

Deep Learning Project #2 Report

2016160311 이재윤

Implementing "ResidualBlock"

In each residual block, we need 3 convolutional layers including two 1x1 convolutions and one 3x3 convolution. If downsampling is needed, one of the convolutional layer should be set by the stride of 2. Refer to the network architecture, first 1x1 conv layer should be set as stride of 2. This part of code is implemented as `conv1x1(in_channels, middle_channels, 2, 0)`. The number of output channel of first 1x1 conv layer and 3x3 conv layer is corresponded to `middle_channels`, and the second 1x1 conv layer's output channel is corresponded to `out_channels`. To maintain the size of tensor, 3x3 conv layer's padding is set as 1. The codes of 3x3 conv layer and the second 1x1 conv layer are implemented as `conv3x3(middle_channels, middle_channels, 1, 1)` and `conv1x1(middle_channels, out_channels, 1, 0)`. If downsampling is not needed, the only difference is first 1x1 conv layer, where stride should be set as 1 instead of 2.

Implementing "ResNet50_layer4"

The answer of the first blank is 10, which is the number of classes in Cifar-10 dataset. In the `layer1`, the answer of five blanks in `Conv2d` parameters should be 3, 64, 7, 2, and 3, because the original image data has 3 channels, and the output channel of 7x7 conv layer is 64, where stride is 2 and padding is 3. The answer of three blanks in `MaxPool2d` parameters should be 3, 2, and 1, because it is 3x3 max pooling layer where stride is 2 and padding is 1. In the `layer2`, the first residual block's input channel is 64 which is the output channel of `layer1`, and the middle channel and output channel are each 64 and 256 due to the network architecture. The last parameter should be set as `False`, because the first and second residual blocks do not need to downsample. The second and third residual blocks' input channel is 256 due to previous blocks' output channel, and the third residual block's last parameter should be set as `True` to enable downsample. The same logic is applied to the `layer3` and `layer4`, but the `layer4` does not use downsample because the output image size is same as layer3's in the network architecture. The fc layer's parameters should be set as 1024 and 10, because 1024 channels after average pooling layer will be fully connected to the final result of 10 classes. The answer of two blanks in `AvgPool2d` parameters are 2 and 2 because we need 2x2 layer to make 2x2 size as 1x1, but the stride seems not an issue whether it is 2 or not because there is no difference between 2 or any other values.

Result and Discussion

The accuracy of the model on the test images was around 81~82% which is quite similar to the expected accuracy. There was an issue with downloading Cifar-10 dataset because of "SSLCertVerificationError", so two lines of code were added on main.py which are `import ssl` and `ssl._create_default_https_context = ssl._create_unverified_context` to solve the problem.