Quiz

Marco Willi

The following questions could be exam questions.

Convolutions

1. A convolutional layer has 64 input activations ($C_{in} = 64$ and $H = 16$, $W = 16$). You want to reduce its spatial dimensionality by half, while doubling the number of channels. How do you parameterize your convolutional layer? Provide an example.
2. In the example above: How many weights do you need to learn?
3. You have very large images (8000×8000 pixels). Your model always crashes with out-of-memory-errors. What options do you have when parameterizing your convolutions?

CNNs

1. Can CNNs be used to count objects? Take a look at the following figure. You want to count in how many quadrants an objects occurs. Justify your answer.

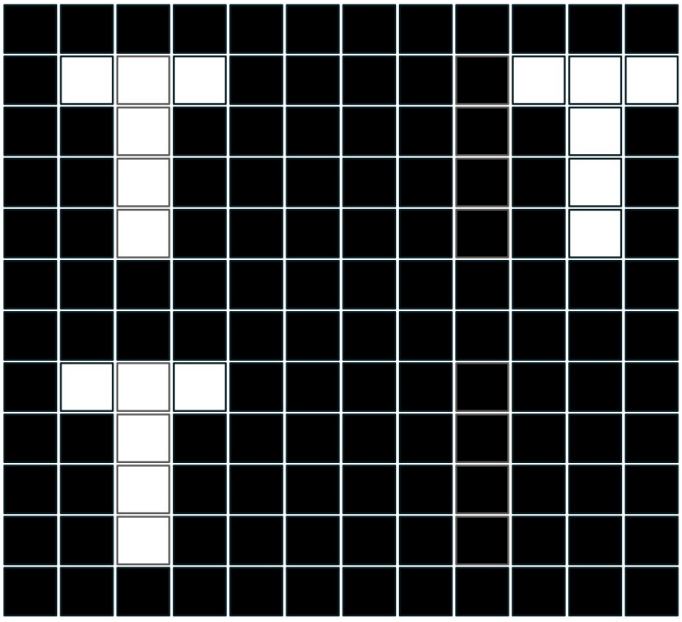

Figure 1: Can CNNs count objects

2. Can CNNs be used to model inputs, e.g. satellite data that are not RGB images, i.e. have more than 3 input channels? Justify your answer.
3. You want to model images which are not square. They have a spatial resolution of 800x400. What is different within a CNN as opposed to if they were square?

Image Classification

1. You trained a model to identify synthetic (fake) images. The model is quite good but not perfect. When deploying the model you have the option to send some images for manual verificaiton by an expert. Which do you choose? Justify your answer.

Foundation Models

1. You applied CLIP on a dataset to identify synthetic / fake images and used the following prompts:

- "A synthetic image".
- "A real image".

When you compare to the ground truth labels that you have collected you see that the model does not perform well. What options do you have to improve the model?