CS344 Exercise 8

Task 1: Advanced convolutional neural networks (CNNs)

Max pooling normally uses parameters that are learned during training.

TRUE FALSE

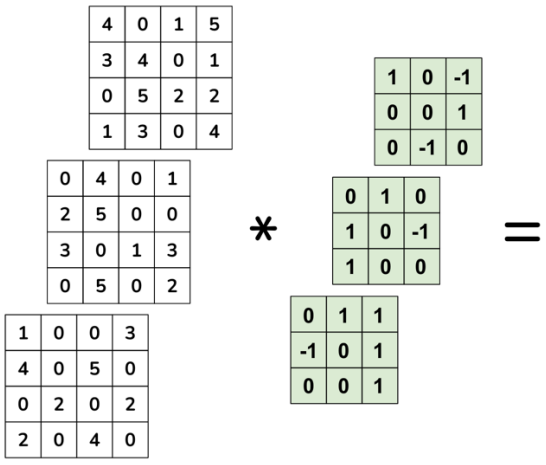
CNNs only use convolution layers.

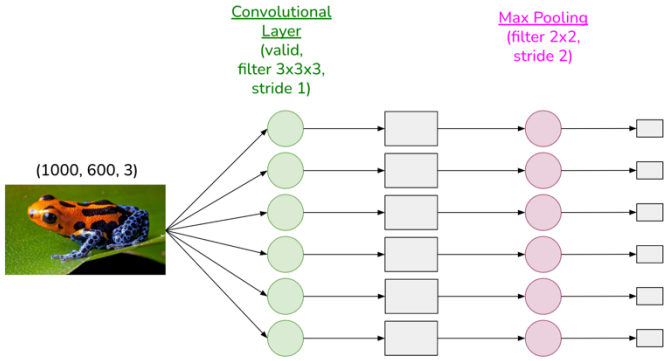
TRUE FALSE

The parameters of a CNN can be optimized during training using gradient descent to minimize a cost function.

TRUE FALSE

Show the result (a 2D array filled with values) of the convolution below between the (4, 4, 3) input and the (3, 3, 3) filter, assuming the convolution is valid and uses a stride of 1.



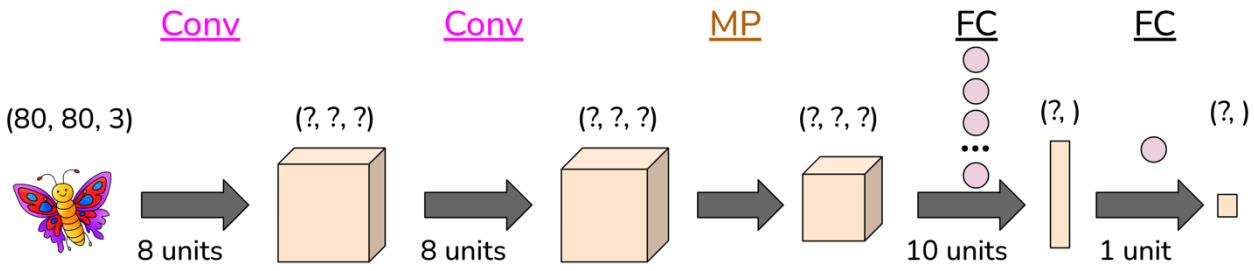


Consider the network architecture shown above where the input has shape (1000, 600, 3). What is the shape of the output from each of the 6 Max Pooling units?

(499, 299)
(500, 300)
(998, 598)
(1000, 600)

Consider the network architecture shown above where the input has shape (1000, 600, 3). How many parameters does the network have that are learned during training?

162
168
186
216



Consider the network architecture shown above where the input has shape (80, 80, 3), the convolution layers are valid and use filters with height and width of 3 and stride of 1, and max pooling use a filter with height and width of 2 and stride of 2.

What is the output shape of the first convolutional layer?

What is the output shape of the second convolutional layer?

What is the output shape of the max pooling?

What is the output shape of the first fully-connected layer?

What is the output shape of the second fully-connected layer?

How many traininable parameters does the first convolutional layer have?

How many traininable parameters does the second convolutional layer have?

How many traininable parameters does max pooling have?

How many traininable parameters does the first fully-connected layer have?

How many traininable parameters does the second fully-connected layer have?

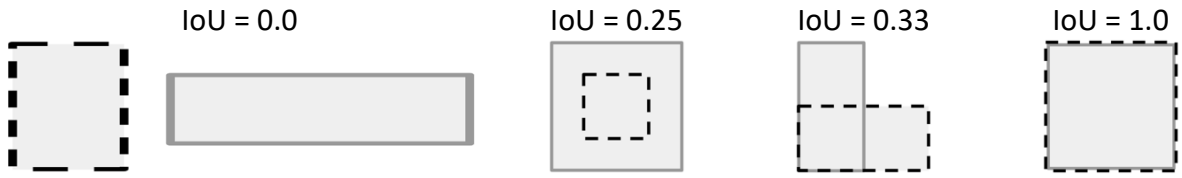
Task 2: Intersection over Union (IoU)

Object localization refers to finding the location of an object in an image, usually by determining a bounding box for the object. If we predict a bounding box for an object using a CNN, how do we quantify (measure) how good our prediction is? Consider the image below, where the solid line box is the "true" bounding box for the car object and the dotted line box is the bounding box predicted by a CNN, which overlaps mostly but not entirely with the "true" box.

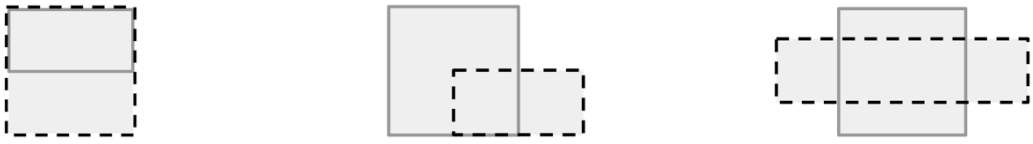
Intersection over Union (IoU), is a way to measure how good a predicted box is, i.e., how well a predicted box corresponds to the true box. IoU is the area of the intersection of the two boxes divided by the area of the union of the two boxes:

$$IoU = \frac{Area\ of\ intersection\ of\ 2\ boxes}{Area\ of\ union\ of\ 2\ boxes}$$

The IoU ranges between 0.0 and 1.0, with 0 indicating the boxes have no overlap and 1 indicated the boxes have perfect overlap. For example, for the four pairs of boxes below, the IoU is indicated.



What is the IoU for each of the three example pairs of boxes below?



Task 3: Coding with advanced convolutional neural networks

Download the Jupyter Notebook for Exercise 8 from the course website. Open the Notebook in your web browser and work through it. As you work through the Notebook, answer the following questions.

Binary or multiclass image classification?

For the **`Pandas`** dataset, what is your model's F1-score on validation data?

For the **`Pandas`** dataset, how many total parameters does your model have?

For the **`Foods`** dataset, what is your model's F1-score on validation data?

For the **`Foods`** dataset, how many total parameters does your model have?

For the **`People`** dataset, what is your model's F1-score on validation data?

For the **`People`** dataset, how many total parameters does your model have?

For the **`LFIW`** dataset, what is your model's F1-score on validation data?

For the **`LFIW`** dataset, how many total parameters does your model have?

Weather

How many parameters does the model have?

Does the model use regularization at all?

What is the F1-score of the model on the testing data?

On what type of weather images does the model perform best? On what type of weather images does the model perform worst?

CS344 Exercise 8 Final Page

In the *TIME* column, please estimate the time you spent on this exercise. Please try to be as accurate as possible; this information will help us to design future exercises.

PART	TIME
Exercise	