Visual Attributes in Video

Update Presentation Week 7
Marielle Morris & Mahdi Kalayeh
July 7, 2017
Last Week

- BeaverDam base system final modifications
- Test cases & bug fixing
- Initial filtering of YT-BB dataset
- Cost estimation based on final video count
This Week

• Hosting BeaverDam with public Amazon domain
  ○ Server issues
  ○ Cross-browser compatibility
  ○ Security and required SSL protocols

• Publishing HITs to MTurk
  ○ Modifications in Sandbox
  ○ Bugs in embedding Django application and public domain into MTurk

• Deep Learning techniques and theory

• Literature review for method
BeaverDam: Status Update

- Tasks published to Amazon Mechanical Turk Sandbox
  - Successful trial run
  - Still need: instructions, GUI fix, script for data collection and task approval
- Plan: pay $0.02 per video task
  - Based on: average 20s video clips
  - Require 5 attribute annotations per video
- Dataset: 20,000 total videos, 10 classes
  - Show 20 attributes per video, of 60 total
  - Assign each video as 3 tasks, to cover all attributes
- Total Cost:
  
  \[ 20,000 \text{ vids} \times 3 \text{ HITs/vid} \times \$0.02/\text{HIT} + 20\% \text{ AMT fee} = \$1,440 \]
Literature Review

- Learning Spatiotemporal Features with 3D Convolutional Networks - Tran et al.

- Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks - Ren et al.

- T-CNN for Action Detection in Videos - Rui Hou et al.
Deep Learning Model

- Review and understand C3D model

- Baseline implementation plan
  - Run C3D on base YT-BB dataset
  - Use model for classification
  - When AMT annotation is finished, modify loss function for attributes
Next Week

- Understand C3D model
- Train for classification on YT-BB, with modification plan
- Publish 100 video subset to AMT
  - Quality control and bugfixing
  - User friendly instructions and UI
- Reduce dataset to 20,000 videos and set up full AMT task