Homework 2: Convolutional Neural Networks

Deep Learning (84100342-0)
Spring 2018
Tsinghua University

1 Introduction

In this homework, we will try to solve an image classification problem by designing and implementing convolutional neural networks (CNNs). A good understanding of convolutional layers, pooling layers, activation functions and regularizers is required. Some optimization techniques might also be leveraged to facilitate the training process.

2 Dataset

Please train your models on two labeled datasets provided by us:

- Each dataset contains 65 categories of real-world images or clip-art images.
- Download link: https://pan.baidu.com/s/10cT-PIYP2QExZGYEfS6ovw.
- Directory structure: "$./dset\{1,2\}/train/label\{0,1,...,64\}/xxx.jpg$".

You might need to split the provided data for training and validation. Finally, you are required to submit two models, one for the real-world dataset and another for the clip-art dataset. Please note that test sets will not be given, but they are quite relevant to the corresponding training sets.

3 Requirements and Evaluations

3.1 Programming Language

Python only.

3.2 Deep Learning Framework

We recommend PyTorch and TensorFlow. If using other frameworks, please contact TA.

3.3 Scoring

- Construct your own CNNs and train your model from scratch (fine-tuning is forbidden) using the recommended deep learning framework. (40%)
- For each layer, you may use well-established libraries, such as tf.layers, tf.slim. Otherwise, it is a bonus (10%) if you implement the convolutional layer with basic arithmetic operators (e.g. tf.add, tf.matmul, tf.concat).
- Train a model for each dataset to achieve a good performance. (30%)
- Use techniques you have learned in the class to prevent over-fitting. (20%)
- Use extra techniques you find in other materials to further improve your model. Please explain why you choose it and how it works. (10%)

3.4 Notification

- Please submit your code, document and trained models as an Archive (zip or tar). The document is supposed to cover your insights of the proposed model, the technical details, the experimental results (including training and cross-validation), and the necessary references. If you use additional data sources, please be sure to specify.
- We will focus on your code and document to decide your score. Your proposed model needs to be different from the main-stream ones, such as AlexNet [2], VGG [3], GoogleNet [4], ResNet [1] and so on. Still, under equal conditions (novelty, code quality, document quality), a higher accuracy along with a reasonable computation efficiency indicates a higher score.

References

[1] K. He, X. Zhang, S. Ren, and J. Sun. Deep residual learning for image recognition. In *The IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, June 2016.

[2] A. Krizhevsky, I. Sutskever, and G. E. Hinton. Imagenet classification with deep convolutional neural networks. In *Advances in Neural Information Processing Systems 25: 26th Annual Conference on Neural Information Processing Systems 2012. Proceedings of a meeting held December 3-6, 2012, Lake Tahoe, Nevada, United States.*, pages 1106–1114, 2012.

[3] K. Simonyan and A. Zisserman. Very deep convolutional networks for large-scale image recognition. *CoRR*, abs/1409.1556, 2014.

[4] C. Szegedy, W. Liu, Y. Jia, P. Sermanet, S. Reed, D. Anguelov, D. Erhan, V. Vanhoucke, and A. Rabinovich. Going deeper with convolutions. In *The IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, June 2015.