

Q3

April 28, 2024

1 a)

Layer	Activation map dimensions	Number of parameters
INPUT	$256 \times 256 \times 3$	0
CONV-3-32	$254 \times 254 \times 32$	896
POOL-2	$127 \times 127 \times 32$	0
CONV-5-64	$123 \times 123 \times 64$	51264
POOL-2	$61 \times 61 \times 64$	0
FC-3	$1 \times 1 \times 3$	714435

2 b)

The MLP model can contains the layers of input layer, hidden layers and output layers. Each neurons in the layers is connected to the neurons in the previous layer by a non-linear activation function taking in the previous neuron value, the weights for each neuron and a bias. The true label value of the output layer can then be used to backward propagate the weight and bias of the model through a stochastic mini-batch gradient descent to obtain a valid model for predicting the output value.

assuming the input is still as $256 \times 256 \times 3$ in dimension, and for simplicity of the model only 1 hidden layer in the model

each node in the input layer will be connect to a hidden layer of dimension H , then to the output layer with dimension 3 (3 nodes as in problem)

The connection between the input and hidden layer will have parameters (assuming there is a bias):

$$(256 \times 256 \times 3 + 1) \times H = 196609H$$

When connecting to the final output layer, the total parameter count of the model will be

$$(196609H + 1) \times 3$$

3 c)

The clear benefits of CNN models is that there is much less parameters to be estimated during training compares of MLP, which can lower the computation power and time needed for developing

a model.

CNN also has another benefits that it can take in the relation of pixel in an image, since it harness the convolution of a tensor of image pixel to predict the required activation value for the next layer. On the other hand MLP hidden take in the whole image as vector input with no discrimination is accounting for the position of the image pixel location. This allow CNN to better extract object features from a large image and can make classification base on part of the features in the picture that is valid, while MLP can only do classification base on the whole image.