DATS_6203 Final Project Proposal

Group 7 Team Members: Junhe Zhang, Ze Gong, Hao Heng

1.Object detection

The problem we choose is to train a real-time object detection model and use it to detect some objects of our interests.

The reason we choose it is because object detection has a lot of applications in current times. It has many meanings in our reality.

2.Dataset

- **Google Open Image Dataset**

  Category : Car, Bicycle wheel, Bus, Traffic light, Jeans, Person, Laptop
  Instance : 1000 for each category
  Total images in our dataset: 7000

- **Is it enough?**

  Person and bus may need more instances to train, we also might try to increase the iteration times according to our cloud platform capability.

3.Network

- **CNN network**

  A standard yolov3 network with a total of 106 layers, which has 53 convolutional layers and 3 output layers at layer 82, 94 and 106.

4.Framework

- **Darknet**

  The best real-time object detection framework so far, easy to implement and support both CPU and GPU for the training process.

5.Metrics

- **Mean Average Precision(mAP)**

  mAP (mean average precision) is a metric used to evaluate accuracy. Firstly, Average Precision is calculated for every class in the custom model. Then, mean of these calculated Average Precisions across all classes gives mAP

6.Schedule

- **research 11.24-11.30**
- **data pre-process 12.1-12.2**
- **train network 12.3-12.5**
- **test and result 12.6-12.7**

7.Reference

[1]Joseph Redmon, Santosh Divvala, Ross Girshick, ALi Farhadi “You Only Look Once: Unified, Real-Time Object Detection”, IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2016.

[2]Open Images Dataset V6 + Extensions. (n.d.). Retrieved December 08, 2020, from https://storage.googleapis.com/openimages/web/index.html

[3]Joseph R. (2013-2016). Darknet: Open Source Neural Networks in C. Retrieved from
http://pjreddie.com/darknet/