ActEV20: Activities in Extended Video (Summary Results)

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

TRECVID 2020 Workshop

December 8-11, 2020

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.

<u>The views and conclusions contained herein are those of the authors</u> 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.

Outline

- ActEV Overview

- TRECVID ActEV20 Evaluation

- ActEV20 Tasks and Measures

- ActEV20 Dataset

- ActEV20 Results and Analyses

- Next Steps

National Institute of Standards and Technology / U.S. Department of Commerce

ActEV Overview

National Institute of Standards and Technology / U.S. Department of Commerce

What is Activities in Extended Video?

"Retrospective detection/localization of activities in long videos"

vehicle_turning_left-00
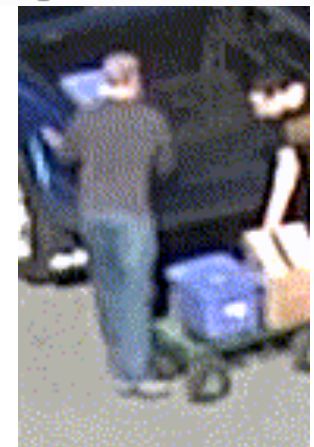
vehicle_turning_left-04

Closing-01

Closing-02

Closing-00

Loading-00

Loading-01

Loading-02

What is ActEV's Goal?

- Goal:
  - To advance 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 are also designed for:
  - Activity detection in a multi-camera environment
  - Temporal (and spatio-temporal) localization of the activity for reasoning

- Changes for the 3rd TRECVid ActEV Evaluation
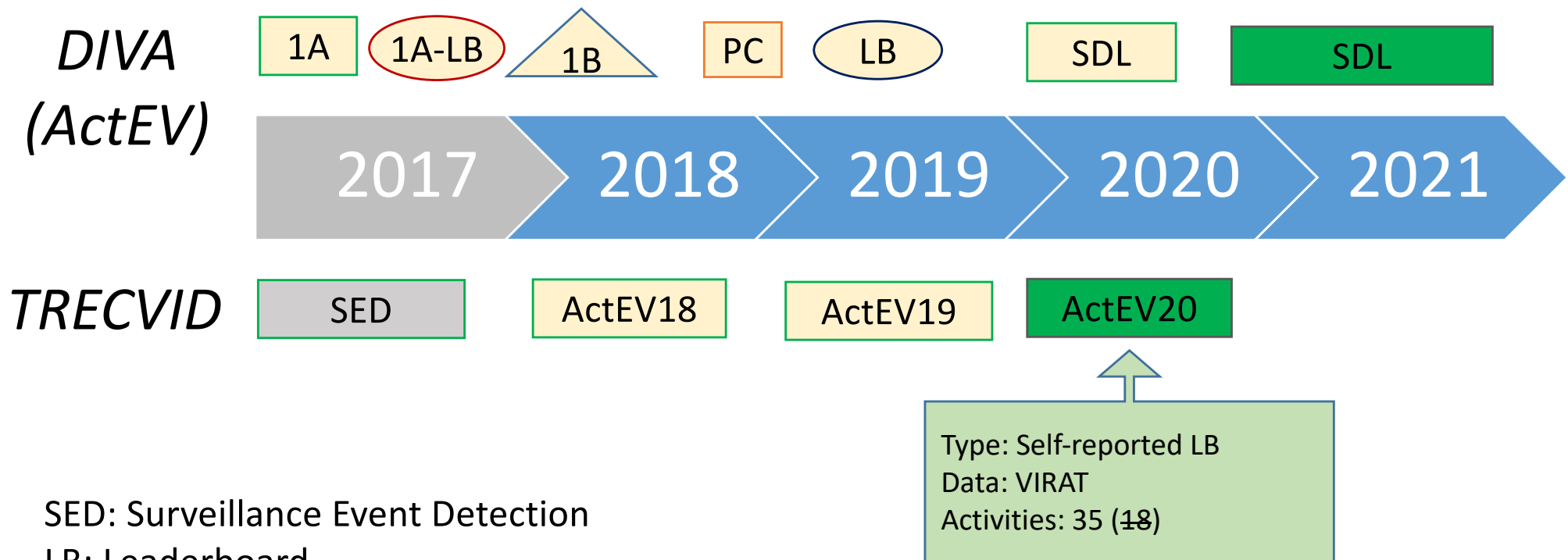 - Added 17 new activities

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's datasets were collected and annotated by Kitware, Inc.

ActEV Series

https://actev.nist.gov/

Type: Sequestered LB
Data:
    Known Facility
    Unknown Facility
Activities:
    37 Known Activities
    10 Surprise Activities

*DIVA
(ActEV)*

| 1A | 1A-LB | 1B | PC | LB | SDL | SDL |

| 2017 | 2018 | 2019 | 2020 | 2021 |

TRECVID

| SED | ActEV18 | ActEV19 | ActEV20 |

Type: Self-reported LB
Data: VIRAT
Activities: 35 (~~18~~)

SED: Surveillance Event Detection
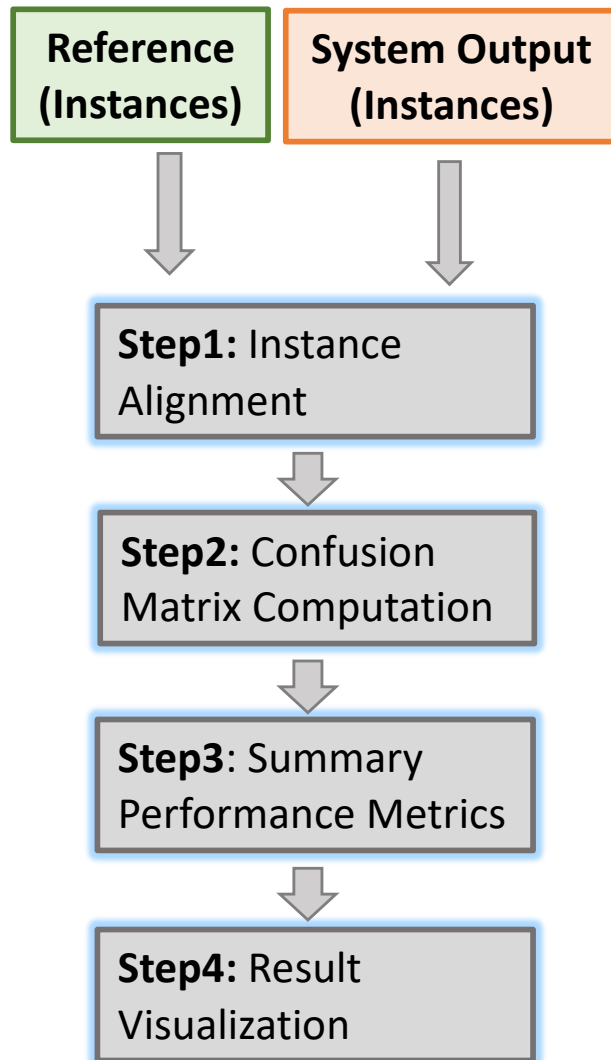LB: Leaderboard
PC: Prize Challenge
SDL: Sequestered Data Leaderboard

ActEV20 Tasks and Measures

Evaluation Tasks (AD)

- "Activity" definition for this evaluation
  - 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 1) detects its presence and then temporally localizes all instances of the activity in video sequences
    - The temporal overlap must fall within a minimal requirement
  - The system output includes:
    - Start and end frames indicating the temporal location of the target activity
    - A presence confidence score that indicates how likely the activity occurred

Performance Metric Calculation

Reference (Instances) → System Output (Instances)

Step1: Instance Alignment

↓

Step2: Confusion Matrix Computation

↓

Step3: Summary Performance Metrics

↓

Step4: Result Visualization

Primary Performance Measures (AD)

$$nAUDC_a = \frac{1}{a}\int_{x=0}^{a} P_{miss}(x)\,dx\,, x = T_{fa}$$

$$P_{miss}(x) = \frac{N_{md}(x)}{N_{TrueInstance}}$$

$$T_{fa} = \frac{1}{NR}\sum_{i=1}^{N_{frames}} max(0, S'_i - R'_i)$$

DET (Detection Error Tradeoff)

P. of Missed Detections

Prob. of Miss Detection

False Alarms

Pmiss@TFA0.02

nAUDC0@TFA0.2

ActEV20 Dataset

National Institute of Standards and Technology / U.S. Department of Commerce
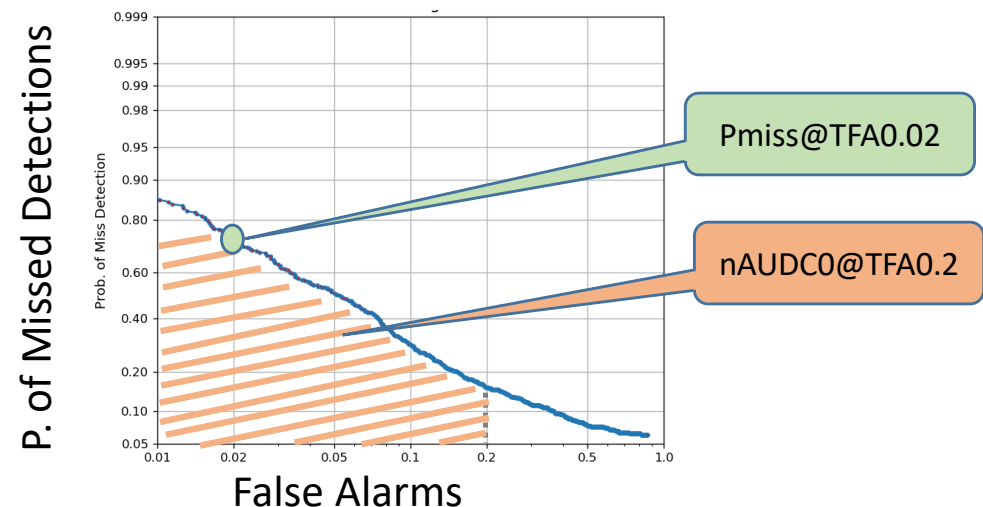
Activities and Number of Instances

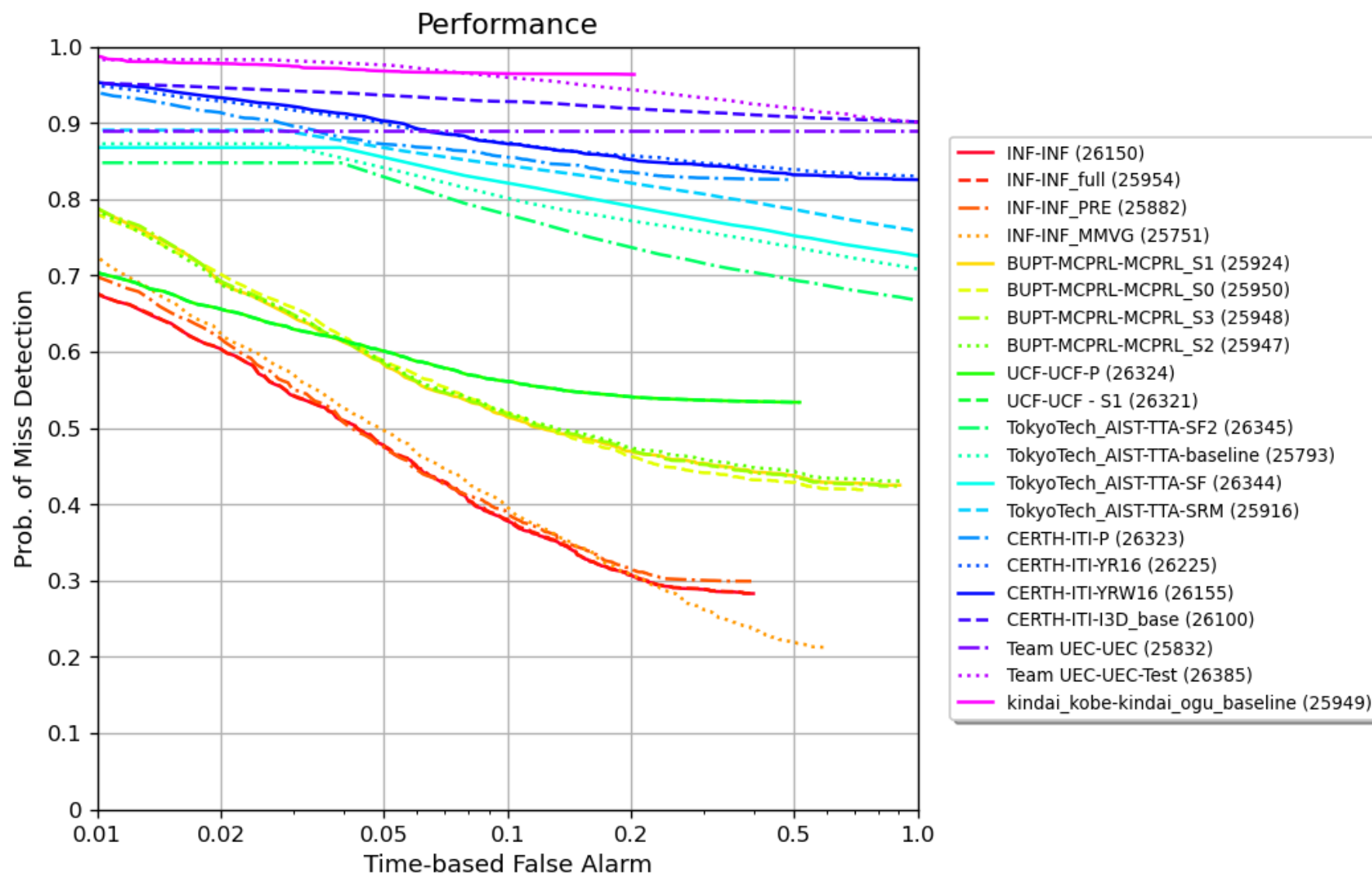| VIRAT18 &19 (18 Activities)	VIRAT20 (35 Activities)	Train	Validate
Closing	person_closes_facility_or_vehicle_door	141	130
Closing_Trunk	person_closes_trunk	21	31
	vehicle_drops_off_person	0	4
Entering	person_enters_facility_or_vehicle	77	70
Exiting	person_exits_facility_or_vehicle	66	72
	person_interacts_object	101	88
Loading	person_loads_vehicle	38	38
Open_Trunk	person_opens_trunk	22	35
Opening	person_opens_facility_or_vehicle_door	137	128
	person_person_interaction	11	17
	person_pickups_object	19	12
	vehicle_picks_up_person	9	5
Pull	person_pulls_object	23	43
	person_pushs_object	4	6
Riding	person_rides_bicycle	22	21
	person_sets_down_object	12	11
Talking	person_talks_to_person	41	67
Transport_HeavyCarry	person_carries_heavy_object	31	44
Unloading	person_unloads_vehicle	32	44
activity_carrying	person_carries_object	237	364
	person_crouches	1	9
	person_gestures	82	148
	person_runs	14	18
	person_sits	21	11
	person_stands	398	819
	person_walks	761	901
specialized_talking_phone	person_talks_on_phone	17	16
specialized_texting_phone	person_texts_on_phone	5	20
	person_uses_tool	7	11
	vehicle_moves	718	797
	vehicle_starts	259	239
	vehicle_stops	292	295
vehicle_turning_left	vehicle_turns_left	152	176
vehicle_turning_right	vehicle_turns_right	149	172
vehicle_u_turn	vehicle_makes_u_turn	9	13

- The Activity names were made more consistent with the MEVA names
- Added 17 more activities (35 total target activities)
- Due to ongoing evaluations, the test counts are not included in the table

ActEV20
Results and Analyses

Leaderboard activity-Averaged DET Curves

35 Activities - As of 12/04/2020

Performance

Legend:
- INF-INF (26150)
- INF-INF_full (25954)
- INF-INF_PRE (25882)
- INF-INF_MMVG (25751)
- BUPT-MCPRL-MCPRL_S1 (25924)
- BUPT-MCPRL-MCPRL_S0 (25950)
- BUPT-MCPRL-MCPRL_S3 (25948)
- BUPT-MCPRL-MCPRL_S2 (25947)
- UCF-UCF-P (26324)
- UCF-UCF - S1 (26321)
- TokyoTech_AIST-TTA-SF2 (26345)
- TokyoTech_AIST-TTA-baseline (25793)
- TokyoTech_AIST-TTA-SF (26344)
- TokyoTech_AIST-TTA-SRM (25916)
- CERTH-ITI-P (26323)
- CERTH-ITI-YR16 (26225)
- CERTH-ITI-YRW16 (26155)
- CERTH-ITI-I3D_base (26100)
- Team UEC-UEC (25832)
- Team UEC-UEC-Test (26385)
- kindai_kobe-kindai_ogu_baseline (25949)

ActEV20 Participants

118 submissions (as of 11/06/2020) from 7 teams from 4 countries (best system result per site)

Team	Organization	nAUDC
INF-CMU	Carnegie Mellon University, USA	0.4230731
BUPT-MCPRL	Beijing University of Posts and Telecommunications, China	0.5551507
UCF	University of Central Florida, USA	0.584852
TokyoTech_AIST	Tokyo Institute of Technology, Japan	0.7975303
CERTH-ITI	Information Technologies Institute, Greece	0.8657579
Team UEC	The University of Electro-Communications, Japan	0.951682
kindai_kobe	Kindai University and Osaka Gakuin University, Japan	0.9682038

NIST National Institute of Standards and Technology / U.S. Department of Commerce
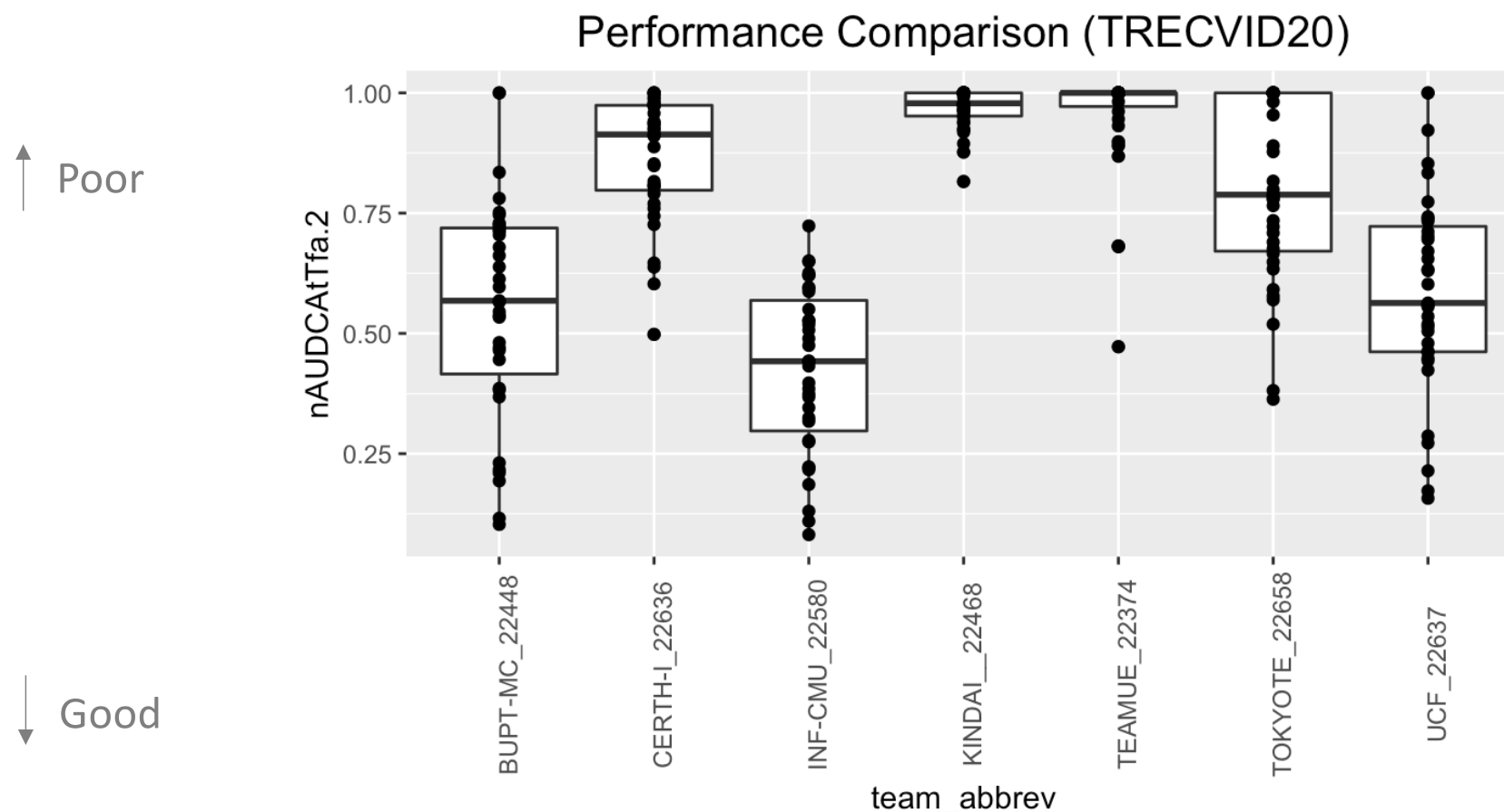
Performance Ranking by System

- The ranked list of system performance (ordered by nAUDC)

Team	nAUDC	mean_p_missAtTfa.15	mean_w_P_missAtRfa.15	mean_P_missAtRfa.15
INF-CMU	0.4230731	0.3324113	0.8096508	0.8046004
BUPT-MCPRL	0.5551507	0.4877911	0.8451903	0.8458609
UCF	0.584852	0.5472993	0.8353953	0.8344754
TokyoTech_AIST	0.7975303	0.7550181	0.8788922	0.878756
CERTH-ITI	0.8657579	0.8445439	0.8823662	0.8951434
Team UEC	0.951682	0.9532886	0.9829977	0.9872394
kindai_kobe	0.9682038	0.9644262	0.9566538	0.9691709

NIST National Institute of Standards and Technology / U.S. Department of Commerce

Performance Ranking (AD)
(35 Activities – Best per Site)

Performance Comparison (TRECVID20)

Observation
- Highest performance on activity detection:
  - INF-CMU (nAUDC: 0.423) followed by BUPT-MC (nAUDC: 0.555)
- A large variance of the 35 activities across for higher performant systems

Activity Ranking (AD)

TRECVID20 Ranked list of activities across systems
Data: VIRAT, Type: TRECVID20 Metric: Ordered by nAUDCAtTfa.2
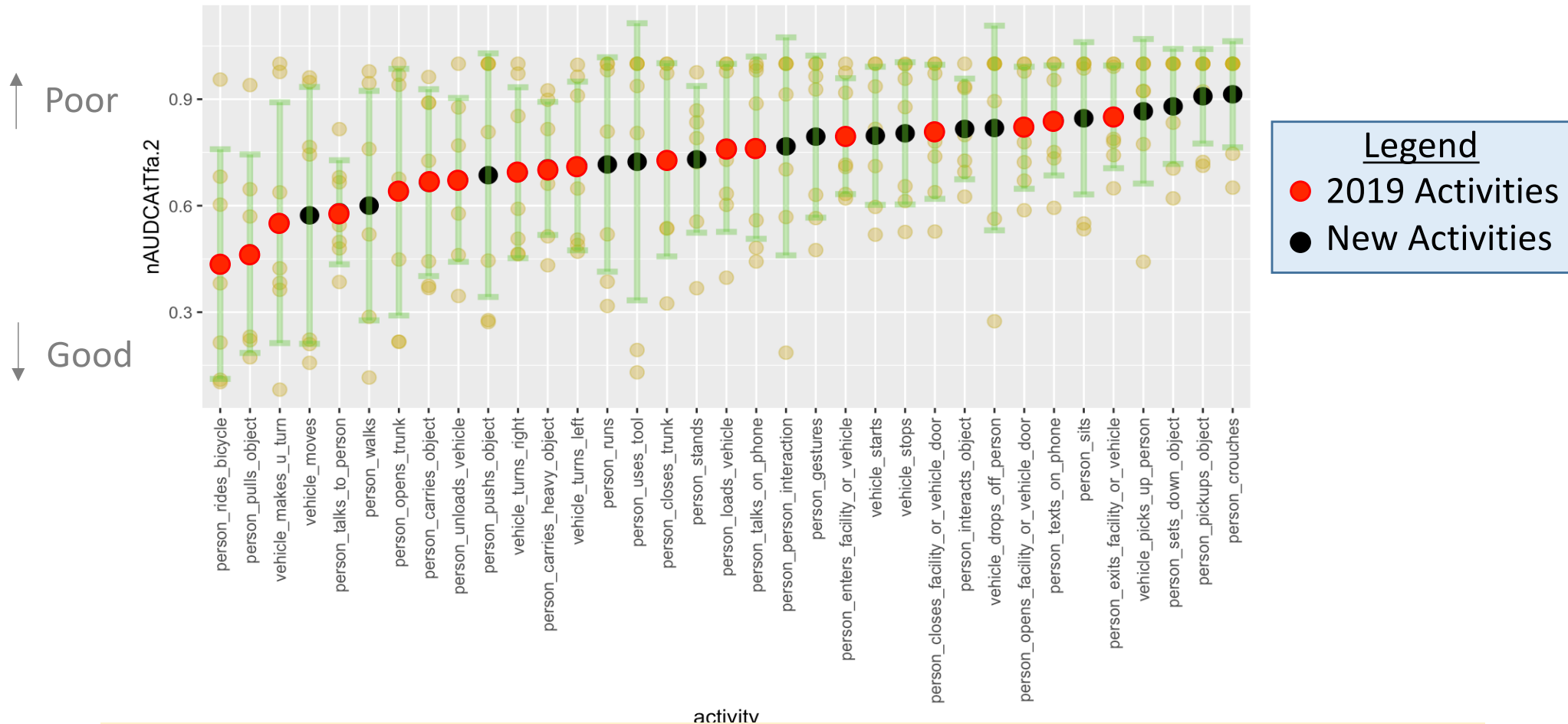
Observation

- Given the dataset and the 35 activities, "person_rides_bicycle" is the easiest to detect while "person_crouches" is the hardest across the

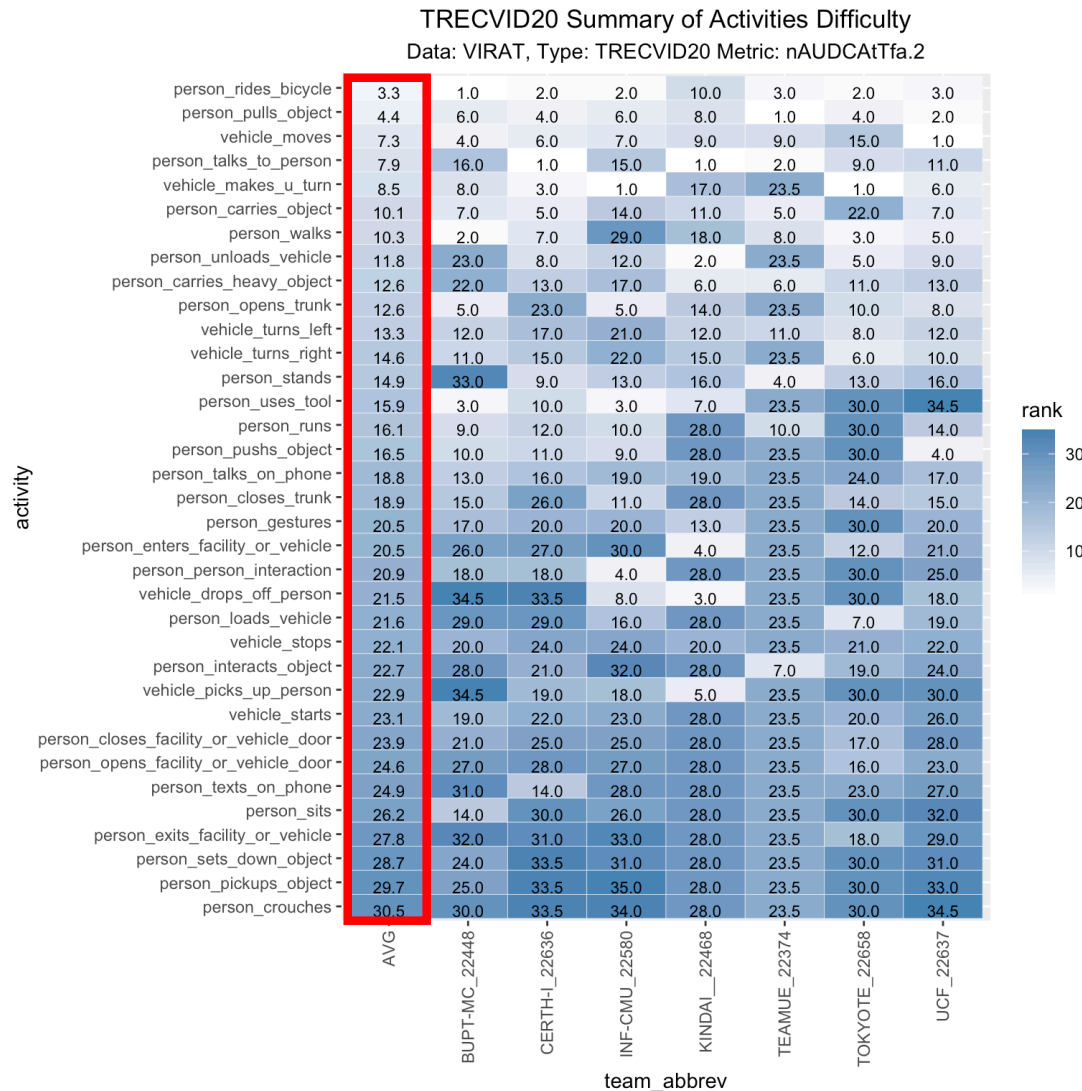# Which activities are easier or more difficult to detect?

TRECVID20 Summary of Activities Difficulty
Data: VIRAT, Type: TRECVID20 Metric: nAUDCAtTfa.2

- X-axis: team names and and average activity ranking (AVG)
- Y-axis: 35 activities - Numbers in the matrix: the ranking of 35 activities per system

activity	AVG	BUPT-MC_22448	CERTH-I_22636	INF-CMU_22580	KINDAI_22468	TEAMUE_22374	TOKYOTE_22658	UCF_22637
person_rides_bicycle	3.3	1.0	2.0	2.0	10.0	3.0	2.0	3.0
person_pulls_object	4.4	6.0	4.0	6.0	8.0	1.0	4.0	2.0
vehicle_moves	7.3	4.0	6.0	7.0	9.0	9.0	15.0	1.0
person_talks_to_person	7.9	16.0	1.0	15.0	1.0	2.0	9.0	11.0
vehicle_makes_u_turn	8.5	8.0	3.0	1.0	17.0	23.5	1.0	6.0
person_carries_object	10.1	7.0	5.0	14.0	11.0	5.0	22.0	7.0
person_walks	10.3	2.0	7.0	29.0	18.0	8.0	3.0	5.0
person_unloads_vehicle	11.8	23.0	8.0	12.0	2.0	23.5	5.0	9.0
person_carries_heavy_object	12.6	22.0	13.0	17.0	6.0	6.0	11.0	13.0
person_opens_trunk	12.6	5.0	23.0	5.0	14.0	23.5	10.0	8.0
vehicle_turns_left	13.3	12.0	17.0	21.0	12.0	11.0	8.0	12.0
vehicle_turns_right	14.6	11.0	15.0	22.0	15.0	23.5	6.0	10.0
person_stands	14.9	33.0	9.0	13.0	16.0	4.0	13.0	16.0
person_uses_tool	15.9	3.0	10.0	3.0	7.0	23.5	30.0	34.5
person_runs	16.1	9.0	12.0	10.0	28.0	10.0	30.0	14.0
person_pushs_object	16.5	10.0	11.0	9.0	28.0	23.5	30.0	4.0
person_talks_on_phone	18.8	13.0	16.0	19.0	19.0	23.5	24.0	17.0
person_closes_trunk	18.9	15.0	26.0	11.0	28.0	23.5	14.0	15.0
person_gestures	20.5	17.0	20.0	20.0	13.0	23.5	30.0	20.0
person_enters_facility_or_vehicle	20.5	26.0	27.0	30.0	4.0	23.5	12.0	21.0
person_person_interaction	20.9	18.0	18.0	4.0	28.0	23.5	30.0	25.0
vehicle_drops_off_person	21.5	34.5	33.5	8.0	3.0	23.5	30.0	18.0
person_loads_vehicle	21.6	29.0	29.0	16.0	28.0	23.5	7.0	19.0
vehicle_stops	22.1	20.0	24.0	24.0	20.0	23.5	21.0	22.0
person_interacts_object	22.7	28.0	21.0	32.0	28.0	7.0	19.0	24.0
vehicle_picks_up_person	22.9	34.5	19.0	18.0	5.0	23.5	30.0	30.0
vehicle_starts	23.1	19.0	22.0	23.0	28.0	23.5	20.0	26.0
person_closes_facility_or_vehicle_door	23.9	21.0	25.0	25.0	28.0	23.5	17.0	28.0
person_opens_facility_or_vehicle_door	24.6	27.0	28.0	27.0	28.0	23.5	16.0	23.0
person_texts_on_phone	24.9	31.0	14.0	28.0	28.0	23.5	23.0	27.0
person_sits	26.2	14.0	30.0	26.0	28.0	23.5	30.0	32.0
person_exits_facility_or_vehicle	27.8	32.0	31.0	33.0	28.0	23.5	18.0	29.0
person_sets_down_object	28.7	24.0	33.5	31.0	28.0	23.5	30.0	31.0
person_pickups_object	29.7	25.0	33.5	35.0	28.0	23.5	30.0	33.0
person_crouches	30.5	30.0	33.5	34.0	28.0	23.5	30.0	34.5

team_abbrev

rank
30
20
10

Observation: the person_rides_bicycle, person_pull_object, and vehicle_moves activities are easier to detect compared to the rest of the other activities

Comparison of ActEV18, ActEV19 and ActEV20 Results

Team	ActEV18		ActEV19		ActEV20	
	Self(12)	LB (19)	LB (18)		LB (18)	
	PR.15↓	PR.15↓	PR.15↓	nAUDC	PR.15	nAUDC
UMD	0.618	x	x	x	x	x
SeuGraph	0.624	x	x	x	x	x
Team_Vision	0.71	0.709	x	x	x	x
UCF	0.759	0.733	0.68	0.491	0.817	0.5188
STR-DIVA Team	0.827	x	x	x	x	x
JHUDIVATeam	0.887	x	x	x	x	x
CMU/INF	0.896	0.844	0.789	0.484	0.788	0.405 ⭐
SRI	0.927	x	x	x	x	x
VANT	0.94	0.882	x	x	x	x
HSMW_TUC	0.961	x	0.951	0.941	x	x
BUPT-MCPRL	0.99	0.749	0.736	0.524	0.807	0.526
USF Bulls	0.991	0.934	x	x	x	x
MKLab (ITI_CERTH)	0.999	x	0.968	0.964	0.867	0.833 ⭐
UTS-CETC	x	0.925	x	x	0.976	0.923
NII_Hitachi_UIT	x	0.925	0.819	0.599	x	x
Fraunhofer IOSB	x	x	0.849	0.827	x	x
NTT_CQUPT	x	x	0.878	0.601	x	x
vireoJD-MM	x	x	0.714	0.601	x	x
TokyoTech_AIST	x	x	x	x	0.821	0.689
Kindai_Kobe	x	x	x	x	0.9499	0.959

Observation:
- System performance on 18 activities improved from ActEV19 to ActEV20 for CMU and ITI_CERTH

Self: Self-reported eval, LB: Leaderboard eval

PR.15: μP_{miss} at $R_{FA} = 0.15$

Summary

- The ActEV20 Activity names were made more consistent with the MEVA names and we added 17 more Activities (35 Target Activities)

- 118 submissions out of 7 teams from 4 countries

- Given the test set and the 35 activities, the person_rides_bicycle, person_pull_object, and vehicle_moves activities are easier to detect compared to the rest of the other activities

- System performance on 18 activities improved from ActEV19 to ActEV20 for CMU and ITI_CERTH

National Institute of Standards and Technology / U.S. Department of Commerce

Next Steps

ActEV Next Steps/Opportunities

- WACV HADCV'21 (Human Activity Detection in multi-camera, Continuous, long-duration Video), 3rd workshop
  - ***Virtual Workshop, Jan 5, 2021:*** https://actev.nist.gov/workshop/hadcv21

- Potential Guest Task at the CVPR ActivityNet 2021 Workshop

- Resources: https://actev.nist.gov (click "Resources")
  - Datasets (training data)
  - Baseline algorithms
  - Annotation Tools

ActEV Sequestered Data Leaderboard (SDL)

- Anyone can **submit their system to NIST**, which will then run the system on **sequestered data** (MEVA), post the results to the leaderboard

- Visit ongoing ActEV SDL Evaluation at https://actev.nist.gov/sdl

- New Sequestered Testing:
  - More activities:
    - 37 Known Activities – training performed before submission
    - 10 Surprise Activities – training at test time from text descriptions and exemplars
  - More Facilities
    - Known Facility – Training/Development MEVA Data (https://mevadata.org/)
      - Indoor and outdoor scenes, night and day, crowds and individuals, EO (Electro-Optical) and IR (Infrared) sensors, Intrinsic/Extrinsic camera calibrations
    - Unknown Facility – No Training/Development data
      - Similar video and metadata to Known Facility never seen before test time

2020 ActEV feedback and 2021 plans

- What do the teams think about the ActEV task ?
- Any feedback on the data repo to download data (VIRAT, MEVA, ..) ?
- Any feedback on the scoring server and different documents?
- Besides the ActEV leaderboard, what additional information would be useful?
- Are there barriers to participation we can remove?
- Is anyone aware of colleagues that work in activity detection that do not participate in TRECVid.

- Current Plan is to continue the ActEV task in 2021
  - Add mAP as a secondary metric
  - We are working on releasing detailed scores automatically updated for teams
  - Potentially leverage MEVA data for TRECVID evaluations
  - Potential cross-team arXiv or journal paper
  - Potential sequestered data evaluation for TRECVID

Questions?

https://actev.nist.gov/

Contact: actev-nist@nist.gov