IASD Master

Deep Learning for Image Analysis
Exam

March 23, 2022

Last name: _____

First name: _____

Total points: 40

1. (2 points) What is the difference between stochastic, batch and mini-batch gradient descent?

2. (2 points) Describe the role of training set, validation set and test set for machine learning methods.

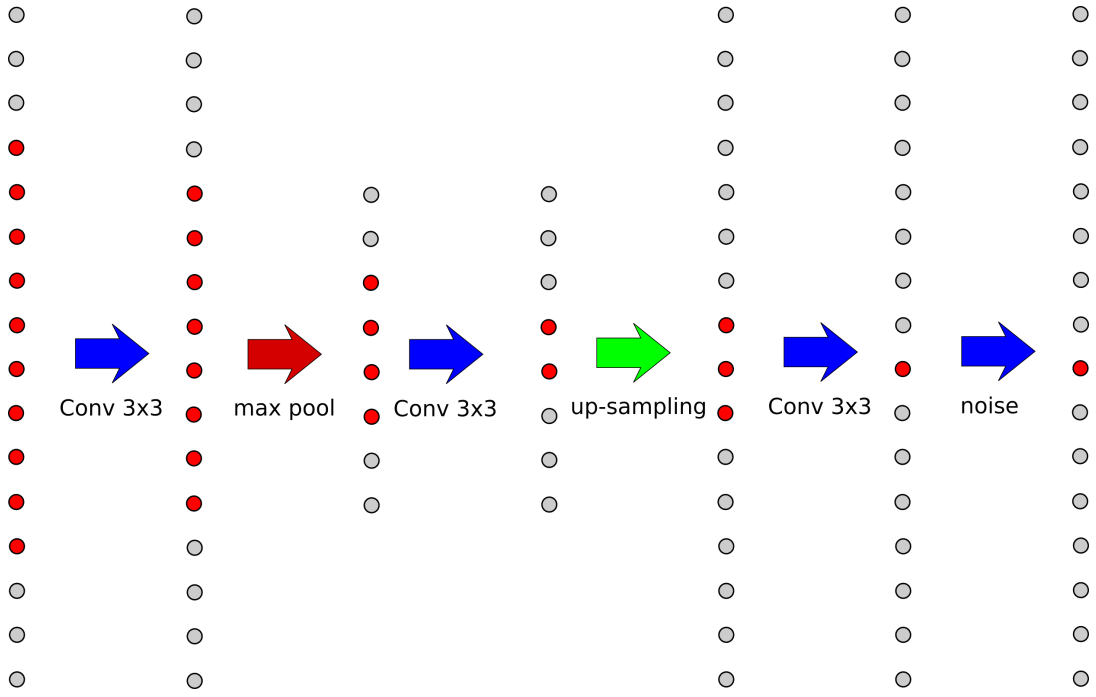
3. (2 points) Describe two different methods of regularization for neural networks.

4. (2 points) Within a convolutional neural network, how can you divide by two the spatial dimensions between two layers?

5. (3 points) In order to learn an image transformation between gray level images of size 1000x800, a fully-convolutional neural network composed of 6 convolutions with kernels of size 3x3 and 8 channels, plus a last convolutional layer with kernel 3x3 and a single channel, is proposed. What is the size of its receptive field?

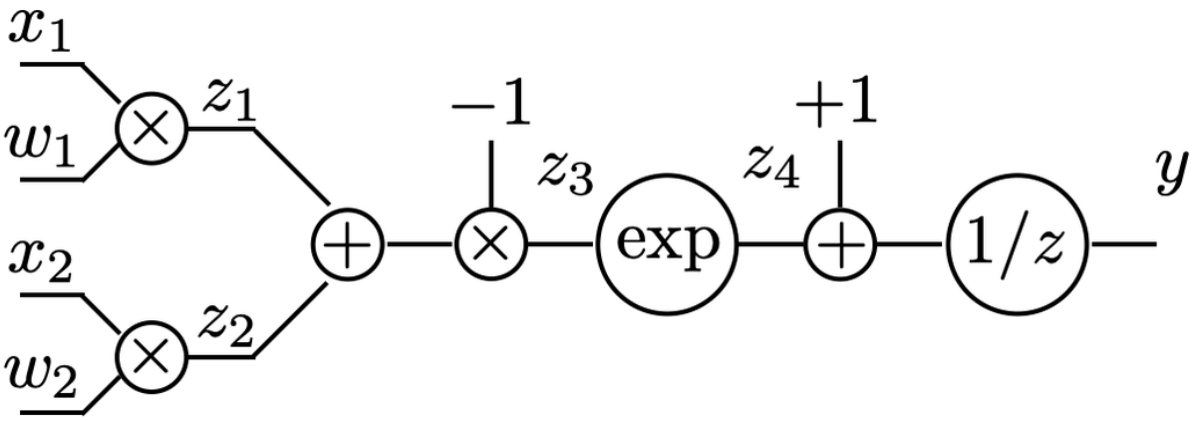
The receptive field is not linked to the number of channels, nor to the image size. Given that we have 7 convolutions with kernels of size 3x3, the width of the receptive field is : 7 x (3-1) +1.

6. (4 points) How many parameters does each layer in the following network contain?



_____ _____ _____ _____ _____ ___

1. (5 points) The computational graph below is given. The current parameters are w1=2, w2=0.5 . There is a new forward pass with the training sample x1=2 and x2=-1. Compute the gradients w.r.t. x1, x2, w1, w2 and all zi by backpropagation.



7. (8 points) A cheap smartphone can only take low resolution images: 640x640 pixels. Our objective here is to develop an artificial neural network that can transform these low resolution images into images of size 2560x2560.
    a. How would you proceed to constitute the database?

Gather high resolution images (2560x2560).
Downsample them to obtain the input low resolution images of size 640x640.

b. What neural network would you choose? How would you adapt it to the task at hand?

A fully-convolutional neural network, such as a U-Net.

Some layers of the decoding branch can be removed to obtain the right output size. The strides can also be adapted to reach this goal.

c. What loss function would you recommend?

L2, L1.

d. Do you anticipate any difficulties? How could you solve them?

The output images might not be sharp enough. A conditional GAN could help treat this problem.

The images constituting our database would be acquired with a different sensor than the one in the smartphone. Therefore the network performance could be disappointing. Image augmentation methods, with different kinds of noise, could help to alleviate this problem.

8. (12 points) Our aim is to build a network to transform images taken at night into daytime images.

Scenario 1: In your database, you have pairs of pictures. For a given pair, both images have been taken from exactly the same viewpoint, but one during daytime, the other during nighttime.

a. Propose a network architecture to tackle this task.
It can be solved as a supervised problem -> fully-convolutional neural network i.e. any Pixel-To-Pixel Network or an UNET architecture.

b. How would you train the network? What loss function would you use?
Classical training using training/validation set for monitoring the overfitting.
Loss Functions: MSE / MAE / Not CrossEntropy.

c. What type of image augmentation method would you use?
Training should include image augmentation on PAIRS of images.
For instance: Rotation / Horizontal Shifts / Horizontal Flips / MixUp

Scenario 2: Contrary to the previous scenario, you have now two separate sets of images, the first acquired during daytime and the second one during nighttime. This time, you DO NOT have pairs of images acquired from the same viewpoint.

a. Propose a network architecture to tackle this task.
Expected answer: Cycle GANs or Conditional GANs or a generative model that can be conditioned for instance Conditional Autoencoder.
Alternative: Another possibility is to train a network by conditioning the feature space to be close enough to an objective. This is an approach used in style transfer.

b. How would you train the network? What loss function would you use?
Expected answer: Training as traditional GANs including cycle consistent. Image augmentation is recommended. Three terms in the loss function of GANs from Domain 1 to Domain 2, Domain 2 to Domain 1 and a cycle consistency term to compare D1(D2) and D2(D1).