Deeper Action

ICCV 2021 Workshop

ICCV DeeperAction Challenge
FineAction Track on Temporal Action Localization

Yi Liu[1] Limin Wang[2] Xiao Ma[1] Yali Wang[1] Yu Qiao[1]

[1] Shenzhen Institutes of Advanced Technology, Chinese Academy of Sciences
[2] State Key Laboratory for Novel Software Technology, Nanjing University

中国科学院深圳先进技术研究院
SHENZHEN INSTITUTE OF ADVANCED TECHNOLOGY
CHINESE ACADEMY OF SCIENCES

南京大学
NANJING UNIVERSITY

Outline

Part 1: Dataset Introduction

Part 2: FineAction Competition

Part 1

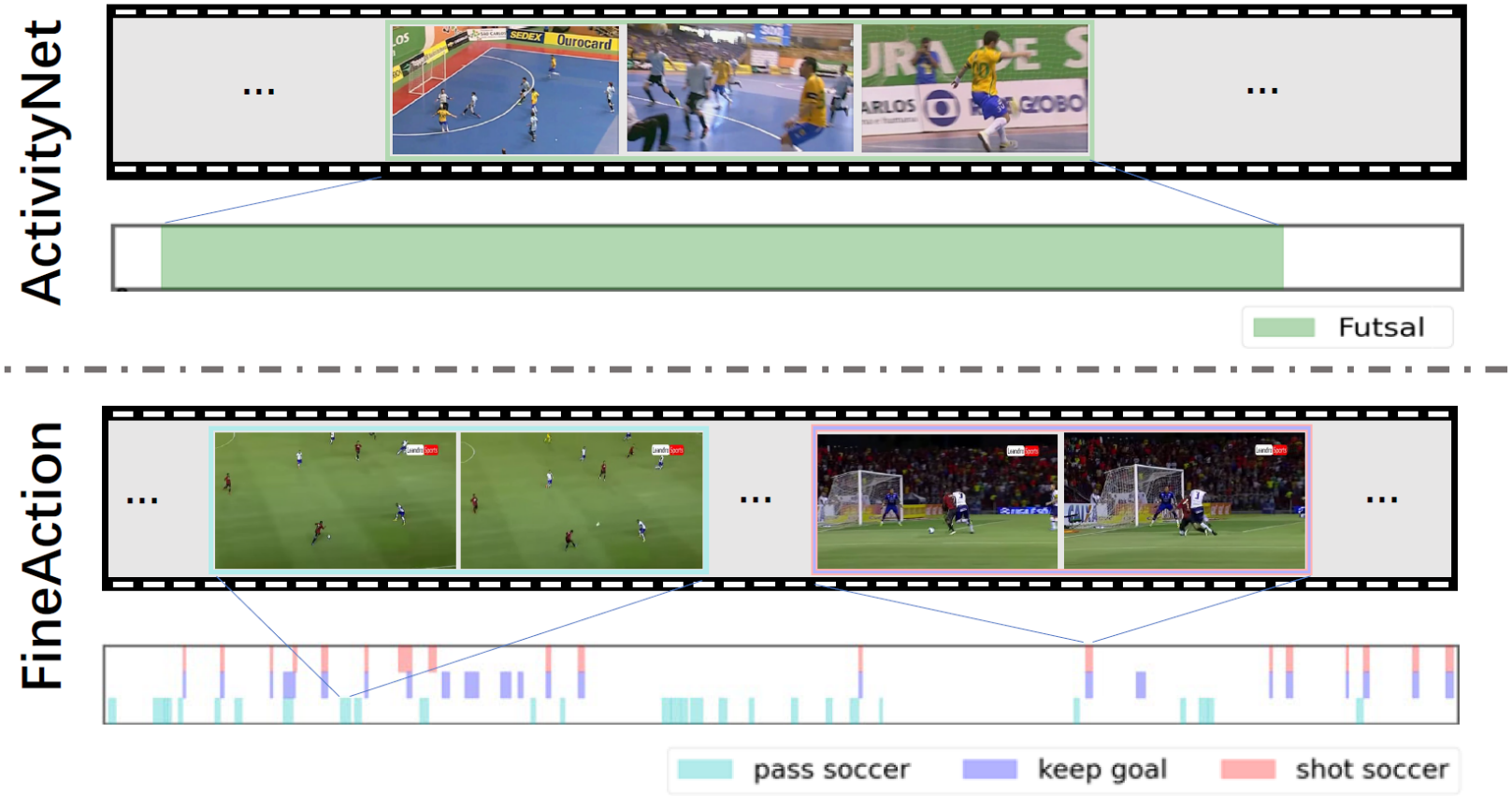
Dataset Introduction

Why to do?

ActivityNet

- ❑ *Sparse annotations*
- ❑ *Coarse-level action*

- ❑ **Dense annotations**
- ❑ **Fine-Grained action**



Why to do?

THUMOS-14

- ❑ *Small scale*
- ❑ *Specific domain*

- ❑ **Large scale**
- ❑ **Diversity domain**

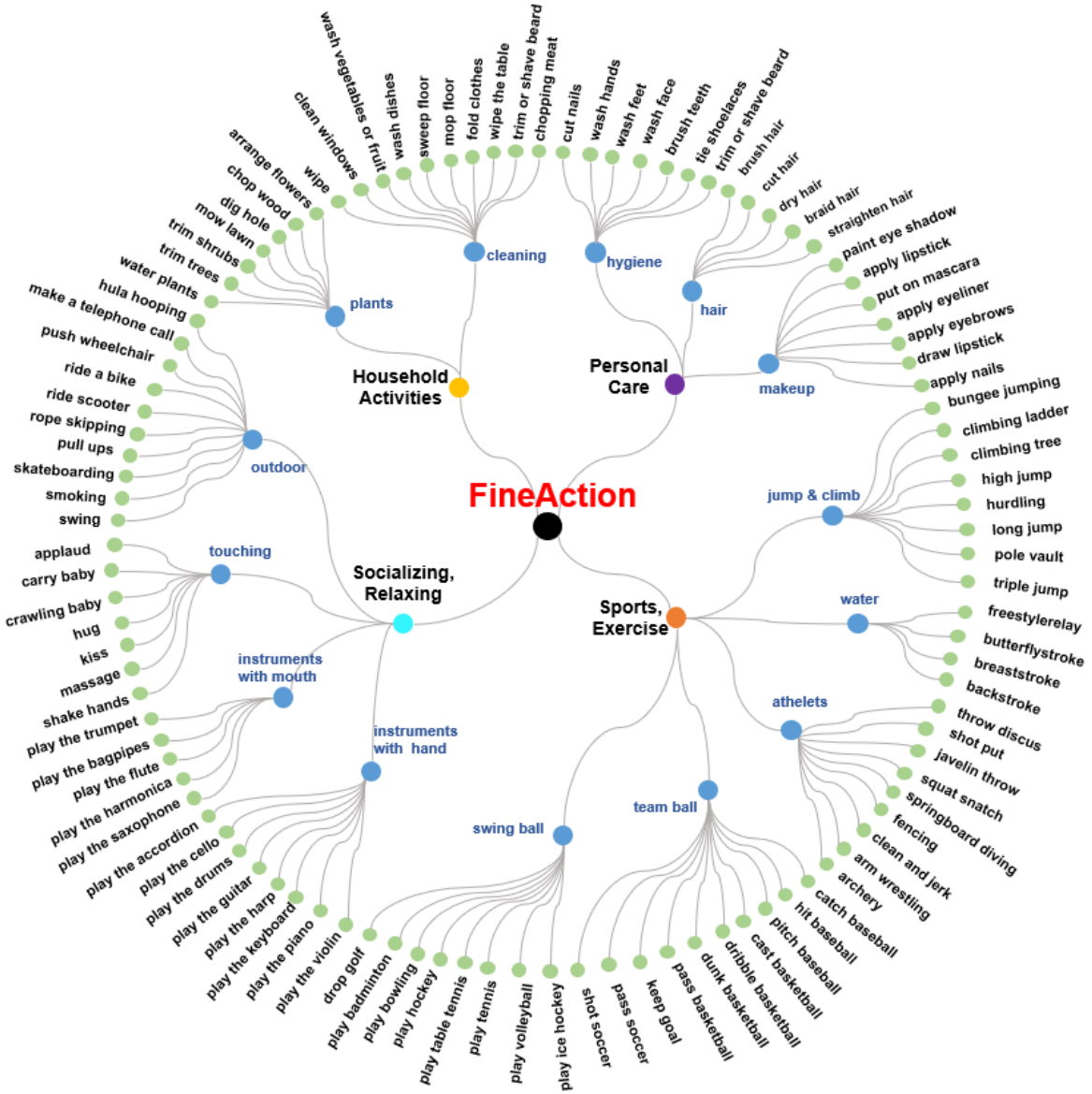
Deeper Action

FineAction: New Taxonomy

ICCV 2021 Workshop

FineAction

- ❑ 4 top-level categories
- ❑ 14 middle-level categories
- ❑ 106 bottom-level categories



FineAction: Video Collection

Video Select

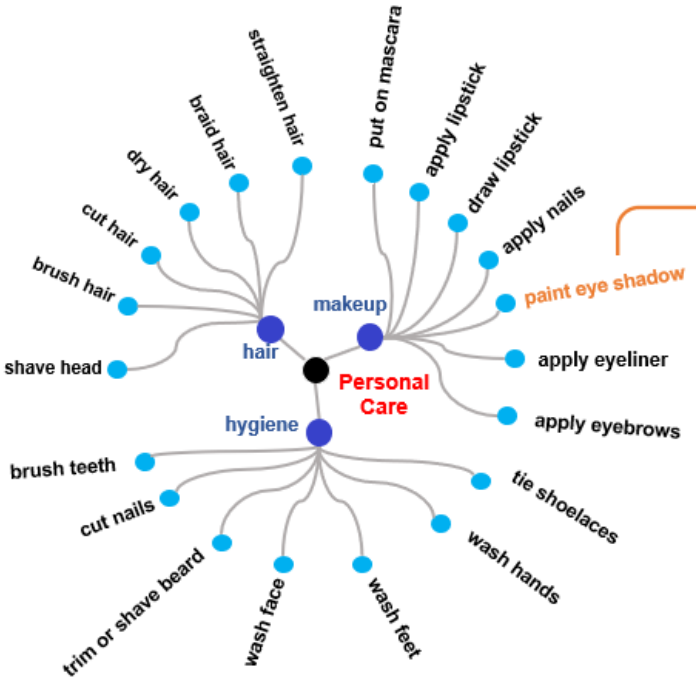
- ❑ **Exisiting video datasets**
  - *YouTube 8M*
  - *Kinetics 400*
  - *FCVID*
- ❑ **Crawl from YouTube videos**

Deeper Action

FineAction: Annotation

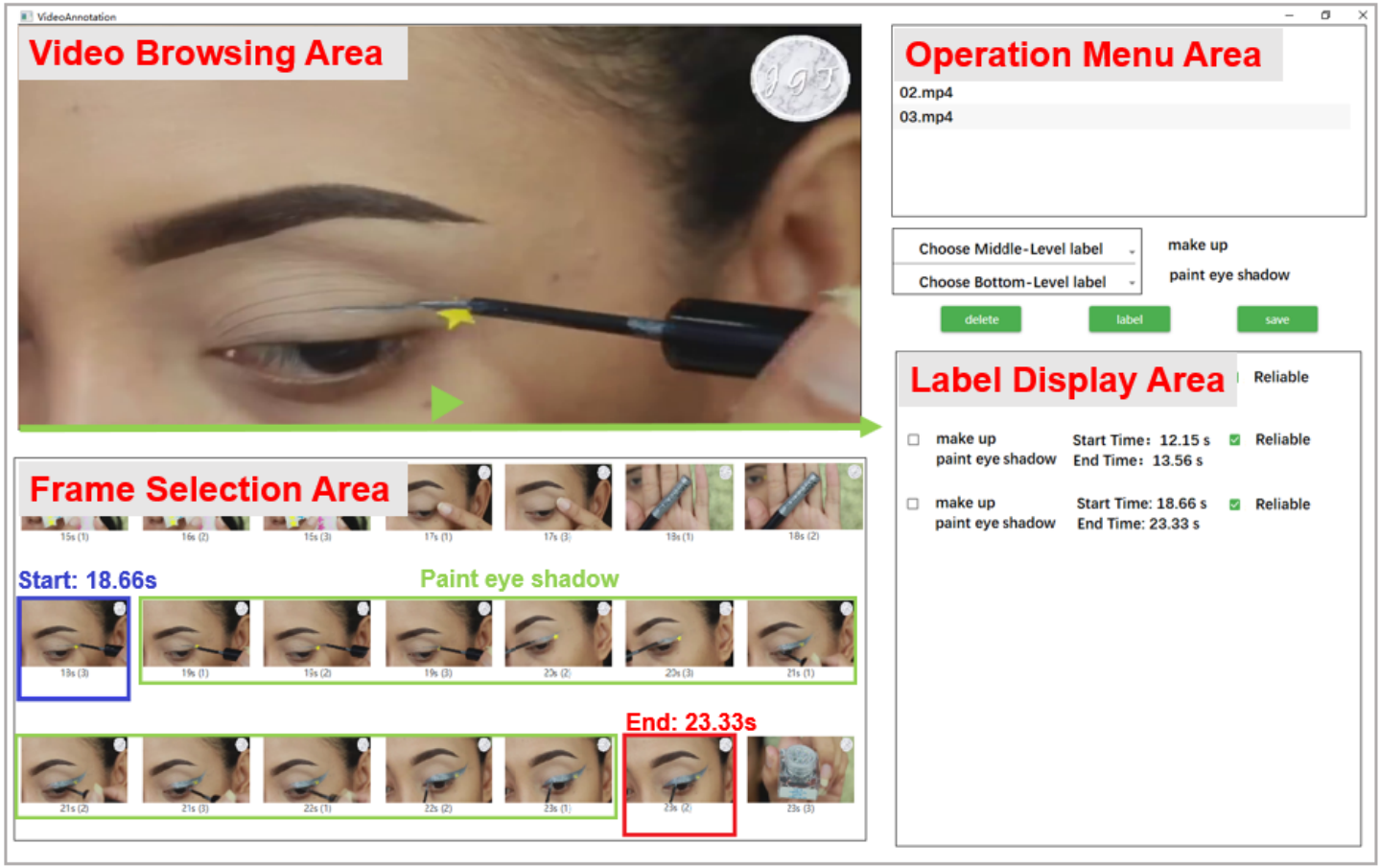
ICCV 2021 Workshop

Framework of Annotation Process



Start Definition:
Draw eye shadow with tools such as brushes.

End Definition:
Take off the tool.



(a) Category Taxonomy

(b) Annotation Guidance

(c) Annotation Tool

FineAction: Statistics

❑ Large-Scale & Fine-Grained

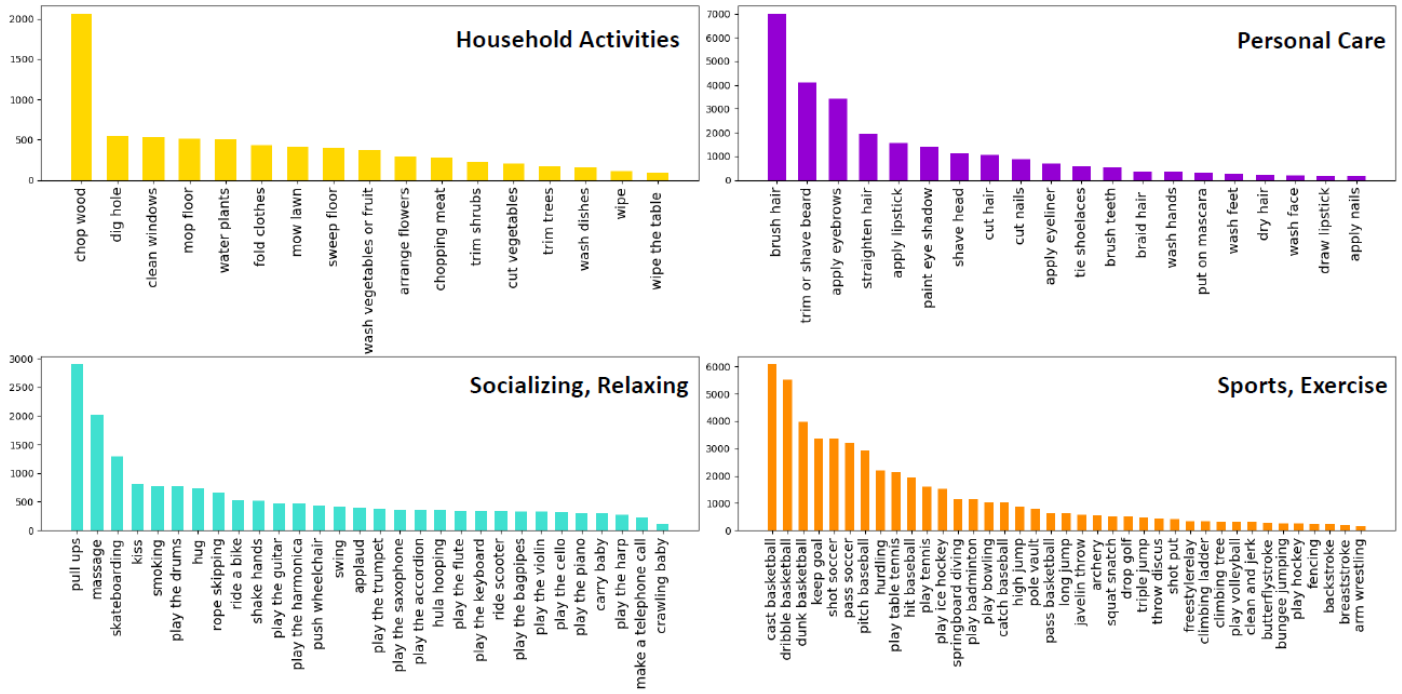
Database	Category	Video	Instance	Overlap	Duration	Action type
MPII Cooking	65	45	5,609	0.1%	11.1 m	kitchens
EPIC-Kitchens	4,025	700	89,979	28.1%	3.1 s	
FineGym V1.0	530	303	32,697	0.0%	1.7 s	sports
THUMOS14	20	413	6,316	17.5%	4.3 s	
ActivityNet	200	19,994	23,064	0.0%	49.2 s	daily events
HACS Segment	200	49,485	122,304	0.0%	33.2 s	
FineAction (Ours)	106	16,732	103,324	11.5%	7.1 s	

Deeper Action

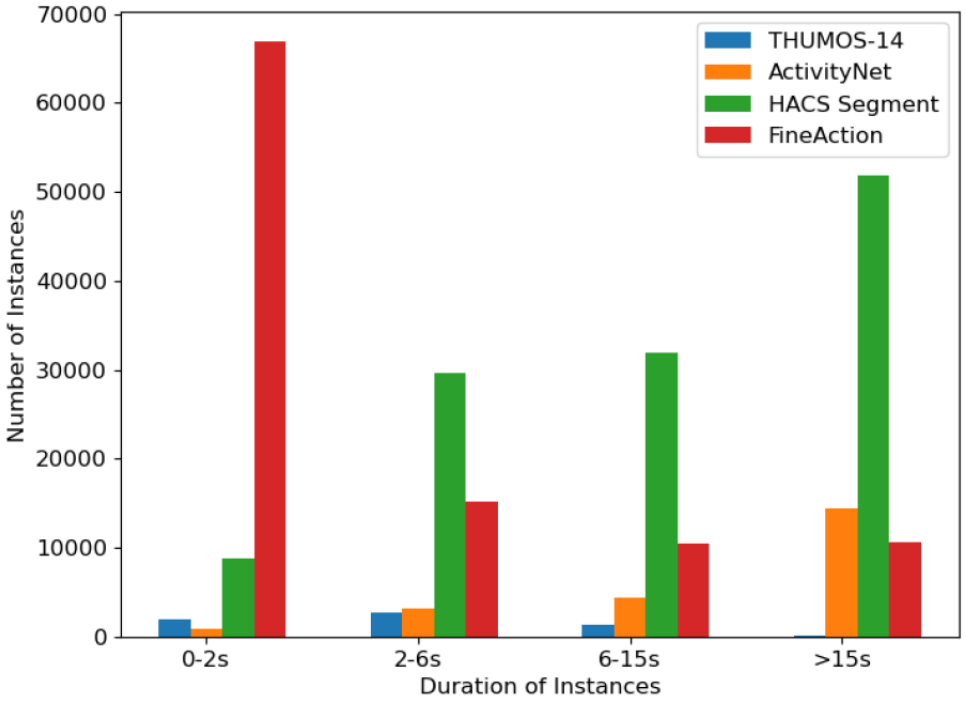
FineAction: Statistics

ICCV 2021 Workshop

❑ Instance distribution



❑ Instance duration



FineAction: Experiment

❑ SOAT methods on FineAction

Method	Modality	Action Proposal Generation				Temporal Action Localization			
		AR@5	AR@10	AR@100	AUC	mAP@0.50	mAP@0.75	mAP@0.95	Avg.mAP
BMN	RGB	8.62	11.20	22.74	17.49	12.56	7.49	2.62	7.86
	Flow	9.85	12.72	24.18	18.94	14.49	8.92	3.19	9.23
	RGB+Flow	9.99	12.84	24.34	19.19	14.44	8.92	3.12	9.25
DBG	RGB	6.82	9.01	21.26	15.48	8.57	5.01	1.93	5.31
	Flow	8.27	10.90	23.37	17.70	11.03	6.95	2.70	7.20
	RGB+Flow	7.82	10.45	23.07	17.24	10.65	6.43	2.50	6.75
G-TAD	RGB	7.96	10.45	20.86	16.06	10.88	6.52	2.19	6.87
	Flow	8.87	11.60	22.01	17.09	12.58	8.18	2.56	8.26
	RGB+Flow	9.02	11.83	22.60	17.65	13.74	8.83	3.06	9.06

Dataset	AR@100	Classifier	mAP@0.5
THUMOS14	47.7	UNet	38.8
ActivityNet-1.3	75.0	TSN	50.1
HACS Segment	70.9	SlowFast	52.5
FineAction	24.3	TSN	14.4

FineAction: Experiment

❑ Different feature length

Length	AR@5	AR@10	AR@100	AUC
100	9.99	12.84	24.34	19.19
150	10.56	13.99	26.89	21.06
200	10.69	14.41	28.57	22.16
250	11.11	14.81	29.56	22.94

❑ Different duration instances

Instance Duration	0-2 s	2-5 s	5-10 s	>10 s
AR@100	7.46	32.58	51.24	71.73

❑ Top-3 worst & Top-3 best

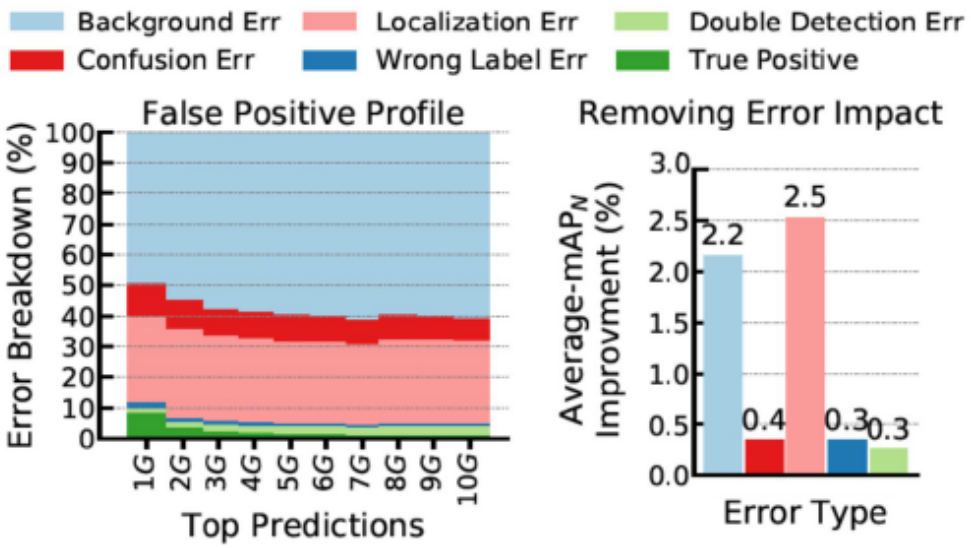
FineAction	AR@100	Avg. Duration	Num
straighten hair	1.6	1.17 s	1934
apply eyebrows	1.7	1.13 s	3428
dig hole	2.2	0.83 s	544
freestyle relay	88.4	53.12 s	345
breaststroke	94.5	53.59 s	211
play the harp	96.2	69.21 s	268

Deeper Action

FineAction: Analysis

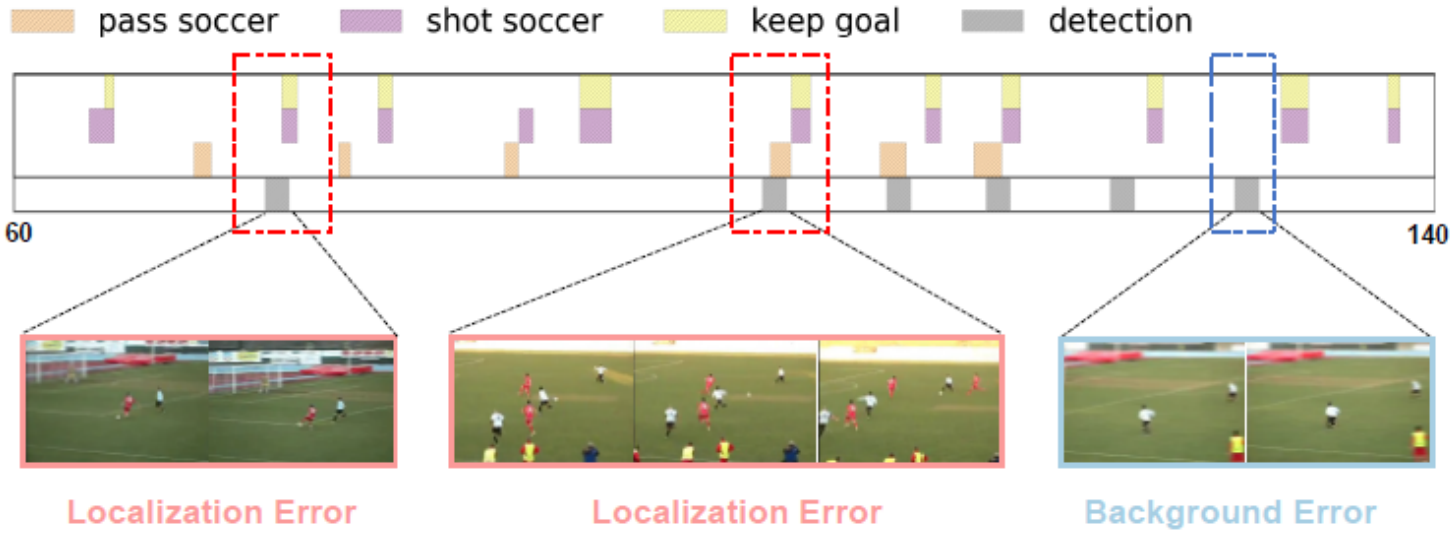
ICCV 2021 Workshop

❑ Error analysis



(a) Error Analysis.

❑ Error visualization



(b) Visualization of Typical Errors.

Part 2

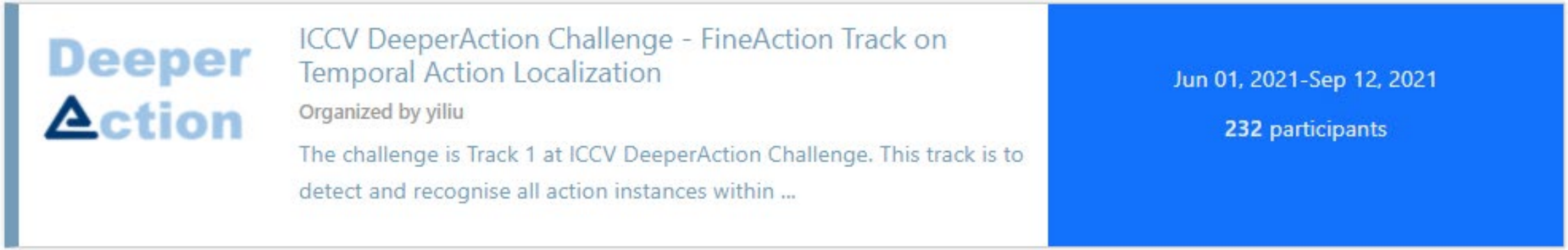
FineAction Competition

Deeper Action

Competition: FineAction Track

ICCV 2021 Workshop

https://competitions.codalab.org/competitions/32363



- **Validation phase:** 2021.06.01-2021.09.01
- **Testing phase:** 2021.09.01-2021.09.12

FineAction Track: Evaluation

❑ Goal

The challenge requires detecting and recognizing all-action instances within an untrimmed video.

❑ Metric

Mean Average Precision (mAP) is a conventional evaluation metric, where Average Precision (AP) is calculated for each action category with tIoU thresholds [0.5:0.05:0.95].

Deeper Action

FineAction Track: Statistics

ICCV 2021 Workshop

- Valid Participants : 228
- Valid Teams: 18 + 14 (Company, University, Institute)

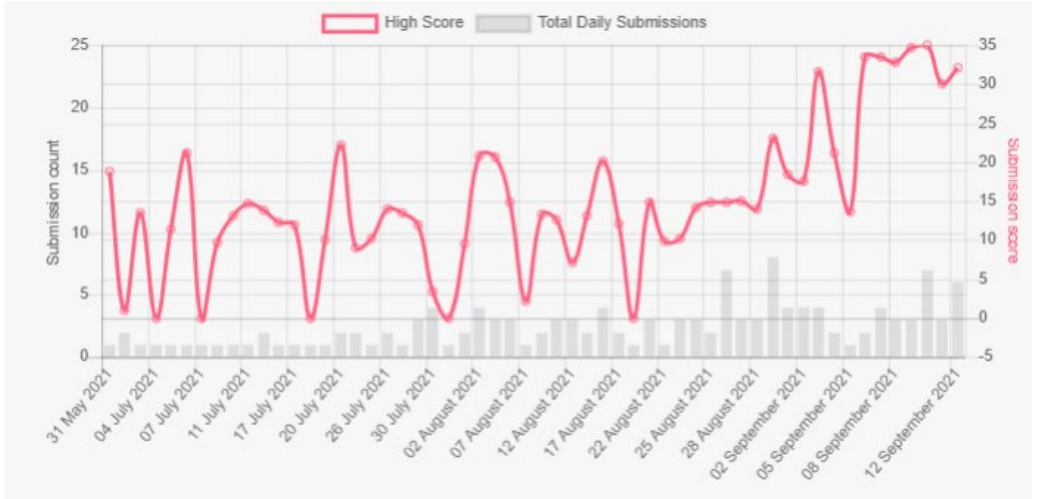
Deeper Action

FineAction Track: Results

ICCV 2021 Workshop

❑ Valid Submission: 95 (Val Phase) + 42 (Test Phase)

Test Set (Mean Average Precision - mAP)													
#	User	Entries	Date of Last Entry	mAP@0.50 ▲	mAP@0.55 ▲	mAP@0.60 ▲	mAP@0.65 ▲	mAP@0.70 ▲	mAP@0.75 ▲	mAP@0.80 ▲	mAP@0.85 ▲	mAP@0.90 ▲	mAP@0.95 ▲	Avg.mAP ▲
1	Dahua_001	10	09/10/21	35.044	32.999	30.599	28.390	26.055	23.348	20.494	17.277	12.707	6.607	23.352 (1)
2	DeepakSridhar	1	09/03/21	20.244	18.854	17.581	16.055	14.305	12.697	11.043	9.059	6.774	3.615	13.023 (2)
3	foolwood	2	09/12/21	19.446	18.050	16.659	15.254	13.845	12.048	10.470	8.775	6.628	4.002	12.518 (3)



Deeper Action

1st Place Winner

ICCV 2021 Workshop

Learning Efficient Feature Representation for Temporal Action Localization

Chenglu Wu*, Xuefeng Yang*, Fuzhi Duan, Yanxun Yu, Yayun Wang, Jun Yin
Zhejiang Dahua Technology Co., Ltd.
Hangzhou City, Zhejiang Province, China
duanfz@whu.edu.cn
{wu_chenglu,yang_xuefeng,yu_yanxun}@dahuatech.com

Deeper Action

2nd Place Winner

ICCV 2021 Workshop

An Empirical Study of Feature Representation for Actionness based Temporal Action Localization

Qiang Wang and Rongliang Cheng
Institute of Automation, Chinese Academy of Sciences.
qiang.wang@nlpr.ia.ac.cn, chengrongliang@hotmail.com

中国科学院
自动化研究所
INSTITUTE OF AUTOMATION
CHINESE ACADEMY OF SCIENCES

Over & Thank you !

Homepage: https://deeperaction.github.io/fineaction/