INFORMATION TECHNOLOGY LABORATORY

ActEV19: Activities in Extended Video (Summary Results)

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

TRECVID 2019 Workshop

November 12-13, 2019

IARPA BE THE FUTURE

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

DIVA

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.

Outline

- ActEV Overview
- TRECVID ActEV19 Evaluation
- ActEV19 Tasks and Measures
- ActEV19 Dataset
- ActEV19 Results and Analyses
- Next Steps

ActEV Overview

What is ActEV?

vehicle_turning_left-00

vehicle_turning_left-04

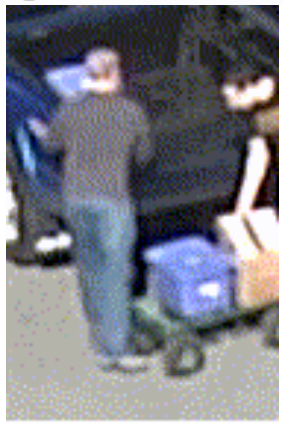
Closing-01

Closing-02

Closing-00

Loading-00

Loading-01



Loading-02

What is ActEV's Goal?

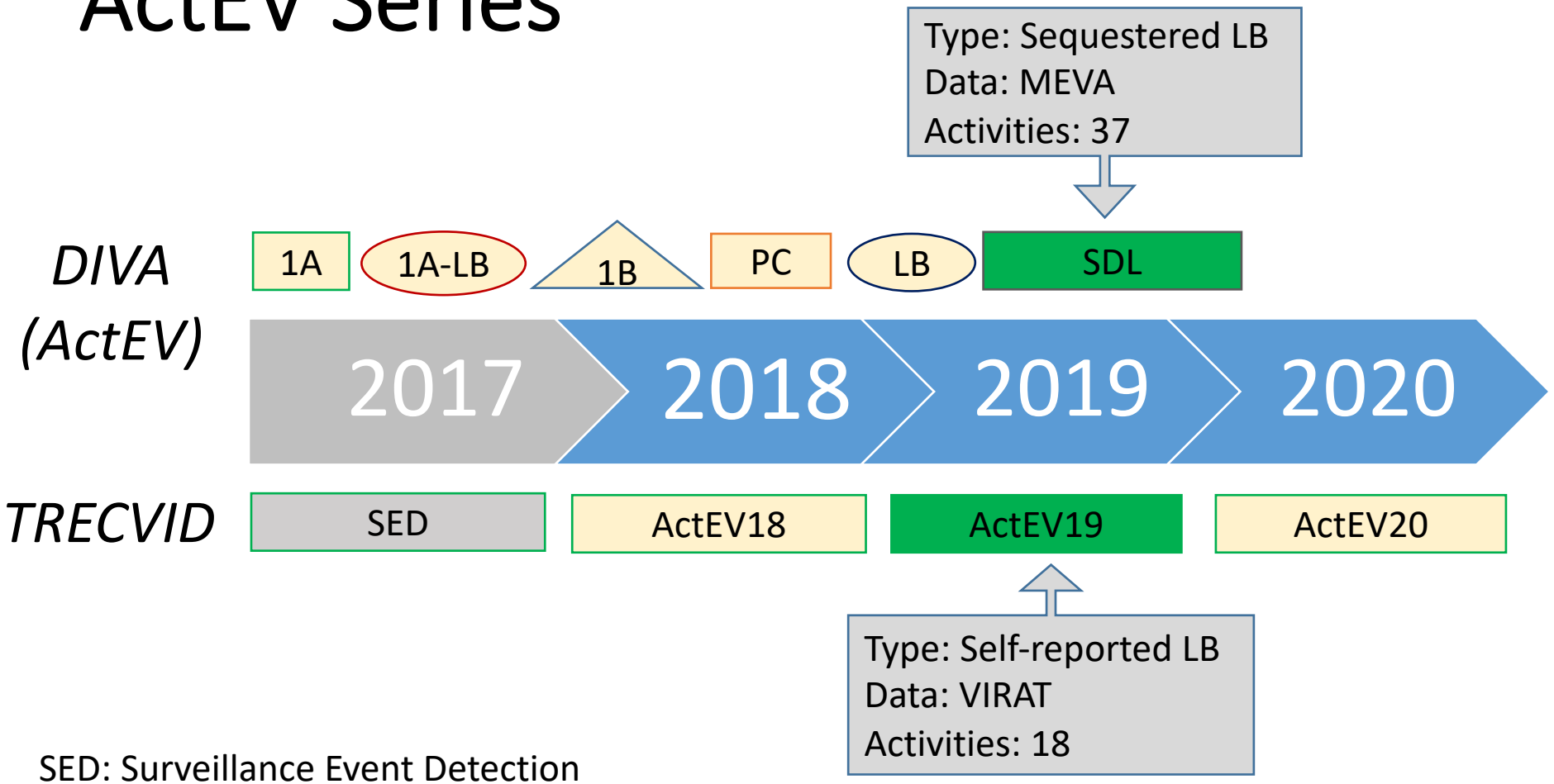
- To advance video analytics technology that can automatically detect a target activity and identify and track objects associated with the activity.
- A series of challenges are also designed for:
  - Activity detection in a multi-camera environment
  - Temporal (and spatio-temporal) localization of the activity for reasoning

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.

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

ActEV Series



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

TRECVID
ActEV19 Evaluation

Evaluation Framework

- Target applications
  - Retrospective analysis of archives (e.g., forensic analytics)
  - Real-time analysis of live video streams (e.g., alerting and monitoring)
- Evaluation Type
  - Self-reported (& take-home) evalulation
    - TRECVID ActEV19
  - Independent (& sequestered) evalulation
    - DIVA ActEV SDL

ActEV19 Tasks and Measures

Evaluation Tasks (AD)

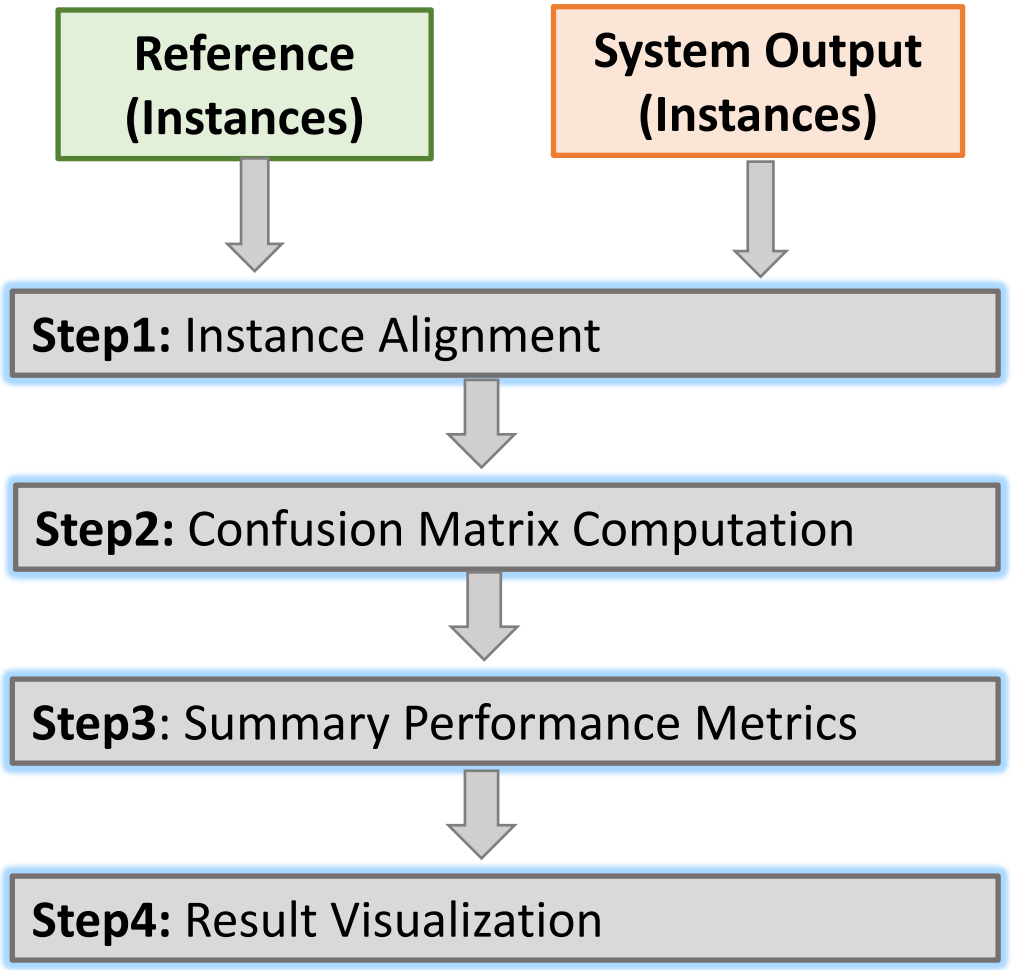
- "Activity" definition for this evaluation
  - One or more people performing a specified movement, or interacting with an object or group of objects (including driving)
- Activity Detection (AD) task
  - Given a target activity, a system automatically 1) detects its presence and then temporally localizes 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

Past Evaluation Tasks (AOD and AODT)

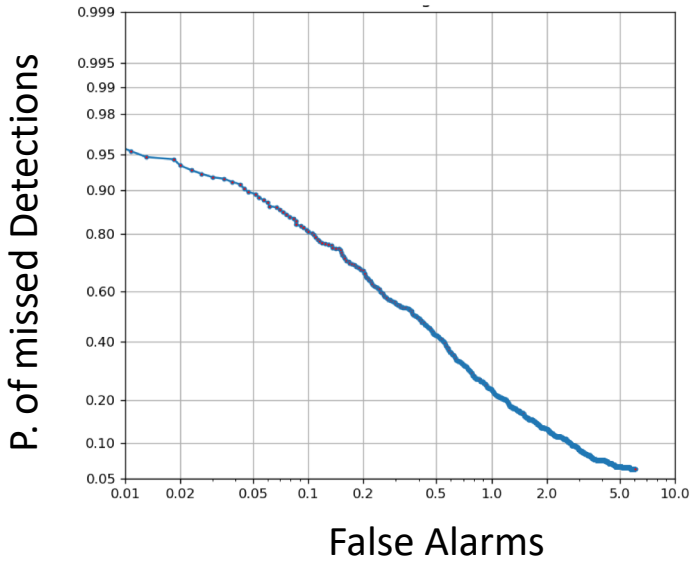
- Activity and Object Detection (AOD)
  - A system not only 1) detects/localizes the target activity, but also 2) detects the presence of required objects and spatially localizes the objects that are associated with the activity
- Activity Object Detection/Tracking (AODT)
  - A system 1) correctly detects/localizes the target activity, 2) correctly detects/localizes the required objects in that activity, and 3) correctly tracks those objects over time.

- The AOD and AODT tasks are NOT addressed in ActEV19 evaluations

Performance Metric Calculation



DET (Detection Error Tradeoff)



Primary Performance Measures (AD)

ActEV18

$$\boldsymbol{P_{miss}} \text{ at } \boldsymbol{R_{fa} = 0.15}$$

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

$$\boldsymbol{R_{fa}(\tau)} = \frac{N_{fa}(\tau)}{VideoDurInMinutes}$$

ActEV19

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

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

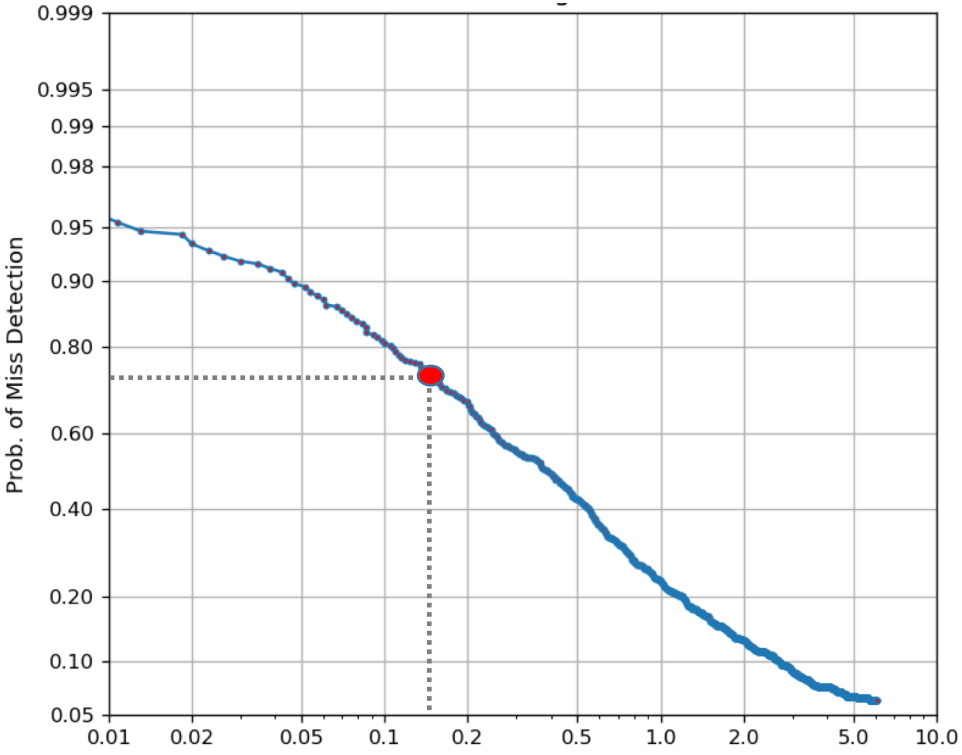
$$\boldsymbol{T_{fa}} = \frac{1}{NR}\sum_{\boldsymbol{i=1}}^{\boldsymbol{N_{frames}}} max(0, S'_i - R'_i)$$

nAUDC (normalized partial Area Under the DET Curve)

Performance Measures (AD)

ActEV18

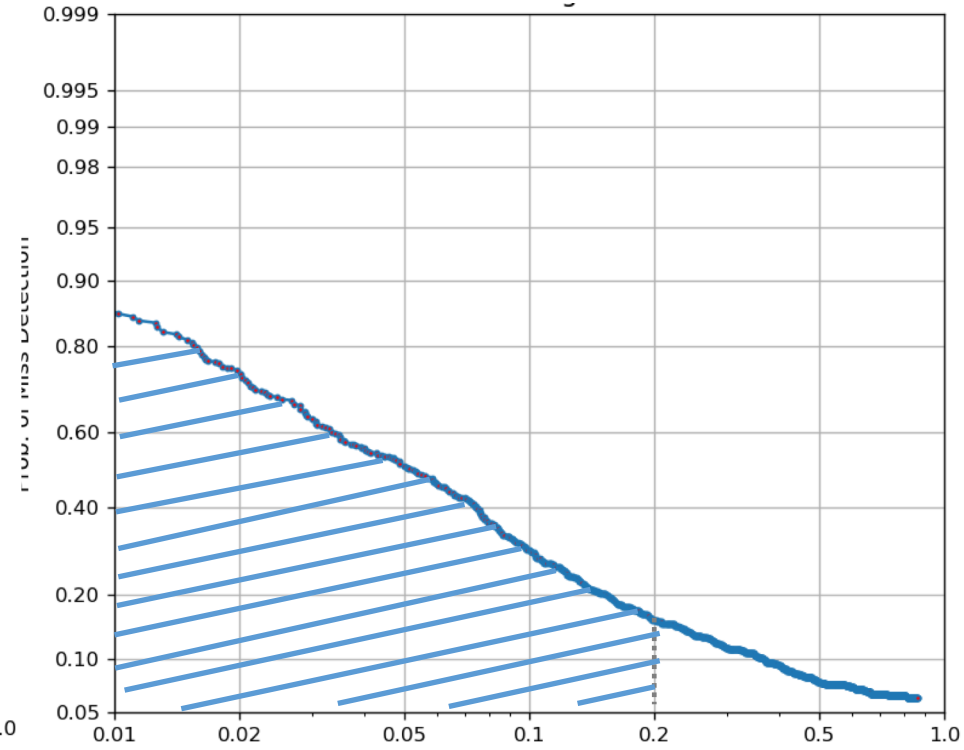
P_{miss} at $R_{FA} = 0.15$



Instance-based Rate of false alarms (R_{fa})

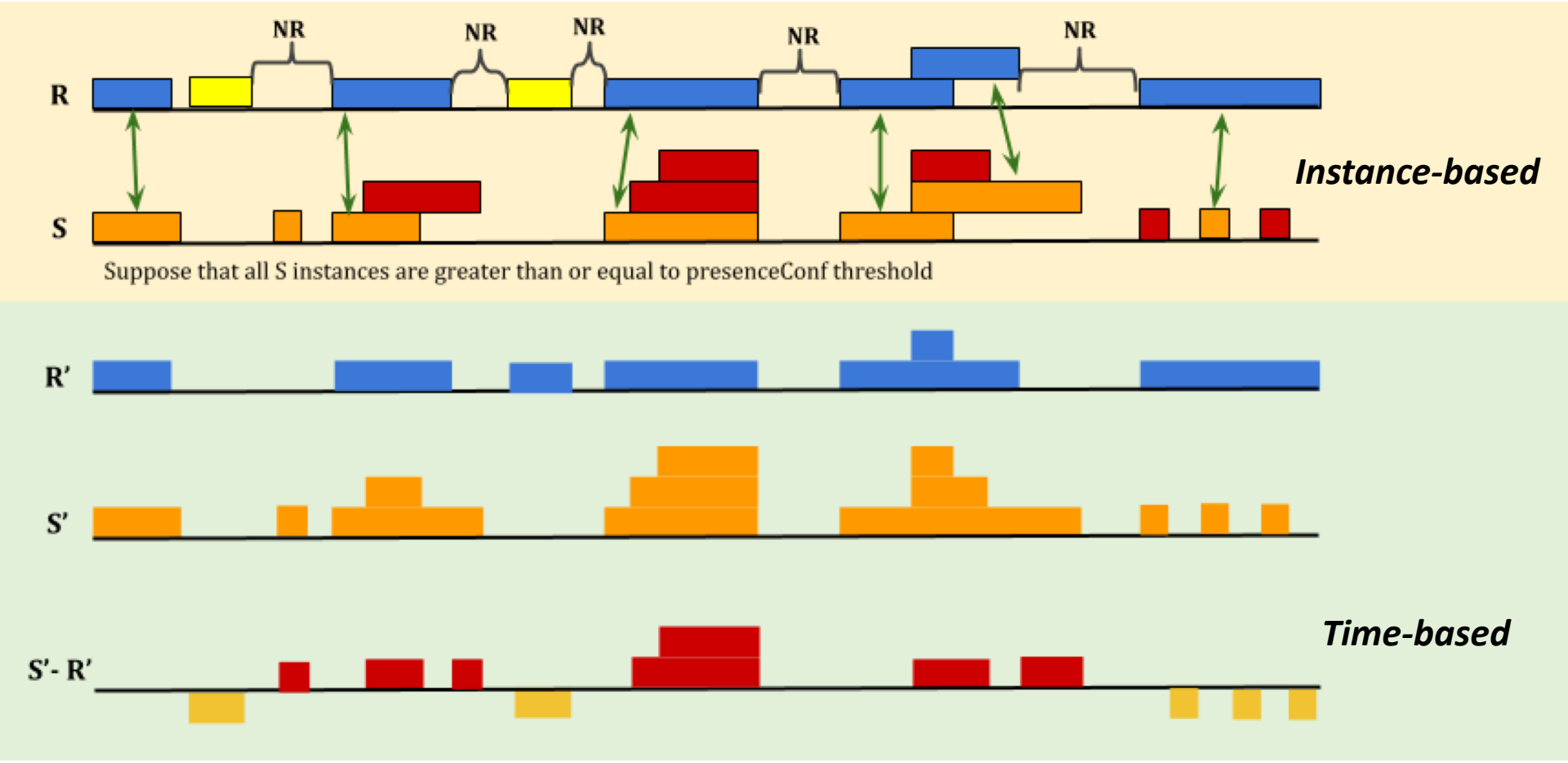
ActEV19

$nAUDC_a, \quad a = 0.2$



Time-based false alarms (T_{fa})

Instance vs Time based False Alarms



NR: Non-Reference

ActEV19 Dataset

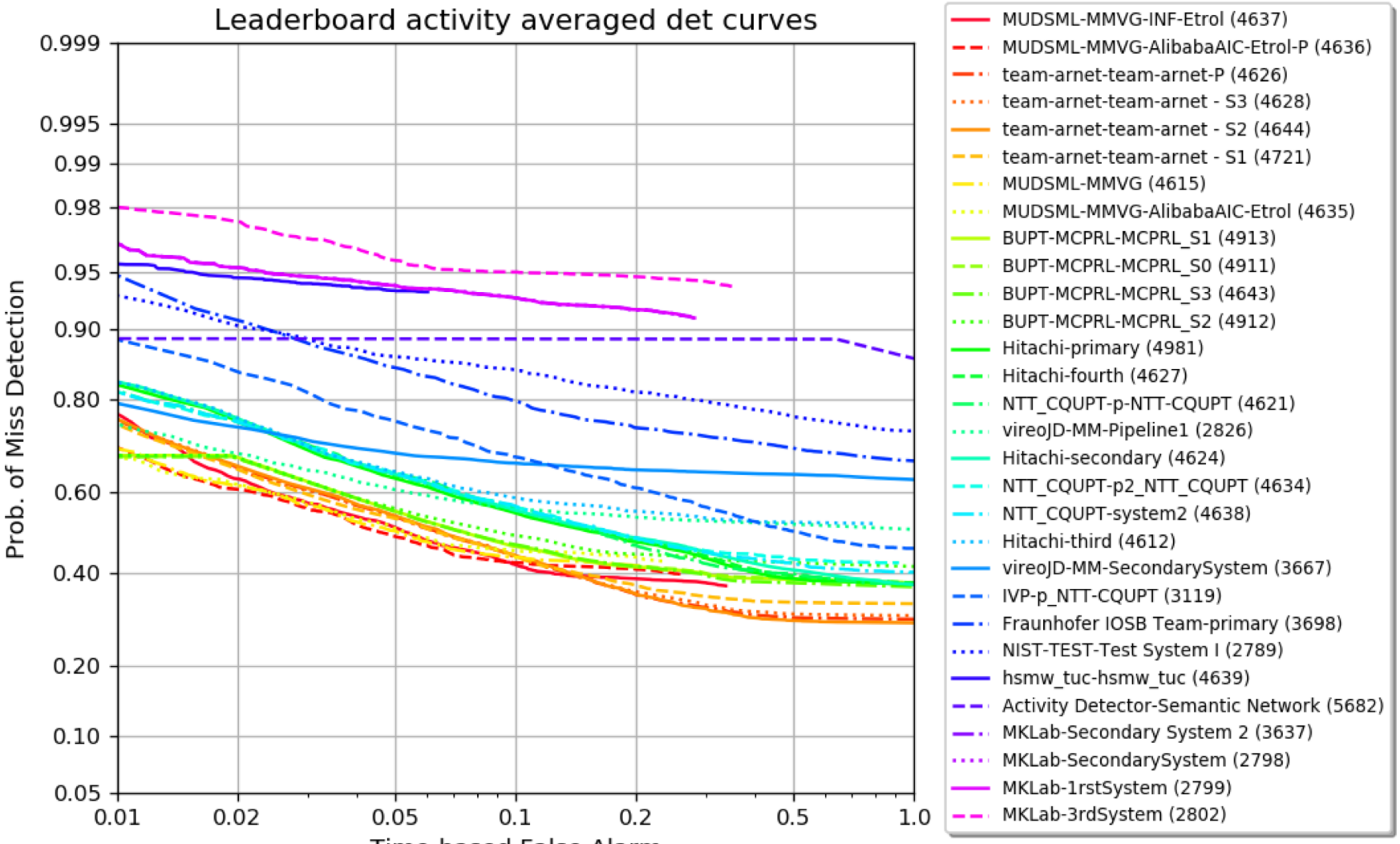
Activities and Number of Instances

Activity Type	ActEV18 (V1)		ActEV19 (V1V2)	
	Train	Validation	Train	Validation
Closing	126	132	126	132
Closing_trunk	31	21	31	21
Entering	70	71	70	71
Exiting	72	65	72	65
Loading	38	37	38	37
Open_Trunk	35	22	35	22
Opening	125	127	125	127
Transport_HeavyCarry	45	31	45	31
Unloading	44	32	44	32
Vehicle_turning_left	152	133	152	133
Vehicle_turning_right	165	137	165	137
Vehicle_u_turn	13	8	13	8
Interacts	88	101	x	x
Pull	21	22	21	22
Riding	21	22	21	22
Talking	67	41	67	41
Activity_carrying	364	237	364	237
Specialized_talking_phone	16	17	16	17
Specialized_texting_phone	20	5	20	5

Due to ongoing evaluations, the test sets are not included in the table

ActEV19
Results and Analyses

As of 11/13/2019

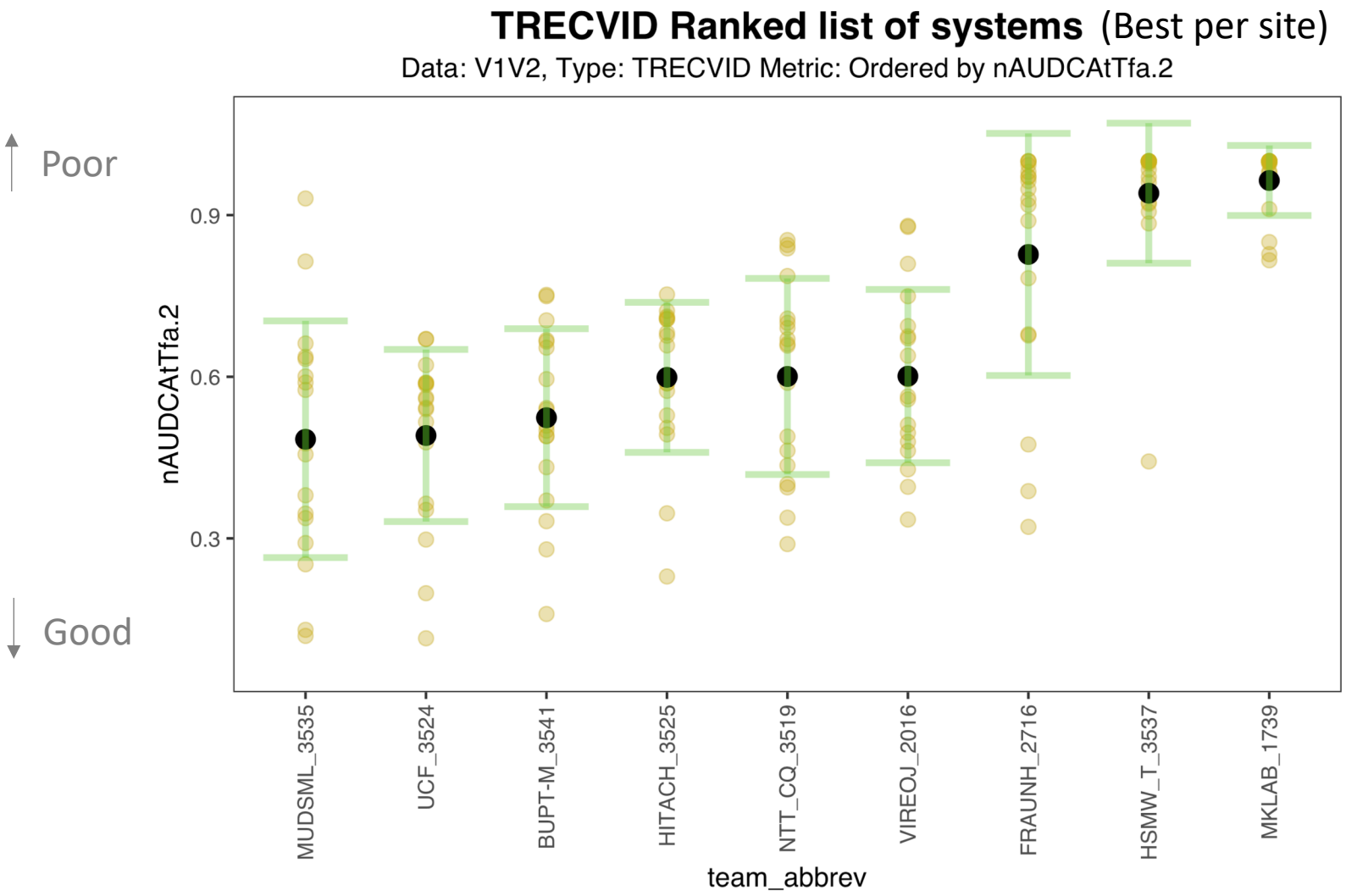


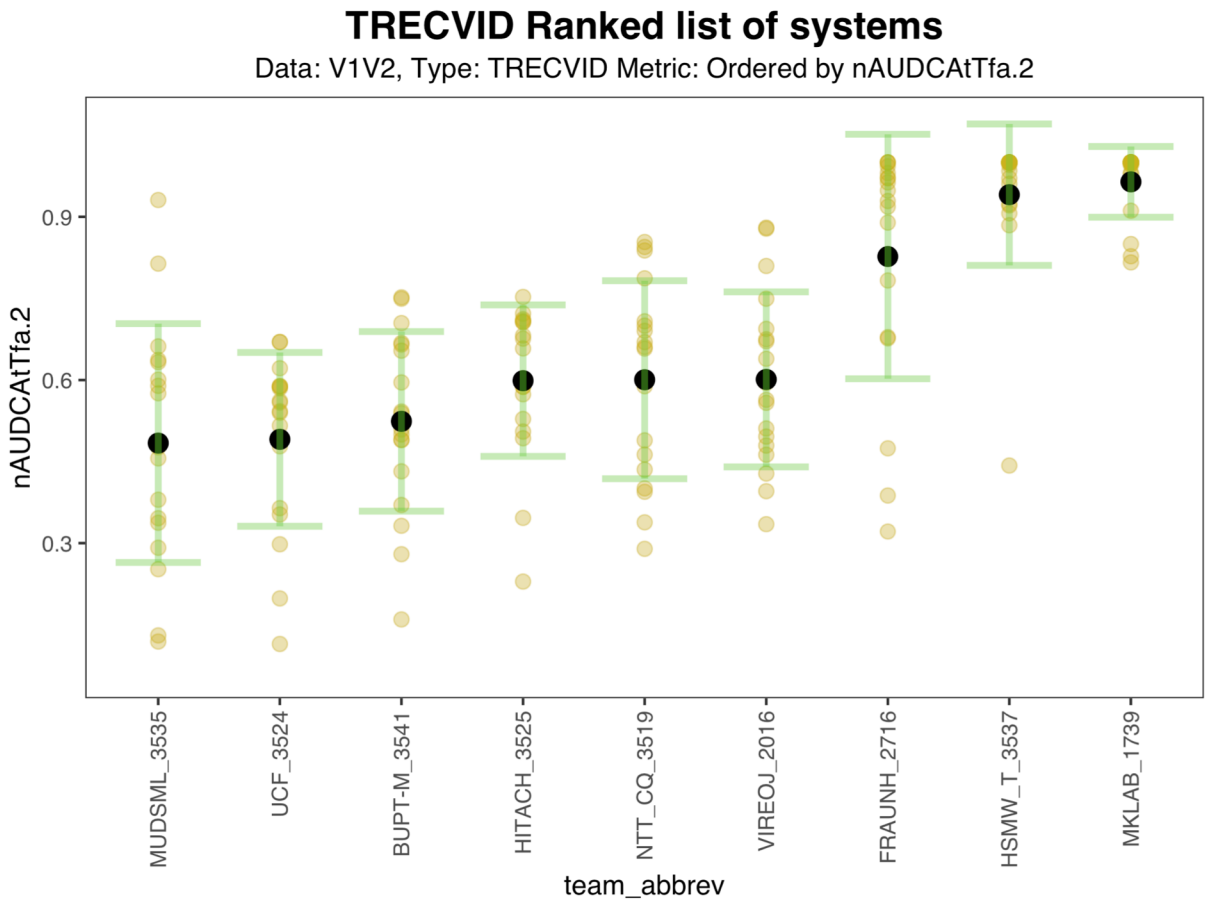
ActEV19 Participants

- 256 submissions (as of 11/1/2019) from 9 teams from 6 countries (best system result per site)

Team	Organization	nAUDC
BUPT-MCPRL	Beijing University of Posts and Telecommunications, China	0.524
Fraunhofer IOSB	Fraunhofer Institute, Germany	0.827
HSMW_TUC	University of Applied Sciences Mittweida and Chemnitz University of Technology, Germany	0.941
MKLab (ITI_CERTH)	Information Technologies Institute, Greece	0.964
MUDSML	Monash University, Australia and Carnegie Mellon University, USA	0.484
NII_Hitachi_UIT	National Institute of Informatics, Japan Hitachi, Ltd., Japan University of Information Technology, Vietnam	0.599
NTT_CQUPT	NTT company & Chongqing University of Posts and Telecommunications, China	0.601
UCF	University of Central Florida, USA	0.491
vireoJD-MM	City University of Hong Kong and JD AI Research, China	0.601

Performance Ranking (AD)

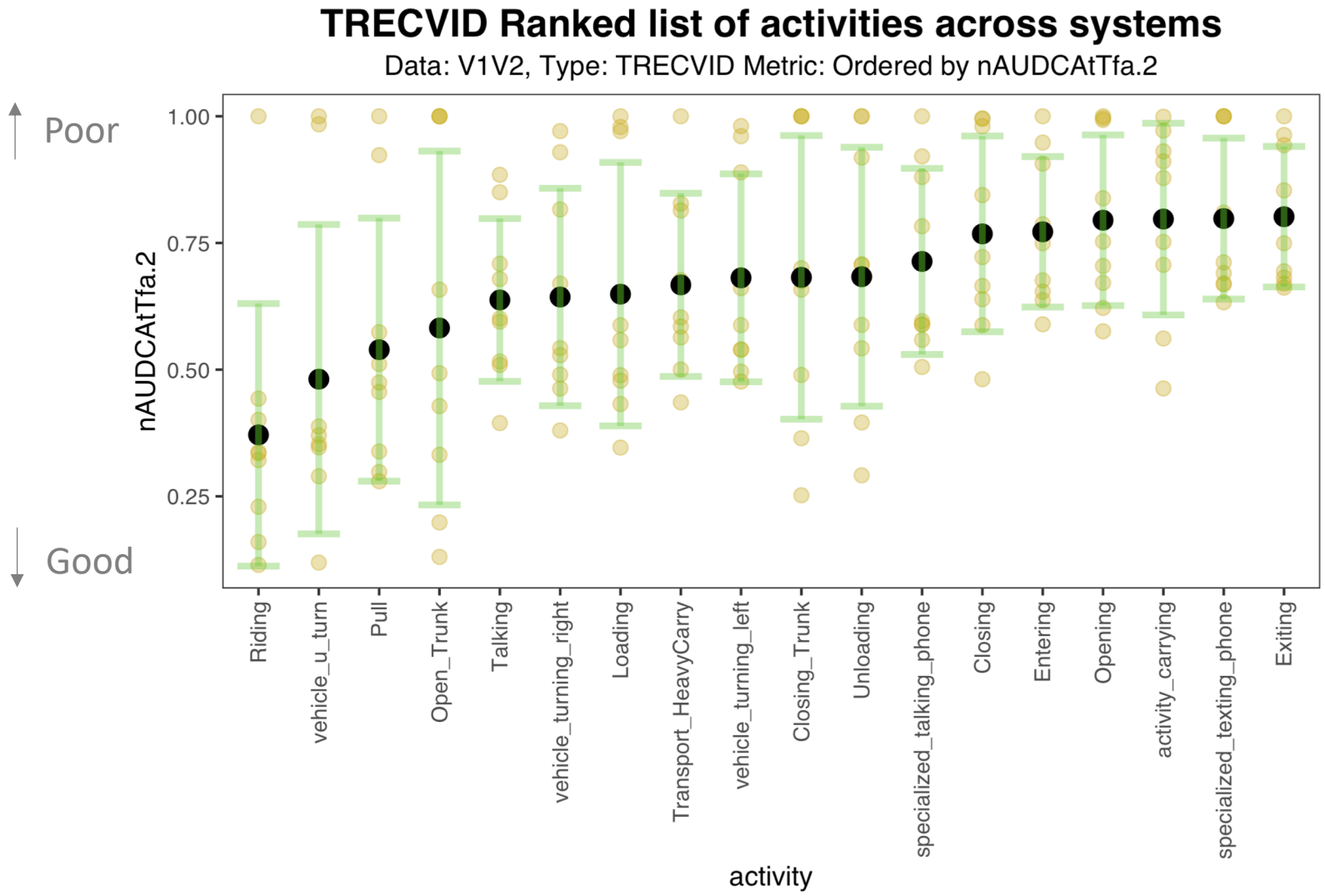


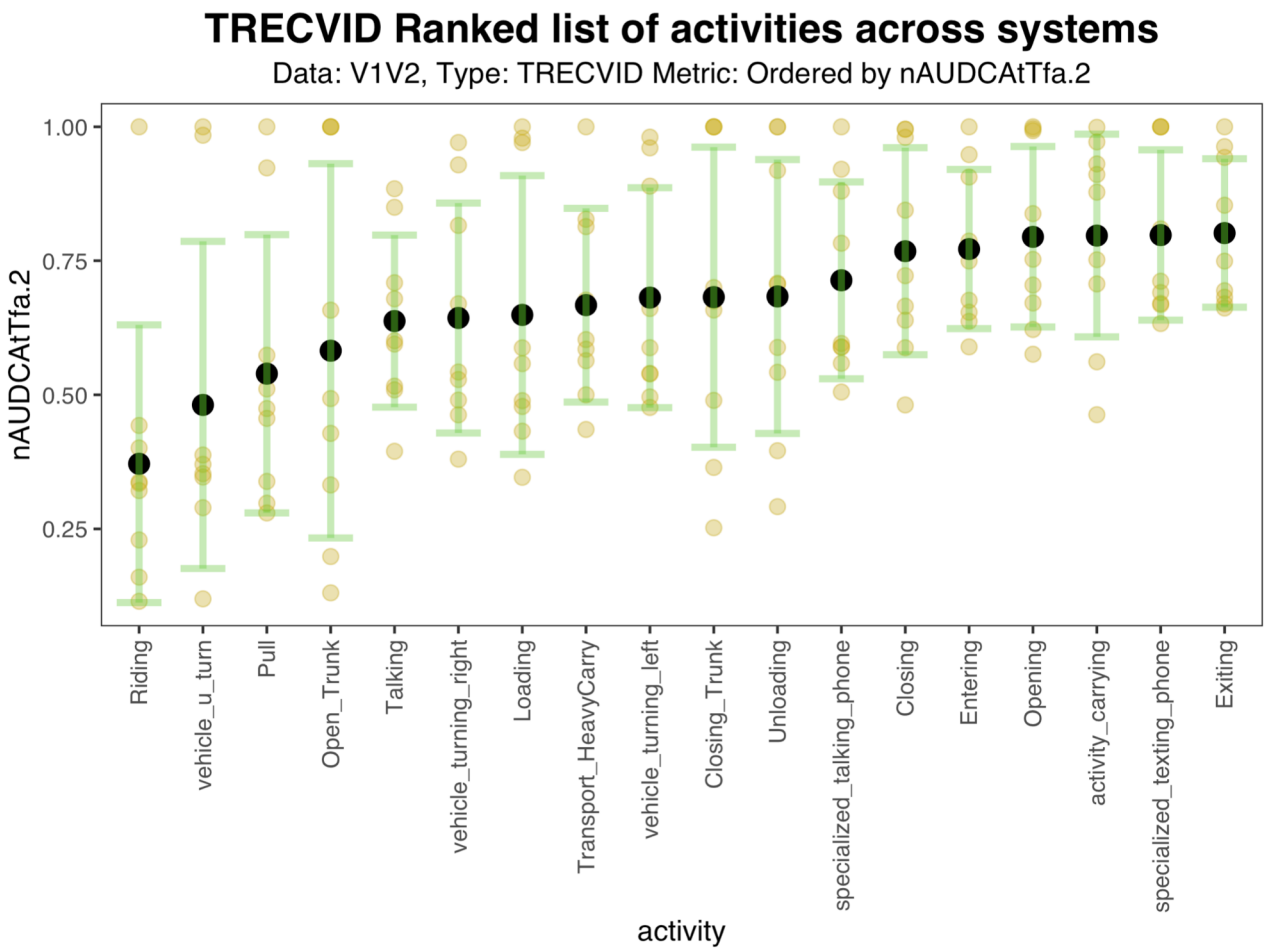


Observation

- Highest performance on activity detection:
  - MUDSML (nAUDC: 48.4%) followed by UCF (nAUDC: 49.1%)
- A large variance of the 18 activities across systems

Activity Ranking (AD)

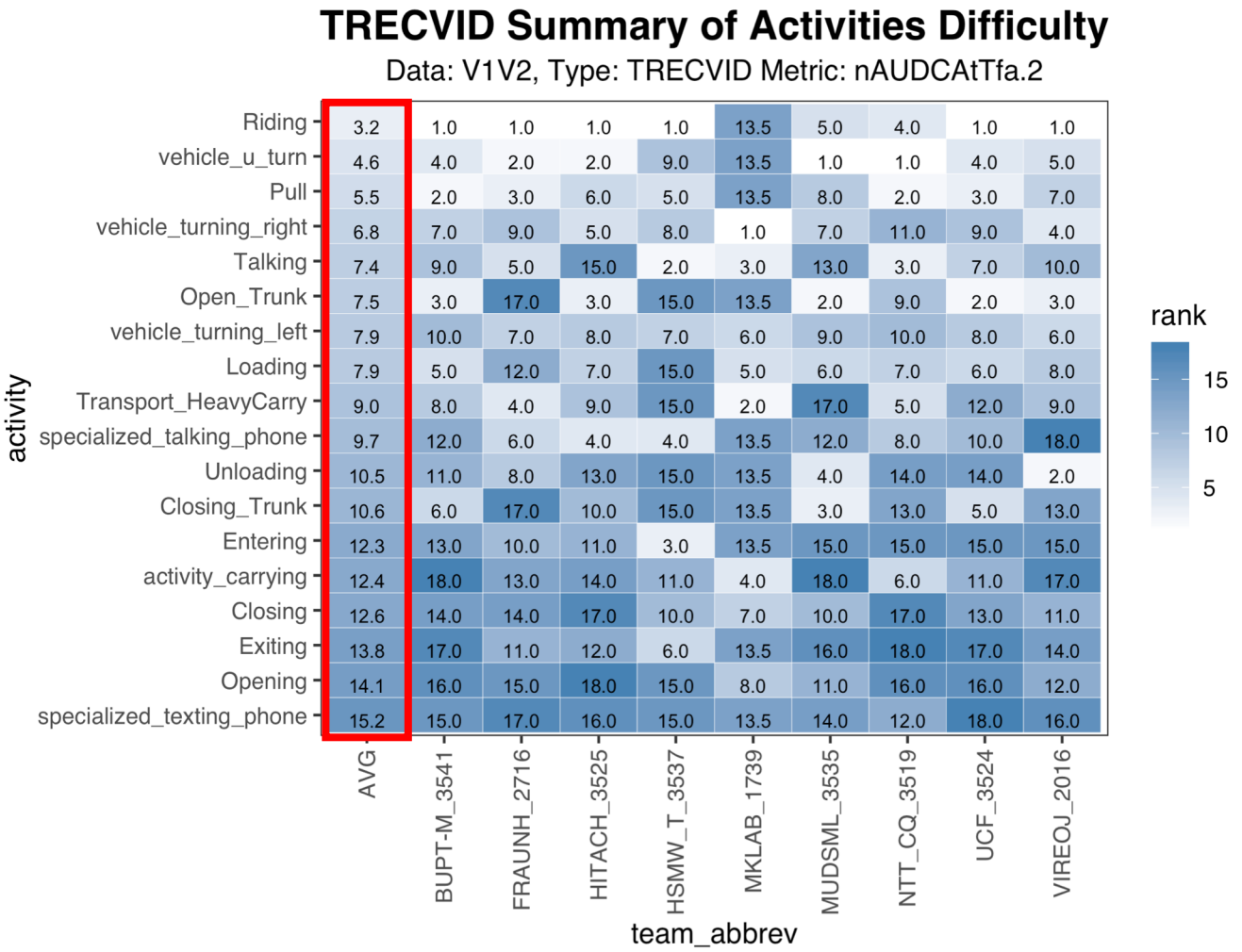




Observation

- Given the dataset and the 18 activities, “Riding” is the easiest to detect while “Exiting” is the hardest across the 9 systems
- “Open_Truck” and “Closing_Truck” have lager variance across systems

Which activities are easier or more difficult to detect?



- X-axis: team names and and average activity ranking (AVG)
- Y-axis:18 activities - Numbers in the matrix: the ranking of 18 activities per system

The activity class was characterized by systems and baseline performance

Observation: the Riding, vehicle_u_turn, and Pull activities are easier to detect compared to the rest of the other activities

Comparison of ActEV18 and ActEV19 Results

Team	ActEV18		ActEV19	
	Self(12)	LB (19)	LB (18)	
	PR.15↓	PR.15↓	PR.15↓	nAUDC
UMD	0.618	x	x	x
SeuGraph	0.624	x	x	x
Team_Vision	0.710	0.709	x	x
UCF	0.759	0.733	0.680	0.491
STR-DIVA Team	0.827	x	x	x
JHUDIVATeam	0.887	x	x	x
MUDSML (INF)	0.896	0.844	0.789	0.484
SRI	0.927	x	x	x
VANT	0.940	0.882	x	x
HSMW_TUC	0.961	x	0.951	0.941
BUPT-MCPRL	0.990	0.749	0.736	0.524
USF Bulls	0.991	0.934	x	x
MKLab (ITI_CERTH)	0.999	x	0.968	0.964
UTS-CETC	x	0.925	x	x
NII_Hitachi_UIT	x	0.925	0.819	0.599
Fraunhofer IOSB	x	x	0.849	0.827
NTT_CQUPT	x	x	0.878	0.601
vireoJD-MM	x	x	0.714	0.601

T: TRECVID, D: DIVA, Self: Self-reported eval, LB: Leaderboard eval

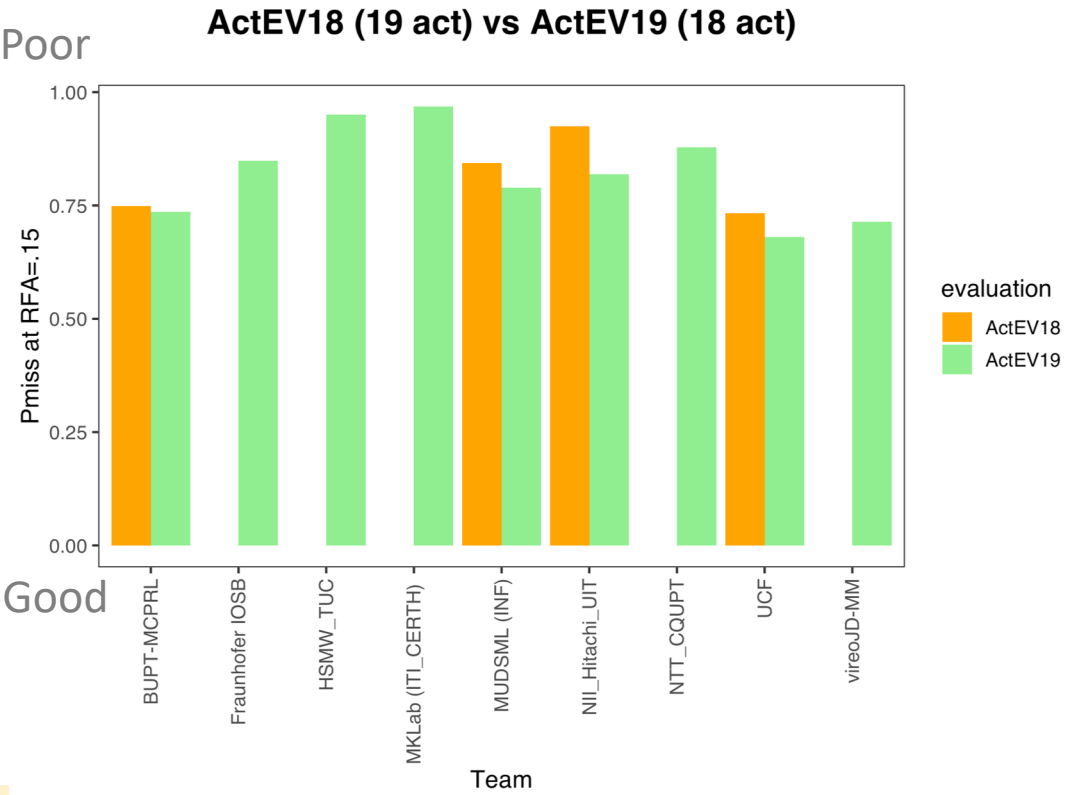
PR.15: μP_{miss} at $R_{FA} = 0.15$

Comparison of ActEV18 vs ActEV19 (Leaderboard only)

	ActEV18	ActEV19
Dataset	VIRAT V1	VIRAT V1V2
# Activities	19	18
Metric	PR.15	PR.15

Team	ActEV18 LB (19)	ActEV19 LB (18)
	PR.15↓	PR.15↓
UCF	0.733	0.680
MUDSML (INF)	0.844	0.789
HSMW_TUC	x	0.951
BUPT-MCPRL	0.749	0.736
MKLab (ITI_CERTH)	x	0.968
NII_Hitachi_UIT	0.925	0.819
Fraunhofer IOSB	x	0.849
NTT_CQUPT	x	0.878
vireoJD-MM	x	0.714

$PR.15$: P_{miss} at $R_{fa} = .15$ ($ActEV18$ $scoring$ $protocol$)



Observation: System performance improved from last year for leaderboard eval. For example, reduced ~12% relative error rate NII_Hitachi_UIT, ~7% for and UCF and MUDSML

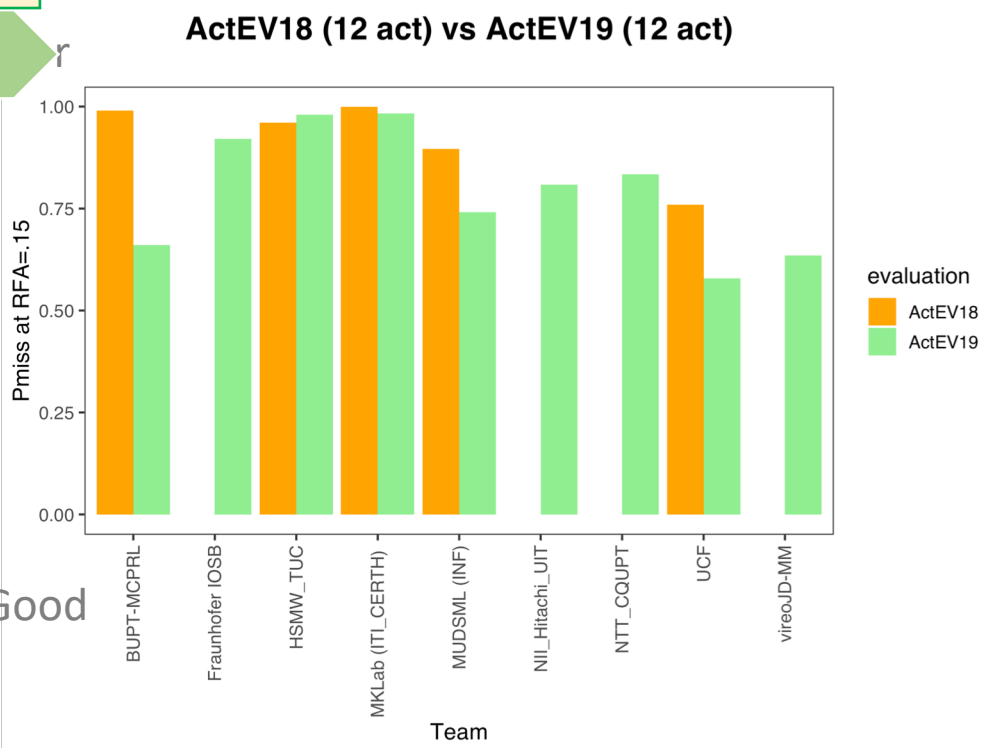
Comparison of ActEV18 vs ActEV19 (12 Activities only)

ActEV18 | ActEV19

Self-reported | Leaderboard | Leaderboard

Team	ActEV18	ActEV19
	PR.15↓	PR.15↓
UCF	0.759	0.605
MUDSML (INF)	0.896	0.731
HSMW_TUC	0.961	0.990
BUPT-MCPRL	0.990	0.683
MKLab (ITI_CERTH)	0.999	0.986
NII_Hitachi_UIT	x	0.827
Fraunhofer IOSB	x	0.921
NTT_CQUPT	x	0.826
vireoJD-MM	x	0.600

Good ↓



Observation:

- System performance on 12 activities improved largely from ActEV18 to ActEV19
- Reduced 31% relative error rate for BUPT-MCPRL, 21% for UCF, and 18% for MUDSML

Summary

- New performance measure to be more relevant to the user cases
- 256 submissions out of 9 teams
- Given the test set and the 18 activities, “Riding” is the easiest while “Exiting” is the hardest across the 9 systems
- Large system improvements this year from last year

Next Steps

Next Steps

- WACV HADCV'20 (Human Activity Detection in multi-camera, Continuous, long-duration Video) workshop (***paper submission deadline: Dec 15, 2019)*** the details at https://wacv20.wacv.net

- Resources: https://actev.nist.gov (click "Resources")
  - Datasets (training data)
  - Baseline algorithms
  - Annotation Tools
- TRECVID ActEV20 plan
  - ActEV Task Discussion Session (including new M1 data release)

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
- *MEVA data (*https://mevadata.org/)
  - *37 activities (72 video hours) : Indoor and outdoor scenes, night and day, crowds and individuals, EO (Electro-Optical) and IR (Infrared) sensors*
  - ***New M1 data release***

9:00 - 9:30 am	Activities in Extended Video (ActEv) Task Overview
9:30 - 10:00 am	BUPT-MCPRL at TRECVID 2019: ActEv ***BUPT_MCPRL Team - Beijing University of Posts and Telecommunications***
10:00 - 10:30 am	AI Surveillance System for Spatial-Temporal Activity Detection in Surveillance Scenarios ***MUDSML + INF Teams - Monash University; Carnegie Mellon University***
10:30 - 11:00 am	Our ActEv approach with object detection and custom tracking algorithm ***HSMW_TUC Team - University of Applied Sciences Mittweida***
11:00 - 11:30 am	**Break with refreshments**
11:30 - 12:00 am	Real-time activity detection in surveillance videos ***UCF Team - University of Central Florida***
12:00 - 12:30 pm	Event Detection with Specialized Object Detectors ***Hitachi Team***
12:30 - 2:00	Lunch
2:00 - 2:30 pm	Traffic Danger Recognition with Surveillance Cameras Without Training Data, *Invited Talk*: ***Lijun Yu, Carnegie Mellon University***
2:30 - 2:50 pm	ActEv Task Discussion

Questions?

https://actev.nist.gov/

Contact: actev-nist@nist.gov

TRECVID ActEV19 Feedback and ActEV20 Discussion

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**
- *MEVA data (*https://mevadata.org/)
  - *37 activities (72 video hours) : Indoor and outdoor scenes, night and day, crowds and individuals, EO (Electro-Optical) and IR (Infrared) sensors*
  - ***New M1 data release***

2019 ActEV feedback and 2020 plans

- What do the teams think about the ActEV task ?
- Any feedback on the new Scoring Metric compared to the 2018 Metric?
- Any feedback on the data repo to download data (VIRAT, MEVA, ..) ?
- Any feedback on the scoring server and different documents?
- Besides the ActEV leaderboard, we have added the ActEV reports (report on next slide), any feedback?
- Current Plan is to continue the ActEV task with the VIRAT dataset with more activities (28 or more activities)

Current TRECVID ActEV reports (at the end of the evaluation)

BUPT-MCPRL-SYS-00293-20191001-113056-4859

	Metadata
Team Name	BUPT-MCPRL
Team Type	TrecVID
System name	MCPRL_S1
Submission id	3540
Submitted Date	2019-10-01 11:30:56 -0400
Submission Target	TRECVID
Evaluation Name	ActEV-2018
Evaluation Dataset	1B-eval