6202_proposal

Group member: Jiajian Chang; Chen-yin Yu; Dongxiao Ma

For our group, we decide to choose the 'Open Image Dataset' from the links of shahabkam's Github page as our final project task. It contains a total of 16M bounding boxes for 600 object classes on 1.9M images in the current version which is enough to train a network. The reason why we choose this topic is because we think this dataset is full of diverse resource and the task how to label these image accurately is challenging (it can let us know the idea about visual recognition). As we all know, media carry massive data in nowadays and we consider label these image can assist people in finding what they look for to save time. Besides, it helps the website firm discover information like their users' favorite kind of image, then recommend the same type of image to them exactly. In this task, we suppose to use the Convolution Neural Networks (CNN) to implement image classification, with the tool package of Tensorflow and Keras. CNN is a kind of standard network. Moreover, we plan to utilize single frame as the framework. Also, during the time we do this project, we will use plenty of reference materials about visual recognition to fulfill our background knowledge such as:

1. Imagenet classification with deep convolutional neural networks
2. Large-scale Video Classification with Convolutional Neural Networks
3. Two-Stream Convolutional Networks

At the end of the project, we will certainly evaluate the performance of our network by the matching accuracy of original human-verified segment labels and our network's. For the schedule, we will read reference and try to figure out how the visual recognition works detailly during July 25-31; we will train the network and finish the project from August 1 to 13; during August 14-19, we will write the report according to the project and submit it to Github on August 20.