



The 10th Visual Object Tracking Challenge New Dataset & Results of VOT-LT2022

Matej Kristan, Aleš Leonardis, Jiri Matas, Joni-Kristian Kämäräinen, **Hyung Jin Chang**, Roman Pflugfelder, Gustavo Fernandez, Luka Čehovin Zajc, Alan Lukežič, Michael Felsberg, Martin Danelljan, Ondrej Drbohlav, Johanna Björklund, Yushan Zhang, **Zhongqun Zhang**, Song Yan, Wenyan Yang, Dingding Cai, Christoph Mayer



UNIVERSITY OF
BIRMINGHAM



li.u LINKÖPING
UNIVERSITY

AIT
AUSTRIAN INSTITUTE
OF TECHNOLOGY

Tampere
University

ETH zürich

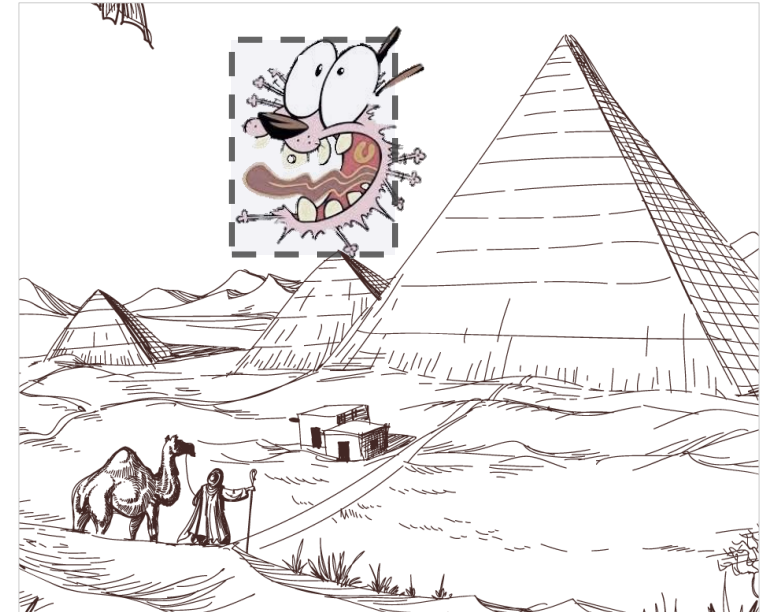
Outline

1. Scope of the VOT2022 LT challenges
2. Results overview (VOT2022 LT)
3. Winner announcement (VOT2022 LT)

The VOT 2022 workshop

VOT2022 LT CHALLENGE: OVERVIEW

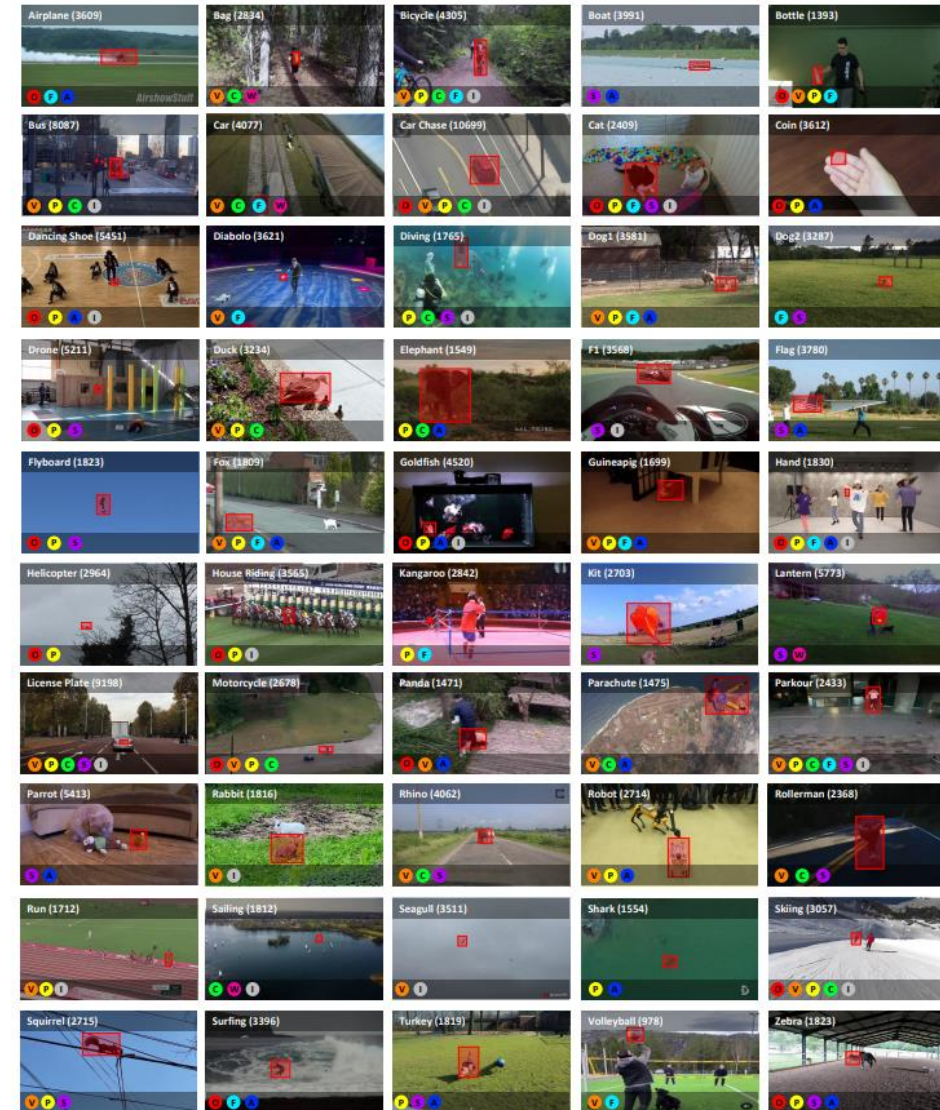
VOT2022 long-term challenge (VOT-LT2022)



- 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

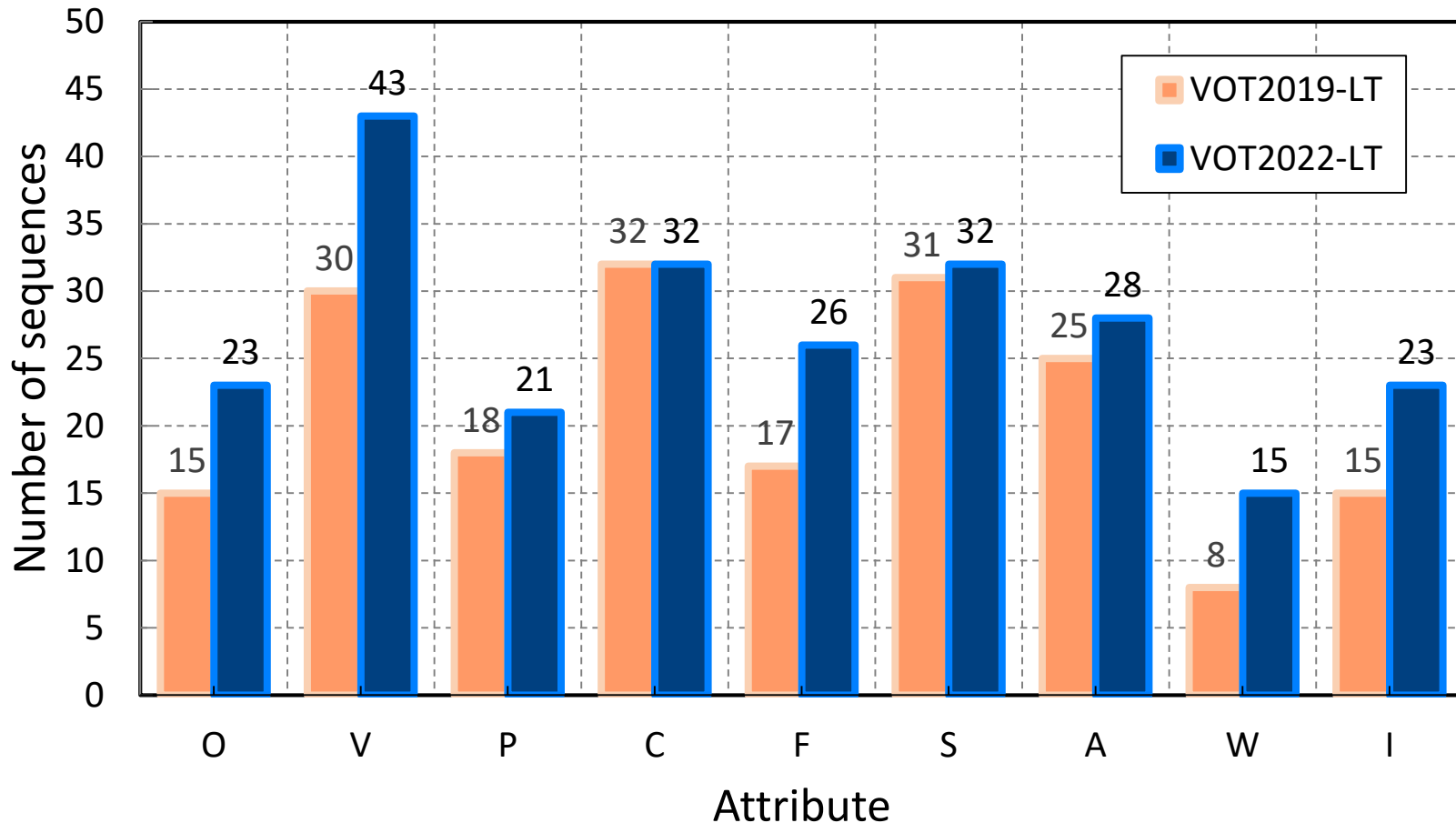
The VOT-LT2022 dataset (50 Sequences)

- VOT-LT2022 introduced new 50 sequences
 - Diverse objects (persons, cars, motorcycles, bicycles, boats, animals, etc.)
 - Total length of **168,282** frames with resolution of **1280x720**
 - Average sequence **length >4k frames**
 - Average **per sequence disappearance: 10**
 - Average **target absence period: 52 frames**
 - Annotated by axis-aligned bounding boxes
- VOT-LT2021 dataset is now available for training (50 Sequences)



¹Lukežič, et al., Now you see me: evaluating performance in long-term visual tracking, TCyb2020

The VOT-LT2022 dataset (9 Attributes)

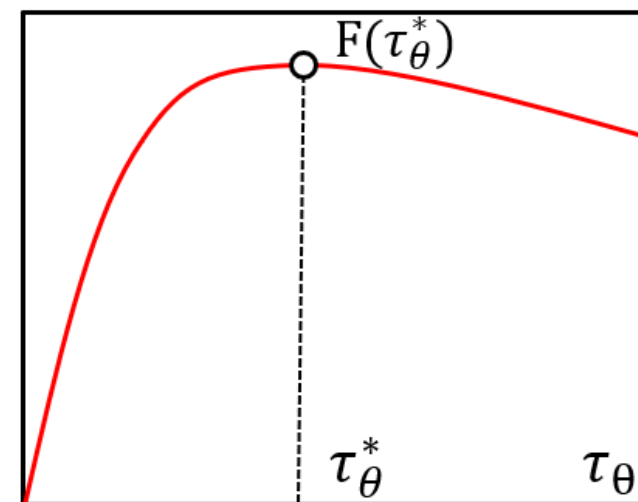
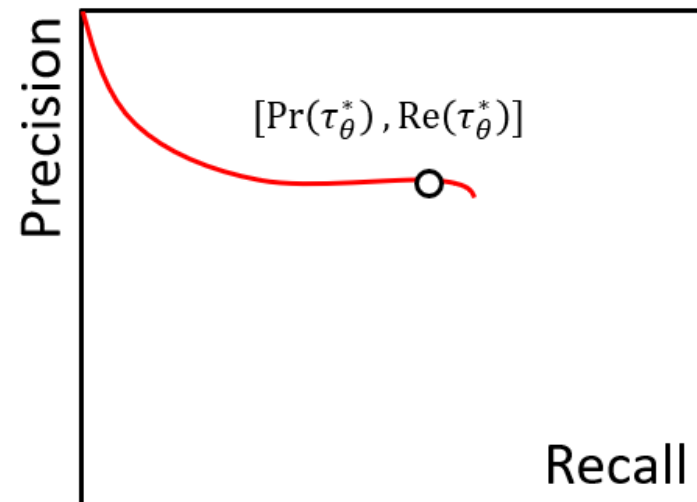


- Visual attributes (9):
 - **O: Full occlusion**
 - V: Out-of-view
 - **P: Partial occlusion**
 - C: Camera motion
 - **F: Fast motion**
 - S: Scale change
 - A: Aspect ratio change
 - **W: Viewpoint change**
 - **I: Similar objects**

*The new VOT-LT dataset is more challenging in **small objects**, **similar objects**, **fast motion**, and **full/partial occlusions***

The VOT-LT2022 evaluation methodology

- Tracking properties measured:
 - Localization, Loss/Presence detection
- Initialized at first frame, no reset at target loss
- Three LT measures from VOT-LT2018:
 - Tracking Precision, Recall & F-score
 - $\text{Pr}(\tau_\theta)$, $\text{Re}(\tau_\theta)$, $F(\tau_\theta)$
 - Evaluated at presence certainty threshold τ_θ^* that maximizes the tracker F-score
- Winner: Top performer in $F(\tau_\theta^*)$



¹Lukežič, et al., Now you see me: evaluating performance in long-term visual tracking, TCyb2020

The VOT 2022 workshop

VOT-LT2022 CHALLENGE RESULTS

VOT-LT2022 challenge overview

- **9** trackers tested (**7** trackers are new submissions and we contribute **2** baseline method)
- All trackers were from LT1 class
 - *Explicit target absence detection and re-detection implemented*
- All trackers extract CNN features.
- Many trackers combined different tracking methods and switched them based on their tracking scores.
- ***Transformer*** are getting more and more popular
 - 4 Trackers applied **Transformer** architecture (**STARK**)
 - 4 Trackers applied **SuperDiMP** structure
 - 3 Trackers applied **KeepTrack** as their auxiliary tracker
 - 2 Trackers used **MDNet-based** verifier to determine their tracking score

VOT-LT2022 challenge overview

- Top 3 trackers: *VITKT_M*, *mixLT*, and *HuntFormer*
 - *Fusion of multiple trackers and motion prediction model*
- Top performance: VITKT_M
 - *Trackers: STARK[1] + KeepTrack[2]*
 - *A simple motion module (~1.2% improved)*
- Second-best (~1.7% Worse) : mixLT
 - *STARK + SuperDiMP[3]*
- Baseline: mlpLT (winner of VOT-LT2021)
 - *4 trackers outperformer the baseline*

Tracker	Pr	Re	F-Score	Year
● VITKT_M	0.629①	0.604②	0.617①	2022
⊕ mixLT	0.608②	0.592③	0.600②	2022
✘ HuntFormer	0.586	0.610①	0.598③	2022
▶ CoCoLoT	0.591③	0.577	0.584	2022
▲ mlpLT	0.568	0.562	0.565	2022
■ KeepTrack	0.572	0.550	0.561	2022
★ D3SLT	0.520	0.516	0.518	2022
● Super_DiMP	0.510	0.496	0.503	2022
⊕ ADiMPLT	0.489	0.514	0.501	2022

[1] Yan et al. ICCV2021

[2] Mayer et al. ICCV2021

[3] Bhat et al. ECCV2020

The VOT 2022 workshop

VOT2022 LT WINNER ANNOUNCEMENTS



Winners talks in Session III
(16:35-18:00, GMT+2 time zone)

VOT-LT2022 Winners:

VITKT_M by: Jie Zhao, Xin Chen, Chang Liu, Houwen Peng, Dong Wang, Huchuan Lu

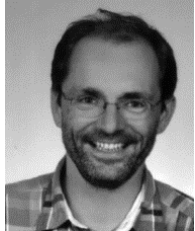
“Fusing VITTrack, KeepTrack and a motion module for Long-term Visual Tracking”

Thanks

- The VOT2022 committee



M. Kristan



J. Matas



A. Leonardis



J. K. Kamarainen



H. J. Chang



R. Pflugfelder



G. Fernandez



L. Čehovin



A. Lukežič



M. Felsberg



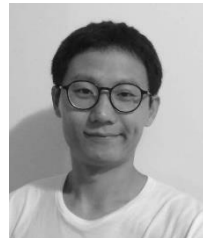
M. Danelljan



O. Drbohlav



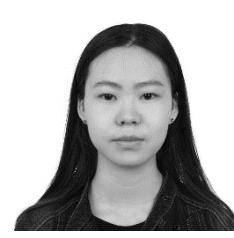
Z. Zhang



Y. Song



Wenyan Yang



Z. Yushan



C. Mayer



Dingding Cai



Johanna Björklund

- Everyone who participated or contributed

Kang Ben18, Goutam Bhat10, Hong Chang24, Guangqi Chen16, Jiaye Chen26, Shengyong Chen43, Xilin Chen24, Xin Chen18, Xiuyi Chen13, Yiwei Chen35, Yu-Hsi Chen12, Zhixing Chen16, Yangming Cheng55, Angelo Ciaramella47, Yutao Cui30, Benjamin D'zubur1, Mohana Murali Dasari22, Qili Deng16, Debajyoti Dhar39, Shangzhe Di14, Emanuel Di Nardo46,47, Daniel K. Du16, Matteo Dunnhofer51, Heng Fan48, Zhenhua Feng50, Zhihong Fu16, Shang Gao41, Rama Krishna Gorthi22, Eric Granger27, Q.H. Gu15, Himanshu Gupta19, Jianfeng He49, Keji He13, Yan Huang13, Deepak Jangid19, Rongrong Ji53, Cheng Jiang30, Yingjie Jiang26, Felix J'aremo Lawin4, Ze Kang26, Madhu Kiran27, Josef Kittler50, Simiao Lai18, Xiangyuan Lan32, Dongwook Lee34, Hyunjeong Lee34, Seohyung Lee34, Hui Li26, Ming Li17, Wangkai Li49, Xi Li55, Xianxian Li20, Xiao Li16, Zhe Li41, Liting Lin37, Haibin Ling40, Bo Liu25, Chang Liu18, Si Liu23, Huchuan Lu18, Rafael M. O. Cruz27, Bingpeng Ma44, Chao Ma36, Jie Ma21, Yinchao Ma49, Niki Martinel51, Alireza Memarmoghadam45, Christian Micheloni51, Payman Moallem45, Le Thanh Nguyen-Meidine27, Siyang Pan35, ChangBeom Park34, Danda Paudel10, Matthieu Paul10, Houwen Peng28, Andreas Robinson4, Litu Rout39, Shiguang Shan24, Kristian Simonato51, Tianhui Song30, Xiaoning Song26, Chao Sun55, Jingna Sun16, Zhangyong Tang26, Radu Timofte10,52, Chi-Yi Tsai42, Luc Van Gool10, Om Prakash Verma19, Dong Wang18, Fei Wang49, Liang Wang13, Liangliang Wang16, Lijun Wang18, Limin Wang30, Qiang Wang35, Gangshan Wu30, Jinlin Wu13, Xiaojun Wu26, Fei Xie38, Tianyang Xu26, Wei Xu16, Yong Xu37, Yuanyou Xu55, Wanli Xue43, Zizheng Xun14, Bin Yan18, Dawei Yang49, Jinyu Yang41, Wankou Yang38, Xiaoyun Yang33, Yi Yang55, Yichun Yang30, Zongxin Yang55, Botao Ye24, Fisher Yu10, Hongyuan Yu13, Jiaqian Yu35, Qianjin Yu49, Weichen Yu13, Kang Ze26, Jiang Zhai38, Chengwei Zhang17, Chunhu Zhang36, Kaihua Zhang29, Tianzhu Zhang49, Wenkang Zhang38, Zhibin Zhang43, Zhipeng Zhang31, Jie Zhao18, Shaochuan Zhao26, Feng Zheng41, Haixia Zheng54, Min Zheng16, Bineng Zhong20, Jiawen Zhu18, Xuefeng Zhu26, and Yueting Zhuang55

- VOT2022 sponsor:



University of Ljubljana
Faculty of Computer and
Information Science