Q3

April 17, 2024

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

[ ]:
```python
import numpy as np
from tensorflow import keras


model = keras.models.Sequential()



#here in the snippet below
#D = 5 (first parameter)
#Stride= (1,1) by default
model.add(keras.layers.Conv2D(32, (3,3), activation='relu', input_shape=(256,␣
↪256, 3)))
model.add(keras.layers.MaxPooling2D(pool_size=(2, 2)))
model.add(keras.layers.Conv2D(64, (5,5), activation='relu'))
model.add(keras.layers.MaxPooling2D(pool_size=(2, 2)))
model.add(keras.layers.Dense(3, activation='relu'))
model.summary()
```

The MLP model can contains the layers of input layer, hidden layers and output layers. Each neurons in the layers is connected to he neurons in the previous layer by an 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 to obtain a valid model for predicting the ouptut value.

assuming the input is still as 256 x 256 x 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

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

It can see that the MLP model ususally contains parameters count of H times to the CNN models, give the hidden layer dimension of H.

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 discrimation is accouting for the postion 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.