

Presentation

...

Data Collection

- 109 Ground Videos
- 134 Aerial Videos
- 15+ Cities (Gaborone, Bandar Seri Begawan, Djibouti, Roseau, San Salvador, Malabo, Tallinn, Brussels, Warsaw, Nottingham, Montreal, Montevideo, Minneapolis, Milan, Chicago, Astana,)

THUMOS

- Objective(s) of THUMOS Challenge:
 - 1) Serve as a benchmark and enable a comparison of different approaches on the tasks of action classification and temporal detection in large-scale realistic video settings
 - 2) Advance the state of the art/accuracy
 - Ex) The accuracy of UCF101 increased from 45% in 2012 to 90% in THUMOS'13



THUMOS

- The THUMOS action classes are from UCF101 and can be divided into five categories (videos are annotated and available on Youtube):
 - Human-Object Interaction
 - Body-Motion Only
 - Human-Human Interaction
 - Playing Musical Instruments
 - Sports

THUMOS Challenge

- Annotation/Verification Procedure
 - If a positive video was found, it was marked as either “Positive” or “Irrelevant” based on the following factors:
 - Slow Motion
 - Sped Up
 - Occlusions/Partial Visibility
 - Motion Blur
 - Clutter/Incorrect Background
 - Unrealistic Instances
 - Animation
- Temporal Annotation
 - Action Boundaries are more abstract than concrete
 - To solve the problem, the 101 action classes were divided into two categories:
 - Instantaneous actions which have a short time span
 - Ex(s): Basketball Dunk
 - Cyclic Actions that are repetitive
 - Biking
 - Playing Guitar

THUMOS

- Positive Videos
 - Uses keywords to find a reasonable set of potential video for actions
 - Some videos are that frequently viewed are “viral videos”
 - Had to blacklist keywords that may lead to viral videos
- Background Videos
 - The best videos are those that share the context of a given action without actually showing instances of the given action
 - An example would be the PlayingPiano Class
 - Shows a piano where the piano is not being played in a background video

THUMOS

- In Action Recognition
 - Each system must produce a real-valued score that shows the confidence of the predicted presence/action in the video
 - Confidence score is between 0 and 1
- Interpolated Average Precision (AP)
 - Used as official measure for evaluating results for each action class
 - Detector Performance for an action class is evaluated by AP

Paper Idea

Use Weather Patterns in Videos to predict the traffic of vehicles