Course Project for Machine Learning

Computer Science Department in SJTU

1 Brief Introduction

1. Two or three students form a group.
2. This project is a group work and divided into three modules:
   A. Complete the public sandbox reasoning experiments on Huawei Shengteng AI platform.
   B. Transfer the given trained model on Huawei Shengteng AI platform for inference.
   C. Build a model for the electron microscopy image segmentation.

Tips

(1). For module A and B, you are required to make an online presentation to TA.
Schedule: 2020.6.13 - 2020.6.14

(2). For module C, you are required to finish it in your free time and submit a report.
Deadline: 2020.6.20

2 Description for Module C

The project is inspired by *ISBI Challenge*[1] and designed for students to learn to build a model for a specific task.

2.1 Data

The ISBI 2012 EM Segmentation Dataset [1] can be downloaded from the course's home page on canvas (project). The data description is same with *ISBI Challenge* except that we split the raw train data set (consist of 30 samples) into two parts: the new train set and new test set. The downloaded data set consists of 30 samples, 25 for train and 5 for test. You are required to train and test your model on the newly split data sets.

[1] http://brainiac2.mit.edu/isbi_challenge/

2.2 Model

This course has a corporation with HUAWEI company, which provides some computation time for each students to train and deploy models on their platform. Therefore, you can train and test your model on GPU in some popular framework like pytorch, tensorflow, keras, etc. With the help of GPU, you are free to use deep learning methods to build your model.

3 Basic Assignments

1. (**30 points**) Complete the public sandbox reasoning experiments on Huawei Shengteng AI platform, an intelligent computing platform. TA will show the details and instructions online.

2. (**30 points**) Transfer the given model on Huawei Shengteng AI platform for inference. The trained model will be provided by TA online.

3. (**40 points**) Train a model by yourself, including:

   a. Learn to build a model for segmentation.

   b. Train the model and tune the hyper-parameters on the **train set**.

   c. Improve your model to gain as higher performance as you can on test set.

   d. Write a report that contains your ideas, methods, algorithms, experimental details and results.

4 Evaluations for the Final Report in Module C

Code takes up $20\% \times 40 = 8$ points while report takes up $80\% \times 40 = 32$ points.

4.1 Codes

Each group is required to submit your codes by providing a link to your github repo. If you do not know how to use github, please visit its tutorial[2] for some advice. Codes will be judged by cleanness and readability, so remember to comment your codes.

4.2 Report

Each students is required to turn in a report for that contains your main ideas, methods, algorithms, experimental settings, and results. You can write your report in word or latex (encouraged), English (encouraged) or Chinese. The details of the four parts is:

[2]https://guides.github.com/activities/hello-world/

Main Ideas A brief introduction of your report, including your proposed methods and the performance.

Methods The proposed method for the project, including the motivation and detailed description of your methods.

Algorithms The algorithms you choose to optimise your classifier (how to train your model or update the parameters), including the motivation, detailed description and pseudo-code if necessary.

Experimental Settings The experimental settings, including the structure of your network (if you have), learning rate, batch size, etc.

Results The performance of your model on test set.

References

[1] Ignacio Arganda-Carreras, Srinivas C Turaga, Daniel R Berger, Dan Cireşan, Alessandro Giusti, Luca M Gambardella, Jürgen Schmidhuber, Dmitry Laptev, Sarvesh Dwivedi, Joachim M Buhmann, et al. Crowdsourcing the creation of image segmentation algorithms for connectomics. *Frontiers in neuroanatomy*, 9:142, 2015.