

Assignment 2

1 Practice: CNN Architecture

As part of this assignment you will implement a Convolutional Neural Network using PyTorch that matches the following architecture:

Input: RGB image of size $64 \times 64 \times 3$

Conv2D with 16 filters, kernel size 3×3 , stride 1, padding 1

ReLU activation

MaxPooling2D with kernel size 2×2 , stride 2

Conv2D with 32 filters, kernel size 3×3 , stride 1, padding 1

ReLU activation

MaxPooling2D with kernel size 2×2 , stride 2

Flatten the output

Fully connected layer with 100 units

ReLU activation

Fully connected layer with 10 units (assume 10 output classes)

2 Practice: Model Deployment

Use the CIFAR10 dataset from class to train a classifier model. Add the classifier to the API implemented as part of Class Activity 1. Commit your code to GitHub.

3 Theory: Arithmetic of CNNs

Answer the following questions.

Question 1

Given an input image of size $32 \times 32 \times 3$ and a convolutional layer with 8 filters of size 5×5 , stride 1, and no padding, what is the output size?

Question 2

How does the output size change if padding is changed to “same”?

Question 3

If you apply a 3×3 filter with stride 2 and no padding to a 64×64 input, what is the output spatial size?

Question 4

You apply a max-pooling layer of size 2×2 with stride 2 on a 16×16 feature map. What is the output size?

Question 5

An image of shape 128×128 is passed through two successive convolutional layers. Each uses a 3×3 kernel, stride 1, and ‘same’ padding. What is the output shape?

Question 6

In the examples in class, before starting the training loop we ran: `model.train()`. What happens if you remove that line?