Question hour & preparing for the exam (+project advice)

Cian Scannell

Project advice (part I)

- Move beyond the minimal solution in the notebook
- State a clear hypothesis for research question (what and why)
- The introduction is not only an introduction of image registration but also the introduction to your project
- Methods should focus on telling us what you did and not too much background (experiments section)
- Results cannot only be tables or figures (need to be described) and think carefully about what to compare
- Discussion – a summary of what you did and then your interpretation of the results (including limitations). Answer your hypothesis
- Think what can be excluded (excessive background)
- But don’t exclude important things (details of what you did)

Project advice (part II)

- Don’t include things like the names of files
- The appendix is not a dump – be selective of what is relevant
- TAs should be able to easily run code with one button (make a README)
- Pay attention to the quality of figures
- Consistent tone
- Make sure details are in correct section (e.g. dataset details not in intro)
- Units
- Captions

Questions??

Exam Qs

1. Consider the following 2D transformation matrices:

$$\mathbf{R} = \begin{bmatrix} \cos(\frac{\pi}{3}) & -\sin(\frac{\pi}{3}) \\ \sin(\frac{\pi}{3}) & \cos(\frac{\pi}{3}) \end{bmatrix}, \quad \mathbf{S} = \begin{bmatrix} 5 & 0 \\ 0 & 2 \end{bmatrix}$$

When applied to an object $\mathbf{G}$, the first transformation matrix performs a counter-clockwise rotation by $\frac{\pi}{3}$ and the second transformation matrix performs scaling by 5 in the x- and scaling by 2 in the y-direction.

(a) (2 points) Write these rotation and scaling matrices in homogeneous form.

(b) (1 point) Construct a homogeneous transformation matrix $\mathbf{T}$ that performs translation by 4 in the x-direction and translation by 3 in the y-direction.

(c) (1 point) How can you combine the $\mathbf{R}$, $\mathbf{S}$, and $\mathbf{T}$ transformation matrices into a new transformation matrix $\mathbf{T}_{\text{combined}}$ that will first scale the object $\mathbf{G}$ that is being transformed, then translate it, and at the end perform a rotation? Write the expression.

(d) (1 point) A transformation can be "undone" or reversed by means of inverse transformation. How can you retrieve the original object $\mathbf{G}$ that has been transformed with $\mathbf{T}_{\text{combined}}$? Write the expression in terms of $\mathbf{R}$, $\mathbf{S}$, and $\mathbf{T}$.

Solution:

(a)

$$\mathbf{R} = \begin{bmatrix} \cos(\frac{\pi}{3}) & -\sin(\frac{\pi}{3}) & 0 \\ \sin(\frac{\pi}{3}) & \cos(\frac{\pi}{3}) & 0 \\ 0 & 0 & 1 \end{bmatrix}, \quad \mathbf{S} = \begin{bmatrix} 5 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

4. Consider the joint histograms of a fixed image and a moving image transformed with two different image transformations in Figure 1. Dark purple indicates low values and yellow/orange high values.

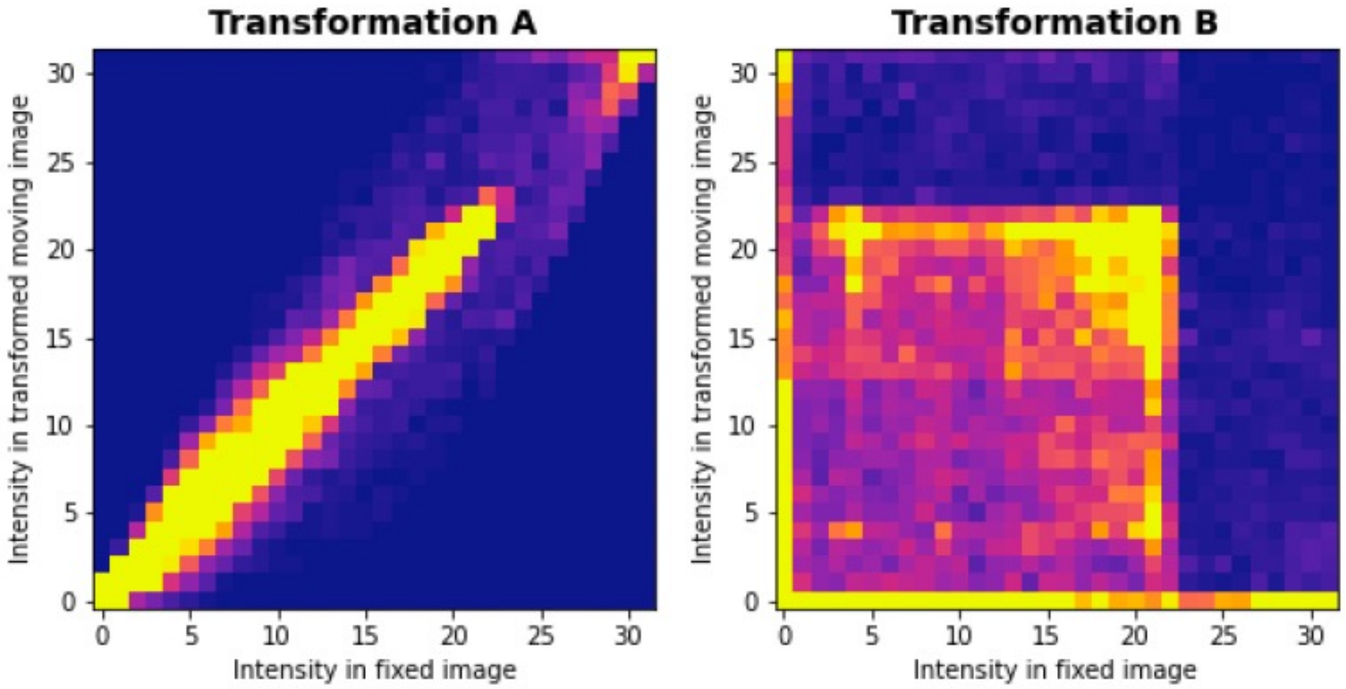

Figure 1: Joint histograms of a fixed and moving image transformed with two different image transformations.

(a) (1 point) Based on the joint histograms, which pair of images is better registered, A or B? Motivate your answer.

(b) (1 point) Which of the two joint histograms will result in a higher value for the mutual information between the fixed and transformed moving image? Why?

12. (2 points) Describe briefly the difference between gradient descent and stochastic gradient descent. When is stochastic gradient decent the preferred option?

16. (2 points) Propose a nonlinear feature transformation that in combination with a linear classifier such as logistic regression can result in the decision boundary in the figure below. Motivate your answer.

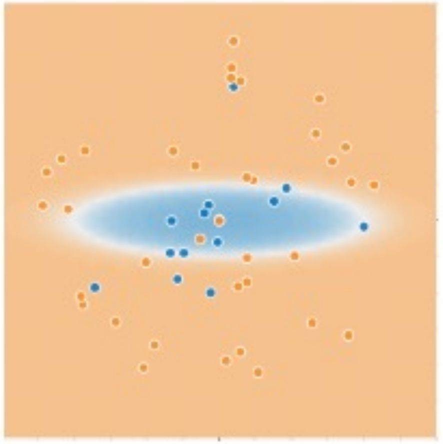

Figure 5: Two-class (blue and orange) classification problem with two features (represented on the horizontal and vertical axis). The decision boundary is produced by a logistic regression classifier trained with nonlinear transformation of the two features.

17. Figure 6 shows a schematic description of a very small convolutional neural network. Nonlinearities and biases are omitted in this question.

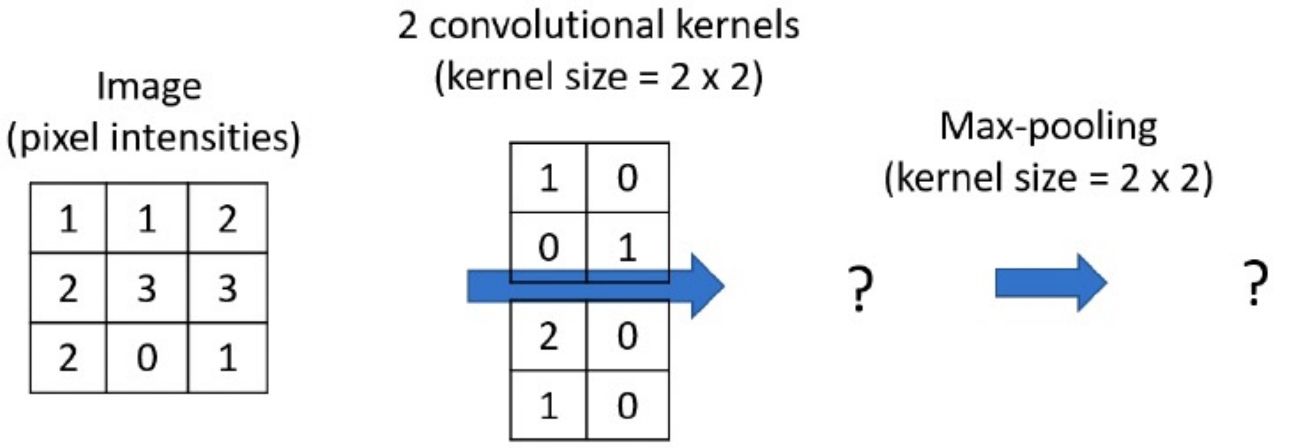

Figure 6: A small convolutional neural network. On the left we see an image consisting of 9 pixels.

(a) (2 points) A convolutional neural network typically consists of convolutional layers and (max) pooling layers. Briefly explain the main **function** of both layers.

(b) (3 points) Consider the network in Figure 1. We first apply two convolutional kernels of size 2 x 2 to an image consisting of 3 x 3 pixels. We use (the typical) stride = 1 and we do not use padding. Then, we apply a max-pooling operation with a kernel of 2 x 2, with (the typical) stride = 2 and without padding. What do we get at the location of the question marks? Show your calculations.

(c) (1 point) What is the receptive field at the location of the second question mark (after max-pooling)? Motivate your answer.

Solution:

(a) Convolutional filters are used to extract (abstract) features from the previous layer (1 point). Pooling layers are used reduce the size of the feature map (1 point). If instead of size reduction, translation invariance is mentioned (only 0.5 point). If students only explain how the layers are applied without describing the function (0 points).

(b) The first question mark = two 2 x 2 feature maps: [[4, 4],[2, 4]] & [[4, 5],[6, 6]]. The second question mark = [4] & [6].

- (1 point) for correct shapes (two 2 x 2 feature maps and two 1 x 1 feature maps)
- (1 point) for correct feature map values
- (1 point) for correct max-pooling values