:: Deep Learning :: COSE474

Project #2. CNN Architecture Implementation

Korea University, Computer Science / 2017320108, Jaeyoung Ko

2022-11-13

This report explains the second assignment, CNN architecture, especially for the VGG model and ResNet model. The goal is to understand VGG-16 model and implement RestNet-50 model.

First of all, I completed the code for 'ResidualBlock' class which works as the bottleneck building block in ResNet-50. The class is divided into two cases : in case of the value of downsample flag. If the flag is True, the three-layered stack which sequentially contains 1x1 conv, 3x3 conv and 1x1 conv should reduce the activation map. For this reason, first 1x1 conv layer should have stride 2 so as to get the output size halved. If False, in order to preserve the size, the stride for every layer should be tuned as 1.

Second, I completed the code for ResNet50_layer4 according to the table shown in ppt slide 14. CIFAR-10 dataset has 10 classes as its name shows, and the first layer gets RGB images applied to 7x7 conv which produces 64 output channels with stride 2. The padding size is tuned as 3 because the input image size gets halved. I completed the rest of layers with residual blocks. Especially for last residual block in second and third layer, the downsample flag should be True because the size should be halved after the layer processing. In contrast, the fourth layer maintains the image size. So its downsample flag value is False as well. In FC layer, 1024 channels produces the 10 classes for CIFAR-10 dataset. Lastly for AvgPool layer, as 2x2 input size gets halved as 1x1 image size, it should have kernel size 2 and stride 1.

The screen captrues for processing this project is shown below. The upper one is the result of VGG-16 model and the bottom one is that of ResNet model.