HW3

Training Convolutional Neural Network (CNN) with different model structures and hyperparameter settings

In this homework, you will train CNN with different network architectures and hyperparameters. Please carefully read the following materials before you start.

https://docs.pytorch.org/tutorials/beginner/blitz/cifar10_tutorial.html#sphx-glr-beginner-blitz-cifa

0.1 Changing the CNN structure in the Pytorch tutorial

Starting from the Pytorch tutorial, retrain the model with the following changes to the CNN structure. Each change below should make independently over the original setting. Report the final training and testing accuracy for each change.

- **Channels in each layer:** In tutorial, there are 2 conv layers. Double the number of "channel" in each layer (the first number in the input, the first input channel is always 3 because an image have 3 channel: red, green and blue);

- **Number of convolutional Layers:** other settings in the tutorial, decrease to 1 conv layer, increase to 3 conv layer. (You can choose the channel of each layer.) Compare results. Notice that when changing number of Conv, you need to change the correspond kernel size and input channel for the later linear layer;

- **Activation functions:** Keep other settings in the tutorial, change the activation function. Try ReLU (the default setting of the tutorial), LeakyReLU, and Tanh. Compare the result;

- **Optional:** Adding residual connection: Try to change each single convolutional layer to a residual block with identity connection. You can design your own residual block structure (i.e. the number of convolutional layer, number of channels, kernel size, etc.). You can learn more about residual block from this paper. You might need to tune the hyperparameters, too. You will get up to 5 extra score for finishing this part. Please clearly describe the network structure you use in your report and report the final training and testing accuracy.

0.2 Changing the training hyperparameters

Starting from the Pytorch tutorial, retrain the model with the following changes to the hyperparameters. Each change below should make independently over the original setting. Report the final training and testing accuracy for each change.

- **Learning rate:** Keep other settings in the tutorial, change learning rate in optimizer definition. Try both smaller and bigger.

- **Optimizer:** Keep other settings in the tutorial, change the optimizer to SGD without momentum and Adam;

- **Large batch size:** Keep other settings in the tutorial, change the batch size to 32;

- **Longer training time:** Keep other settings in the tutorial, change the total number of training epochs to 5;

Please report the training and testing accuracy for each setting listed above in your final report. Please clearly state how you implement each setting in your report. (You can screenshot the important blocks of your codes, e.g. network structure, hyper-parameter, etc. and paste them in your report). For submitting the code, you only need to attach the runable code for one of settings listed above.