College of Science - Computer Science

CSCM77 Computer Vision and Deep Learning

Release Time: **9.20am Wednesday 20/05/2020 (Time Zone: BST)**

Deadline: **5.20pm Wednesday 20/05/2020 (Time Zone: BST)**

Alternative Assessment Information

- *This is an open-book test. This means you may consult you notes, textbooks, and other resources, including calculators, as you see fit.*
- *This assessment is designed to take 2 hours to complete (maybe a little longer to account for typing speed). The deadline has been set to give you a longer window than necessary to allow you time to deal with technical issues, provide some flexibility of starting times, and to help students with disability access plans that require rest breaks and extra time.*
- *It is suggested that you use Microsoft Word (or any other editor of your choice) to type your answers, then save as PDF when you are ready to submit. All submitted text must be word-processed, but you may include images (or photos of hand drawn images) as part of the document.*
- *This is an individual assessment. Under no circumstances are you to discuss any aspect of this assessment with anyone; nor are you allowed to share this document, ideas or solutions with others using email, social media, instant messaging, websites, or any other means. Your attempts at these questions must be entirely your own work. Those found to have collaborated with others will receive a mark of 0.*
- ***Attempt all questions**.*

Submission Instructions

- Please submit a single PDF file named as your student number (e.g. 123456.pdf) via the submission link located on the module page in Blackboard.

By submitting, electronically and/or hardcopy, you state that you fully understand and are complying with the university's policy on Academic Integrity and Academic Misconduct. The policy can be found at https://myuni.swansea.ac.uk/academic-life/academic-misconduct.

Originator: ***Prof. X. Xie***

In case of queries email <u>both</u>: x.xie@swansea.ac.uk and j.deng@swansea.ac.uk

CSCM77 Computer Vision and Deep Learning (May/June 2020)

(answer ALL questions)

Question 1.

(a) The concept of randomisation is fundamental in the Random Forests method for supervised learning. Explain two randomisation techniques used in Random Forests in order to cope with over-fitting.

[2 marks]

(b) Explain how Microsoft Kinect uses structured light to estimate depth from camera.

[2 marks]

(c) Camera Model: Identify the intrinsic and extrinsic parameters for a single camera.

[2 marks]

(d) Stereo:

i. Describe the Essential Matrix and the Fundamental Matrix of a stereo system (no need to write down the formulae).

[2 marks]

ii. Briefly describe the use of triangulation 3D reconstruction. Draw a diagram to show the triangulation.

[3 marks]

(e) Motion analysis in many ways is considered similar to the stereo vision problem in that it often involves computing reconstruction after establishing correspondences from two or more views of a scene. Discuss the important differences between motion analysis and reconstruction using stereo vision.

[2 marks]

Question continues on next page.

(f) Tracking

i. What are the key issues in designing a motion tracking algorithm?

[2 marks]

ii. The following equations describe a trajectory model used in motion tracking

$$x_k = x_{k-1} + v_{k-1}\Delta t + a_{k-1}(\Delta t)^2/2$$
$$v_k = v_{k-1} + a_{k-1}\Delta t$$
$$a_k = a_{k-1}$$

where x denotes position, v is velocity, a denotes acceleration, Δt represents time interval, and subscripts k and $k-1$ denote adjacent image frames.

Identify which trajectory model is applied in this motion tracking.

[2 marks]

(g) The following equation defines a convolutional kernel for image filtering.

$$G_\sigma(x, y) = \frac{1}{2\pi\sigma^2} e^{-\frac{x^2+y^2}{2\sigma^2}}$$

i. Explain the role of parameter σ in this filtering processing.

[2 marks]

ii. Compare the performance in removing "pepper and salt" type images noise using this filter with filtering using median filter.

[2 marks]

(h) AdaBoost learning:

i. What is the fundamental concept in AdaBoost based learning?

[2 marks]

ii. In applying AdaBoost to face detection, would you choose the strongest classifier first to filter out negatives? If not, why?

[2 marks]

Question 2.

(a) Discuss the relationship between Bagging and Boosting. The discussion should include the difference and similarity between these two approaches.

[2 marks]

(b) The following diagram illustrates a convolution-pooling structure in a Convolutional Neural Network (CNN). List 3 commonly used pooling techniques and discuss why pooling is useful in constructing deep neural networks.

Input Image Convolutional Layer Pooling Layer

[2 marks]

(c) What are the advantages of a CNN over a fully connected Deep Neural Network for image classification?

[3 marks]

(d) Consider a CNN composed of three convolutional layers, each with 3 × 3 kernels, a stride of 2, and zero padding. The lowest layer outputs 200 feature maps, the middle one outputs 300, and the top one outputs 400. The input images are RGB images of 200 × 300 pixels. What is the total number of parameters in the CNN?

[3 marks]

(e) Discuss the advantages and disadvantages of using Sigmoid Function as the activation function in deep neural network.

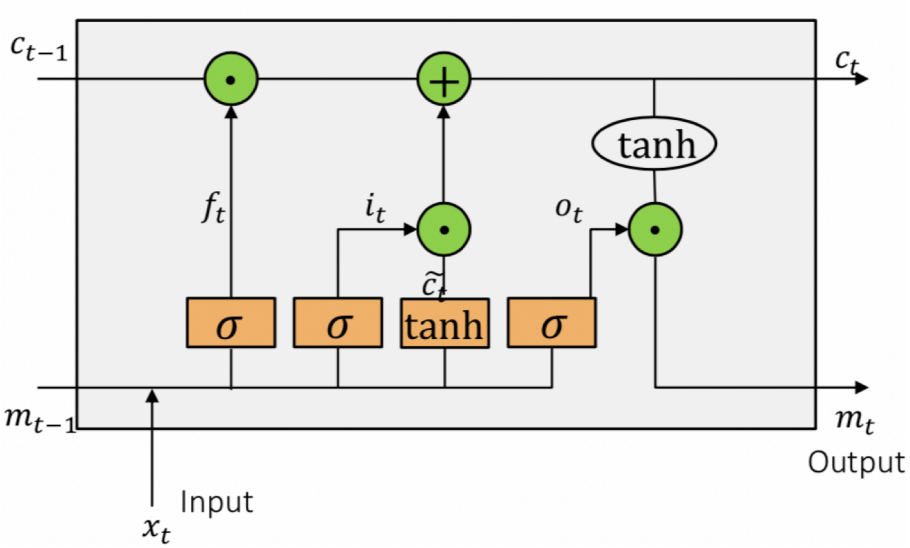
[2 marks]

(f) Deep neural network models have the potential to provide features with semantic meanings and have shown superior performances in many applications compared to, for instance, conventional fully connected neural networks. What are the neural network training strategies available to deep neural networks in order to tackle the overfitting problem?

[2 marks]

Question continues on next page.

(g) The figure below illustrates a memory cell of a (Long Short-Term Memory) LSTM network. Identify and explain the purpose of the following gates: input gate, output gate, and forget gate. You may use diagrams in your answer.



[3 marks]

(h) Name two applications each for sequence-to-sequence Recurrent Neural Network (RNN), sequence-to-vector RNN, and vector-to-sequence RNN.

[2 marks]

(i) For video classification, how can a CNN be combined with an RNN?

[3 marks]

(j) What advantages do LSTM networks have compared to conventional RNNs?

[3 marks]

End of exam paper.