

# THUMOS 2015

## Action Localization Track - Summary

Amir R. Zamir

Alexander Gorban, Haroon Idrees, Yu-Gang Jiang, Ivan Laptev, Mubarak Shah, Rahul Sukthankar



# Action Localization

- Where in the video the action takes place
  - Temporal
  - Spatio-temporal

# Action Localization

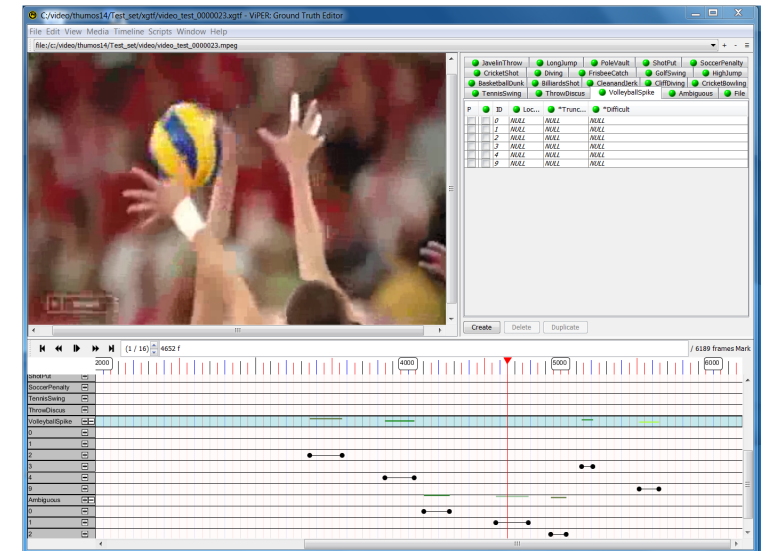
- Where in the video the action takes place
  - Temporal (THUMOS14, 15)
  - Spatio-temporal (THUMOS13)



Trimmed Spatio-temporal (TH13)



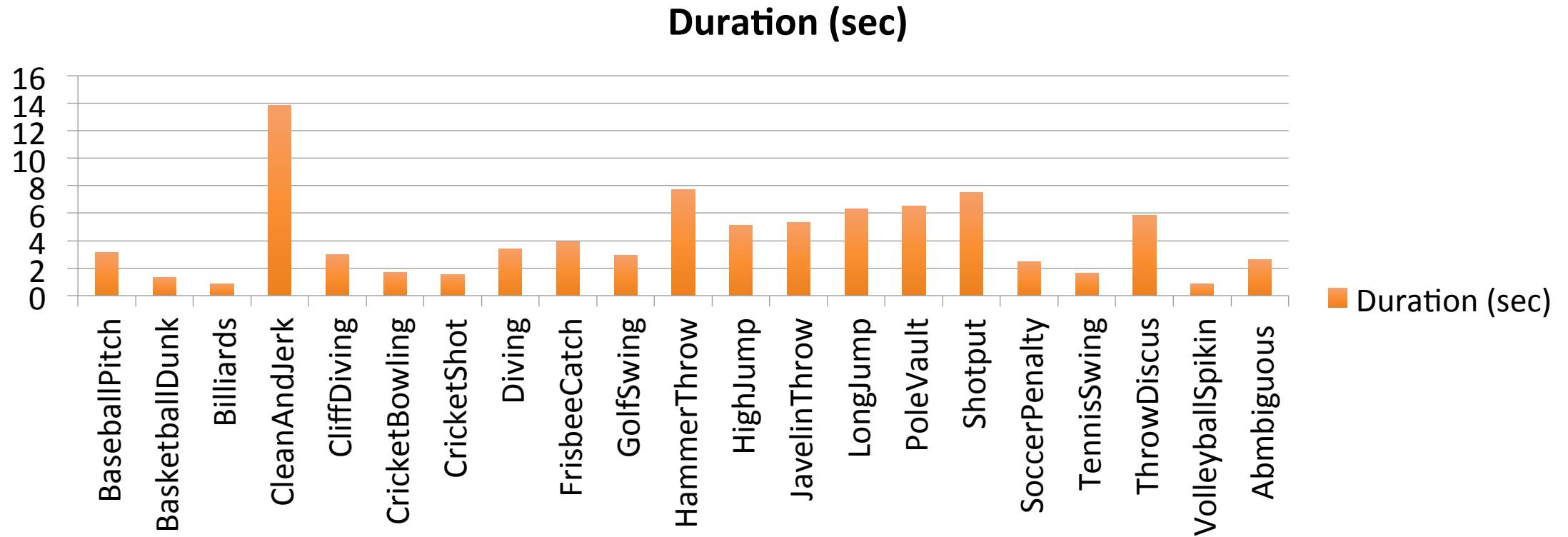
Untrimmed (TH14 & 15)



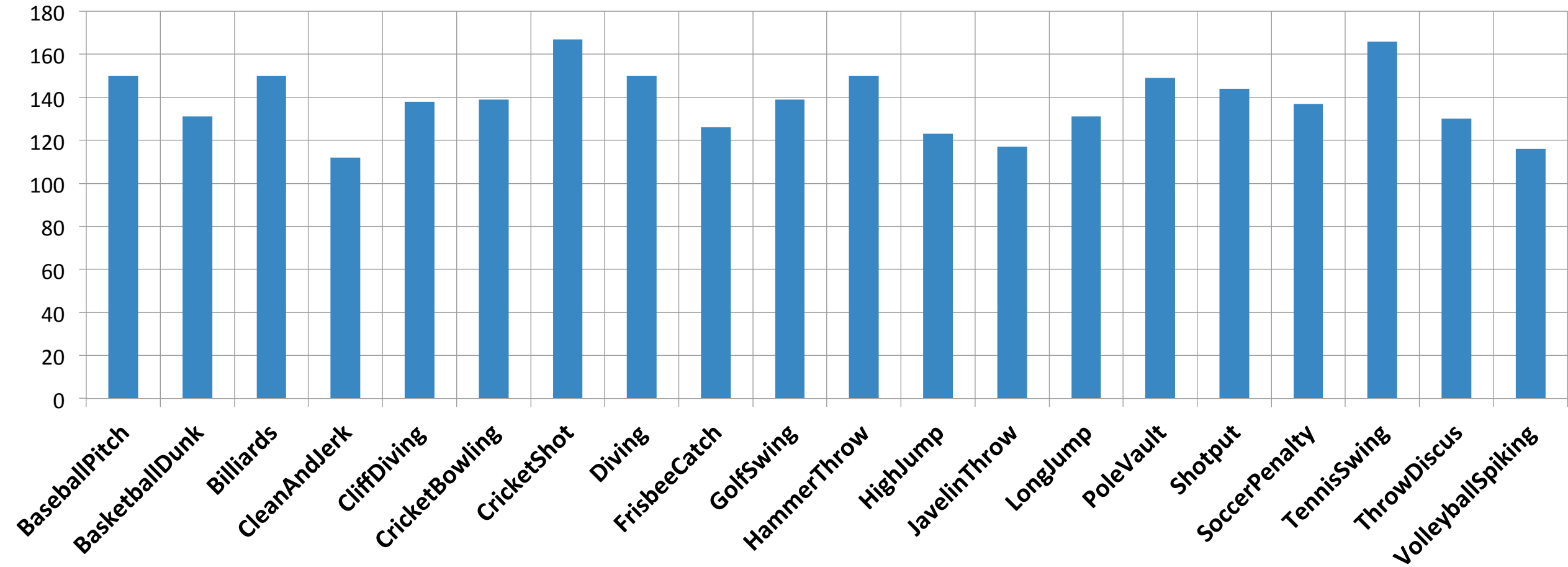
# THUMOS15

## Action Localization Statistics

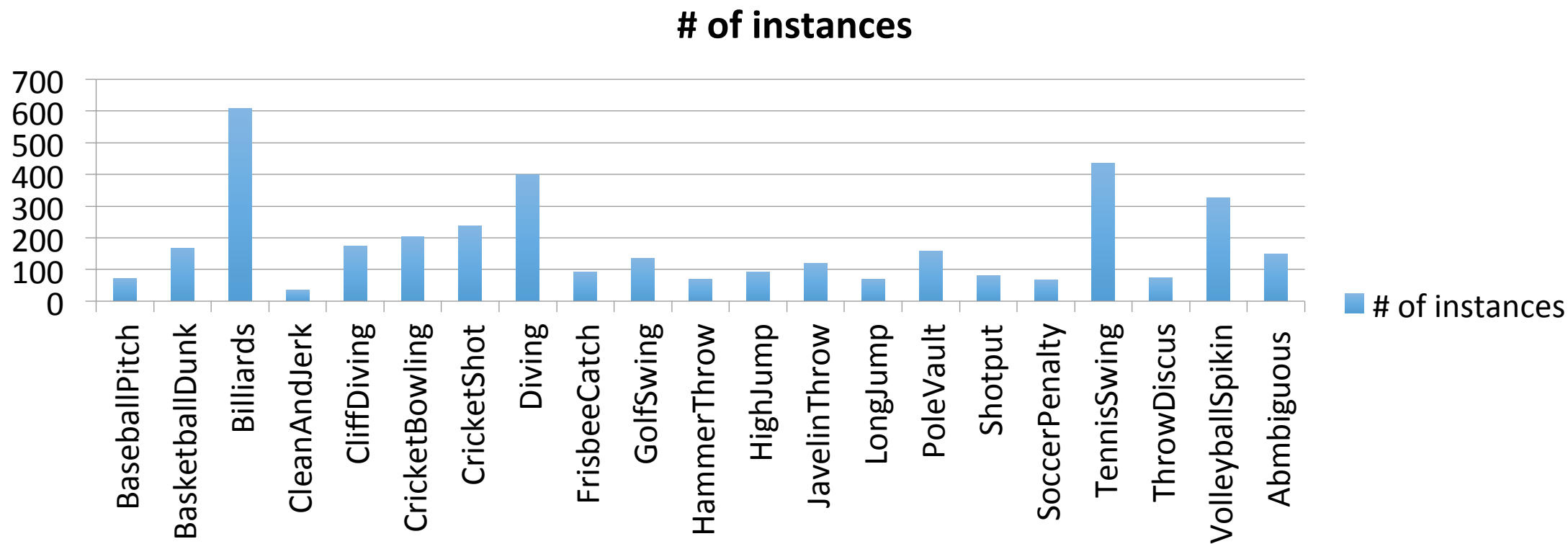
# Instance Duration (Validation)



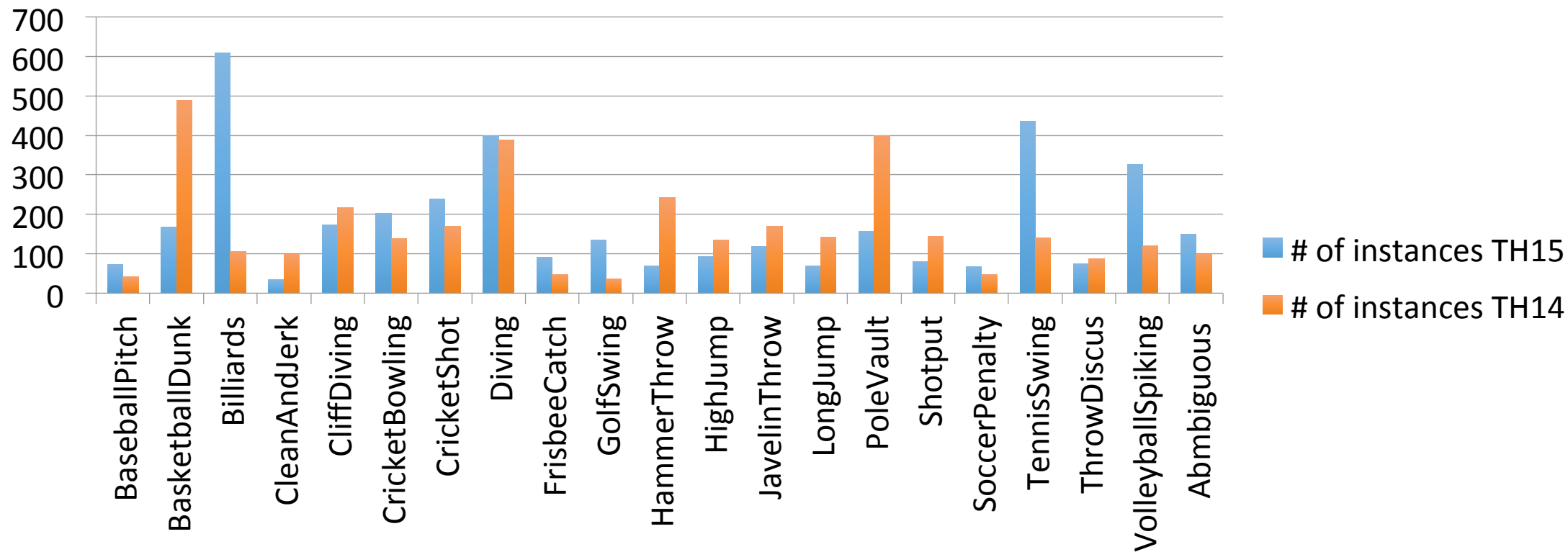
# Number of Clips Per Class (Training)



# Number of Clips Per Class (Validation)

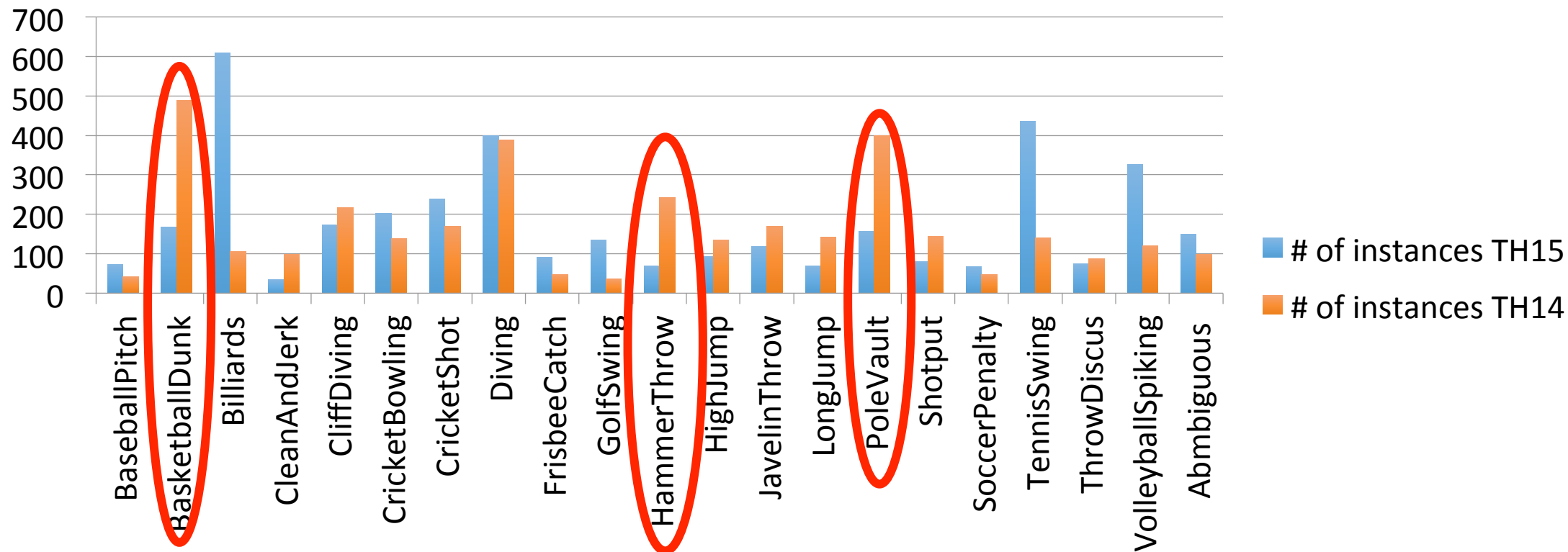


# Number of Clips Per Class (Validation)



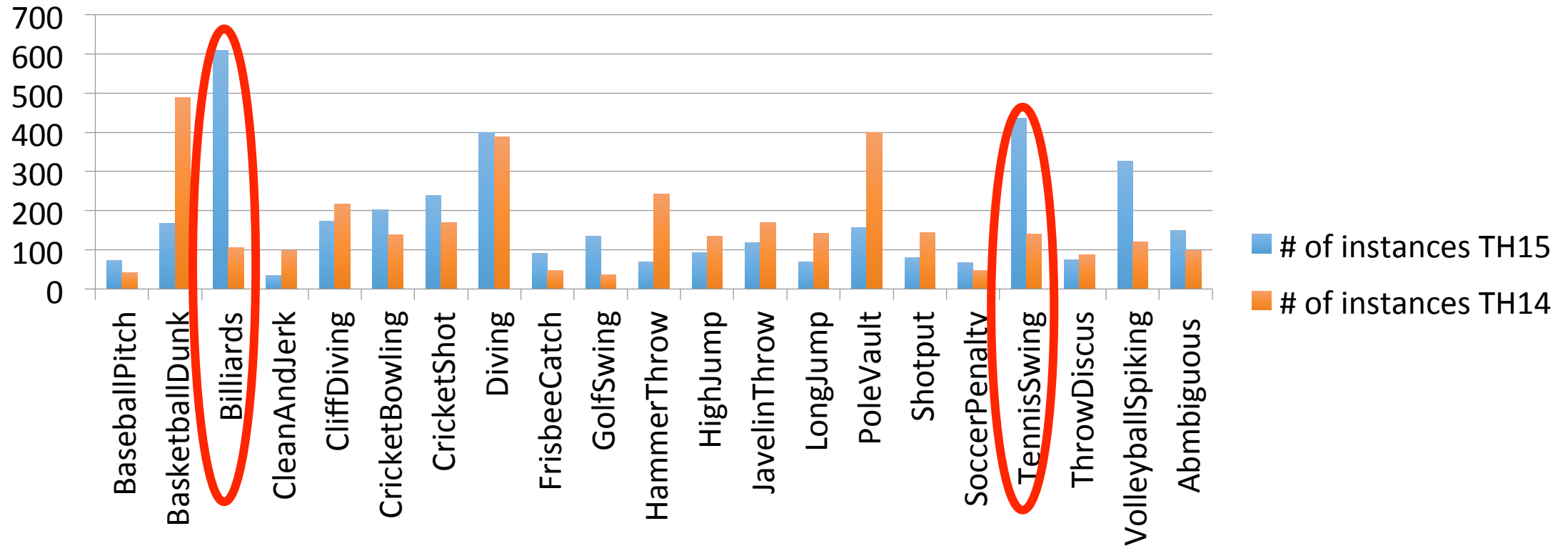


# Number of Clips Per Class (Validation)



**Compilation videos!**

# Number of Clips Per Class (Validation)



**“Burst action” videos!**

# “Burst Action” videos!

The screenshot displays a video annotation software window. The main video player shows a pool player in a blue shirt and black pants, leaning over a pool table. The background features logos for 'KAMMI', 'Mezz Cuis', 'Delta', and 'Rasson'. The software interface includes a menu bar (File, Edit, View, Media, Timeline, Scripts, Window, Help) and a toolbar with various action labels such as PoleVault, Shotput, SoccerPenalty, TennisSwing, ThrowDiscus, VolleyballSpiking, Ambiguous, File, CricketShot, Diving, FrisbeeCatch, GolfSwing, HammerThrow, HighJump, JavelinThrow, LongJump, BaseballPitch, BasketballDunk, Billiards, CleanAndDerk, CliffDiving, and CricketBowling. A table on the right side of the window lists 23 rows of data, with the 9th row checked. The table has columns for ID, Loc..., \*Trunc..., and \*Difficult. Below the table are buttons for 'Create', 'Delete', and 'Duplicate'. At the bottom, a timeline shows frame markers and a red vertical line indicating the current frame position at 1307 f.

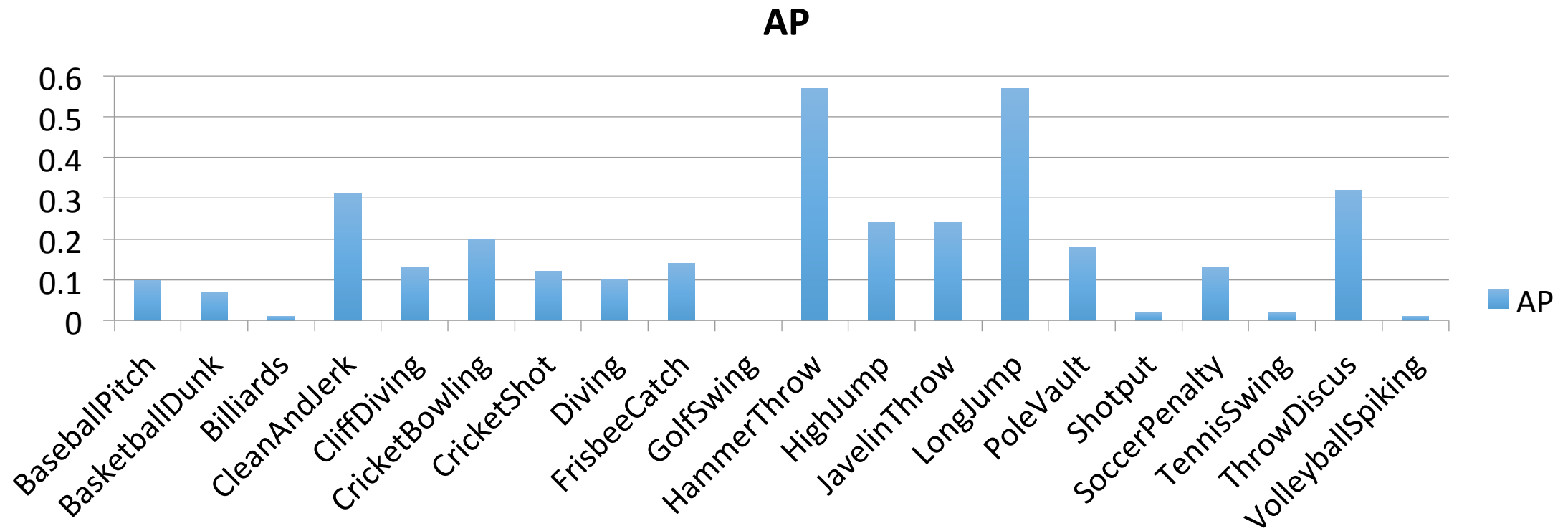
P	ID	Loc...	*Trunc...	*Difficult
<input type="checkbox"/>	0	NULL	NULL	NULL
<input type="checkbox"/>	1	NULL	NULL	NULL
<input type="checkbox"/>	2	NULL	NULL	NULL
<input type="checkbox"/>	3	NULL	NULL	NULL
<input type="checkbox"/>	5	NULL	NULL	NULL
<input type="checkbox"/>	6	NULL	NULL	NULL
<input type="checkbox"/>	7	NULL	NULL	NULL
<input type="checkbox"/>	8	NULL	NULL	NULL
<input checked="" type="checkbox"/>	9	NULL	NULL	NULL
<input type="checkbox"/>	10	NULL	NULL	NULL
<input type="checkbox"/>	11	NULL	NULL	NULL
<input type="checkbox"/>	12	NULL	NULL	NULL
<input type="checkbox"/>	13	NULL	NULL	NULL
<input type="checkbox"/>	14	NULL	NULL	NULL
<input type="checkbox"/>	15	NULL	NULL	NULL
<input type="checkbox"/>	16	NULL	NULL	NULL
<input type="checkbox"/>	17	NULL	NULL	NULL
<input type="checkbox"/>	18	NULL	NULL	NULL
<input type="checkbox"/>	19	NULL	NULL	NULL
<input type="checkbox"/>	20	NULL	NULL	NULL
<input type="checkbox"/>	21	NULL	NULL	NULL
<input type="checkbox"/>	22	NULL	NULL	NULL
<input type="checkbox"/>	23	NULL	NULL	NULL

Handled using “ambiguous” label.

# Results (Advanced Digital Sciences Center, Singapore)

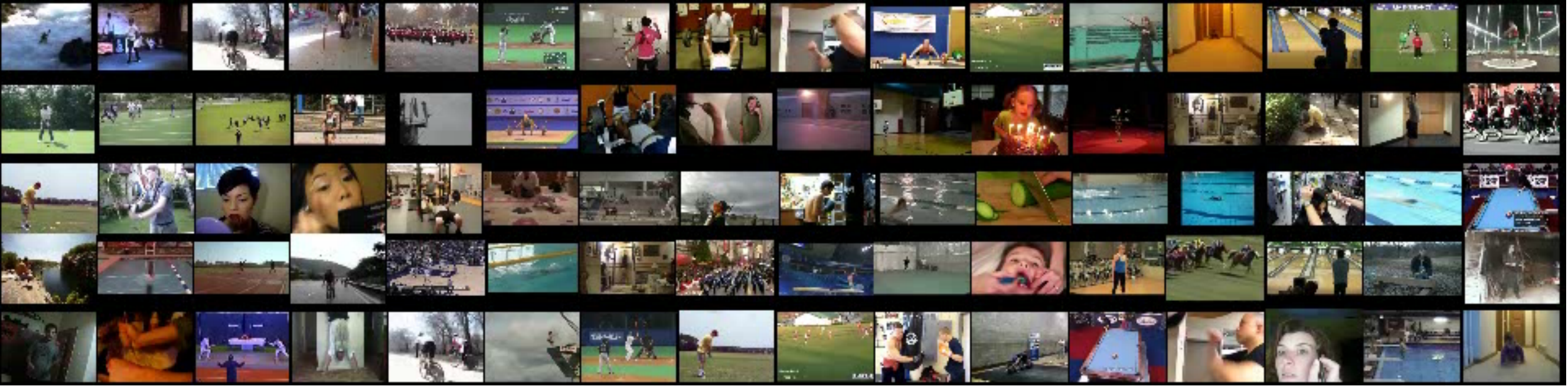
<b>Run/mAP</b>	<b>overlap=0.1</b>	<b>overlap=0.2</b>	<b>overlap=0.3</b>	<b>overlap=0.4</b>	<b>overlap=0.5</b>
<b>run1</b>	<b>0.408623126</b>	<b>0.362934338</b>	<b>0.307624655</b>	<b>0.235190774</b>	<b>0.182980846</b>
<b>run1</b>	0.111525203	0.091928701	0.077222284	0.059726235	0.042117036
<b>run2</b>	0.161122549	0.134905616	0.107279919	0.082979258	0.056195248
<b>run3</b>	0.157680431	0.134554141	0.111701514	0.088168933	0.065211717
<b>run4</b>	0.138596713	0.115371517	0.093920344	0.072763489	0.050972089
<b>run5</b>	0.141321244	0.118000274	0.098005909	0.077266321	0.05524482

# Class AP (Advanced Digital Sciences Center, Singapore)



# Discussion

- **Question: why not action localization?**
  - Temporal vs spatio-temporal?
  - Action localization is boring?
  - Too computationally demanding?
    - Need features?
  - Not enough time?
    - Release data earlier?
    - Move deadline?
  - Want more data for your ConvNets?



# THUMOS 2015

## Action Localization Track – Summary

Amir R. Zamir ([zamir@cs.stanford.edu](mailto:zamir@cs.stanford.edu))

Alexander Gorban, Haroon Idrees, Yu-Gang Jiang, Ivan Laptev, Mubarak Shah, Rahul Sukthankar

