Machine Learning for Image Analysis, SoSe 2024
Exercise Sheet #5: Convolutional Neural Networks
Due date: May 21, 2024, before 11 am

Problem 1 (CNN Layers) (15 pt).
You want to classify images into 10 predefined classes. You build a CNN taking as input 100x100 RGB-images which consists of the following layers:

- Two consecutive convolutional layers, each of them having filter dimensions 5x5, padding size 2 and stride 3. There are 8 filters in the first conv-layer and 16 in the second.
- An average pooling layer with filter size 2x2 and stride 2 (no padding).
- And lastly a fully connected layer to the final output vector containing the scores for each class.

(a) (6pt) What are the ouput sizes (width, height, channels) of the intermediate layers?

(b) (6pt) How many trainable weights are present in this architecture?

(c) (3pt) Compute the size (width, height, channels) of the receptive fields in each layer up to the input image for a neuron receiving outputs from the pooling layer.

Problem 2 (Normalization) (10 pt).
Consider a (quite strange) training set with only two images A and B both having size 4x4. Image A is uniformly cyan colored, i.e. each pixel has the RGB value $(0, 255, 255)$ and image B is uniformly magenta colored, i.e. each i.e. each pixel has the RGB value $(255, 0, 255)$.

(a) (3pt) Consider **batch** normalization on that training set. What is the resulting mean vector?

(b) (3pt) Consider **layer** normalization on that training set. What is the resulting mean vector?

(c) (3pt) Consider **instance** normalization on that training set. What are the resulting mean vectors?

(d) (1pt) What is the receptive field of a neuron in the output after applying **batch** normalization? Specify number of channels, width and height of that field.