Much Ado About Time:
Exhaustive Annotation of Temporal Data

Gunnar A. Sigurdsson (CMU), Olga Russakovsky (CMU), Ali Farhadi (UW, AI2), Ivan Laptev (INRIA), Abhinav Gupta (CMU)

http://allenai.org/plato/charades/

Published in HCOMP’16

Goal: explore cost-effective strategies for annotating temporal data (e.g., videos)

Dataset: Charades video dataset of [Sigurdsson ECCV16]

Experiment #1: Run one round of annotation
Annotate 157 actions in 140 videos on Amazon Mechanical Turk

Experiment #2: Improve and re-evaluate the UI
- Improving time/cost: Keep questions consistent within a HIT
  - 13.6% reduction in annotation time
- Improving precision: Use the available annotations to inject more positive questions into the HIT
  - 2.3% improvement in precision
- Improving recall try #1: First request a 20-word video summary
  - No effect on recall but 40% slower annotation
- Improving recall try #2: Force yes/no response to every question
  - Recall actually drops (annoys workers?)
- Improving recall for real: Run multiple rounds of annotation
  - Recall improves from 58.0% to 83.3% after 3 rounds

Experiment #3: Bringing it all together

Conclusion: Asking many questions per video is a much more effective use of worker time than asking just a few questions

Annotated data: 9,848 videos exhaustively annotated
- 66,963 action instances from 1,310,014 collected annotations
- Increased density of annotation from 3.7 to 9.0 labels per video
- Additionally collected temporal annotations
- All publicly available at http://allenai.org/plato/charades/

How many questions about a video can we ask simultaneously?
Few questions: questions can be kept in worker’s short-term memory, increasing annotation recall
Many questions: video watching time can be amortized, reducing annotation cost

Instance:
Defined as a link to a video of one or two people, please watch each video and answer the questions:
- You will receive multiple videos, each followed by few questions.
- The number of videos and questions is balanced such that the task should take 5 minutes.
- Make sure you fully and carefully watch each video so you do not miss anything. This is important.
- It is possible that many of the actions in the HIT do not exist. It is important to verify an action is indeed present in the video.
- Check all that apply if there are any doubts, check it anyway for good measure.
- Read each and every question carefully. Do not take shortcuts, it will cause you to miss something.

Annotate 157 actions in 140 videos on Amazon Mechanical Turk

Analyze worker behavior during annotation
- Worker enjoyment doesn’t depend on number of questions
- Workers interact with the video more when asked more questions
- Workers annotate fewer actions when asked more questions

Average Question Time

Video Questions

Average time to annotate a video [min]

Average time to annotate a video [sec]