IMPERIAL COLLEGE OF SCIENCE, TECHNOLOGY AND MEDICINE

EXAMINATIONS 2023

MEng Honours Degree in Electronic and Information Engineering Part IV
MEng Honours Degree in Mathematics and Computer Science Part IV
MEng Honours Degrees in Computing Part IV
MSc in Advanced Computing
MSc in Artifical Intelligence
MSc in Computing Science (Specialist)
MRes in Artificial Intelligence and Machine Learning
for Internal Students of the Imperial College of Science, Technology and Medicine

This paper is also taken for the relevant examinations for the Associateship of the City and Guilds of London Institute

PAPER COMP70014=COMP97105

MACHINE LEARNING FOR IMAGING

Wednesday 22nd March 2023, 10:00
Duration: 120 minutes

Answer ALL TWO questions

Paper contains 2 questions
Calculators required

1 a ML basics

i) Explain how regularisation can be used to control the bias/variance trade-off in machine learning models?

ii) How many parameters does a linear regression model have that takes greyscale images of size 32×16 as inputs?

b Forward passes in CNNs

Given the following CNN architecture:

```
conv1 = Conv2d(in=3, out=4, kernel=5, stride=1, pad=0)
conv2 = Conv2d(in=4, out=6, kernel=3, stride=1, pad=0)
conv3 = Conv2d(in=6, out=8, kernel=3, stride=1, pad=1)
conv4 = Conv2d(in=8, out=4, kernel=2, stride=2, pad=0)
fc1 = Linear(in=900, out=64)
fc2 = Linear(in=64, out=10)
```

i) What is the size of the receptive field for the following forward pass?

```
conv4(conv3(conv4(conv2(conv4(conv1(x))))))
```

ii) What is the valid input image size for this CNN given the following forward pass?

```
x = max_pool(conv1(x), kernel=2)
x = max_pool(conv2(x), kernel=2)
y = conv3(x)
x = conv4(y)
x = fc1(linearize(x))
x = fc2(x)
```

iii) What is the size of the tensor `y` above?

iv) Using the same architecture and forward pass from (ii), what change is required such that the CNN can process images that are twice as large as the original valid input size?

v) What change to the original architecture given at the top is required such that the CNN can process input images of any size equal or bigger than the original valid input size?

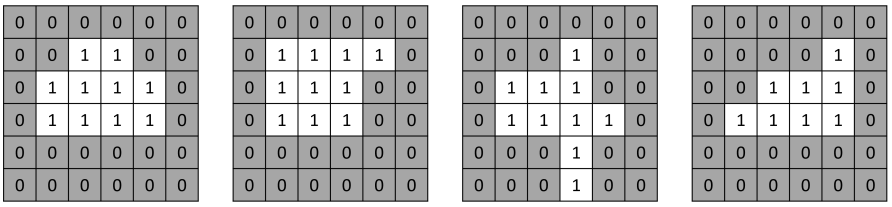
vi) What is the size of the output for this new architecture from (v) for an input image with a width and height of 224×224?

c Segmentation

i) A research paper compares two models for image segmentation. For model A, the reported precision is 75%, the sensitivity is 80%, and the specificity is 98%. For model B, the reported precision is 85%, the sensitivity is 70%, and the specificity is 99%.

Which of the two models has better segmentation performance?

ii) Given the following reference segmentation and three predictions A, B, and C. Which prediction is the best, and why? Provide numbers.



d Registration

Design a fully convolutional spatial transformer network with two convolutional layers for estimating the parameters of a 2D affine transformation for RGB image inputs with a width and height of 16×16.

Assume the following forward pass:

```
x = max_pool(conv1(x), kernel=2)
x = conv2(x)
```

e Unsupervised learning

i) What do the latent variables in a Gaussian mixture model represent?

ii) Given a collection of rigidly registered 3D shapes of the spinal anatomy of 200 patients. After applying PCA to the dataset, what anatomical variation does the first principal component most likely encode?

iii) In Generative Adversarial Networks, what is the role of the discriminator?

The five parts carry, respectively, 10%, 35%, 30%, 10%, and 15% of the marks.

2a Convolutional Neural Networks

i) What is the main difference between a fully connected layer and a convolutional layer in a CNN? (List all that apply)

A) A fully connected layer connects all neurons in the previous layer to all neurons in the current layer, while a convolutional layer only connects neurons within a small region of the previous layer to the current layer.

B) A fully connected layer only connects neurons within a small region of the previous layer to the current layer, while a convolutional layer connects all neurons in the previous layer to all neurons in the current layer.

C) A fully connected layer performs regression tasks, while a convolutional layer performs classification tasks.

D) A fully connected layer performs classification tasks, while a convolutional layer performs regression tasks.

ii) Which of the following is true about Batch Normalization? (List all that apply)

A) Batch Normalization layers are skipped at test time because a single test example cannot be normalized.

B) Its learnable parameters can only be learned using gradient descent or mini-batch gradient descent, but not other optimization algorithms.

C) It helps speed up learning in the network.

D) It introduces noise to a hidden layer's activation, because the mean and the standard deviation are estimated with a mini-batch of data.

iii) What is the purpose of using 1x1 convolution?

iv) Why is scaling (γ) and shifting (β) often applied after the standard normalization in the batch normalization layer?

v) A convolutional neural network has 4 consecutive layers as follows: (1) 3x3 convolutions with stride 2, (2) 2x2 pooling, (3) 3x3 convolutions with stride 2 and (4) 2x2 pooling. How large is the receptive field (the set of image pixels which activate) of a neuron in the 4th layer of this network?

vi) The input to a block in a CNN is of size is $64 \times 64 \times 8$. The block consists of the following layers:

* A convolutional layer with 32 filters with height and width 3 and no padding. Each filter has a weight and a bias.
* A 2×2 max-pooling layer with stride 2 and no padding.
* A batch normalization layer.

Compute the output activation volume dimensions and number of parameters for each of the layers.

b Object detection

The YOLO algorithm is an efficient method for object detection that makes extensive use of convolutional neural networks (CNNs).

i) A key element in the YOLO object detection is the introduction of a fixed size parameterisation of the output of the CNNs in form of a 3D grid of cells, i.e. a tensor. Explain how this tensor is formed and the meaning of the different cells within the tensor.

ii) You are asked to design a YOLO-based organ detector for localising the following organs in a 3D CT scan of the upper body: Kidney, liver, left and right lungs, pancreas and spleen. Explain how the tensor (the output of the CNN) would need to change for this application and why?

iii) Discuss the advantages of YOLO compared to approaches such as R-CNN.

c Federated learning

i) Training large, well-generalizable models often require large volumes of carefully curated training data. Such data is often difficult to obtain due to various data governance and data protection regulations. As a result, new methods for collaborative machine learning have been proposed, one of which is called *federated learning* (FL). Describe the FL paradigm in detail. You may (but do not have to) use a diagram to aid your answer.

ii) Explain four problems which may occur when training neural networks in a federated learning setting but which do not occur in training in the standard learning setting.

The three parts carry, respectively, 45%, 30%, and 25% of the marks.