Using the application, the participants will photograph the markers placed on each door to obtain information about the current presentation. The conference center has two blocks: A and B. Each block has 8 rooms. The markers, containing room and block IDs, are designed as an 11x11 matrix area in a bold square as given in the figure.



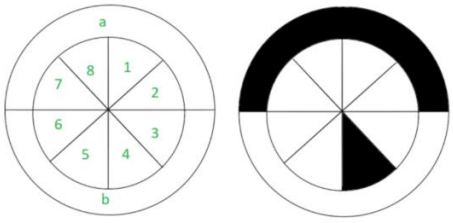
Q8-1) Suppose that the participants photographed the markers with perfect orientation as given in the figure below. Describe the steps to find the room ID from the photo in the application. (Use maximum 200 characters.)

Answer: There are multiple solutions. We can detect the edges inside the image, find a square-shaped structure and check each cell of the marker.

Q8-2) Suppose that, some of the older researchers are not very successful to photograph the markers in a good orientation. An example photo is given below. How should you change your algorithm in the previous part to catch markers in these kind of photographs? (Use maximum 200 characters.)

Answer: There are multiple solutions. After detecting the lines which represents a polygon, a perspective transform may be applied to transfer it into the square form. Then, the same procedure applies.

Q8-3) The organizers decided to change the marker with an alternative having two nested circles. The outer circle gives the block and inner circle gives the room. An example marker for room A4 is given below. Write the steps to find the room ID from this kind of markers.

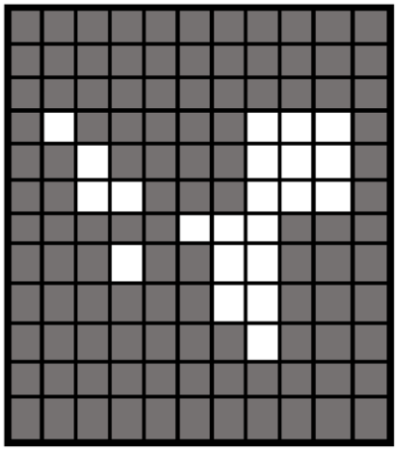


Answer: There are multiple solutions. Using Hough circle detection, two circles are found. Then, the pixel values for each part are controlled.

Q8-4) What is the main problem of the circular markers?

Answer: There are multiple answers. The markers can be upside down. Hough circle detection will not work if the camera is not perfectly oriented etc.

Q9) Using a 3x3 filter [1 1 1; 1 1 1; 1 1 1], if we want to apply some dilation and erosion operations to the following image. Each step of our operations are applied one after another.



(Step 1 - 20 pts) Apply a dilation operation. How many white cells are obtained?

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

68 cells

(Step 2 - 20 pts) Apply an erosion operation afterwards. How many white cells are obtained?

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

26 cells

(Step 3 - 20 pts) Apply another erosion operation afterwards. How many white cells are obtained?

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 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 0 0 0 0 0 0
0 0 0 0 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 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 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 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 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 0 0 0 0 0

1 cells

(Step 4 - 40 pts) How many connected components remains after the last step?

Only one connected component

Q10) Early in the morning...

In this quiz, we will work on world-famous PC game: Sea of Thieves. In the game, the player, as a pirate, is in search of a buried treasure chest.

Compass

Part 1) Suppose that, the pirate landed on an island and he can check his compass to find his direction. In an AI approach, first thing we should do is to detect the compass circle. Assuming the compass is the only circular object in the screen, which approach can be used to detect the compass?

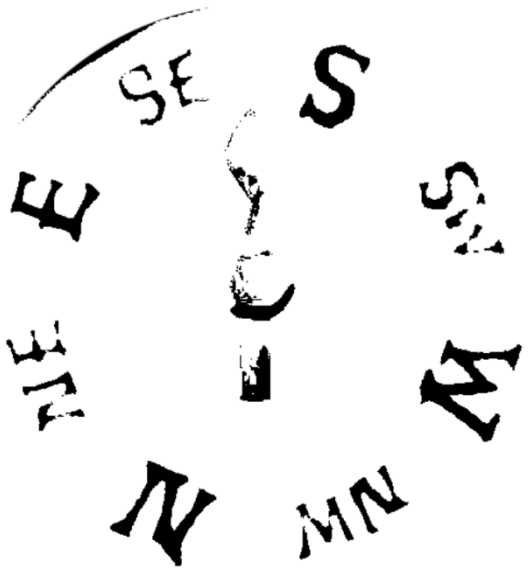
Hough Circle detection

The Compass, alone

Part 2) Now we have a perfectly cut view of the compass. However, we do not have a good sight of letters for four main directions because of the background. What should we do?

Thresholding

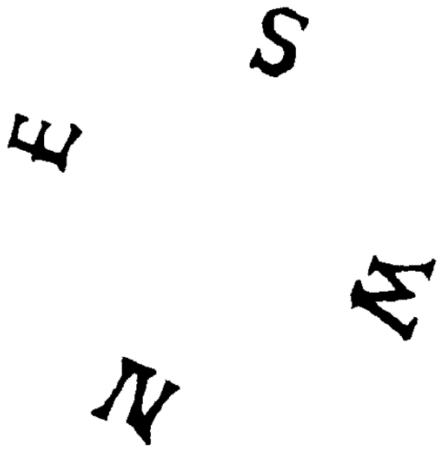
The compass after the operation



Part 3) Good job cabin-boy! Now we got rid of the background. So let's try to obtain only the four main directions. The easiest way is to find the 4 largest letters. What can we do?

Selecting 4 largest connected component, or doing morphological operations.

"I only watch the NEWS"



Part 4) Here we have the main directions. But still we did not decide on our currect direction. To do this, we should analyze the letters. Previously, I obtained many images containing N, S, W, E images and maybe I can try to find a similarity between them. What can I use in this step?

Template matching

Part 5) "Instead of all these tricks, it may be easier if we got a Convolutional Neural Network, trained to make _____"

Object detection/classification etc.

After finding the treasure chest, the pirate boarded on his ship. Now, he should sail to another land!

Q11) (50 pts) What operation should be applied to the video at https://youtu.be/DSONH8MB64g, to obtain the video at https://youtu.be/UyH3a0SyDlc ?

Frame differencing

(50 pts) What operation should be applied to obtain the video at https://www.youtube.com/watch?v=PUZajW-Jq8s ?

Motion estimation/detection. Both could be used.

Q12) For the given video at https://youtu.be/8MHwcXHc2pw compare the reliability of the following three points when calculating motion vectors.

Most reliable point: A

Least reliable point: C