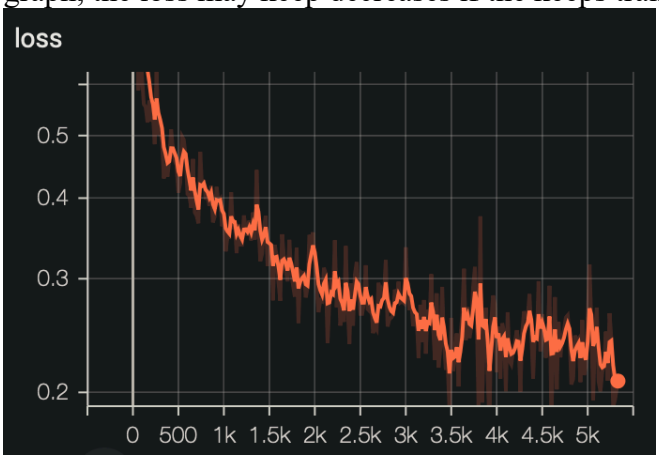
CSC249 Final Project
Hecheng Sun
Task 2: Actor-Action Detection

Method description:

- For this task, I used the provided "A2DCOCO.py" program to preprocess my data.
- The network architecture I used for this task is Faster-RCNN, which is backed by ResNet50. I also initialize the network with the MS-COCO pretrained model.
- The optimization method is SGD optimization, with a learning rate of 0.0002. There are 6 different losses: total loss, bounding box loss, actor classification loss, action classification loss, RPN classification loss, and RPN bounding box loss.
- All other hyperparameters should be the default values in the given skeleton code, except the number of workers is changed from 0 to 8. The total iteration of the best model is 5332. The batch size is 3, and the segment length is 2. According to the tensor board graph, the loss may keep decreases if the keeps training with more iterations.



Novelty of the method:

- There are a few novelties of this implementation.
- First, it uses the technique of transfer learning to initialize the network, making the convergence easier and faster.
- Second, it uses neighbor frames to get segments of the original videos and calculates optical flows by Gunnar-Farneback optical flow estimation algorithm.

Performance on validation set:

- Actor mAP: 0.7012652697412004
- Action mAP: 0.2546470109356425
- Actor-action mAP: 0.25886174809280044