TRECVID21 ActEV Task
(Activity Detection in Extended Video)

Presenter: Yooyoung Lee
Jon Fiscus, Andrew Delgado, Afzal Godil,
Baptiste Chocot, ,Lukas Diduch,
Eliot Godard, Jim Golden

**Day 2
December 8
7 a.m. – 11 a.m.(ET)**

**trecvid-conference
trecvidco-9sb4982.slack.com
Channel:
#activities-in-extended-videos**

Disclaimer

Certain commercial equipment, instruments, software, or materials are identified in this paper to specify the experimental procedure adequately. Such identification is not intended to imply recommendation or endorsement by NIST, nor necessarily the best available for the purpose.

The views and conclusions contained herein are those of the authors and should not be interpreted as necessarily representing the official policies or endorsements, either expressed or implied, of IARPA, NIST, or the U.S. Government.

ActEV Overview

What is ActEV? (Activity Detection in Extended Videos)
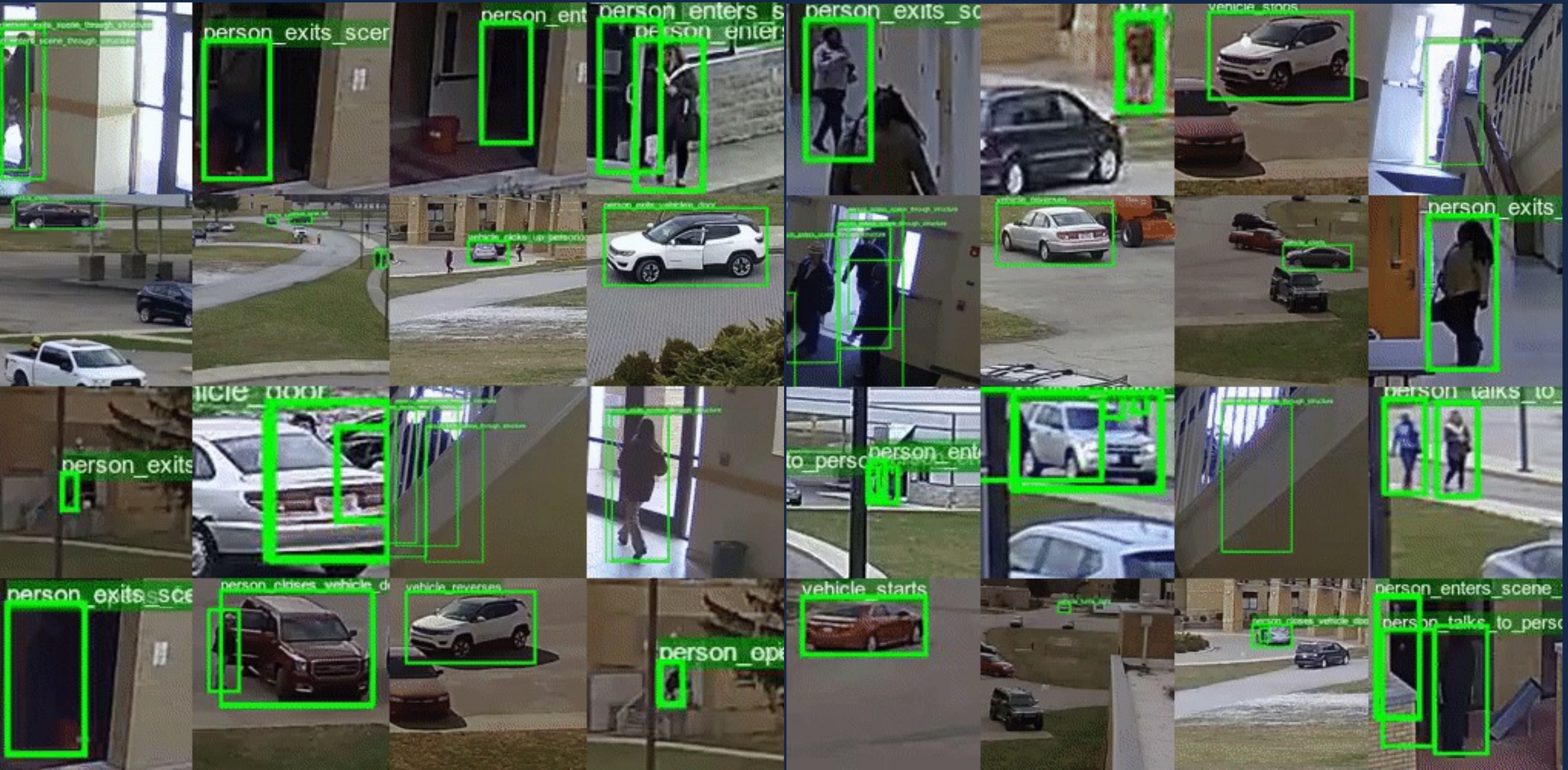
What is ActEV's Goal?

- To promote video analytics technology that can automatically <u>detect a target activity</u> and <u>identify and track objects</u> associated with the activity.

- A series of challenges is also designed for:
  - Robust detection of known/surprise activities in
    - Known/unknown facilities and
    - Multi-camera environment
  - Temporal (and spatio-temporal) localization of the activity for reasoning
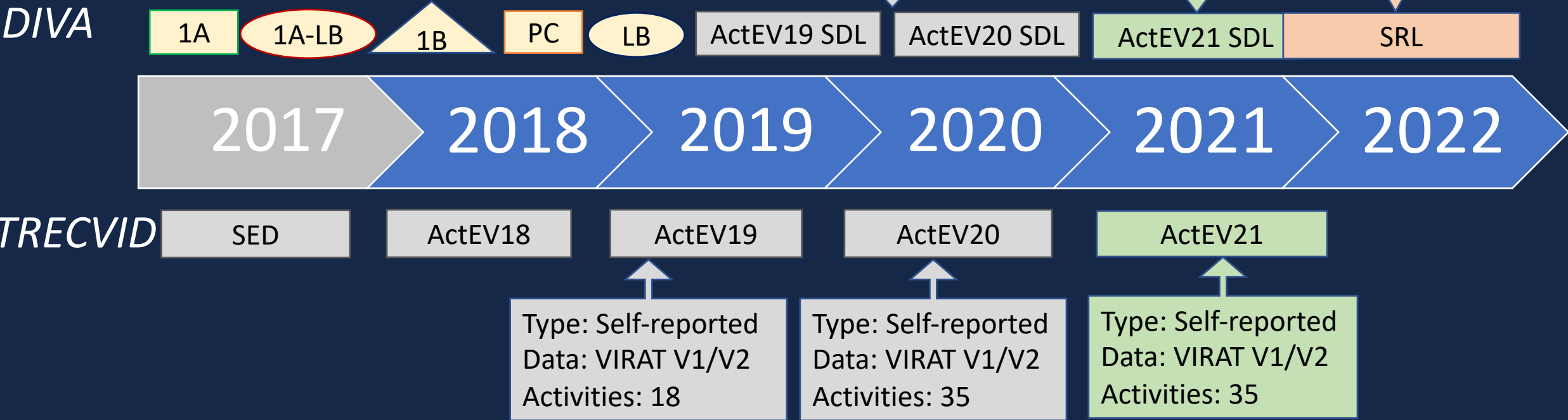
DIVA (NIST, IARPA, and Kitware)

- NIST developed the ActEV evaluation series to support the metrology needs of the Intelligence Advanced Research Projects Activity (IARPA) Deep Intermodal Video Analytics (DIVA) Program
- The ActEV datasets were collected and annotated by Kitware, Inc.

ActEV Series

SED: Surveillance Event Detection
LB: Leaderboard
PC: Prize Challenge
SDL: Sequestered Data Leaderboard
1A: Phase1-A, 1B: Phase1-B

Type: Self-reported
Data: MEVA Test 5
Known Facilities
Known Activities

Type: Sequestered LB
Data: MEVA Test 3
Known Activities: 37

Type: Sequestered LB
Data: MEVA Test 4
Known/Unknown Facilities
Known/Surprise Activities

DIVA

| 1A | 1A-LB | 1B | PC | LB | ActEV19 SDL | ActEV20 SDL | | ActEV21 SDL | SRL |

| 2017 | 2018 | 2019 | 2020 | 2021 | 2022 |

TRECVID

| SED | ActEV18 | ActEV19 | ActEV20 | ActEV21 |

Type: Self-reported
Data: VIRAT V1/V2
Activities: 18

Type: Self-reported
Data: VIRAT V1/V2
Activities: 35

Type: Self-reported
Data: VIRAT V1/V2
Activities: 35

Evaluation
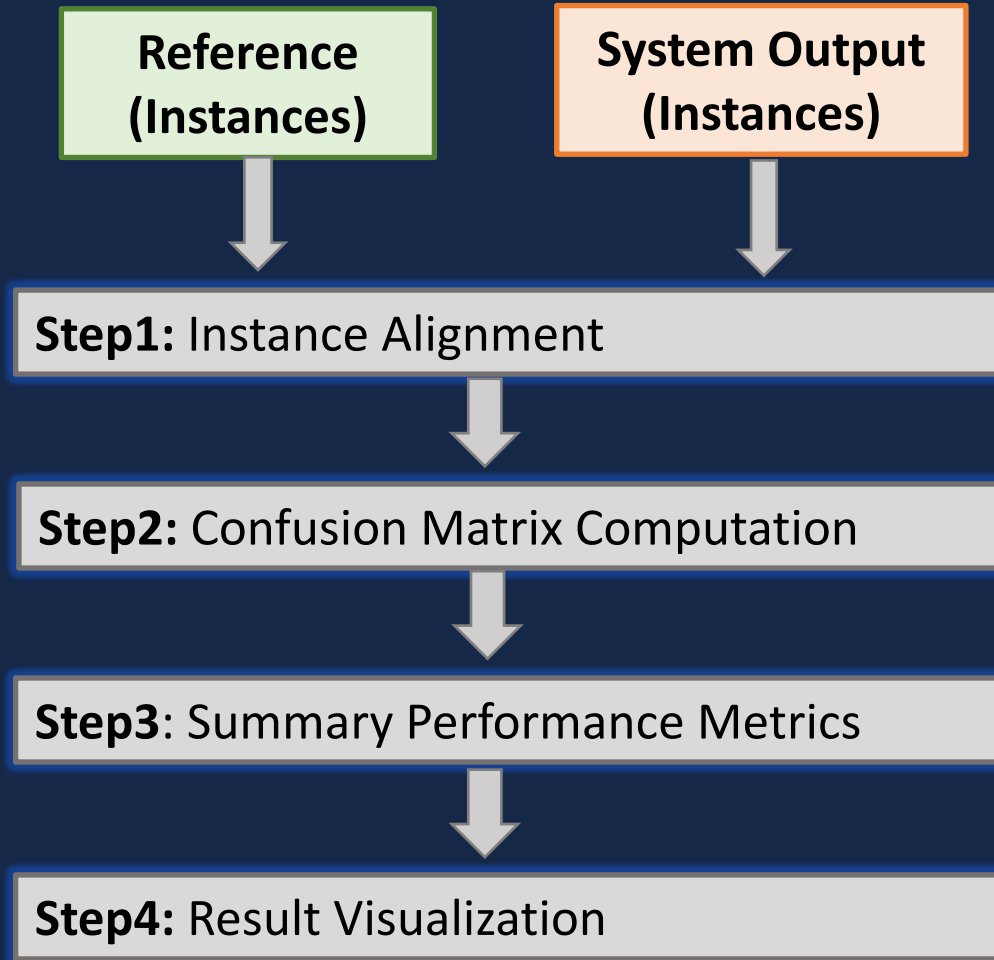
ActEV Evaluation Framework

- Target applications
  - <u>Retrospective analysis of archives</u> (e.g., forensic analytics)
  - Real-time analysis of live video streams (e.g., alerting and monitoring)
- Evaluation Type
  - <u>Self-reported (or take-home) evaluation</u>
    - Participants download the test dataset, run their systems on their machines, and submit the system outputs to NIST
    - TRECVID ActEV'21, DIVA SRL (Self-Reported Leaderboard)
  - Independent (or sequestered) evaluation
    - Submit their systems to run on the sequestered test dataset at NIST
    - DIVA SDL (Sequestered Data Leaderboard)
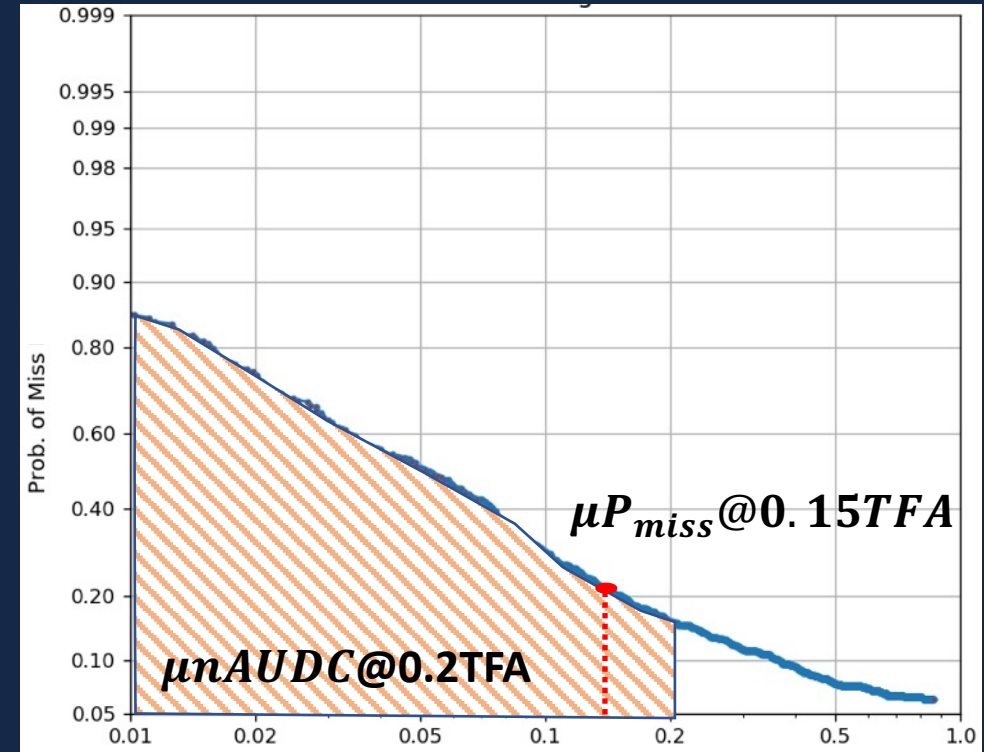
Tasks and Measures

Evaluation Tasks (AD)

- "Activity" definition
  - One or more people performing a specified movement, or interacting with an object or group of objects (including driving)

- Activity Detection (AD) task
  - Given a target activity, a system automatically detects its presence and then temporally localizes all instances of the activity in video sequences
  - Please visit https://actev.nist.gov

Performance Measure Calculation

Reference (Instances)

System Output (Instances)

Step1: Instance Alignment

Step2: Confusion Matrix Computation

Step3: Summary Performance Metrics

Step4: Result Visualization

DET (Detection Error Tradeoff)

$\mu P_{miss}@0.15TFA$

$\mu nAUDC@0.2TFA$

Time-based False Alarm (TFA) Rate

Primary metric: $\mu nAUDC@0.02TFA$

Mean Normalized partial Area Under the DET Curve ($\mu nAUDC$)
Please see details at https://actev.nist.gov

Datasets & Activities

Approved by Institutional Review Board (IRB)

#ITL-17-0037

VIRAT (10-hr annotated videos with 35 activities)

Activities	Train	Validation	Activities	Train	Validation
person_closes_facility_or_vehicle_door	141	130	person_unloads_vehicle	32	44
person_closes_trunk	21	31	person_carries_object	237	364
vehicle_drops_off_person	0	4	person_crouches	1	9
person_enters_facility_or_vehicle	77	70	person_gestures	82	148
person_exits_facility_or_vehicle	66	72	person_runs	14	18
person_interacts_object	101	88	person_sits	21	11
person_loads_vehicle	38	38	person_stands	398	819
person_opens_trunk	22	35	person_walks	761	901
person_opens_facility_or_vehicle_door	137	128	person_talks_on_phone	17	16
person_person_interaction	11	17	person_texts_on_phone	5	20
person_pickups_object	19	12	person_uses_tool	7	11
vehicle_picks_up_person	9	5	vehicle_moves	718	797
person_pulls_object	23	43	vehicle_starts	259	239
person_pushs_object	4	6	vehicle_stops	292	295
person_rides_bicycle	22	21	vehicle_turns_left	152	176
person_sets_down_object	12	11	vehicle_turns_right	149	172
person_talks_to_person	41	67	vehicle_makes_u_turn	9	13
person_carries_heavy_object	31	44			

Results and Analyses

Leaderboard as of today

TRECVID21

Show 10 entries

RANK	SUBMISSION ID	TEAM NAME	SYSTEM NAME	PARTIAL AUDC*	MEAN-P_MISS@0.15TFA
1	26562	INF	INF_PRE	0.39607	0.30622
2	26542	BUPT-MCPRL	MCPRL_S0	0.40853	0.32489
3	26539	BUPT-MCPRL	MCPRL_S2	0.40947	0.32919
4	26540	BUPT-MCPRL	MCPRL_S3	0.41305	0.32866
5	26546	UCF	UCF - S1	0.43059	0.34080
6	26543	UCF	UCF - S2	0.43271	0.34207
7	26534	UCF	UCF-P	0.43562	0.34466
8	26532	INF	INF_full	0.44436	0.35079
9	26388	INF	INF	0.45115	0.35161
10	26544	UCF	UCF - S3	0.45700	0.36994

Missed the deadline slightly

ActEV'21 Participants Ranking

16 submissions as of 10/8/21

6 teams (best system result per team)

Team_Name	Team_Abbrev	Organization	$\mu nAUDC$	μP_{miss} @ $.15TFA$
BUPT-MCPRL	BUPT-MC_26542	Beijing University of Posts and Telecommunications, China	0.409	0.325
UCF	UCF_26546	University of Central Florida, USA	0.431	0.341
INF	INF_26532	Carnegie Mellon University, USA	0.444	0.351
M4D_2021	M4D_26467	Information Technologies Institute, Greece	0.847	0.794
TokyoTech_AIST	TOKYOTE_26508	Tokyo Institute of Technology, Japan	0.852	0.820
Team UEC	TEAMUE_26530	The University of Electro-Communications, Japan	0.964	0.950

Activity-level Comparison

Activities Difficulty · Instance Count

activity	BUPT-MCPRL	UCF	INF	AVG	Training	Validation
person_uses_tool	3.0	1.0	1.0	1.7	7	11
person_rides_bicycle	1.0	3.0	3.0	2.3	22	21
vehicle_makes_u_turn	2.0	5.0	2.0	3.0	9	13
vehicle_moves	4.0	2.0	4.0	3.3	718	797
person_pulls_object	6.0	8.0	5.0	6.3	23	43
person_opens_trunk	8.0	6.0	7.0	7.0	22	35
person_walks	5.0	7.0	10.0	7.3	761	901
person_pushs_object	10.0	4.0	13.0	9.0	4	6
vehicle_drops_off_person	15.0	9.0	8.0	10.7	0	4
person_talks_to_person	19.0	11.0	9.0	13.0	41	67
person_carries_heavy_object	14.0	16.0	11.0	13.7	31	44
person_carries_object	13.0	14.0	15.0	14.0	237	364
person_runs	12.0	12.0	19.0	14.3	14	18
person_stands	16.0	10.0	18.0	14.7	398	819
vehicle_turns_left	9.0	15.0	21.0	15.0	152	176
person_person_interaction	11.0	20.0	14.0	15.0	11	17
vehicle_turns_right	7.0	13.0	26.0	15.3	149	172
vehicle_picks_up_person	20.0	25.0	6.0	17.0	9	5
person_unloads_vehicle	18.0	17.0	16.0	17.0	32	44
person_closes_trunk	17.0	26.0	12.0	18.3	21	31
person_loads_vehicle	25.0	18.0	17.0	20.0	38	38
person_gestures	23.0	24.0	20.0	22.3	82	148
person_talks_on_phone	24.0	23.0	24.0	23.7	17	16
person_sits	21.0	29.0	22.0	24.0	21	11
person_opens_facility_or_vehicle_door	29.0	22.0	31.0	27.3	137	128
vehicle_stops	27.0	33.0	23.0	27.7	292	295
person_enters_facility_or_vehicle	30.0	19.0	34.0	27.7	77	70
person_interacts_object	31.0	28.0	25.0	28.0	101	88
person_texts_on_phone	34.0	21.0	30.0	28.3	5	20
vehicle_starts	26.0	32.0	28.0	28.7	259	239
person_sets_down_object	28.0	31.0	29.0	29.3	12	11
person_crouches	22.0	34.0	35.0	30.3	1	9
person_exits_facility_or_vehicle	33.0	27.0	33.0	31.0	66	72
person_closes_facility_or_vehicle_door	32.0	30.0	32.0	31.3	141	130
person_pickups_object	35.0	35.0	27.0	32.3	19	12

Easiest

Difficult

Hardest

Count: 750, 500, 250, 0
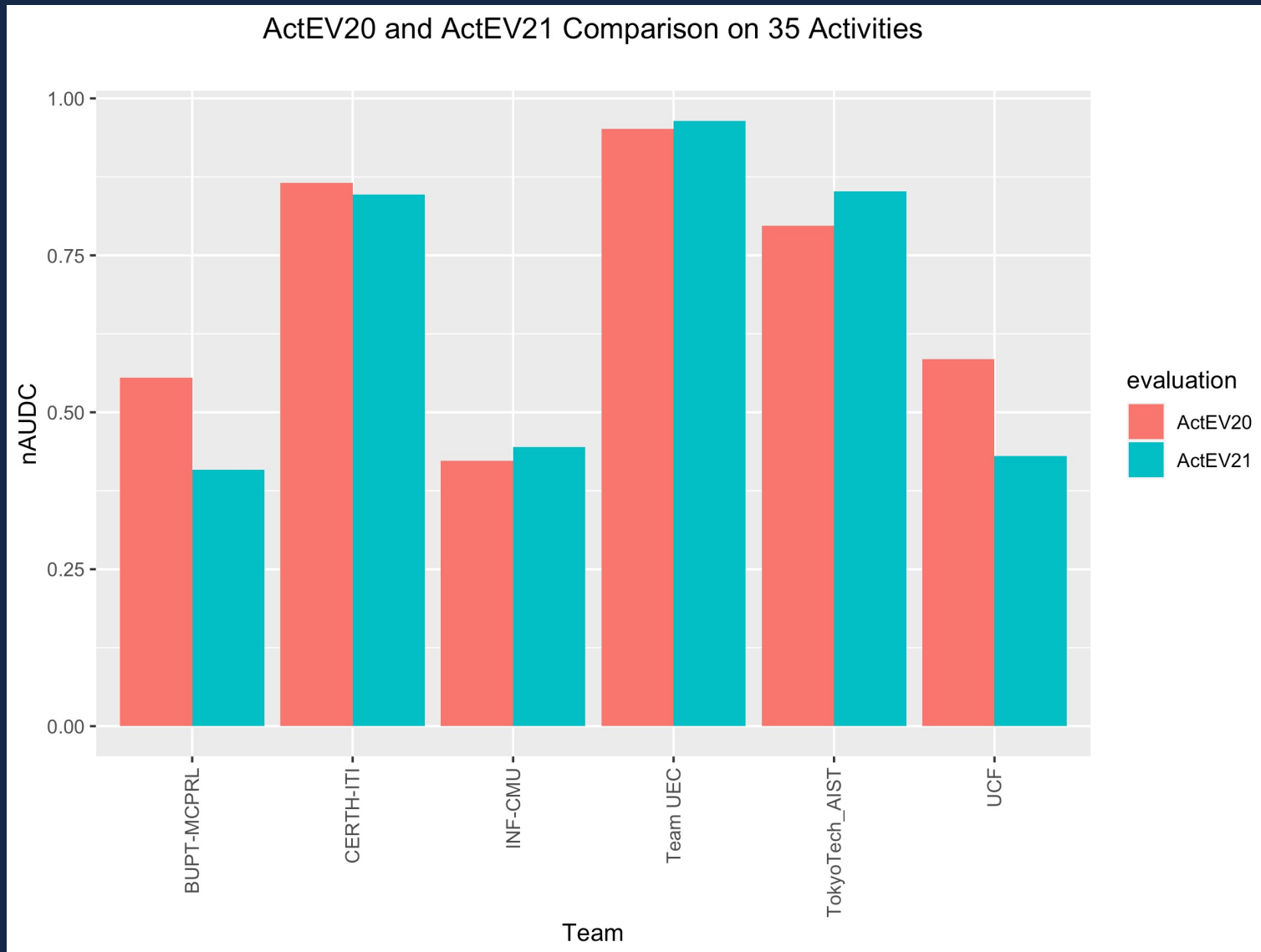
1/27/22

20
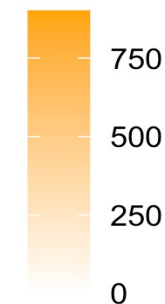
Comparison of ActEV20 vs ActEV21
(35 Activities)

ActEV20 and ActEV21 Comparison on 35 Activities

Summary

- We observed that each system has their own detection strength depending activity classes
  - Fusion among different systems can improve overall performance for activity detection
- For activity difficulty, "person_uses_tool" and"person_rides_bicycle" are easier while "person_pickup_objects" is more difficult to detect
- Based on the submission deadline of 10/08/2021, we invited the top-three teams to present their work:
  - BUPT (Beijing University of Posts and Telecommunications, China)
  - UCF (University of Central Florida, USA)
  - INF (Carnegie Mellon University, USA)

What is Next?

- TRECVID 2022
  - Potentially leverage MEVA data for TRECVID in addition to the VIRAT dataset
  - SRL (Self-Reported Leaderboard) Challenge, https://actev.nist.gov/srl
    - Release a subset of *sequestered data (MEVA KF-KA) to public*
    - Include **Activity and Object Detection (AOD) in addition to Activity Detection (AD)**
    - Anyone can *submit their system output to NIST* and their results are displayed on the leaderboard
    - Support a deeper understanding of system behavior via interactive data analysis

Thank you!
Any Questions?

Contact: actev@nist.gov

TRECVID Workshop 2021
December 7-10, 2021

Day 2
December 8

7 a.m. – 11 a.m.(ET)

Our Slack channel:
#activities-in-extended-videos

Please disconnect from any VPN for better connection.
All participant microphones and video will be muted
upon entering the event.

7:00	Keynote Talk: Learning Representations from Sign Language Videos
7:40	Activities in Extended Video - Task Overview
8:00	CMU Informedia at TRECVID 2021: Towards Better Spatial-Temporal Activity Detection, Wenhe Liu, Carnegie Mellon University, USA
8:20	215AD: A Comprehensive Activity Detection for Surveillance Videos in TRECVID 2021 ActEV Challenge, Binyu Zhang, Beijing University of Posts and Telecommunications, China
8:40	Break
9:00	GabriellaV2: UCF DIVA system, Zacchaeus Scheffer, University of Central Florida, USA
9:20	Multi-label activity recognition in extended videos using objects' Spatio-temporal boundaries, Konstantinos Gkountakos and Despoina Touska, Centre for Research and Technology Hellas, Greece
9:40	Activities in Extended Video - Task Discussion
10:10~	OpenMFC

For technical support please contact conferences@nist.gov

NIST

ActEV Task Discussion

- What are the bottlenecks for the ActEV task?

- How can we help participants better with submission errors?

- What are thoughts on performance metrics?
  - Current metrics: nAUDC, Pmiss@TFA=0.15
  - Potential metrics (e.g., mAP (mean average precision)

- What are thoughts on data?
  - Is VIRAT data suitable for continued testing in TRECVID?
  - Potential leverage a subset of MEVA (https://mevadata.org)
  - Fine-grained activities collected from Consented Activities of People (CAP) using handheld devices (https://visym.com/collector)

- How can we help promote activity detection technologies in research community?