ECCV DeeperAction Challenge

SportsMOT Track on Multi-actor Tracking

Yutao Cui, Xiaoyu Zhao, Chenkai Zeng, Yichun Yang, Gangshan Wu, Limin Wang

State Key Laboratory for Novel Software Technology, Nanjing University, China

Track 3, DeeperAction, ECCV 2022

Deeper Action

Outline

ECCV TEL AVIV 2022

Part 1: Motivation

Part 2: SportsMOT Dataset

Part 3: SportsMOT Challenge

Part 1: Motivation

Motivation: MOT

- **Multi-Object Tracking**
  - Foundational task for high-level action recognition

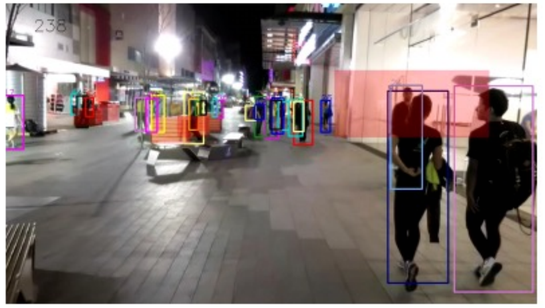
- **Input: Video**
- **Output: Tracks (Bounding Boxes & IDs)**

Deeper Action

MOT Benchmarks

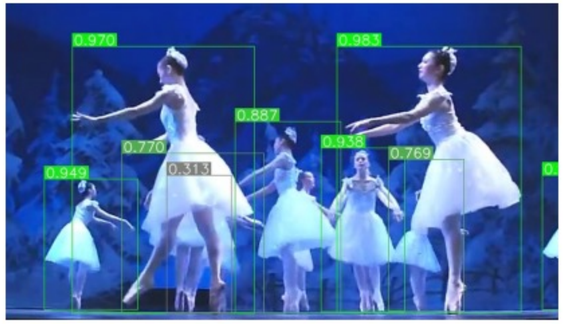
ECCV TEL AVIV 2022

• MOT Challenge[1][2]



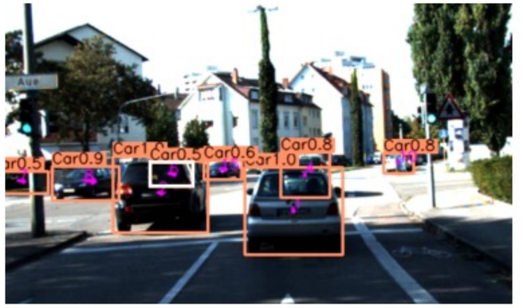
- Pedestrians
- Varying scales
- Simple motion
- Few videos

• DanceTrack[3]



- Dancers
- Nearly static camera
- Uniform appearance
- Complex motion

• KITTI[4]



- Pedestrians & Vehicles
- Driving scenes
- Limited motion
- Stereo & laser information

[1] MILAN A, LEAL-TAIXÉ L, REID I, et al. MOT16: A benchmark for multiobject tracking[J]. arXiv preprint arXiv:1603.00831, 2016.
[2] DENDORFER P, REZATOFIGHI H, MILAN A, et al. Mot20: A benchmark for multi object tracking in crowded scenes[J]. arXiv preprint arXiv:2003.09003, 2020.
[3] SUN P, CAO J, JIANG Y, et al. DanceTrack: Multi-Object Tracking in Uniform Appearance and Diverse Motion[J]. arXiv preprint arXiv:2111.14690, 2021.
[4] GEIGER A, LENZ P, URTASUN R. Are we ready for autonomous driving? the kitti vision benchmark suite[C]//2012 IEEE conference on computer vision and pattern recognition. 2012: 3354-3361.

Motivation

- **Applications**
  - MOT in sports scenes
  - Sports scenarios: player tracking, strategy analysis, etc.
- **Focus: Tracking in Sports Scenes**
  - Visually comfusable
  - Complex motion
  - Fine annotations
- **Encourage the Community**
  - Concentrate more on the complicated sports scenes

Part 2: SportsMOT Dataset

Deeper Action

SportsMOT Dataset

ECCV TEL AVIV 2022

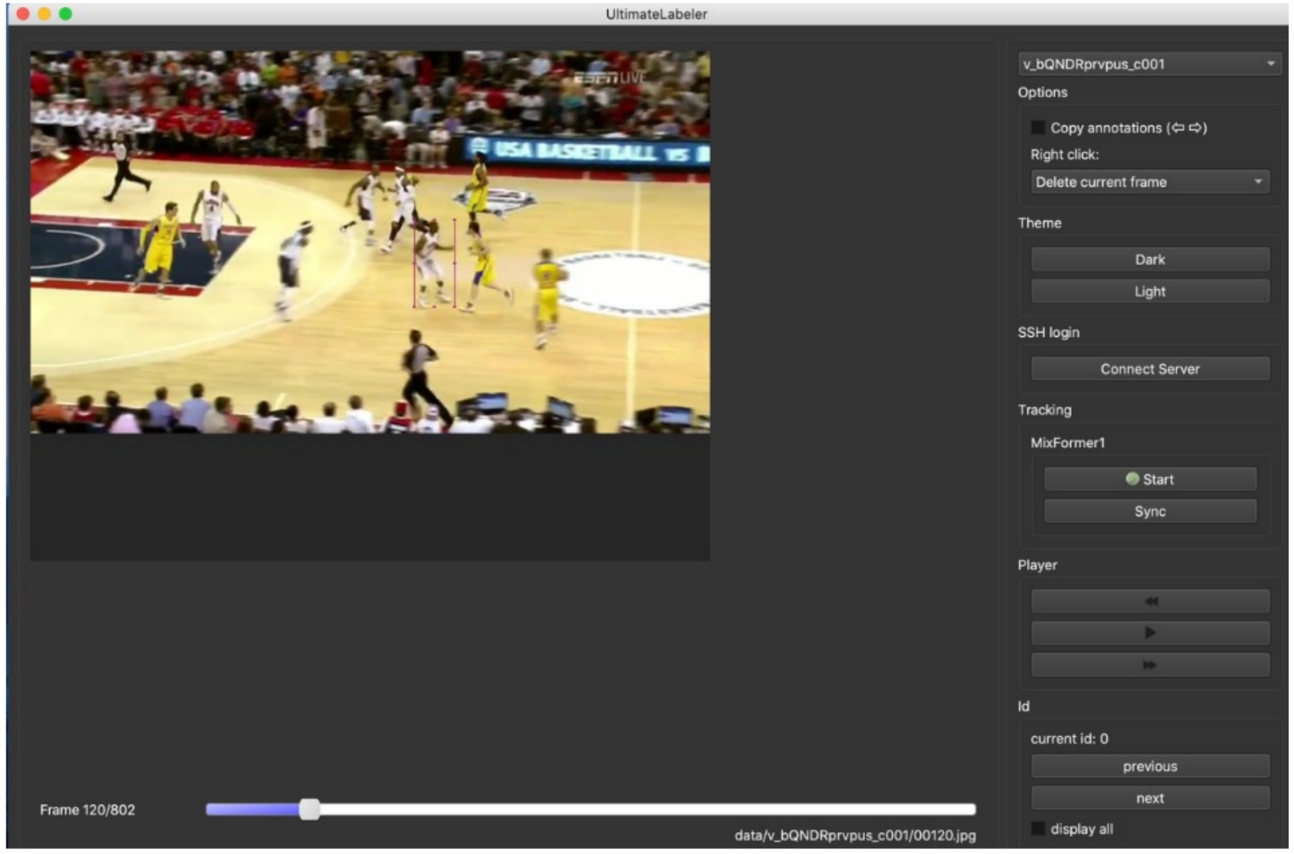
- **A Large-Scale Multi-Object Tracking Dataset in Sports Scenes**

Deeper Action

Data Annotation

ECCV TEL AVIV 2022

- **Collect Videos**
  - No shooting switch
  - Camera angle and position change
- **Fine Annotations**
  - KCF-based labeling tool
- **Manual Check**
  - Bounding box accurary
  - Consistent ID



Deeper Action

Overview

ECCV TEL AVIV 2022

- **Features**
  - ***1.6M+*** bounding boxes in different sports scenes
    - Total #frames ≈ MOT20 > DanceTrack/ MOT17
  - High resolution(1280*720) & High FPS(25)
  - Confusable appearance
  - Complex motion

Dataset	Videos	Total len.(s)	Avg. len.(s)	Tracks	Images	Bounding boxes
MOT17	14	463	35.4	1342	11235	292733
MOT20	8	535	**66.8**	**3456**	13410	**1652040**
DanceTrack	100	5292	52.9	990	105855	-
SportsMOT	**240**	**6015**	25.1	3401	**150379**	1629490

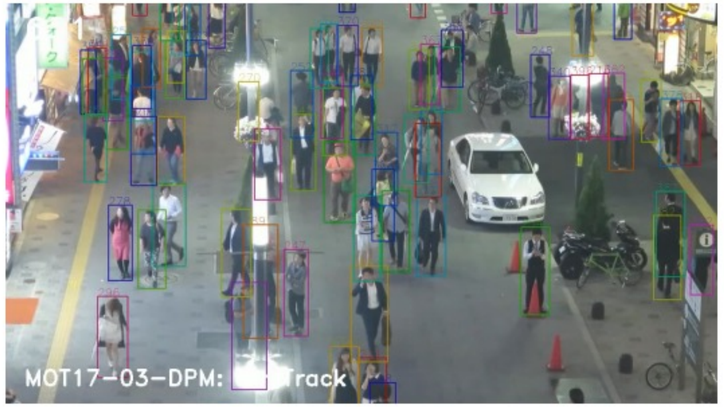
SportsMOT	Avg
#Frames	626.6
#Tracks	14.2
Track gap length	96.6
Track length	479.1
#Bboxes per frame	10.8

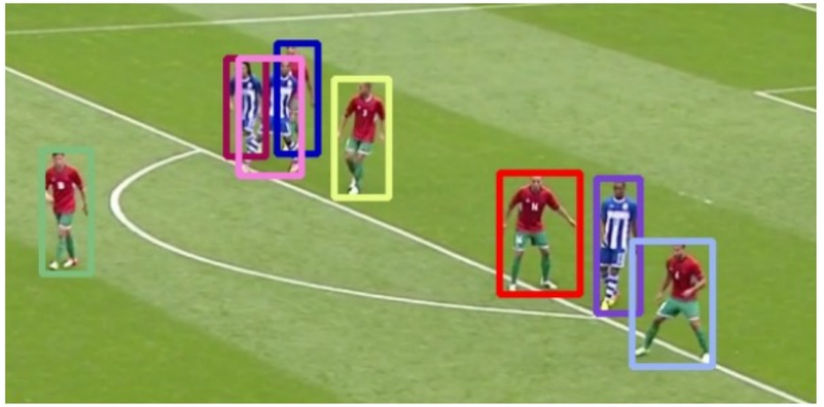
Deeper Action

Analysis: Appearance

ECCV TEL AVIV 2022

- **Re-Identification Feature Space**





MOT 17

SportsMOT

Deeper Action

Analysis: Motion

ECCV TEL AVIV 2022

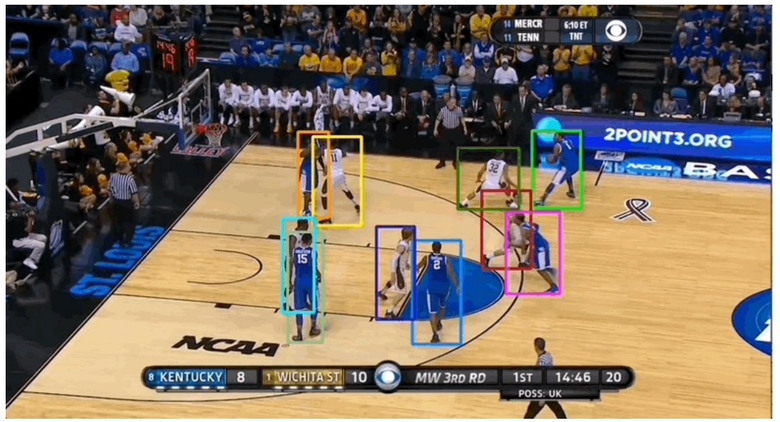
- **Diverse Speed Distribution**
  - Non-uniform speed
  - Significant variation

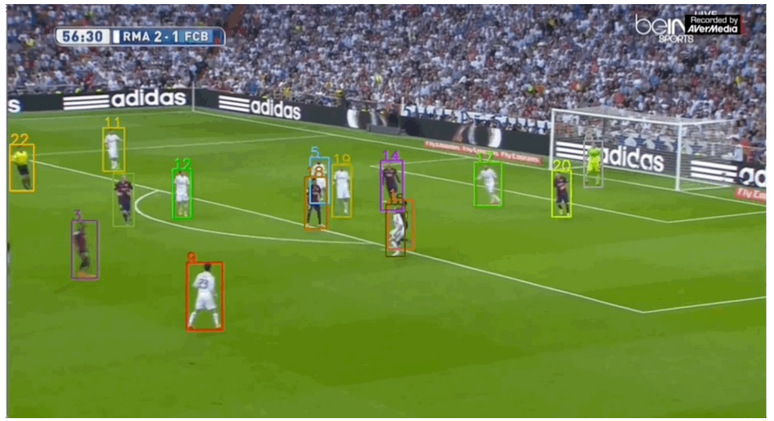
	Basketball	Volleyball	Football
Avg. speed	6.15	9.31	6.64
Std. speed	2.39	5.58	3.91

- **Severe Articulation**

$$\text{deformationRate}(\mathbf{b}_i) = \frac{w_{max} - w_i}{w_{min}} + \frac{h_{max} - h_i}{h_{min}}$$

Category	Basketball	Volleyball	Football	Total avg.
Avg. def. rate	2.17	1.68	2.62	2.26



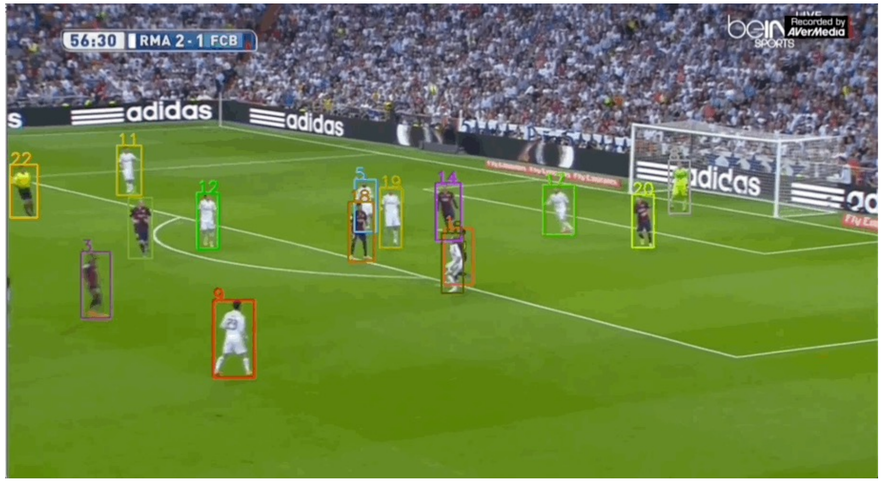


- **Complex Motion Patterns**
  - Discourage strong motion prior
  - Encourage adaptive motion modeling

Deeper Action

Easy Cases

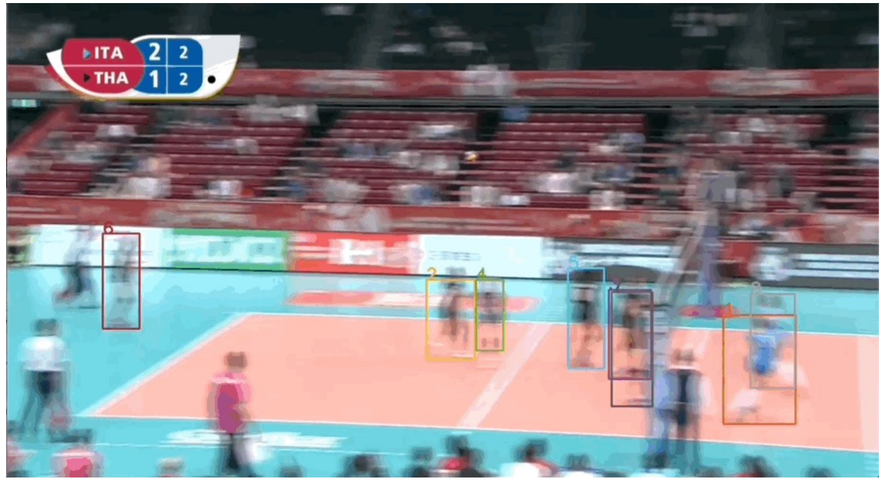
ECCV TEL AVIV 2022

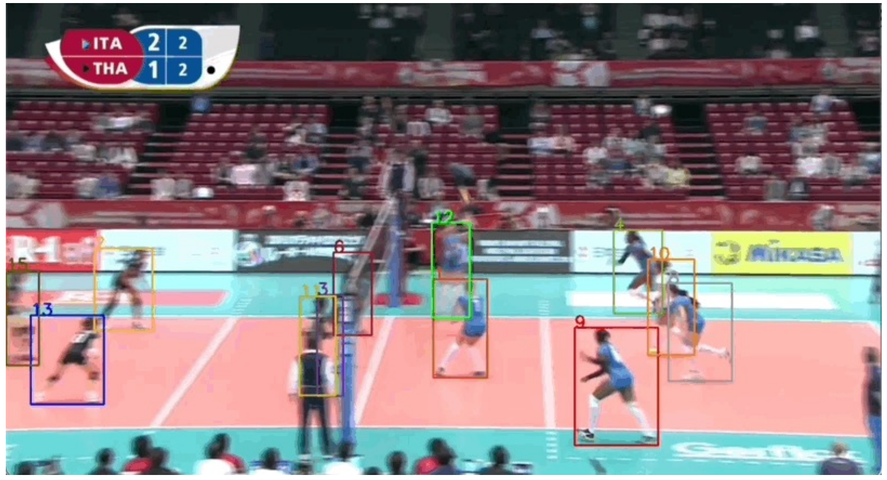


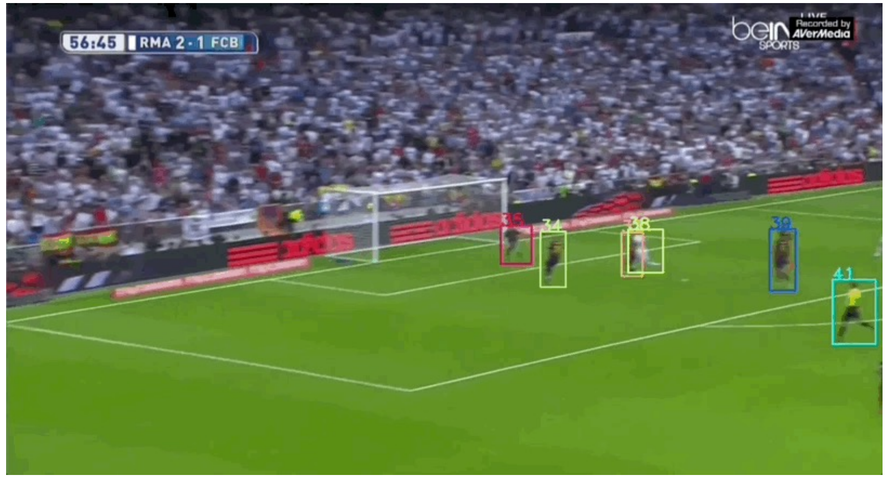
Deeper Action

Hard Cases

ECCV TEL AVIV 2022



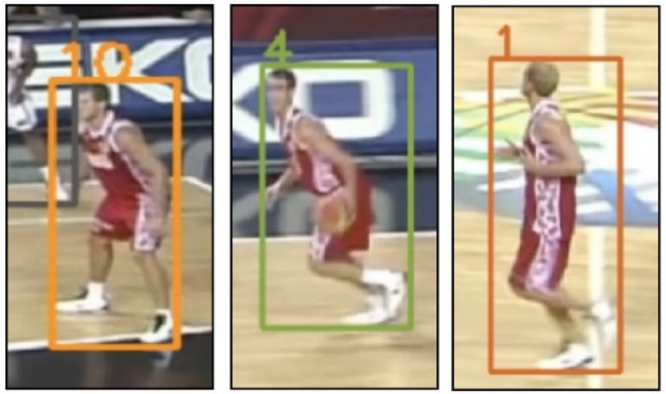


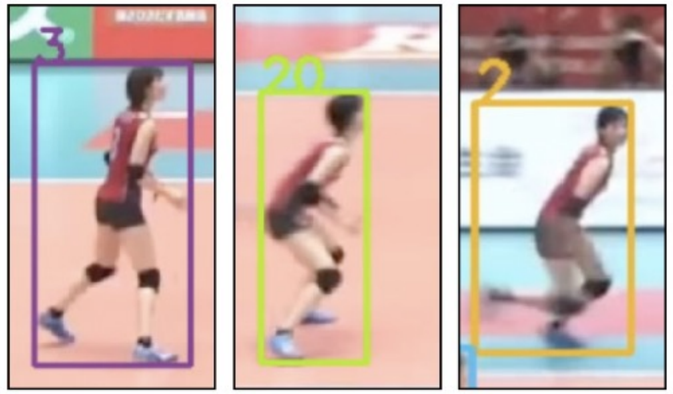


Deeper Action

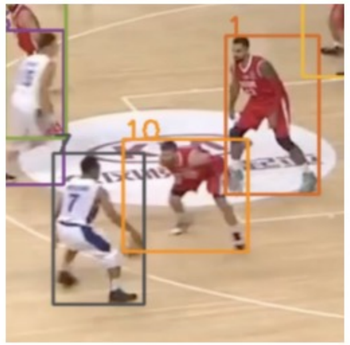
Common Errors

ECCV TEL AVIV 2022





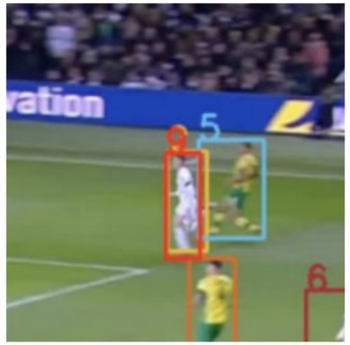
Similar appearance



Inaccurate bbox

Missed detection

ID switch



Duplicate bboxes

False detection

Part 3: SportsMOT Challenge

Metrics

- **HOTA**[5]
  - DetA: detection accuracy
  - AssA: association accuracy
- **MOTA**[6]
- **Identity-based**[7]
  - IDF1
  - IDs
  - Frag

$$\mathrm{DetA}_\alpha = \frac{|\mathrm{TP}|}{|\mathrm{TP}| + |\mathrm{FN}| + |\mathrm{FP}|},$$

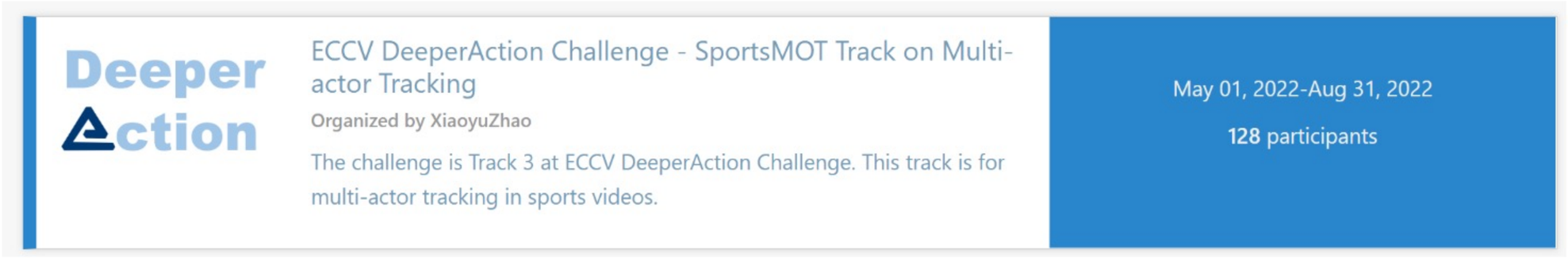
$$\mathrm{AssA}_\alpha = \frac{1}{|\mathrm{TP}|} \sum_{c \in \{\mathrm{TP}\}} \mathcal{A}(c),$$

$$\mathcal{A}(c) = \frac{|\,\mathrm{TPA}(c)|}{|\,\mathrm{TPA}(c)| + |\,\mathrm{FNA}(c)| + |\,\mathrm{FPA}(c)|}.$$

$$\mathrm{HOTA}_\alpha = \sqrt{\frac{\sum_{c \in \{\mathrm{TP}\}} \mathcal{A}(c)}{|\mathrm{TP}| + |\mathrm{FN}| + |\mathrm{FP}|}} = \sqrt{\mathrm{DetA}_\alpha \cdot \mathrm{AssA}_\alpha}.$$

[5] LUITEN J, OSEP A, DENDORFER P, et al. Hota: A higher order metric for evaluating multi-object tracking[J]. International journal of computer vision, 2021, 129(2): 548-578.
[6] K. Bernardin and R. Stiefelhagen, “Evaluating Multiple Object Tracking Performance: The CLEAR MOT Metrics,” EURASIP Journal on Image and Video Processing, vol. 2008, pp. 1–10, 2008, doi: 10.1155/2008/246309.
[7] E. Ristani et al., “Performance Measures and a Data Set for Multi-Target, Multi-Camera Tracking,” arXiv:1609.01775 [cs], Sep. 2016, Accessed: Apr. 12, 2022. [Online].

SporsMOT Challenge



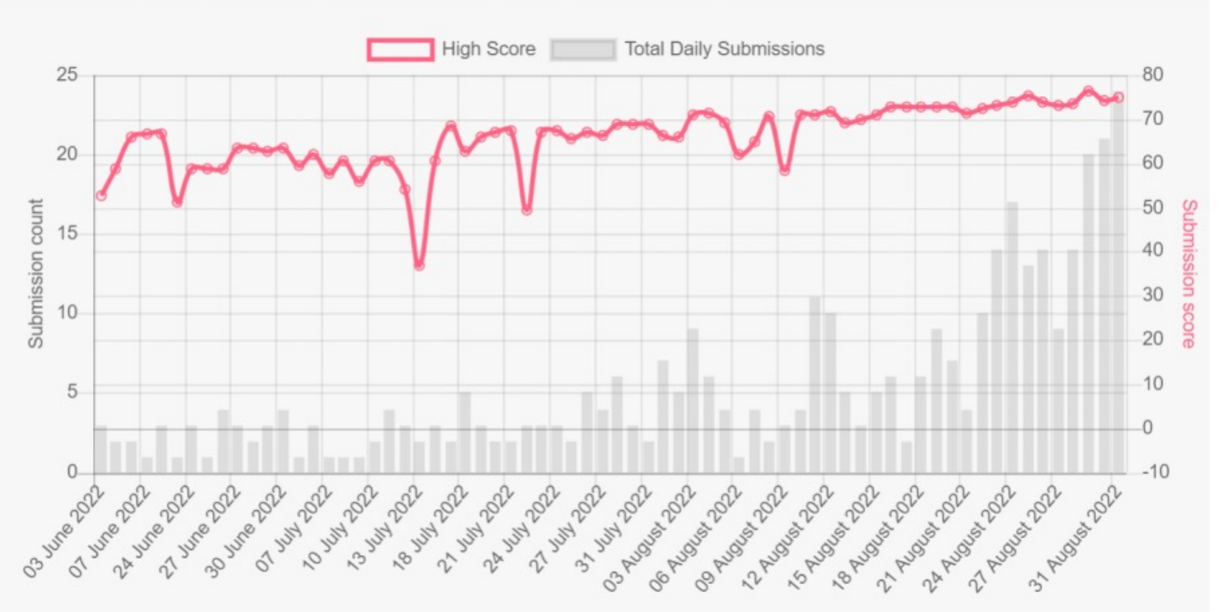
- **Valid participants**
  - 129
  - Industry, academia

Deeper Action

SporsMOT Challenge

ECCV TEL AVIV 2022

- **Development**
  - 228 submissions
- **Test**
  - 210 submissions



Test Set of SportsMOT (Ranked by HOTA)

#	User	Entries	Date of Last Entry	Team Name	HOTA ▲	AssA ▲	DetA ▲	MOTA ▲	IDF1 ▲	IDs ▲	Frag ▲
1	BOE_AIoT_CTO	8	08/29/22	BOE_AIoT_CTO	76.264 (1)	73.538 (1)	79.180 (6)	89.316 (8)	84.453 (1)	2567.0 (1)	6104.0 (5)
2	PingPingPangPangBangBangBang	16	08/31/22	IPIU	74.899 (2)	64.592 (2)	86.968 (2)	95.590 (2)	78.342 (2)	4853.0 (6)	4536.0 (4)
3	hsiangwei0903	10	08/31/22	UWIPL_ETRI	73.968 (3)	63.460 (3)	86.316 (3)	94.832 (3)	78.271 (3)	2754.0 (2)	3592.0 (1)

Deeper Action

1st Place Winner

ECCV TEL AVIV 2022

SportsTrack

Jie Wang*, Xiaodong Yang, Pengyu Zhou, Ting Wang, Yanming Zhang
BOE AIoT CTO
**Corresponding to: bluetornado@zju.edu.cn*

BOE

Deeper Action

2nd Place Winner

ECCV TEL AVIV 2022

A Technical Report for SportsMOT Track on Multi-actor Tracking

Jiahao Wang*, Chang Meng, Donghao Li, Hao Wang, Yuting Yang, Licheng Jiao, Fang Liu

School of Artificial Intelligence, Xidian University

*Corresponding to: jh_wang1024@163.com

西安電子科技大學

XIDIAN UNIVERSITY

Deeper Action

3rd Place Winner

ECCV TEL AVIV 2022

Observation Centric and Central Distance Recovery on Sports Player Tracking

Hsiang-Wei Huang, Cheng-Yen Yang, Jenq-Neng Hwang
Pyong-Kun Kim, Kwangju Kim, Kyoungoh Lee

University of Washington, Seattle
Electronics and Telecommunications Research Institute, Korea

UNIVERSITY of WASHINGTON

ETRI 한국전자통신연구원 Electronics and Telecommunications Research Institute

Deeper Action

ECCV TEL AVIV 2022

Thanks!

Homepage: https://deeperaction.github.io/tracks/sportsmot

Github: https://github.com/MCG-NJU/SportsMOT