

The 9th Visual Object Tracking Challenge Results

VOT-ST2021, VOT-RT2021

Matej Kristan, Jiri Matas, Aleš Leonardis, Joni-Kristian Kämäräinen, Hyung Jin Chang, Roman Pflugfelder, Gustavo Fernandez, Luka Čehovin Zajc, Alan Lukežič, Michael Felsberg, Martin Danelljan, Ondrej Drbohlav, Zhongqun Zhang, Jinyu Yang, Song Yan, Gustav Hager, Jani Kapyla, Christoph Mayer

Outline

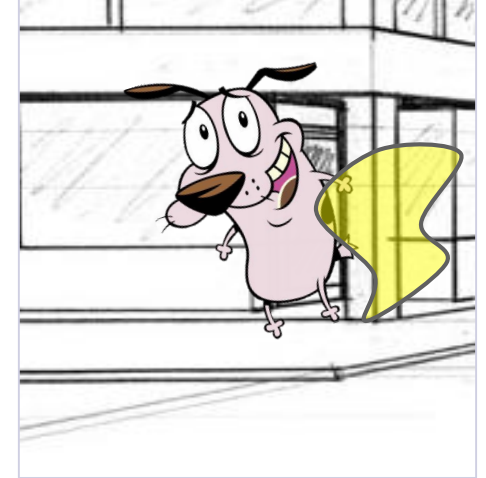
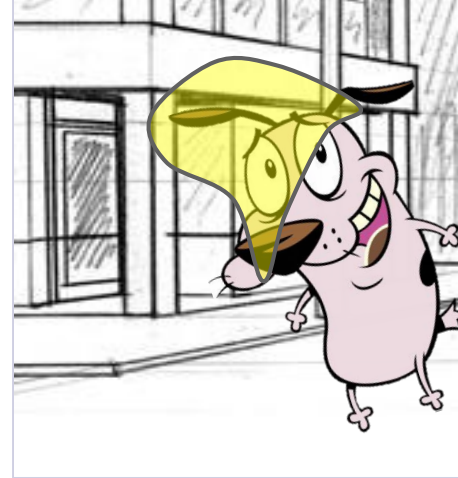
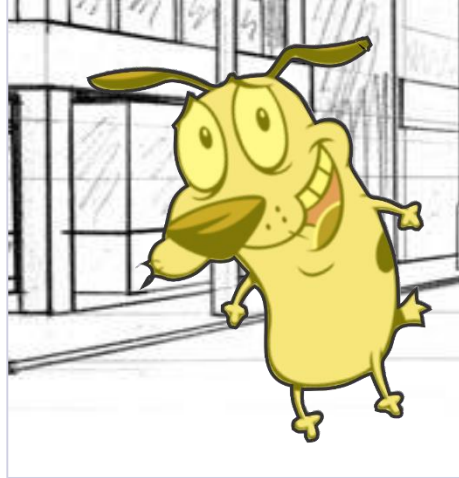
1. Scope of the VOT2021 ST/RT challenges
2. Results overview (VOT2021 ST/RT)
3. Winner announcement (VOT2021 ST/RT)

The VOT 2021 workshop

VOT2021 ST/RT CHALLENGES: OVERVIEW

VOT-ST2021 challenge scope

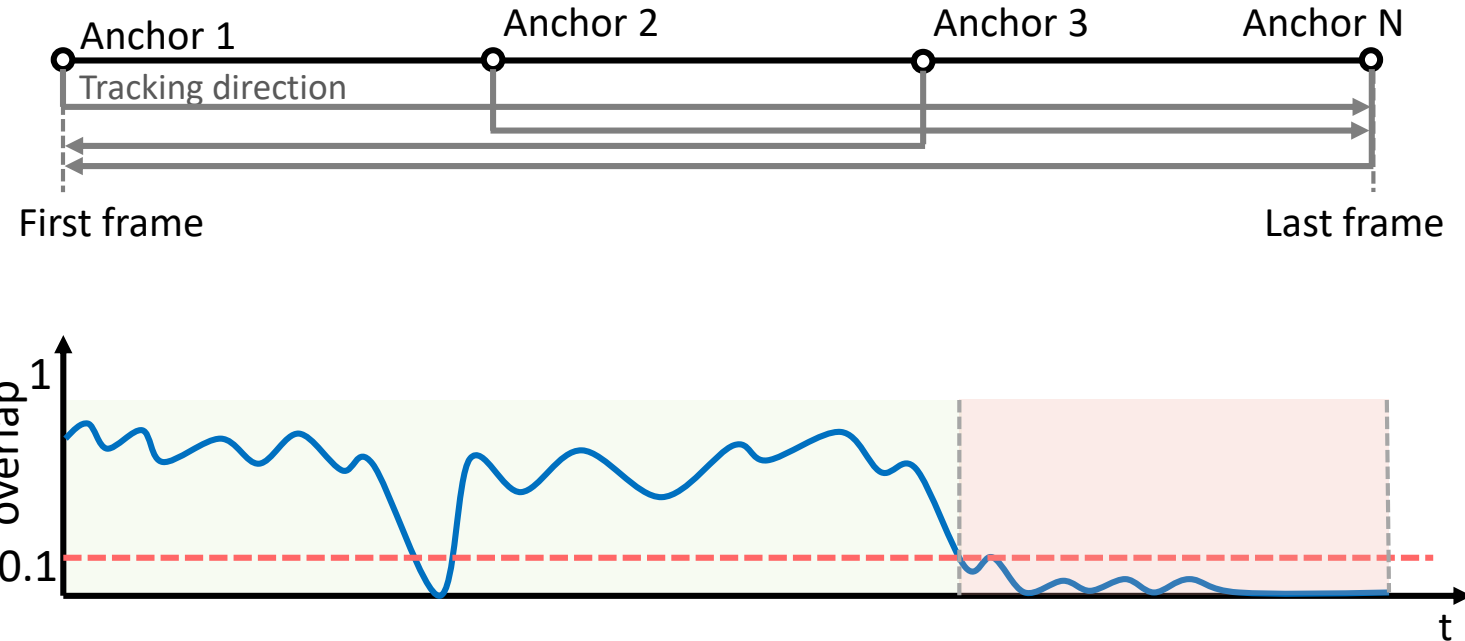
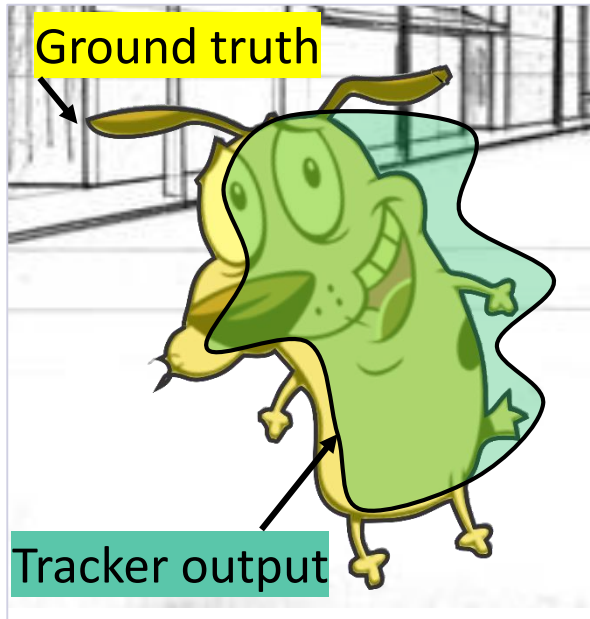
- Short-term, single-target, causal trackers, segmentation output
- Initialize on the first frame and report position on the rest



- Mission: promote trackers that track for as long as possible
- No redetection: drifting off the target is considered a failure

VOT anchor-based evaluation protocol

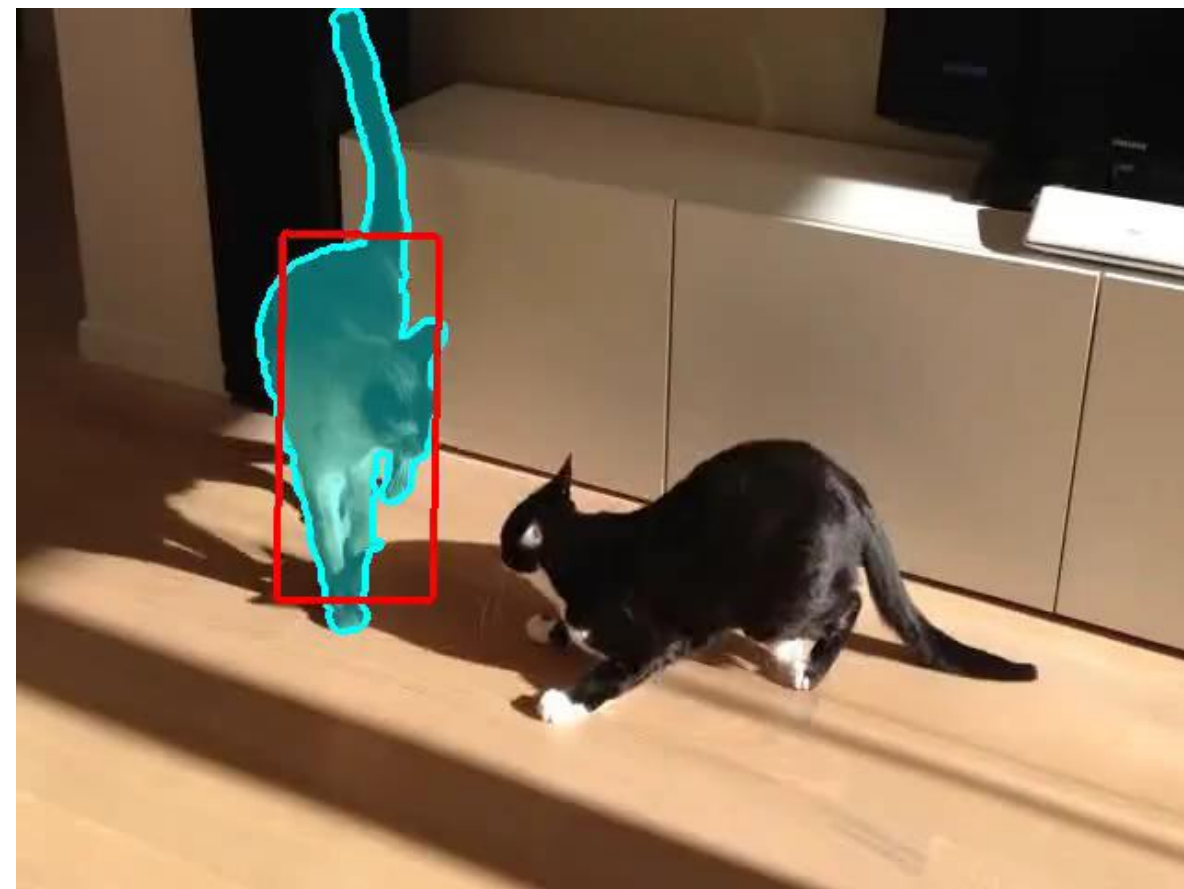
- Multiple initialization points



- Performance measures
 - Accuracy** (overlap between the ground truth mask and tracker output)
 - Robustness** (extent of the successfully tracked parts of sub-sequences)
 - EAO** (combination of both basic measures)

The VOT-ST2021 dataset

- Public dataset (60 sequences) + Sequestered dataset (60 sequences)
Winner identified on *sequestered dataset*
- Both datasets refreshed (10%)
 - A challenging sequence added to each
- All frames manually segmented!
- Each frame annotated by 6 attributes:
Occlusion, Illumination change , Object motion, Object size change, Camera motion, Unassigned

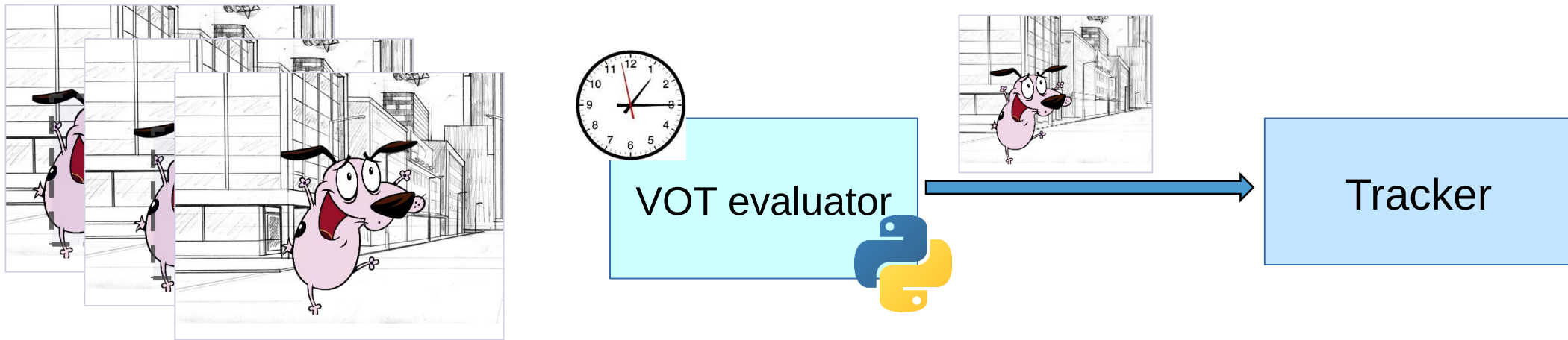


Red – VOT2019 annotation by a bounding box

Blue – VOT2020 annotation by a segmentation mask

The VOT2021 ST real-time challenge (VOT-RT2021)

- Introduced in VOT2017
- Same performance evaluation protocol and measures as VOT-ST2021
- Required to process sequences at ~ 20 fps



- Winner identification protocol:
Top EAO on the VOT-ST2021 public dataset

The VOT 2021 workshop

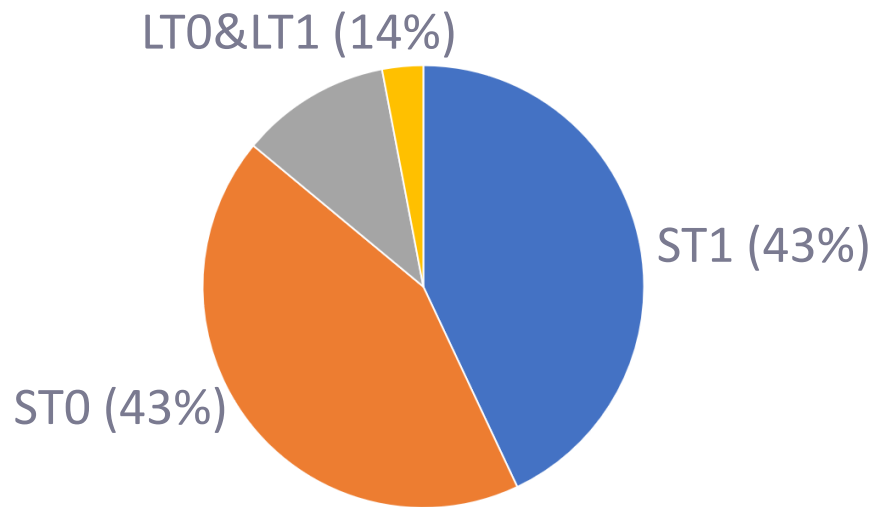
VOT-ST2021 & VOT-RT2021 CHALLENGE RESULTS

VOT-ST2021, VOT-RT2021: 53 trackers tested

Target state reporting:

64% of trackers reported a [segmentation mask](#)

ST/LT category:



Tracking approach:

Deep features: 90%

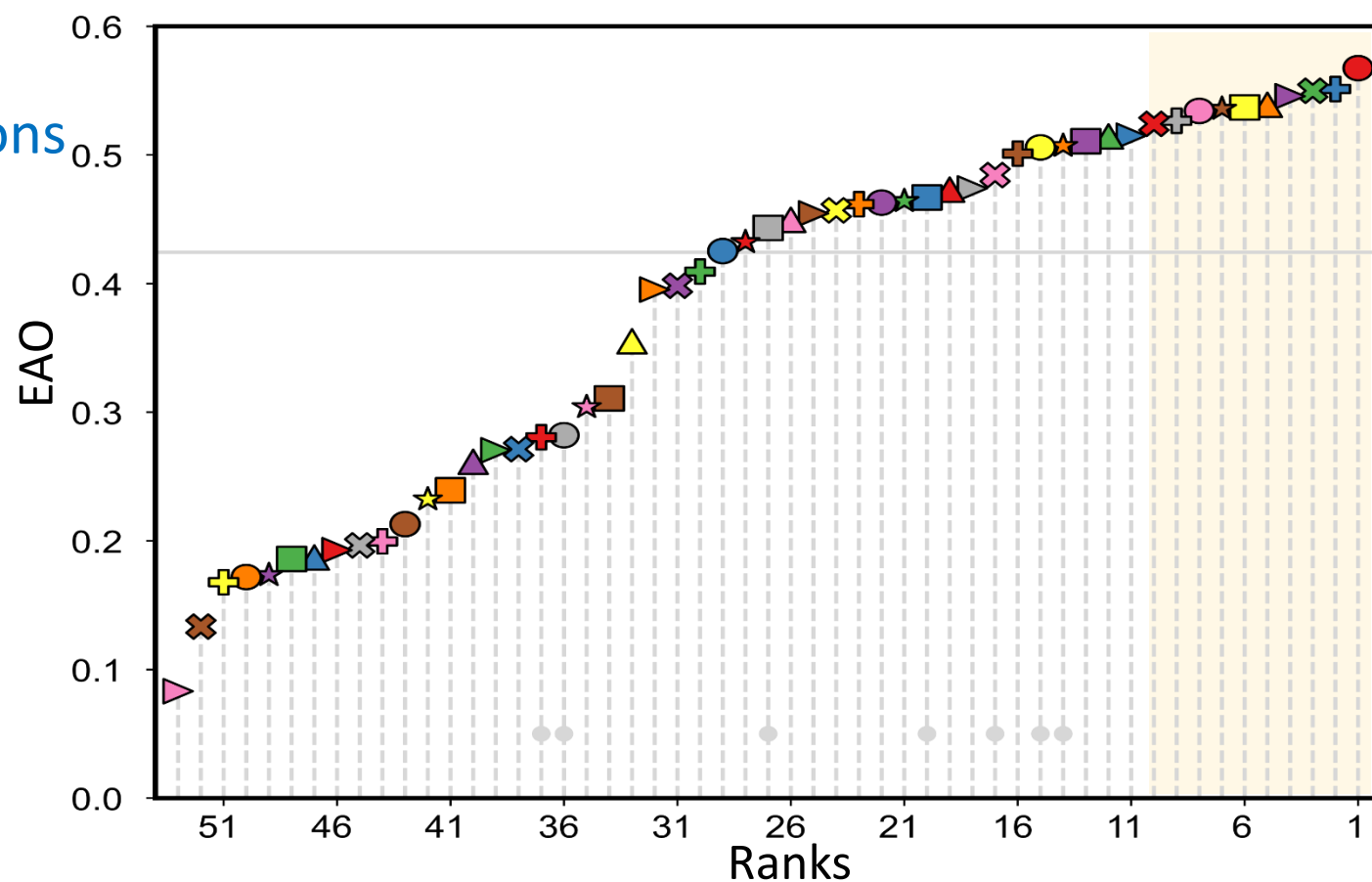
DCF: 56%

Siamese: 25%

Transformers: 17% [NEW!]

VOT-ST2021 results on public dataset

- Top trackers: (1) RPTMask, (2) CFRPT, (3) TransT_M, (4) TregPlus, (5) DualTFRon, (6) F_TregPlus, (7) DualTFRst, (8) STARK_RT, (9) DualTFR, (10) RPT
- All top trackers are deep NNs:
 - VOT-ST2020 winner + 2 extensions
 - 7 Transformers-based
- Segmentation:
 - 5 localize by bounding box & segment by Alpharef¹
 - 3 apply transformer segmentation



¹Alpharef [YanCVPR2021] winner of VOT-RT2020

VOT-ST2021 results on public dataset

- Top trackers in EAO that stand out from the rest:

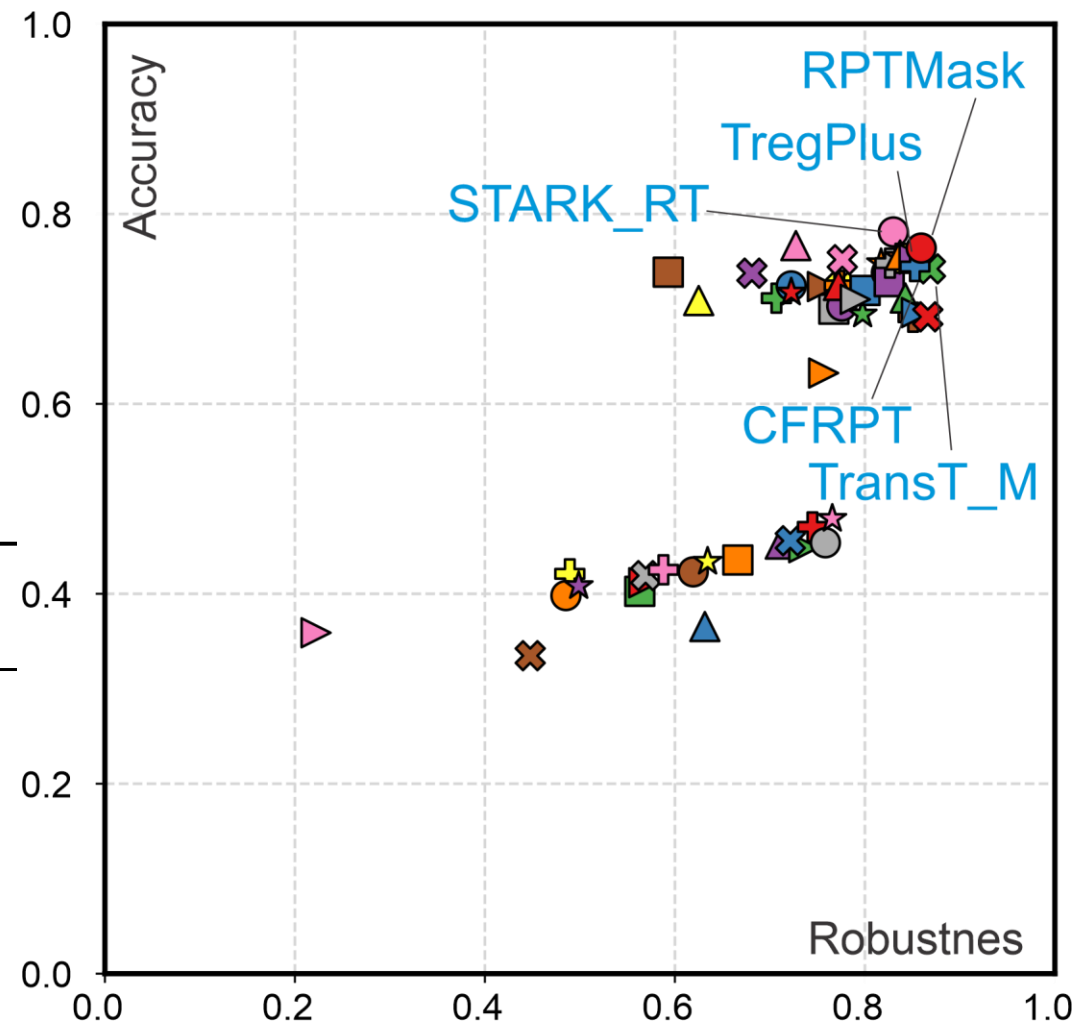
- RPTMask, CFRPT, TransT_M, TregPlus
- Excel both in robustness and accuracy

- Top in accuracy: STARK_RT

- Per-attribute analysis:

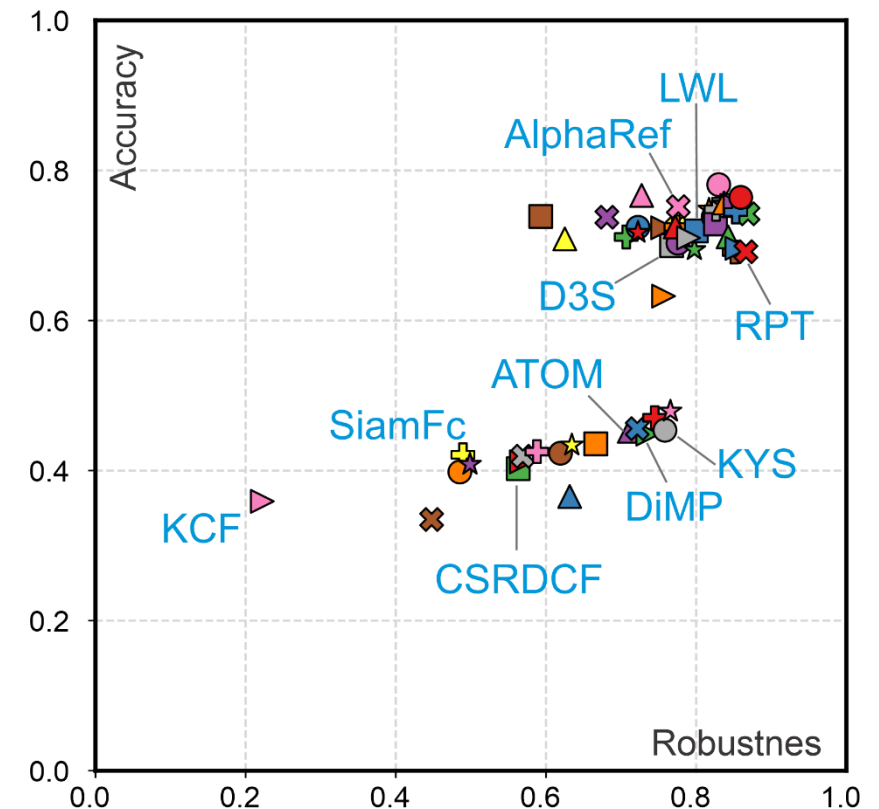
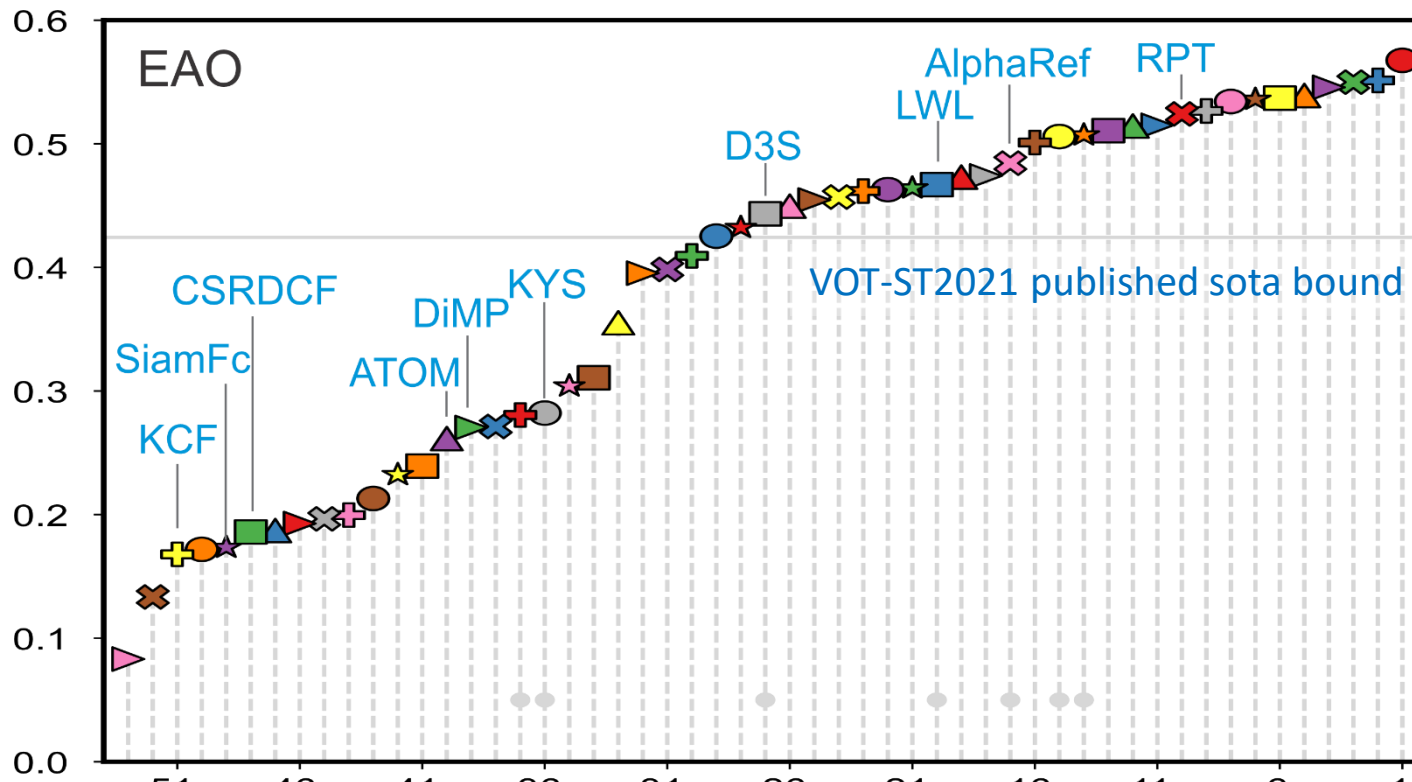
	CM	IC	OC	SC	MC
Accuracy	0.59②	0.59③	0.50①	0.60	0.60
Robustness	0.75	0.71②	0.66①	0.74	0.72③

- **Occlusion** causes most failures and most strongly reduces the accuracy



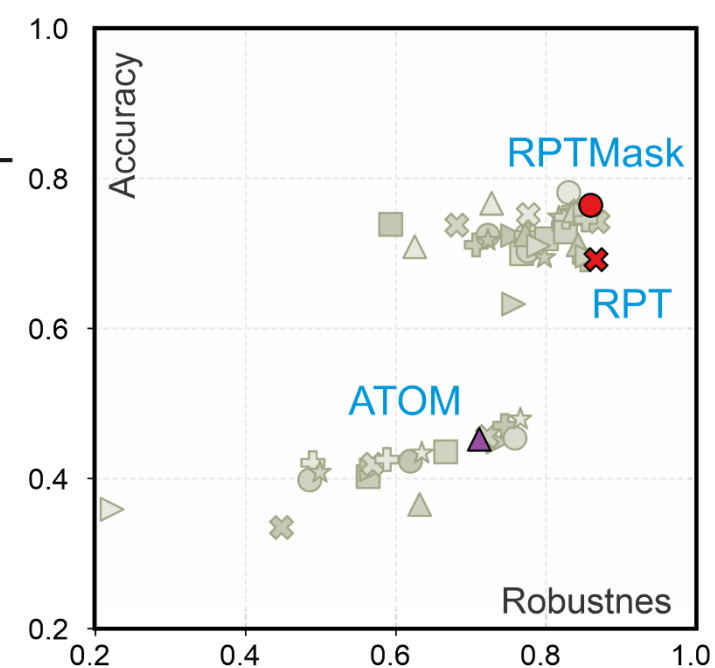
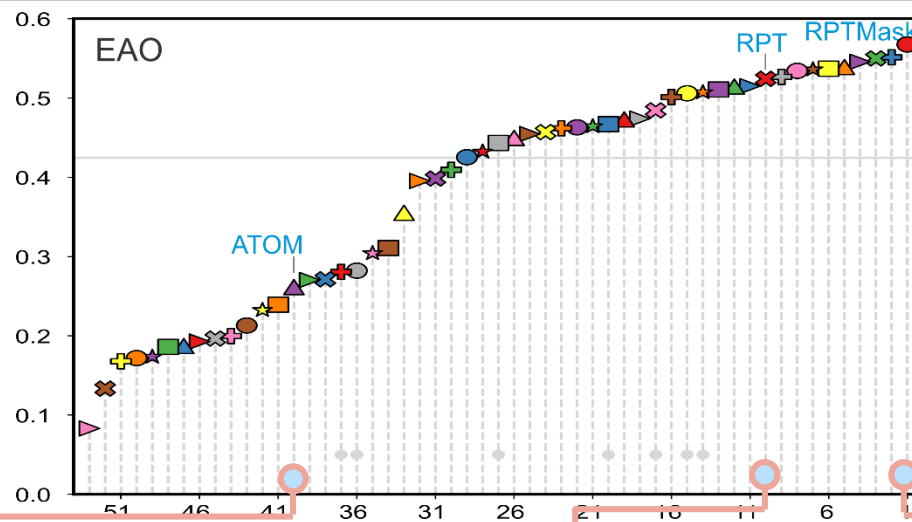
Tracking keeps advancing

- Baselines & 2019 SOTA ranked at the tail of the benchmark
 - 54% submissions above the VOT2021 sota bound
- 9 trackers outperform the VOT-ST2020 winner RPT



Tracking keeps advancing

„Although tracking itself is by and large a solved problem...“
-- Jianbo Shi & Carlo Tomasi
CVPR1994 --



ATOM [Danelljan et al. CVPR2019]

RPT [Ma et al. VOT2020]

RPTMask [Fu et al. VOT2021]



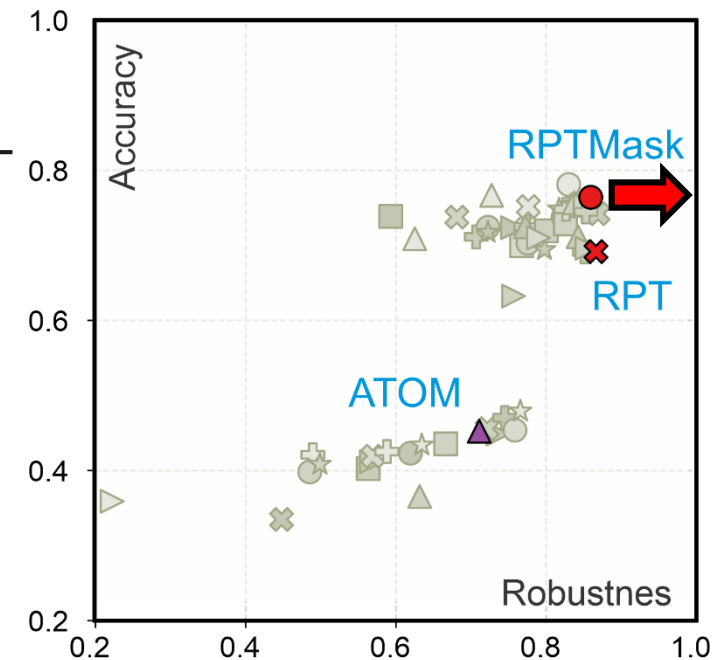
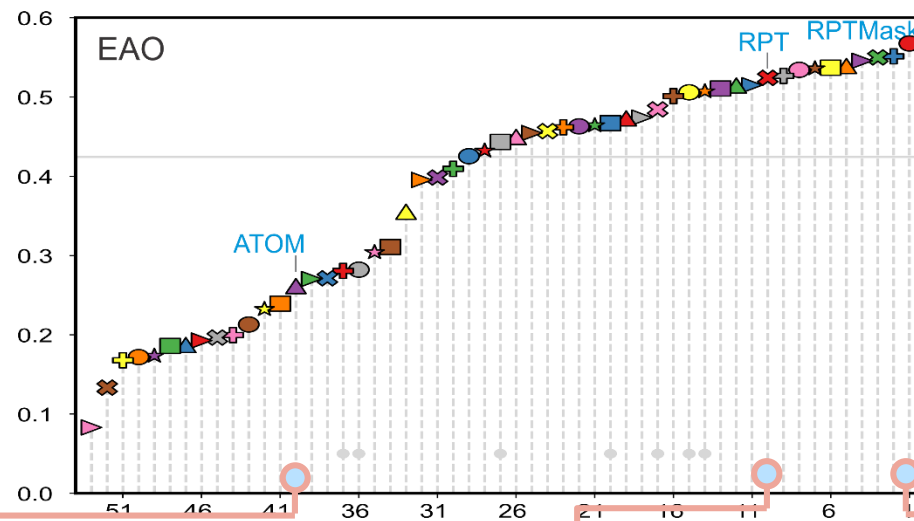
Widely used in the top-performers 2019 -

VOT-ST2020 winner

Top on VOT-ST2021 public set

Close, but no cigar...

~~„Although tracking itself is by
and large a solved problem...“;
-- Jianbo Shi & Carlo Tomasi
CVPR 1994 --~~



ATOM [Danelljan et al. CVPR2019]



RPT [Ma et al. VOT2020]



RPTMask [Fu et al. VOT2021]



Widely used in the top-performers 2019 -

VOT-ST2020 winner

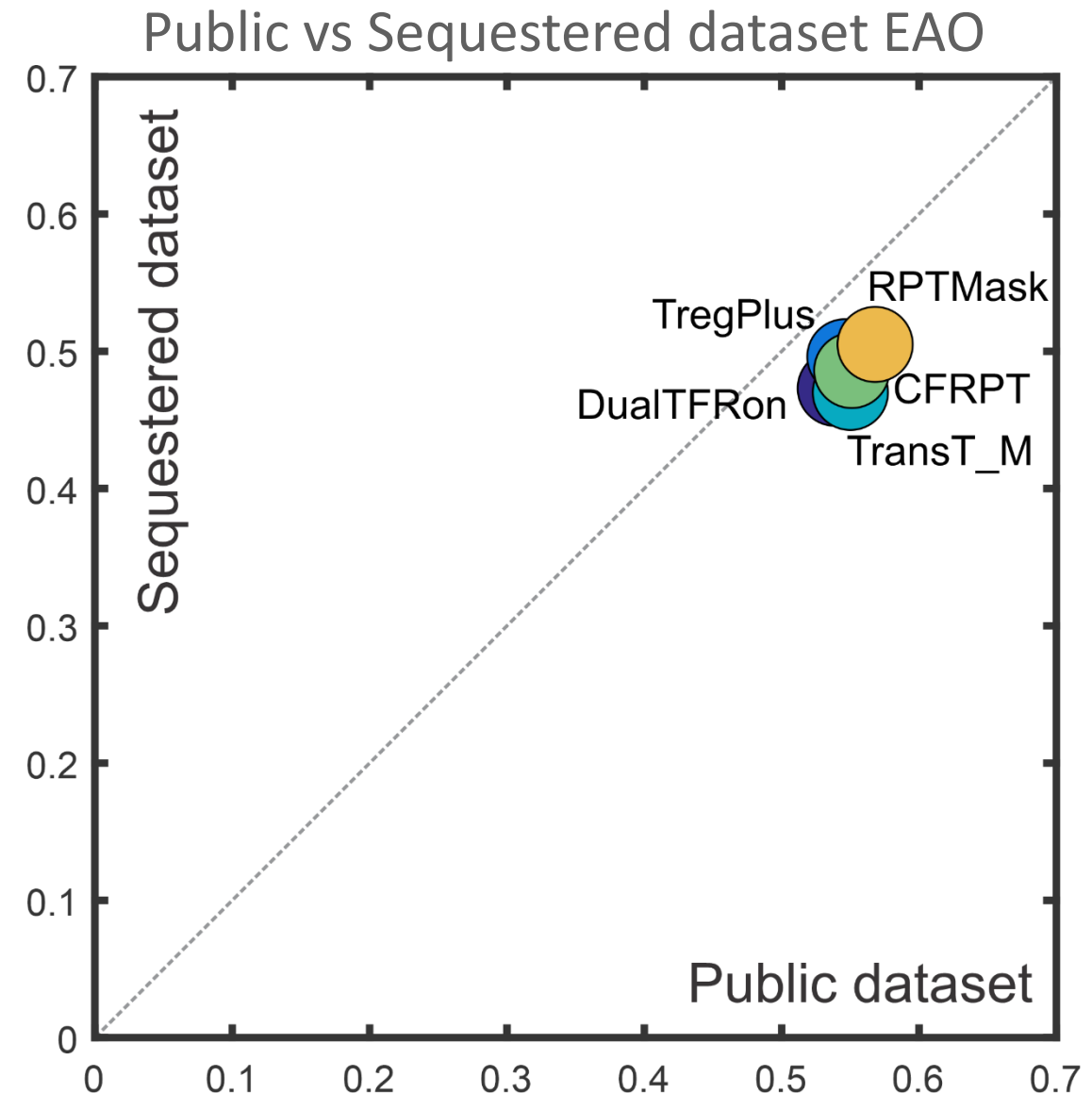
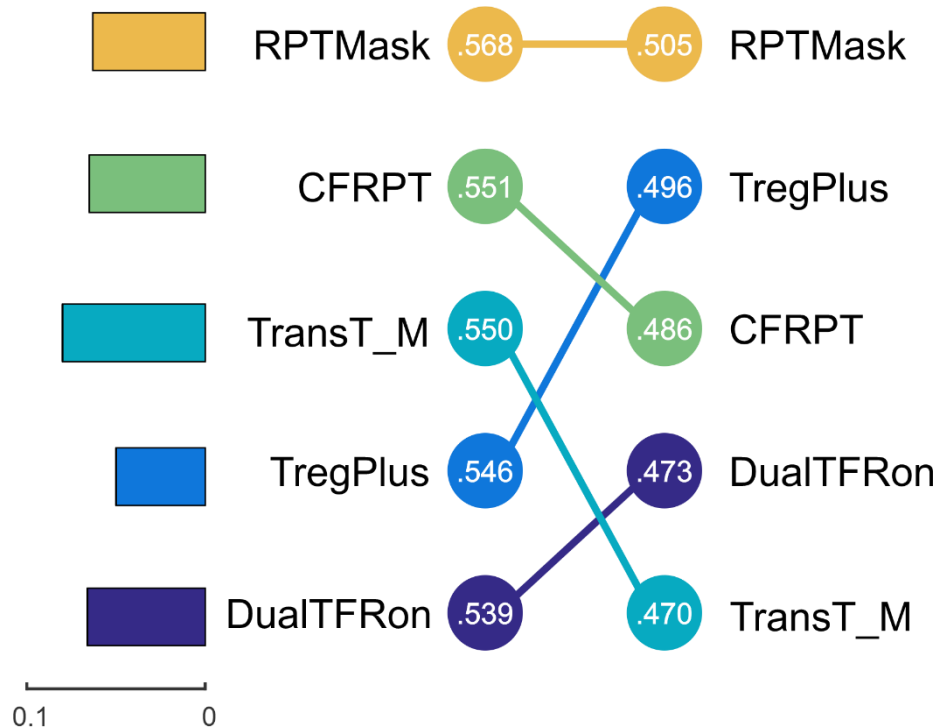
Top on VOT-ST2021 public set

VOT-ST2021 results on sequestered dataset

- Comparable results between public and sequestered set

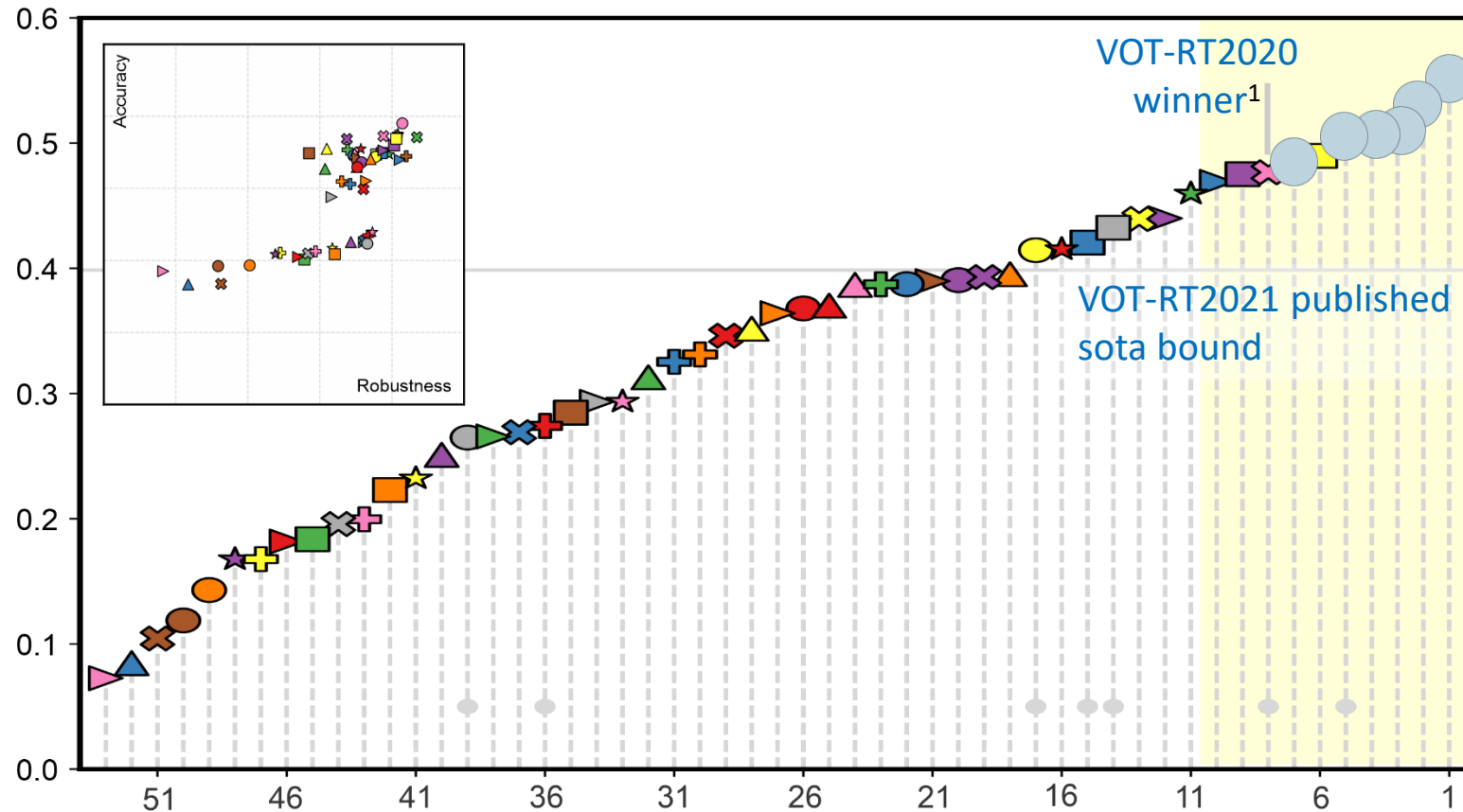
- Biggest performance drop: TransT_M

- Clearly stands out: RPTMask



VOT realtime challenge (VOT-RT2021) results

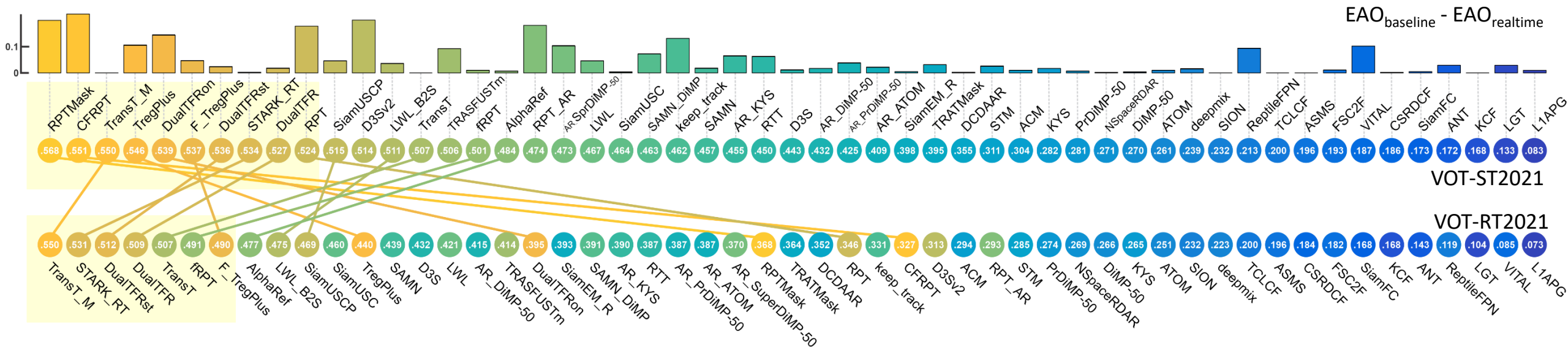
- Top 10: (1) TransT_M, (2) STARK_RT, (3) DualTFRst, (4) DualTFR, (5) TransT, (6) fRPT, (7) F_TregPlus, (8) AlphaRef, (9) LWL_B2S, (10) SiamUSCP



- 7 outperform the VOT-RT2020 winner¹
- 28% of submissions outperform VOT-RT2021 sota bound
- 6 transformer-based

¹AlphaRef [Yan et al., CVPR2021]

VOT2021 Realtime vs Baseline results



- 5 out of 7 top RT trackers among top 10 on VOT-ST2021 challenge!
- The top RT tracker **does not loose any performance** under RT constraint
- 6 of top 7 RT trackers are **Transformers** – *a new class indicating RT potential apart from excellent performance*

VOT2021 ST/RT challenges summary

- VOT-ST2021:
 - Deep DCF and Siamese correlation the dominant methodology
 - Rise of new methodology – Transformers
 - Half of top-performers two stage & apply AlphaRef for segmentation
 - Performance improvements since VOT2020 are primarily due to segmentation accuracy
- VOT-RT2021:
 - Emergence of transformers in real-time trackers
 - Top RT trackers among top ST trackers – performance not sacrificed for speed

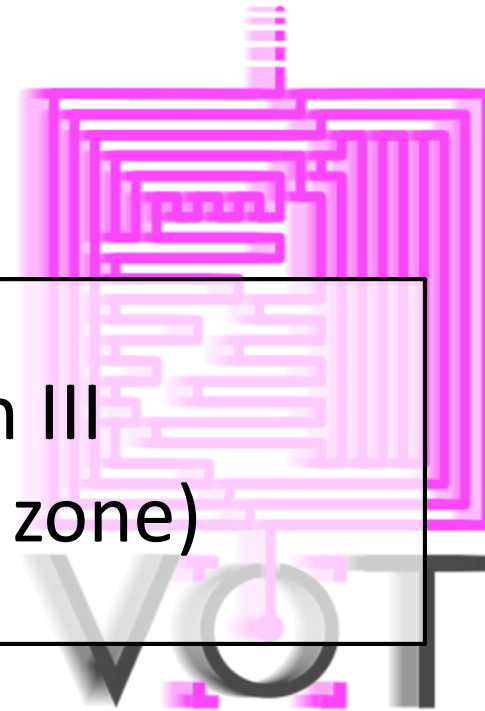
The VOT 2021 workshop

VOT2021 ST/RT WINNER ANNOUNCEMENTS



visual object tracking challenge

Winners talks in Session III
(15:30-16:15, GMT-4 time zone)



visual object tracking challenge

VOT-ST2021 Winners:

RPTMask by: [Zhihong Fu](#),
[Liangliang Wang](#), [Qili Deng](#), [Kang Du](#),
[Qingjie Liu](#), [Yunhong Wang](#)

VOT-RT2021 Winners:

TransT_M by: [Xin Chen](#), [Jiawen Zhu](#),
[Bin Yan](#), [Dong Wang](#),
[Huchuan Lu](#), [Xiaoyun Yang](#)

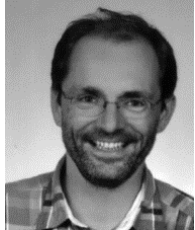
“MultiTemplate Transformer Tracking”

Thanks

- The VOT2021 committee



M. Kristan



J. Matas



A. Leonardis



J. K. Kamarainen



H. J. Chang



R. Pflugfelder



G. Fernandez



L. Čehovin



A. Lukežič



M. Felsberg



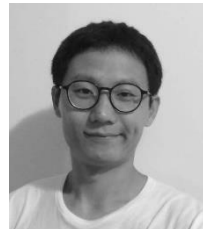
M. Danelljan



O. Drbohlav



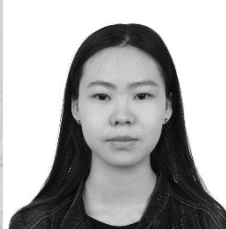
Z. Zhang



Y. Song



Y. Jinyu



Z. Yushan



C. Mayer



G. Häger



J. Kapyla

- Everyone who participated or contributed

Mohamed Abdelpakey⁴⁴, Goutam Bhat⁸, Llukman Cerkezi²³, Hakan Cevikalp¹⁵, Shengyong Chen⁴², Xin Chen¹⁴, Miao Cheng⁵², Ziyi Cheng²⁷, Yu-Chen Chiu⁴¹, Ozgun Cirakman²³, Yutao Cui³¹, Kenan Dai¹⁴, Mohana Murali Dasari²², Qili Deng¹², Xingping Dong²¹, Daniel K. Du¹², Matteo Dunnhofer⁵⁰, Zhen-Hua Feng⁴⁹, Zhiyong Feng⁴³, Zhihong Fu¹⁰, Shiming Ge⁴⁵, Rama Krishna Gorthi²², Yuzhang Gu³⁸, Bilge Günsel²³, Qing Guo³², Filiz Gurkan²³, Wencheng Han¹¹, Yanyan Huang¹⁷, Felix J'aremo Lawin⁴, Shang-Jhih Jhang⁴¹, Rongrong Ji⁵¹, Cheng Jiang³¹, Yingjie Jiang²⁵, Felix Juefei-Xu⁹, Yin Jun⁵², Xiao Ke¹⁷, Fahad Shahbaz Khan²⁹, Byeong Hak Kim²⁶, Josef Kittler⁴⁹, Xiangyuan Lan¹⁹, Jun Ha Lee²⁶, Bastian Leibe³⁶, Hui Li²⁵, Jianhua Li¹⁴, Xianxian Li¹⁸, Yuezhou Li¹⁷, Bo Liu²⁴, Chang Liu¹⁴, Jingen Liu²⁴, Li Liu³⁷, Qingjie Liu¹⁰, Huchuan Lu^{14,33}, Wei Lu⁵², Jonathon Luiten³⁶, Jie Ma²⁰, Ziang Ma⁵², Niki Martinel⁵⁰, Christoph Mayer⁸, Alireza Memarmoghadam⁴⁶, Christian Micheloni⁵⁰, Yuzhen Niu¹⁷, Danda Paudel⁸, Houwen Peng²⁸, Shoumeng Qiu³⁸, Aravindh Rajiv²², Muhammad Rana⁴⁹, Andreas Robinson⁴, Hasan Saribas¹⁶, Ling Shao²¹, Mohamed Shehata⁴⁴, Furao Shen³¹, Jianbing Shen²¹, Kristian Simonato⁵⁰, Xiaoning Song²⁵, Zhangyong Tang²⁵, Radu Timofte⁸, Philip Torr⁴⁷, Chi-Yi Tsai⁴¹, Bedirhan Uzun¹⁵, Luc Van Gool⁸, Paul Voigtlaender³⁶, Dong Wang¹⁴, Guangting Wang⁴⁸, Liangliang Wang¹², Lijun Wang¹⁴, Limin Wang³¹, Linyuan Wang⁵², Yong Wang⁴⁰, Yunhong Wang¹⁰, Chenyan Wu³⁴, Gangshan Wu³¹, Xiao-Jun Wu²⁵, Fei Xie³⁹, Tianyang Xu^{25,49}, Xiang Xu³¹, Wanli Xue⁴², Bin Yan¹⁴, Wankou Yang³⁹, Xiaoyun Yang³⁵, Yu Ye¹⁷, Jun Yin⁵², Chengwei Zhang¹³, Chunhui Zhang⁴⁵, Haitao Zhang⁵², Kaihua Zhang³⁰, Kangkai Zhang⁴⁵, Xiaohan Zhang¹⁴, Xiaolin Zhang³⁸, Xinyu Zhang¹⁴, Zhibin Zhang⁴², Shaochuan Zhao²⁵, Ming Zhen¹², Bineng Zhong¹⁸, Jiawen Zhu¹⁴, and Xue-Feng Zhu²⁵

- VOT2021 sponsor:



University of Ljubljana
Faculty of Computer and
Information Science