CRCV HSAP PRESENTATION

EMILY PARK
May 13, 2021

DATA COLLECTION FOR KRISHNA

Ground Videos: 80-90 new (~640 total)

Aerial Videos: 80-90 new (~648 total)

Total: 200 new videos (~1488 total videos collected)

+ 10 total cities

RESEARCH PAPER READING (1)

Computer Vision and Image Understanding: The THUMOS challenge on action recognition for videos "in the wild"

- **Action recognition**
  - For action recognition, each system is expected to output a real valued score for confidence, for which they calculated with the equation of AP(c)
- **Temporal Detection:**
  - Evaluated for twenty classes of instantaneous actions in all test videos; the system is expected to give a real-valued score indicating the confidence of the prediction, as well as the starting and ending time for the given action

RESEARCH PAPER READING (2)

Computer Vision and Image Understanding: The THUMOS challenge on action recognition for videos "in the wild"

- <u>Methodology:</u>
  - ***1) Classification***
    - Most teams adopted the 2 features:
      - 1. Deep Learning
      - 2. Improved Dense Trajectories (iDT): most powerful hand-crafted feature for video classification
    - Average Fusion = most popular option due to its simplicity and good generalizability
  - ***2) Temporal Detection***
    - New problem that was recently introduced in THUMOS

RESEARCH PAPER READING (3)

Computer Vision and Image Understanding: The THUMOS challenge on action recognition for videos "in the wild"

- <u>Temporal Detection</u>
  - Make layers through sets extractions and normalizations
  - Features of pyramid score distributions
    - 9 windows for each frame
    - Hypothesis = the scores at the correct window length should be the highest, and should vary smoothly for neighboring temporal resolutions
    - Each windows are normalized → actions are computed → median filtering on output labels for smoothness

Thank you!

Any questions?