The Visual Object Tracking Challenge Results VOT-ST2019, VOT-RT2019, VOT-LT2019

Matej Kristan, Aleš Leonardis, Jiri Matas, Michael Felsberg, Roman Pflugfelder, Joni-Kristian Kämäräinen, Luka Čehovin Zajc, Gustavo Fernandez, Alan Lukežič, Ondrej Drbohlav, Amanda Berg, Abdelrahman Eldesokey, et al.
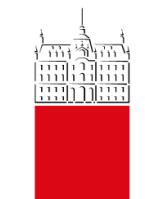
Outline

1. Scope of the VOT2019 ST/RT/LT challenges

2. Results overview (VOT2019 ST/RT/LT)

3. Winner announcement (VOT2019 ST/RT/LT)

The VOT 2019 workshop

VOT2019 ST/RT CHALLENGES: OVERVIEW

VOT2019 short-term challenge (VOT-ST2019)

- Short-term, single-target, causal trackers

- Tracker reports the target state as a rotated bounding box
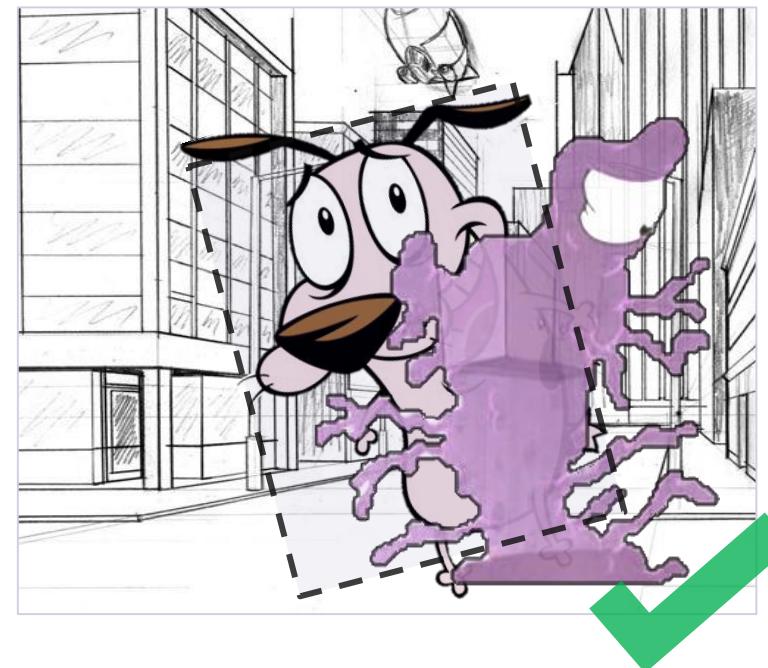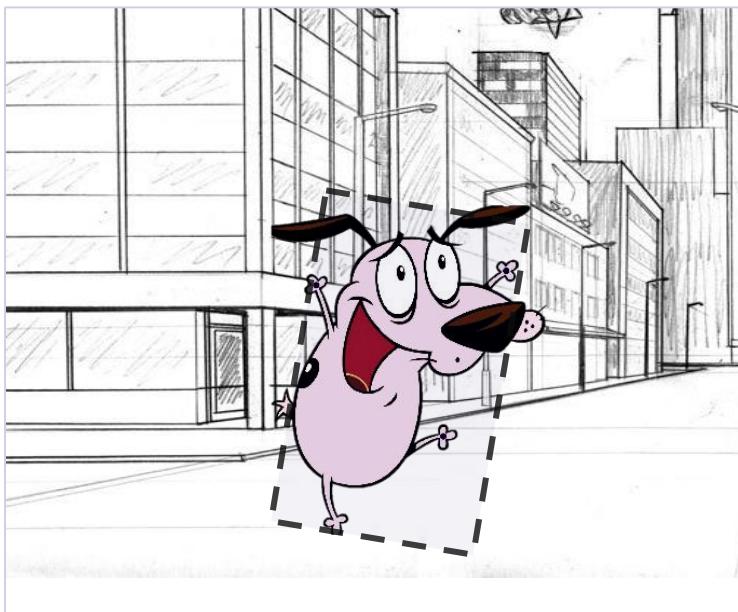
- No redetection: drift is considered a failure and tracker is reset

The VOT-ST2019 dataset

- Public dataset (60 sequences) +
  Sequestered dataset (60 sequences)

- The VOT sequence selection protocol used
  to refresh the VOT2018 dataset

- 20% of VOT2018 public dataset replaced,
  5% of VOT2018 sequestered dataset replaced

- Rotated bounding box automatically
  computed from pre-segmented image

- Each image annotated by 6 attributes:
  Occlusion, Illumination change , Object motion, Object size
 change, Camera motion, Unassigned
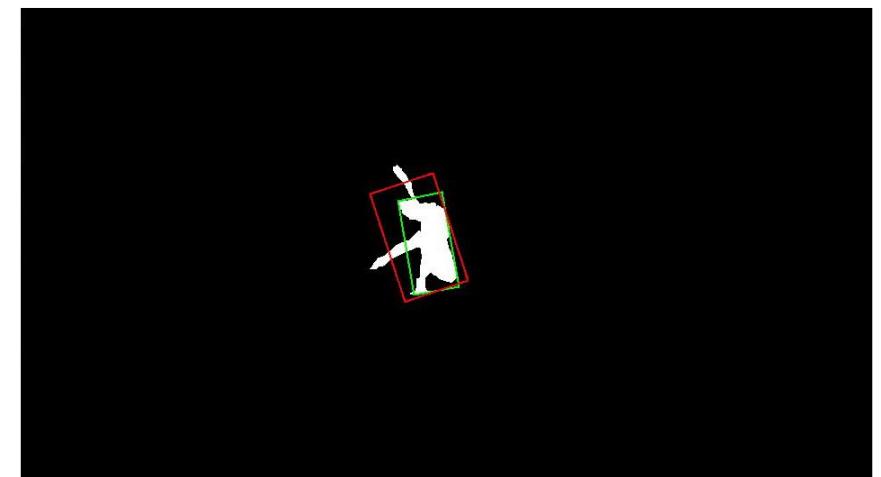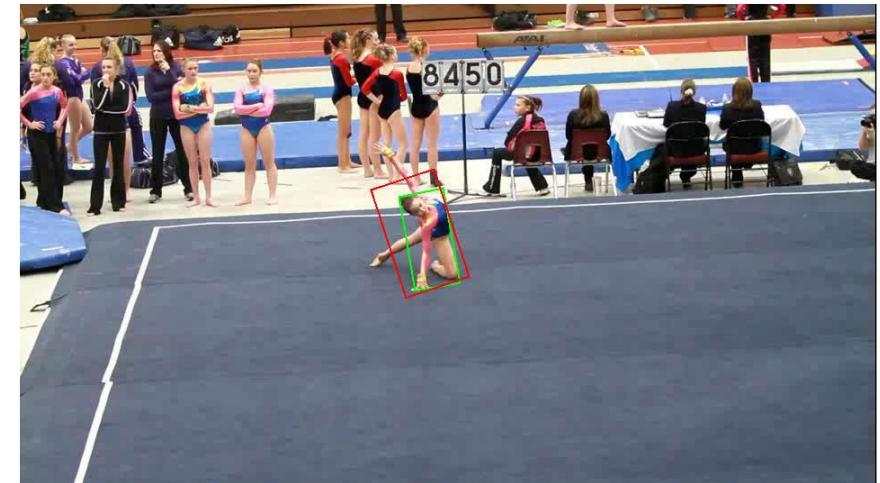
The VOT-ST2019 evaluation methodology

Performance measure correlation analysis[1]

- Two weakly correlated measures[2] chosen according to[1]:

  - Robustness (number of times a is reinitialized)

  - Accuracy (average overlap while tracking)

  - + Combination of basic measures (EAO)

- Winner: Top EAO on the sequestered dataset
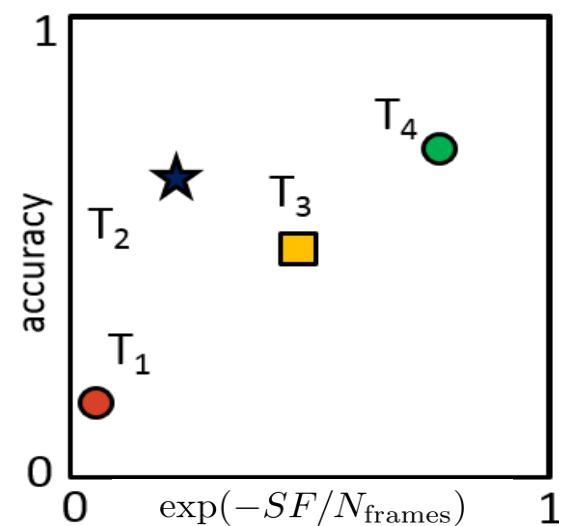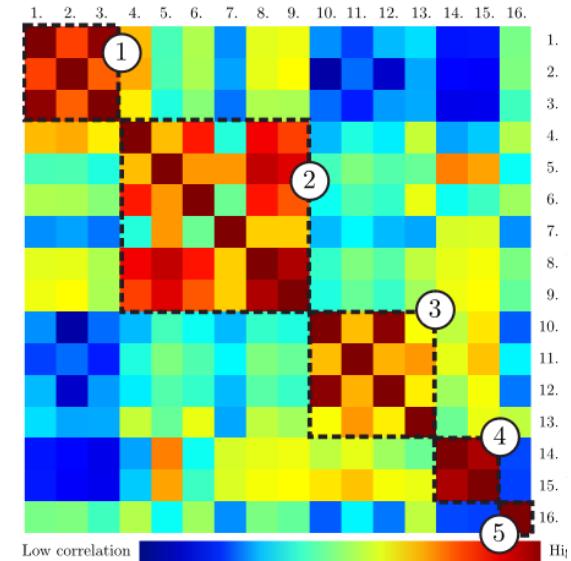
[1]Čehovin, Leonardis, Kristan. *Visual object tracking performance measures revisited*, IEEETIP 2016
[2]Kristan et al., A Novel Performance Evaluation Methodology for Single-Target Trackers, IEEETPAMI 2016

The VOT2019 ST real-time challenge (VOT-RT2019)

- Introduced in VOT2017

- Required to process sequences at ~20 fps
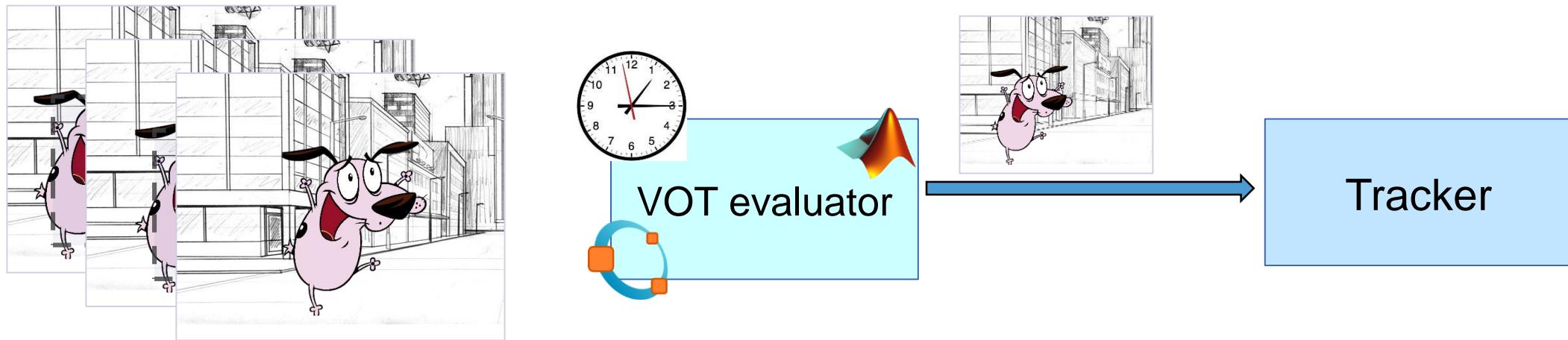
- Same performance evaluation protocol and measures as VOT-ST2019

- The VOT-ST2019 public dataset used

- Winner: Top EAO on the public dataset

The VOT 2019 workshop

VOT2019 LT CHALLENGE: OVERVIEW

VOT2019 long-term challenge (VOT-LT2019)
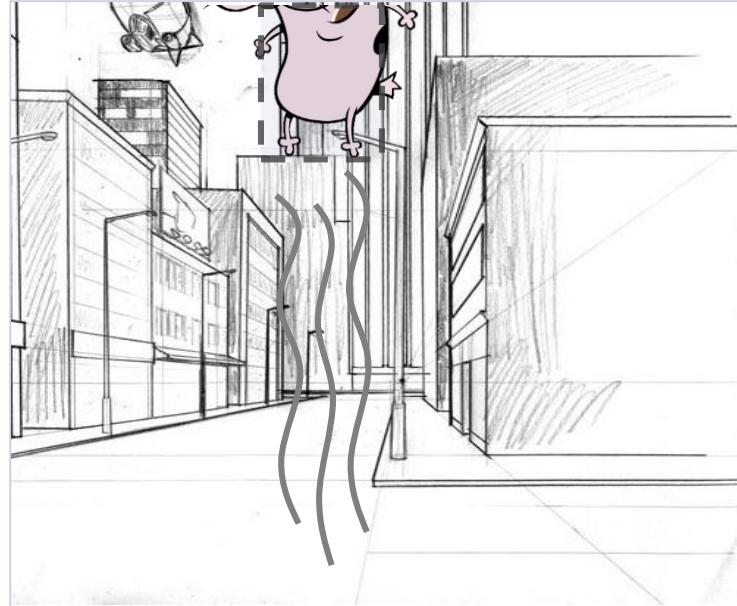
- Required long-term tracker properties:

  - Determine whether the target has been lost (or disappeared)

  - Re-detect the target when it reappears

- Tracker output at each frame: bounding box + certainty score

Short-term vs long-term spectrum[1]

ST/LT levels	Position reported	Determines target lost?	Target re-detection
ST_0: Basic ST	each frame	no	no
ST_1: Basic ST with conservative updating	each frame	not explicitly, selective update of visual model	no
LT_0: Pseudo LT	only when visible	yes	no
LT_1: Re-detecting LT	only when visible	yes	yes

- ST_0 (e.g., KCF[2], MS[3])

- ST_1 (e.g., MDNet[4], ECO[5]) -> easily converted to LT_0

- LT_1 (e.g., TLD[5])

[1]Lukežič, Čehovin, Vojir, Matas, Kristan, *Now you see me: evaluating performance in long-term visual tracking*, arXiv2018
[2]Enriques et al. PAMI 2015 ; [2]Comaniciu et al. PAMI 2002; [3]Nam et al. *CVPR2016*;
[4]Danelljan et al. *CVPR2017*; [5]Kalal et al. PAMI 2011

The VOT-LT2019 dataset (50 sequences)

- VOT-LT2018 extended by 15 sequences
  (average sequence length >4k frames)

- Average per sequence disappearance: 10

- Average target absence period: 50 frames

- Axis-aligned bounding boxes

- Nine per-sequence attributes:
  (1) full occlusion, (2) out-of-view motion,
  (3) partial occlusion, (4) camera motion,
  (5) fast motion, (6) scale change,
  (7) aspect ratio change, (8) viewpoint
  change, (9) similar objects

The VOT-LT2019 evaluation methodology

- Tracking properties measured:
  Localization, Loss/Presence detection

- Initialized at first frame, no reset at target loss

- Three LT measures from VOT-LT2018[1]:

  - Tracking Precision, Recall & F-score:
  $$\mathrm{Pr}(\tau_\theta), Re(\tau_\theta), F(\tau_\theta)$$
  (depend on target presence certainty threshold τ_θ)

  - Evaluated at presence certainty threshold τ_θ^* that maximizes the tracker F-score

- Winner: Top performer in $F(\tau_\theta^*)$
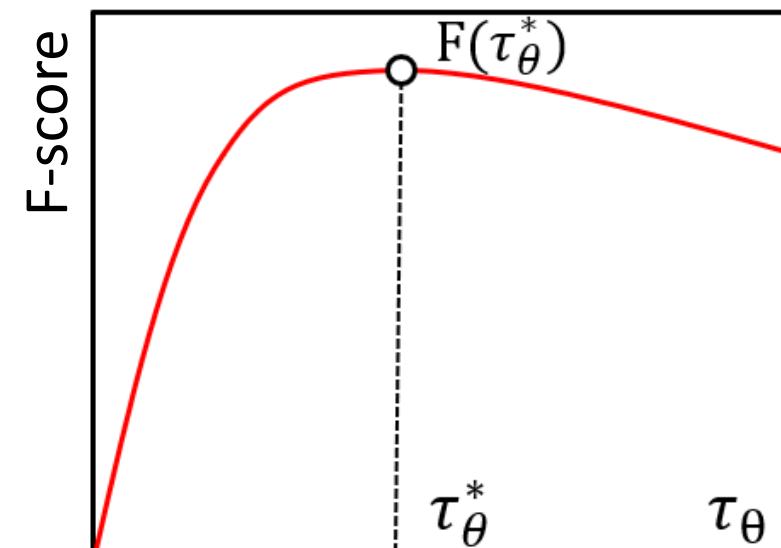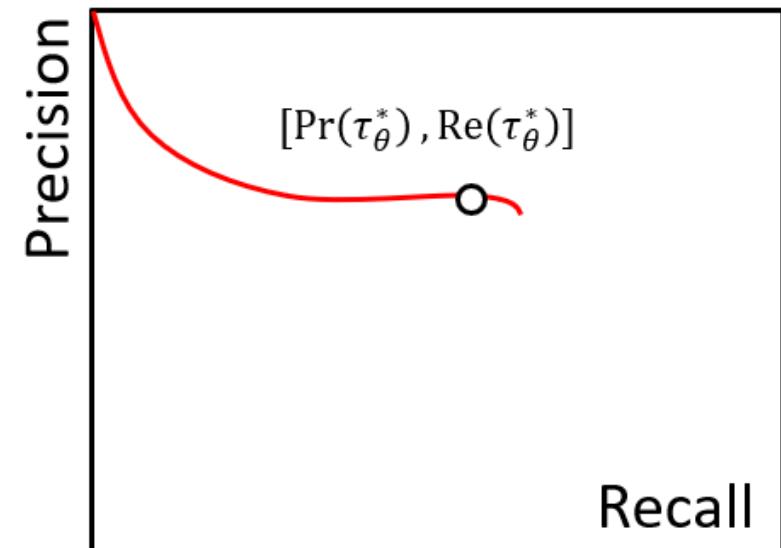
[1]Lukežič, et al., Now you see me: evaluating performance in long-term visual tracking, Arxiv2018

Matej Kristan (matej.kristan@fri.uni-lj.si)

The VOT 2019 workshop

VOT-ST2019 & VOT-RT2019 CHALLENGE RESULTS

VOT-ST2019, VOT-RT2019: 57 trackers tested

Tracking approach:
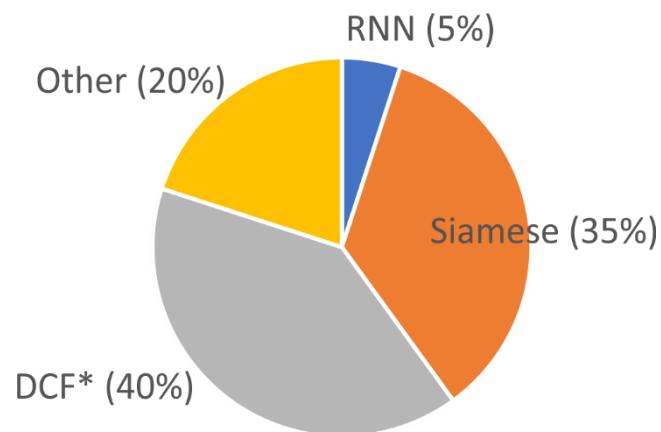
ST/LT category:
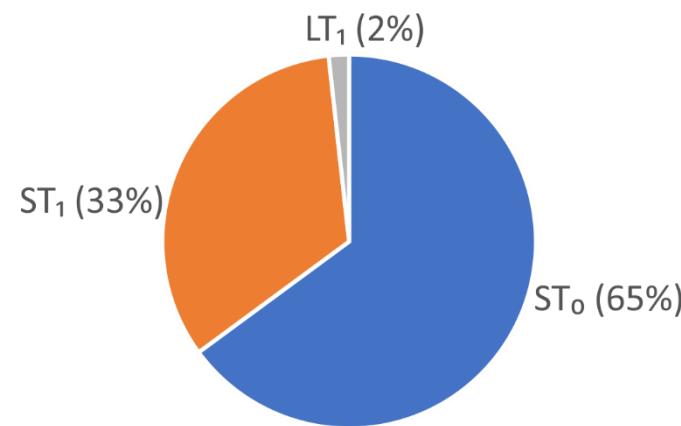
Target model:
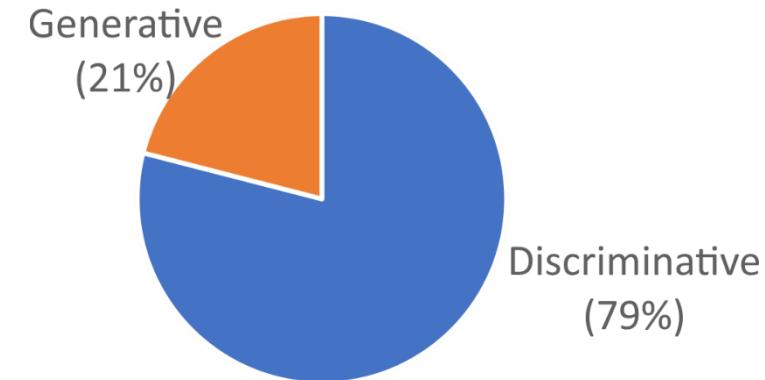
VOT-ST2019 results on public dataset

- **Top trackers:** (1) DRNet, (2) Trackyou, (3) ATP, (4) DiMP, (5) Cola, (6) ACNT, (7) SiamMargin, (8) DCFST, (9) SiamFCOT, (10) SiamCRF

- **All top trackers are deep trackers:**

  7 deep DCF (ATOM[1]),

  4 Siamese (SiamMask[2], SiamRPN[3])

  [1]Danelljan et al. CVPR2019, [2]Wang et al. CVPR2019,
  [3]Li et al. CVPR2018

- **Localization:**

  - Mostly correlation by a template (Discriminative/Generative)

  - Position refinement by a regression network or segmentation

- Top trackers are among the most robust trackers

  (1) DRNet, (2) Trackyou, (3) DiMP, (3) ACNT

- Top in accuracy:

  (1) ATP, (2) MPAT, (3) ACNT

- Per-attribute analysis:

	CM	IC		MC		OC		SC	
Accuracy	0.53	0.48	③	0.51		0.44	①	0.48	②
Robustness	0.63	1.18	③	1.44	①	1.20	②	0.56	

- Most failures due to: Motion change

- Mostly affects accuracy: Occlusion
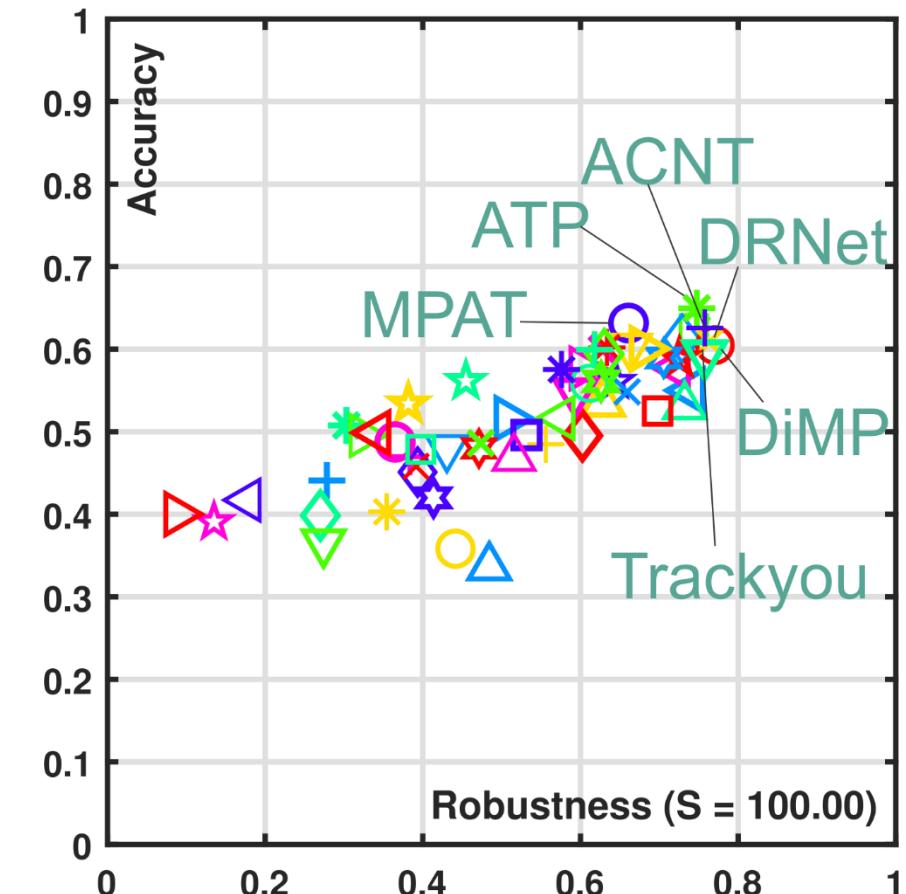
VOT-ST2019 results on public dataset

- Baselines ranked at the very tail of the benchmark

- 11 trackers published at major CV venues (≥2018)

  - Their average performance: VOT2019 sota bound

  - Over 38% submissions exceed this bound
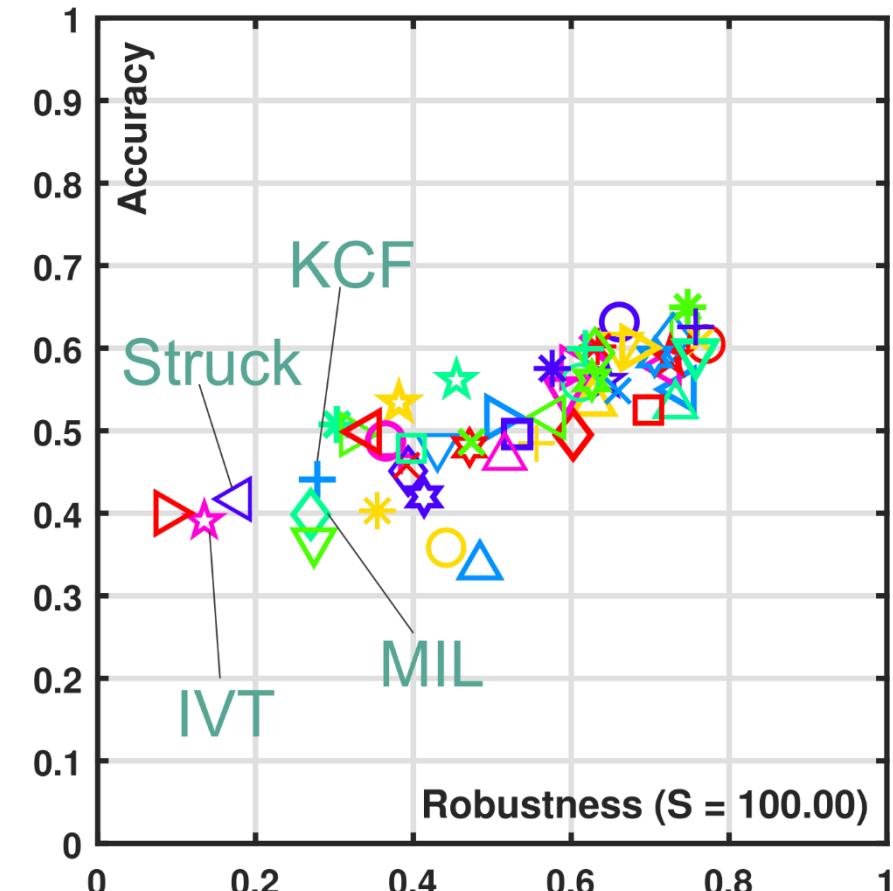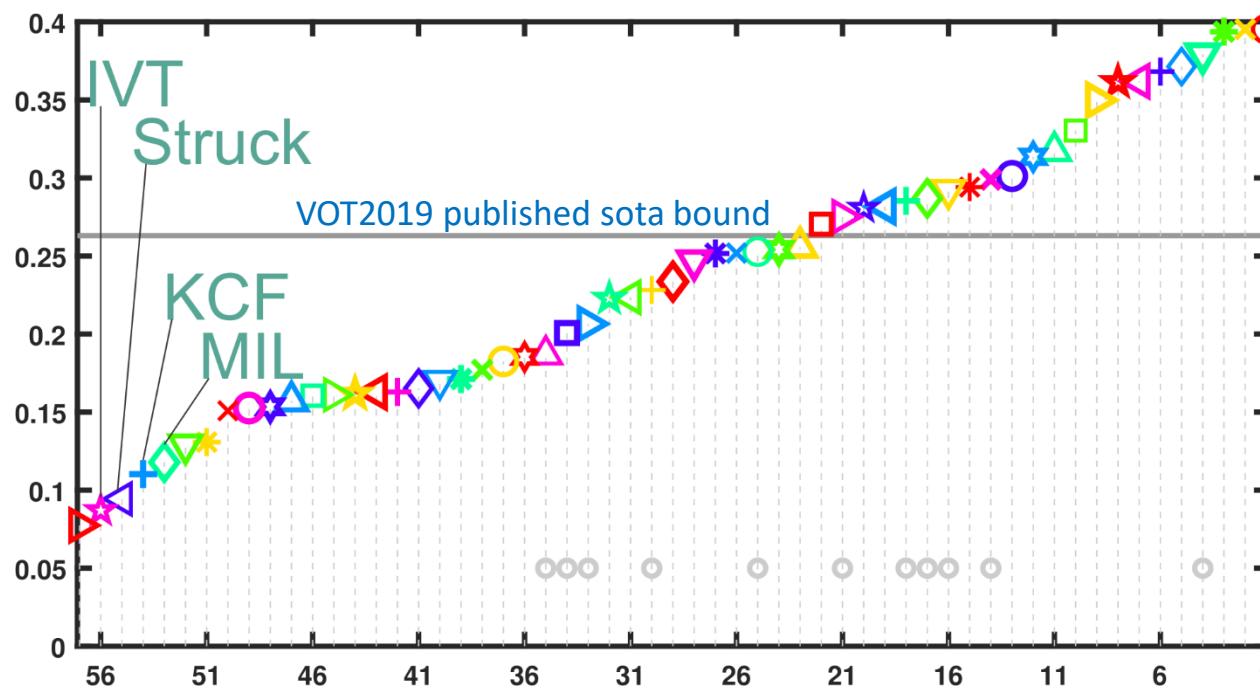
VOT-ST2019 results on sequestered dataset

- Large EAO value drop (39%)

- 1.6 times increase in failures, accuracy comparable

- Smallest change: ATP
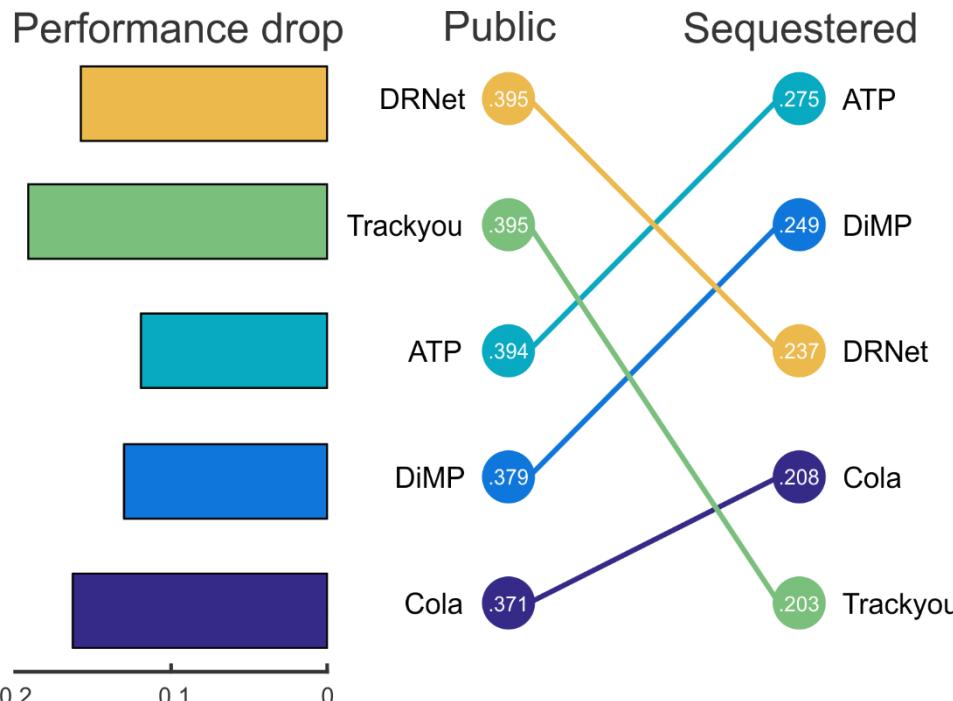
Public vs Sequestered dataset EAO
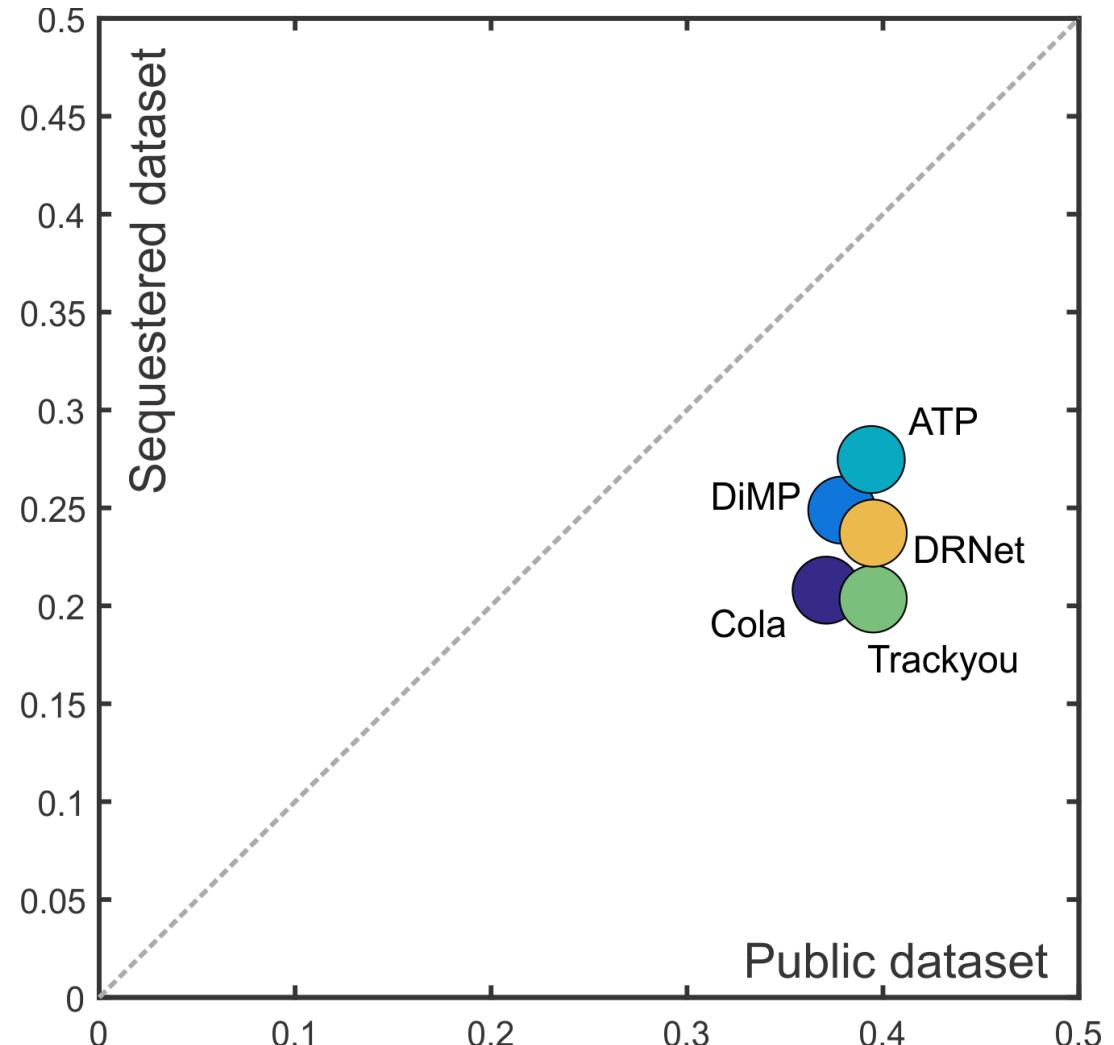
VOT realtime challenge (VOT-RT2019) results

- **Top 10:** (1) SiamMargin, (2) SiamFCOT, (3) DiMP, (4) DCFST, (5) SiamDW-ST, (6) SRTCS, (7) SiamMask, (8) SiamRPNpp, (9) SPM, (10) SiamCRF-RT
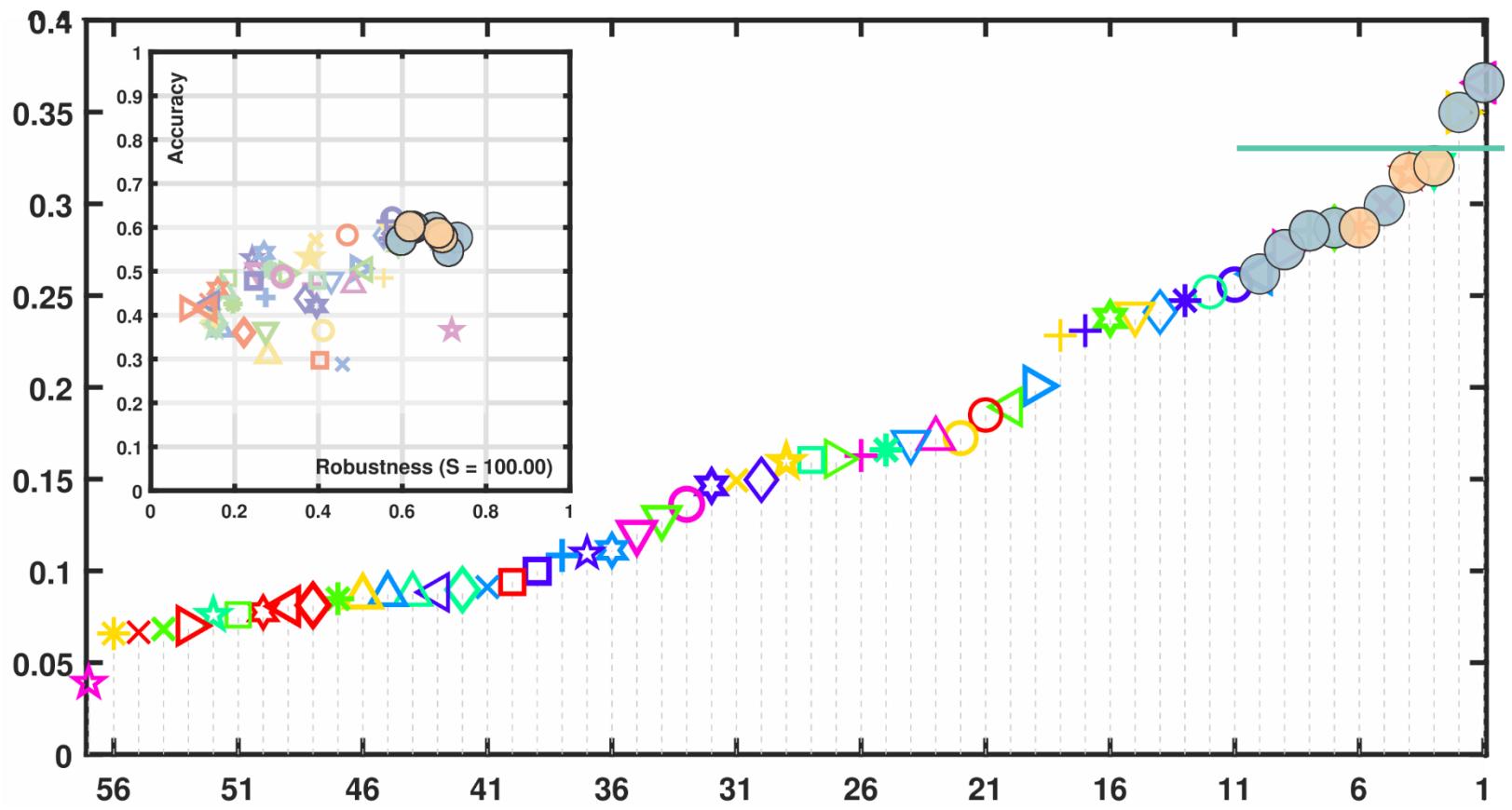
Two classes:

Approach:

Siamese correlation
Bounding box regression
(e.g., SiamRPN[1] , SiamMask[2])
GPU-based

Approach:

deep DCF correlation
Bounding box regression
(e.g., ATOM[3])
GPU-based

[1]Li et al. CVPR2018, [2]Wang et al. CVPR2019, [3]Danelljan et al. CVPR2019

VOT2019 Realtime vs Baseline results
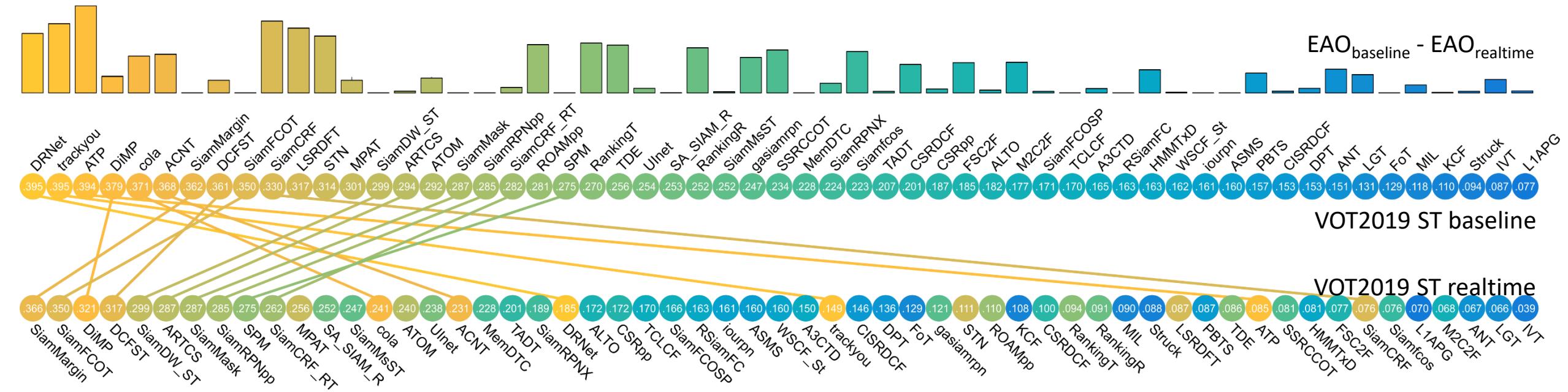
- A lot of the top baseline performers drop with real-time constraint

- The drop is smaller for real-time trackers on the baseline ST challenge

- Some achieve top real-time performance AND perform well on the baseline

The VOT 2019 workshop

VOT-LT2019 CHALLENGE RESULTS

VOT-LT2019 challenge overview

- 9 trackers tested

- All trackers were from LT_1 class:

  Explicit target absence detection and re-detection implemented

Features:

Handcrafted (11%)
CNN (89%)

Model update:

Mixed (11%)
Never (22%)
When conf. (44%)
Exponential (22%)

Architecture:

Classical DCF (11%)
CNN classifier (22%)
Siamese (67%)
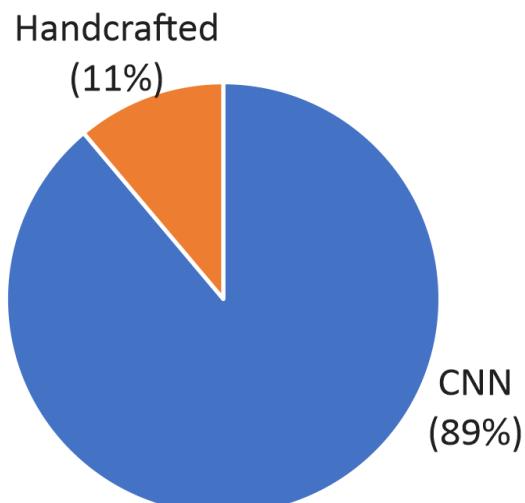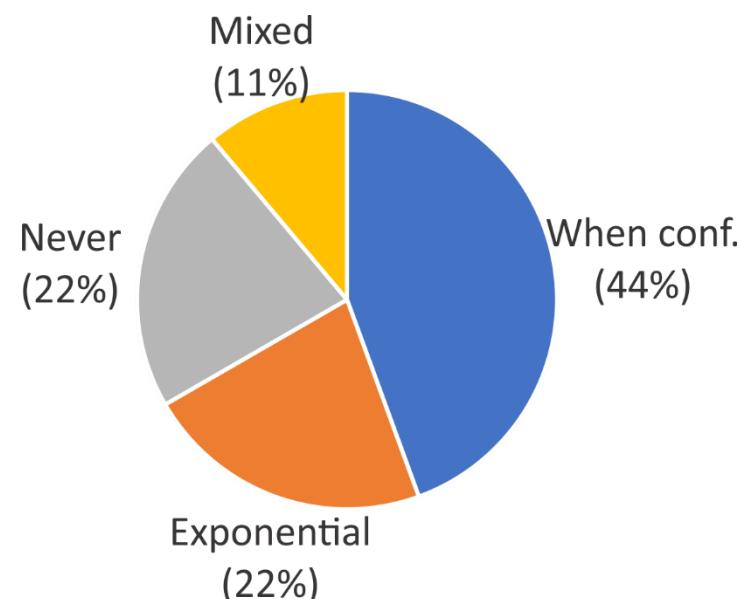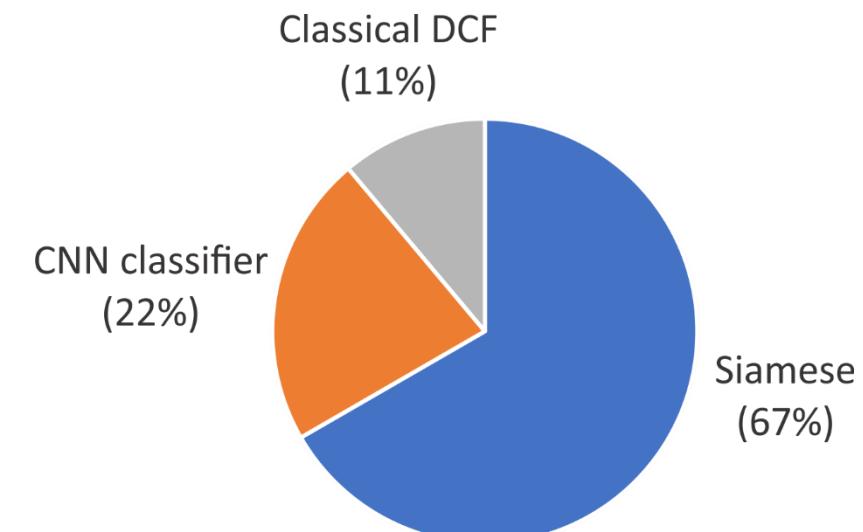
VOT-LT2019 challenge results

- **Properties of top 3 trackers:**

  Apply a Short-term tracker *+ Detector*

  ST: deep DCF[1] or Siamese template[2]

  Absence det.: MDNet[3], localization score

  Re-Det: Region proposal nets (e.g., RPN[4])

- **Top-performer:** LT-DSE

  ST: ATOM[1] + SiamMask[5]

  Absence det.: MDNet[3] (winner of VOT-ST2015)

  Re-Det: Region proposal net from MBMD[6]

  (winner of VOT-LT2018)

LT_DSE (0.695)
CLGS (0.674)
SiamDW_LT (0.665)
mbdet (0.567)
SiamRPNsLT (0.556)
Siamfcos (0.520)
CooSiam (0.508)
ASINT (0.505)
FuCoLoT (0.411)
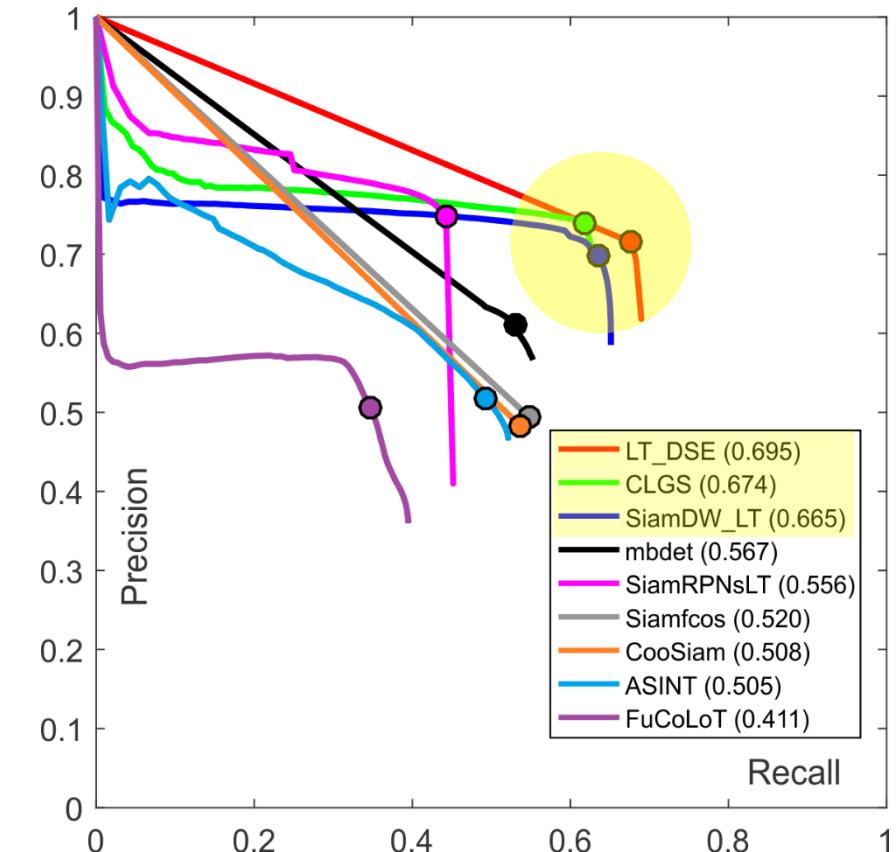
[1]Danelljan et al. CVPR2019,
[2]Bertinetto et al. VOT2016, [3]Nam et al CVPR2016,
[4]Li et al. CVPR2018, [5]Wang et al. CVPR2019,
[6]Zhang et al VOT-LT2018 winner

VOT-LT2019 attribute analysis

Most challenging:

- Out of view (target absent)

- View point change (appearance)

- Similar objects (identity switch)
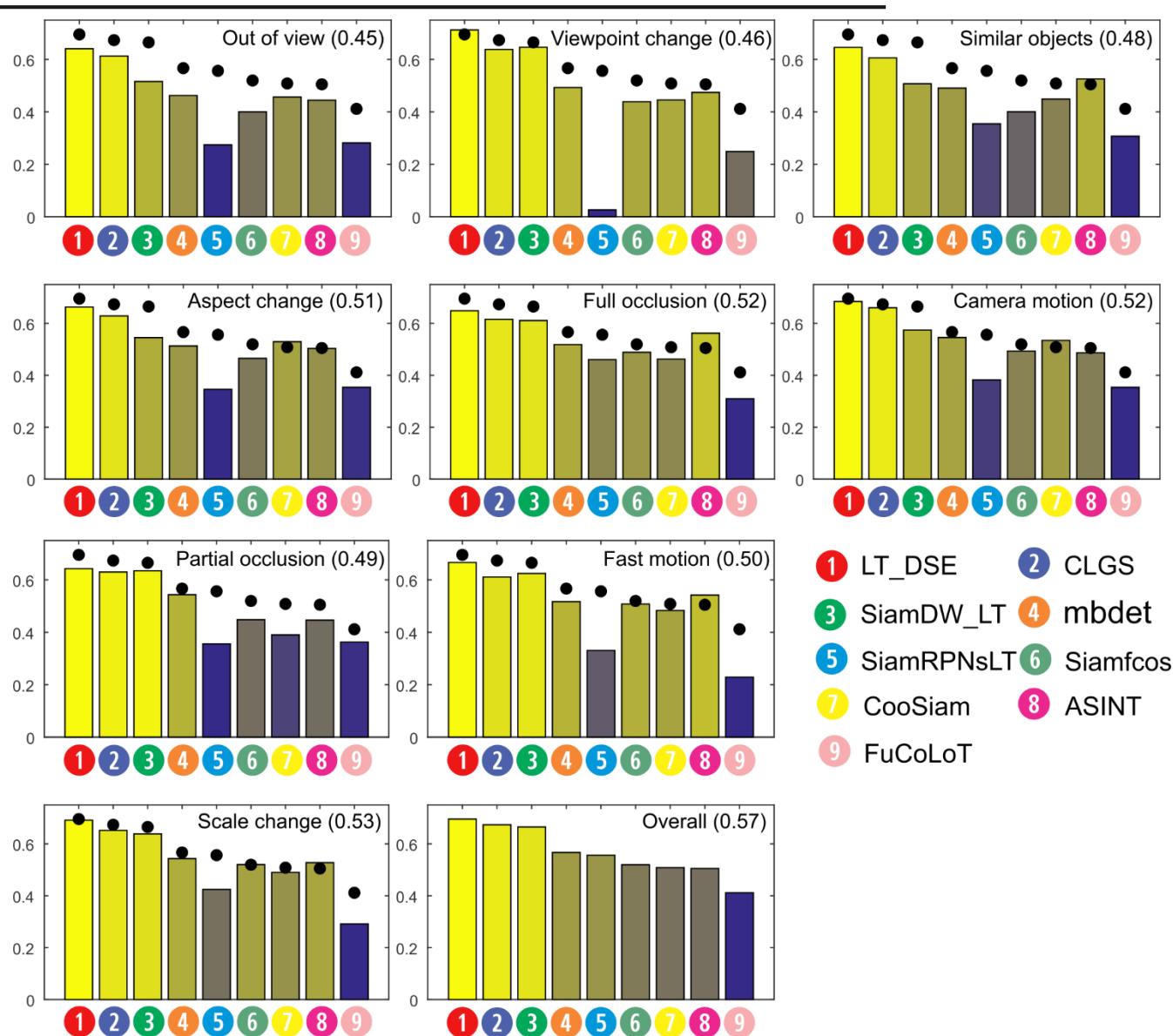
VOT2019 ST/RT/LT challenges summary

- VOT-ST2019:

  - Deep DCF and Siamese correlation the dominant methodology

  - Adoption of bounding box regression networks improves accuracy

- VOT-RT2019:

  - Siamese correlation and Deep DCF the dominant methodology (switched places)

  - Some of the fastest trackers are among top-10 on VOT-ST2019

- VOT-LT2019:

  - Explicit object detection integrated

  - Top performers: deep ST component, deep detector component

The VOT 2019 workshop

VOT2019 ST/RT/LT WINNER ANNOUNCEMENTS

VOT-ST2019 Winners

Winners of the VOT2019 short-term challenge:

(The talk up next!)

VOT-RT2019 Winners

Winners of the VOT2019 ST real-time challenge:

5.

(The talk up next!)
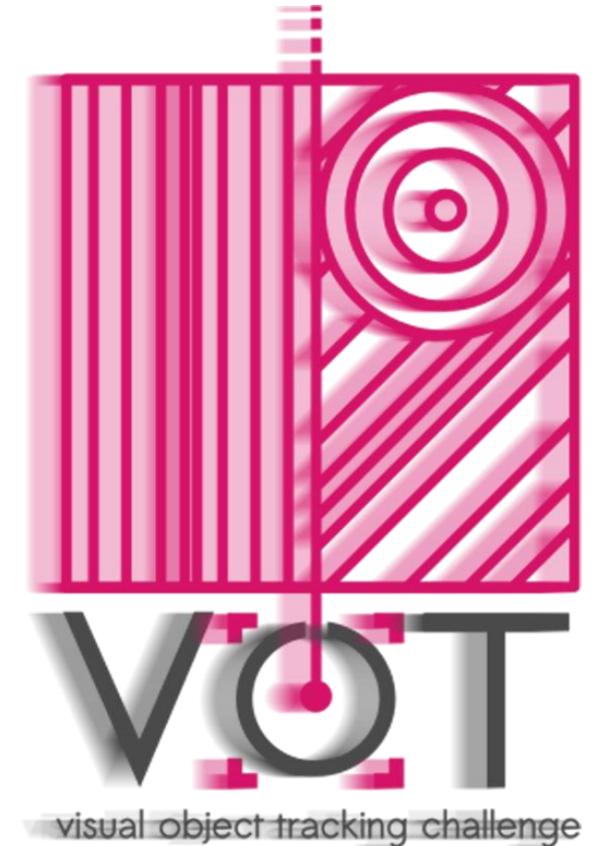
VOT-LT2019 Winners

Winners of the VOT2019 long-term challenge:

(The talk up next!)
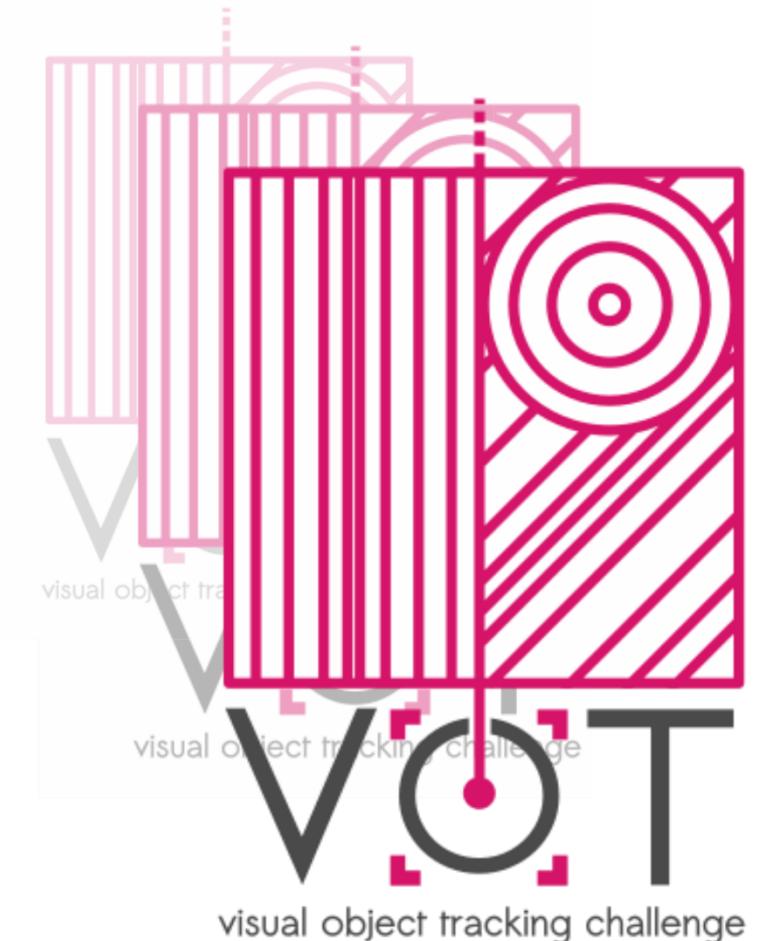
Thanks

- ## The VOT2019 committee

M. Kristan   J. Matas   A. Leonardis   M. Felsberg   R. Pflugfelder   J. K. Kamarainen   G. Fernandez   L. Čehovin   A. Lukežič   A. Eldesokey
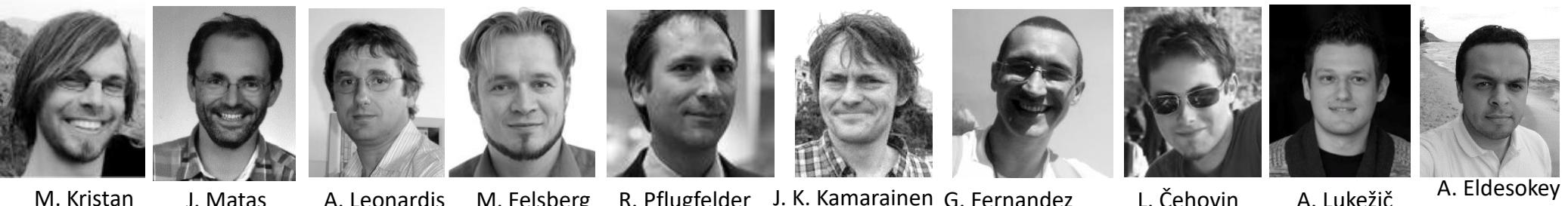
- ## Everyone who participated or contributed

Matej Kristan1, Jiˇr´ı Matas2, Aleˇs Leonardis3, Michael Felsberg4, Roman Pflugfelder5,6, Joni-Kristian Kamarainen7, Luka Cˇ ehovin Zajc1, Ondrej Drbohlav2, Alan Lukezˇicˇ1, Amanda Berg4,8, Abdelrahman Eldesokey4, Jani K¨apyl¨a7, Gustavo Fern´andez5, Abel Gonzalez-Garcia18, Alireza Memarmoghadam50, Andong Lu9, Anfeng He52, Anton Varfolomieiev37, Antoni Chan17, Ardhendu Shekhar Tripathi23, Arnold Smeulders45, Bala Suraj Pedasingu29, Bao Xin Chen58, Baopeng Zhang12, Baoyuan Wu43, Bi Li28, Bin He10, Bin Yan19, Bing Bai20, Bing Li16, Bo Li40, Byeong Hak Kim25,33, Chao Ma41, Chen Fang35, Chen Qian40, Cheng Chen38, Chenglong Li9, Chengquan Zhang10, Chi-Yi Tsai42, Chong Luo34, Christian Micheloni55, Chunhui Zhang16, Dacheng Tao54, Deepak Gupta45, Dejia Song28, Dong Wang19, Efstratios Gavves45, Eunu Yi25, Fahad Shahbaz Khan4,30, Fangyi Zhang16, Fei Wang40, Fei Zhao16, George De Ath49, Goutam Bhat23, Guangqi Chen40, Guangting Wang52, Guoxuan Li40, Hakan Cevikalp21, Hao Du34, Haojie Zhao19, Hasan Saribas22, Ho Min Jung33, Hongliang Bai11, Hongyuan Yu16,34, Houwen Peng34, Huchuan Lu19, Hui Li32, Jiakun Li12, Jianhua Li19, Jianlong Fu34, Jie Chen57, Jie Gao57, Jie Zhao19, Jin Tang9, Jing Li26, Jingjing Wu27, Jingtuo Liu10, Jinqiao Wang16, Jinqing Qi19, Jinyue Zhang57, John K. Tsotsos58, Jong Hyuk Lee33, Joost van de Weijer18, Josef Kittler53, Jun Ha Lee33, Junfei Zhuang13, Kangkai Zhang16, Kangkang Wang10, Kenan Dai19, Lei Chen40, Lei Liu9, Leida Guo59, Li Zhang51, Liang Wang16, Liangliang Wang28, Lichao Zhang18, Lijun Wang19, Lijun Zhou48, Linyu Zheng16, Litu Rout39, Luc Van Gool23, Luca Bertinetto24, Martin Danelljan23, Matteo Dunnhofer55, Meng Ni19, Min Young Kim33, Ming Tang16, Ming-Hsuan Yang46, Naveen Paluru29, Niki Martinel55, Pengfei Xu20, Pengfei Zhang54, Pengkun Zheng38, Pengyu Zhang19, Philip H.S. Torr51, Qi Zhang , Qiang Wang16,31, Qing Guo44, Radu Timofte23, Rama Krishna Gorthi29, Richard Everson49, Ruize Han44, Ruohan Zhang57, Shan You40, Shao-Chuan Zhao32, Shengwei Zhao16, Shihu Li10, Shikun Li16, Shiming Ge16, Shuai Bai13, Shuosen Guan59, Tengfei Xing20, Tianyang Xu32, Tianyu Yang17, Ting Zhang14, Tom´aˇs Voj´ıˇr47, Wei Feng44, Weiming Hu16, Weizhao Wang38, Wenjie Tang14, Wenjun Zeng34, Wenyu Liu28, Xi Chen60, Xi Qiu56, Xiang Bai28, Xiao-Jun Wu32, Xiao-Jun Wu32, Xiaoyun Yang15, Xier Chen57, Xin Li26, Xing Sun59, Xingyu Chen16, Xinmei Tian52, Xu Tang10, Xue-Feng Zhu32, Yan Huang16, Yanan Chen57, Yanchao Lian57, Yang Gu20, Yang Liu36, Yanjie Chen40, Yi Zhang59, Yinda Xu60, Yingming Wang19, Yingping Li57, Yu Zhou28, Yuan Dong13, Yufei Xu52, Yunhua Zhang19, Yunkun Li32, Zeyu Wang , Zhao Luo16, Zhaoliang Zhang14, Zhen-Hua Feng53, Zhenyu He26, Zhichao Song20, Zhihao Chen44, Zhipeng Zhang16, Zhirong Wu34, Zhiwei Xiong52, Zhongjian Huang57, Zhu Teng12, and Zihan Ni10

- ## VOT2019 sponsor:

University *of Ljubljana*
Faculty *of Computer and*
Information Science