

Unsupervised and Semi-Supervised Domain Adaptation for Action Recognition from Drones



Jinwoo Choi^{1,2}

Gaurav Sharma²

Manmohan Chandraker²

Jia-Bin Huang¹

¹Virginia Tech

²NEC Labs America

Dataset available at:
<http://bit.ly/NEC-Drone>



Why Action Recognition from Drones?

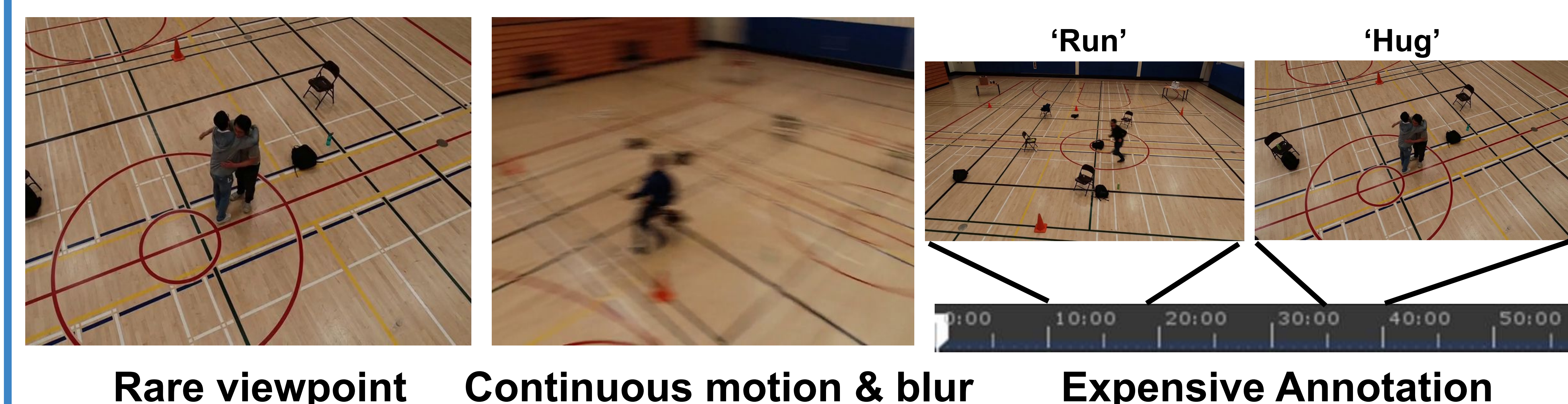


NEC-Drone Dataset



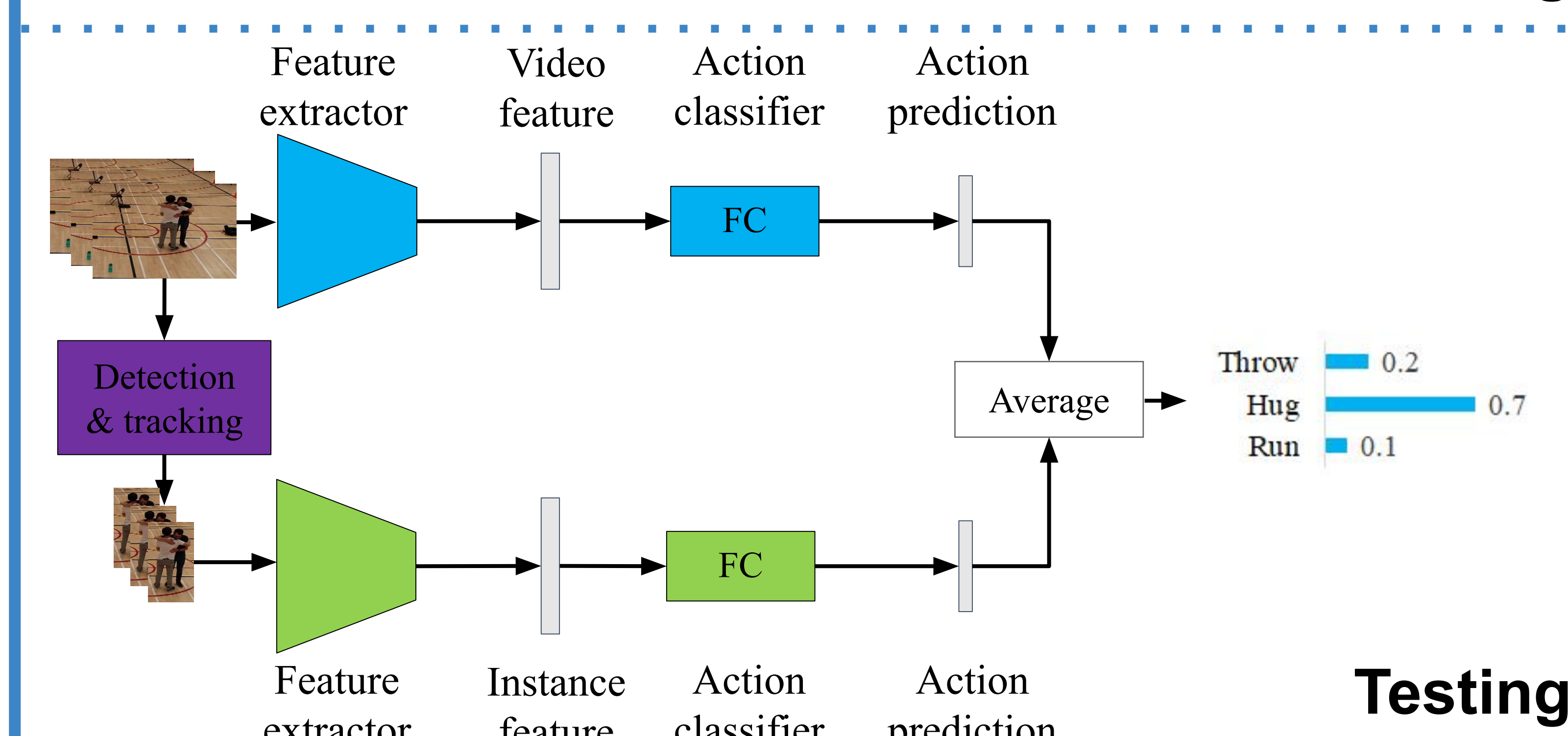
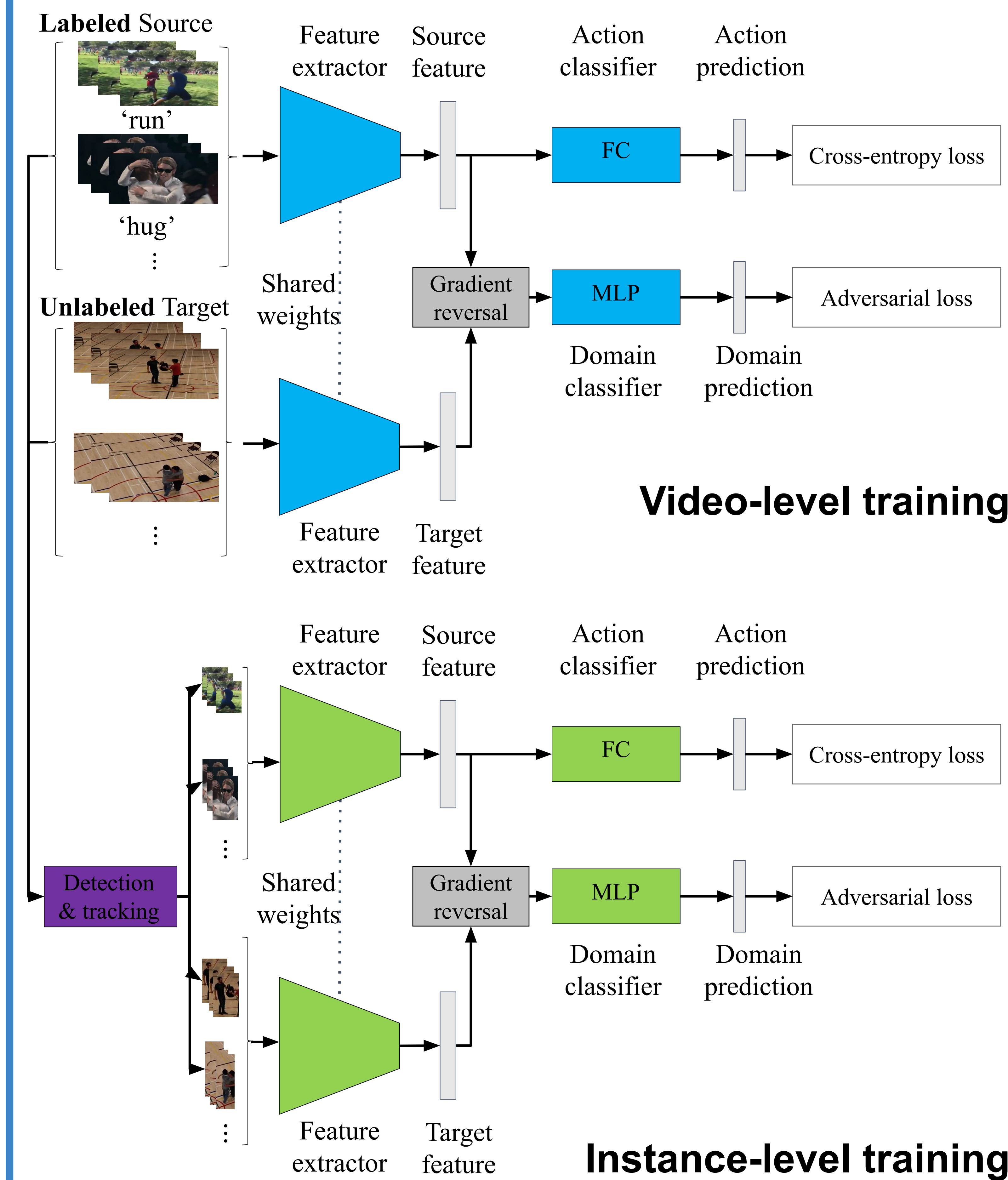
- 5,250 full HD videos
 - 1,188 train
 - 437 val
 - 454 test
 - 3,161 unlabeled
- Captured by two drones
- 16 action categories
- 19 actors

Challenges

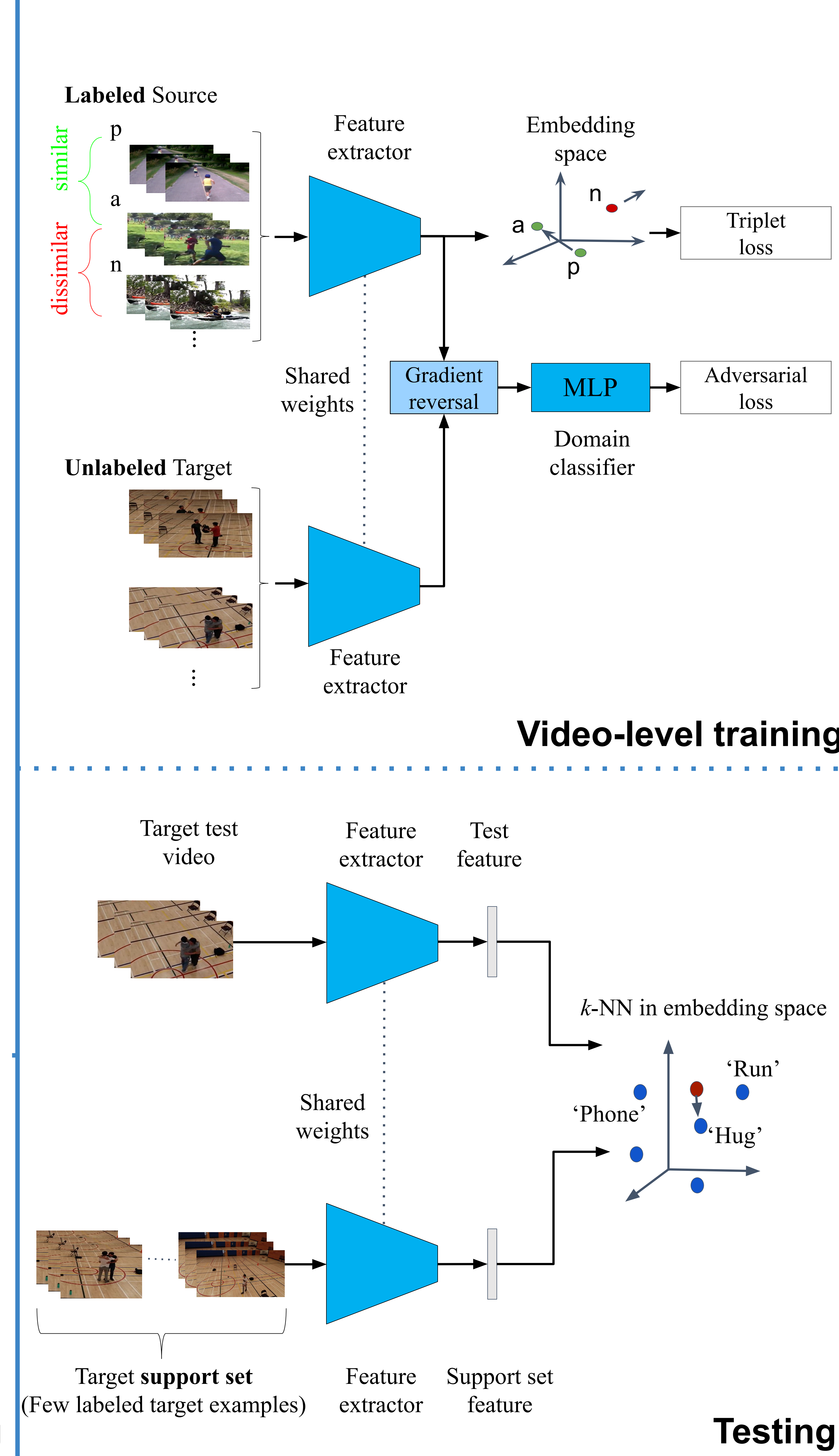


Proposed Method

Same source & target label sets



Different source & target label sets



Experimental Results

Same source & target label sets

Main results	
Method	Accuracy
Source only	13.6% (reference)
Video-based DA*	27.2% (+13.6)
Instance-based DA*	29.4% (+15.8)
Both video & instance DA*	32.0% (+18.4)
Fully supervised on target	69.3%

*Using five target labeled examples per class

Ablation: # of target labeled examples per class used

	m = 0	m = 3	m = 5	m = 10	m = 20
Instance-based w/o. DA	12.6	31.1	35.4	43.2	49.5
Instance-based w. DA	16.5	31.6	39.3	41.3	52.4
Video-based w. DA	13.6	24.3	35.4	53.9	52.9
Both video & instance DA	15.1	32.0	41.8	54.9	58.3

Different source & target label sets

Main results	
Method	Accuracy
Source only	8.2% (reference)
Video-based DA*	18.7% (+10.5)
Instance-based DA*	21.6% (+13.4)
Both video & instance DA*	22.5% (+14.3)
Fully supervised on target	68.3%

*Using five target labeled examples per class

Ablation: # of target labeled examples per class used

	m = 3	m = 5	m = 10	m = 20
Instance-based w. DA	15.9	21.6	31.3	34.6
Video-based w. DA	12.8	18.7	29.1	34.4
Both video & instance DA	18.1	22.5	36.1	39.7