Calculators may be used in this examination provided they are not capable of being used to store alphabetical information other than hexadecimal numbers

UNIVERSITY OF BIRMINGHAM

School of Computer Science

Robot Vision

Main Summer Examinations 2024

Time allowed: 2 hours

[Answer all questions]

Turn Over

Note

Answer ALL questions. Each question will be marked out of 20. The paper will be marked out of 80, which will be rescaled to a mark out of 100.

Question 1

Satellite images can be used for monitoring river banks, which can provide early warning for natural hazards, such as riverbank erosion, flooding, and land instability. One such image is given in Figure 1. These images can be collected over time and processed to detect changes in the river bed.



Figure 1: Satellite image of a river

(a) Close visual inspection reveals that a some random noise is present in these images. Propose a noise removal technique, mentioning the size of the kernel. **[2 marks]**

(b) Provide a clear justification for your selection, including the size of the kernel. **[2 marks]**

(c) In this application, accurate localisation of continuous boundary of the river is critical. With this in mind, which edge detection algorithm would you propose? **[2 marks]**

(d) Briefly explain using illustrations, the two (02) steps in the algorithm proposed above, which ensures both accurate localisation and getting a continuous boundary of the river. **[6 marks]**

(e) For registration of multiple similar images collected over time, a landmark based approach was proposed. Name two (02) factors the user needs to consider when selecting control points (landmarks) for obtaining an accurate registration? **[2 marks]**

(f) Sum of Squared Differences (SSD) was proposed for the similarity metric. In your opinion, is this a feasible selection? If yes, provide justification. If no, provide an alternative for the proposed metric. **[2 marks]**

(g) Propose a suitable algorithm to segment out the river from the image shown in Figure 1. Provide justification for your selection. **[2 marks]**

(h) At times, you need to use morphological operators to post process the segmentation outcome. Name two (02) such operators outlining their purpose. **[2 marks]**

Question 2

To perform classification of RGB images containing different categories of objects, a convolutional neural network has been chosen. Its description follows the convention:

- CONV-K-N-p-s stands for a convolutional layer with N filters (kernels), each of them of the size $K \times K$, p denotes the padding, and s the stride.
- POOL-K-s denotes a $K \times K$ pooling layer with a stride s and no padding.

(a) For each layer determine the dimension of the activation maps (please write your answer as a multiplication, e.g. $128 \times 128 \times 128$), the number of weights and the number of biases.

Layer	Activation map dimension	Number of weights	Number of biases
INPUT	$224 \times 224 \times 3$	N/A	N/A
CONV-3-32-1-1			
POOL-2-2			
CONV-3-64-1-1			
POOL-2-2			
CONV-3-128			
POOL-2-2			

[12 marks]

(b) The network is intended to classify four different classes. Specify the final layer L that has to follow the last pooling layer. For the designed layer L specify the activation map dimensions, the number of weights and the number of biases. **[5 marks]**

(c) Which activation function must be applied to the layer L? Given a vector $[0.4, 0.4, 0.4, 0.4]$, what is the result of using that activation function?

[3 marks]

Question 3

(a) Explain single object detection in the context of deep learning for computer vision. **[2 marks]**

(b) Give the dimension of the network output for localising a single object in a given image. Describe the meaning of each of the dimensions. **[3 marks]**

(c) R-CNN, Fast-RCNN, and Faster-RCNN are three deep learning-based object detection methods. Given the following properties (1-3), please connect them (by filling the numbers in the table below) to the correct method that has the corresponding unique property different from others (i.e. the property that was introduced by this method).

1. Region proposal from the convolutional feature map using selective search, and RoI pooling.
2. Region proposal network, followed by RoI pooling.
3. Region proposals extraction on the image, and selective search.

Methods	Properties
R-CNN	
Fast-RCNN	
Faster-RCNN	

Then, please describe what are the RoI pooling, region proposal network, and selective search.

[9 marks]

(d) Please explain what YOLO is, and its main differences (at least two differences) from the above three object detection methods. **[3 marks]**

(e) Please explain what Mask-RCNN is, and its key difference from the above RCNN series. **[3 marks]**

Question 4

(a) **Robot Ballet Choreography**: Imagine you are programming a team of robots equipped with cameras to perform a synchronised ballet routine. Each robot has a different viewpoint of the stage, and your task is to develop a multiple view geometry algorithm to coordinate their movements seamlessly.

(i) Identify and elaborate on the key challenges (at least two) in achieving synchronisation and coordination using multiple view geometry in the context of the robot ballet. **[4 marks]**

(ii) Propose a comprehensive algorithmic framework to represent the spatial relationships between the robots based on their different viewpoints. Consider factors such as camera calibration, perspective projection, and scene geometry. **[4 marks]**

(iii) Discuss the potential sources of errors in your proposed model and suggest strategies to mitigate or minimise these errors. **[4 marks]**

(iv) Explore potential TWO more real-world applications beyond robot ballet where your multiple view geometry algorithm could be applied, discussing both opportunities and challenges. **[4 marks]**

(b) Consider a mobile robot equipped with two identical pinhole model cameras with the focal length of 5 cm, both with optical axes lying on a horizontal plane and parallel to each other. The baseline distance between the cameras is 12 cm. Initially, a point P in space has a disparity of 24 mm, which increases to 40 mm when the robot moves closer to point P. Determine how much the robot moves forward and provide your calculations. **[4 marks]**

End of Paper

Do not complete the attendance slip, fill in the front of the answer book or turn over the question paper until you are told to do so

Important Reminders

- Coats/outwear should be placed in the designated area.
- Unauthorised materials (e.g. notes or Tippex) must be placed in the designated area.
- Check that you do not have any unauthorised materials with you (e.g. in your pockets, pencil case).
- Mobile phones and smart watches **must** be switched off and placed in the designated area or under your desk. They must not be left on your person or in your pockets.
- You are not permitted to use a mobile phone as a clock. If you have difficulty seeing a clock, please alert an Invigilator.
- You are not permitted to have writing on your hand, arm or other body part.
- Check that you do not have writing on your hand, arm or other body part – if you do, you must inform an Invigilator immediately
- Alert an Invigilator immediately if you find any unauthorised item upon you during the examination.

Any students found with non-permitted items upon their person during the examination, or who fail to comply with Examination rules may be subject to Student Conduct procedures.