Learning a Robust Society of Tracking Parts using Co-occurrence Constraints

Elena Burceanu [1,2,3] and Marius Leordeanu[1,2,4]

eburceanu@bitdefender.com, marius.leordeanu@imar.ro

[1] Bitdefender, Romania, [2] Institute of Mathematics of the Romanian Academy

[3] Mathematics and Computer Science, University of Bucharest, Romania

[4]Automatic Control and Computer Science, University Politehnica of Bucharest, Romania

Motivation

- **Stability through steadiness**: Each pathway has golden members (classifiers and first frames). Their vote always count and they never loose their status.
- **Robust Adaptation**: Parts promote to a higher role or update, only if they are validated in time, based on continuous co-occurrences supervision signal.
- **Robust frame to frame tracking**: Each move is based on strong co-occurrences of votes at a single location.

Contribution

1. **Co-occurrences** guided supervision. Train only on **Highly Confident Frames**.
2. Different **reliability roles** for FilterParts classifiers that prove their value in time.
3. **Society of Tracking Parts (STP)**: combining votes from a robust conservative pathway with a more adaptive and up-to-date one.
4. **Efficient one sample vs. all** strategy for learning FilterParts classifiers.

1. ConvNetPart

- Deep end-to-end architecture
- Segmentation task for obj center prediction
- Provides **robust adaptability** to change

Training

- On **Highly Confident Frames** (HCFs): frames where the different and complementary pathways agree.
- Update on the first initial (**gold**) frames too.

2. FilterParts

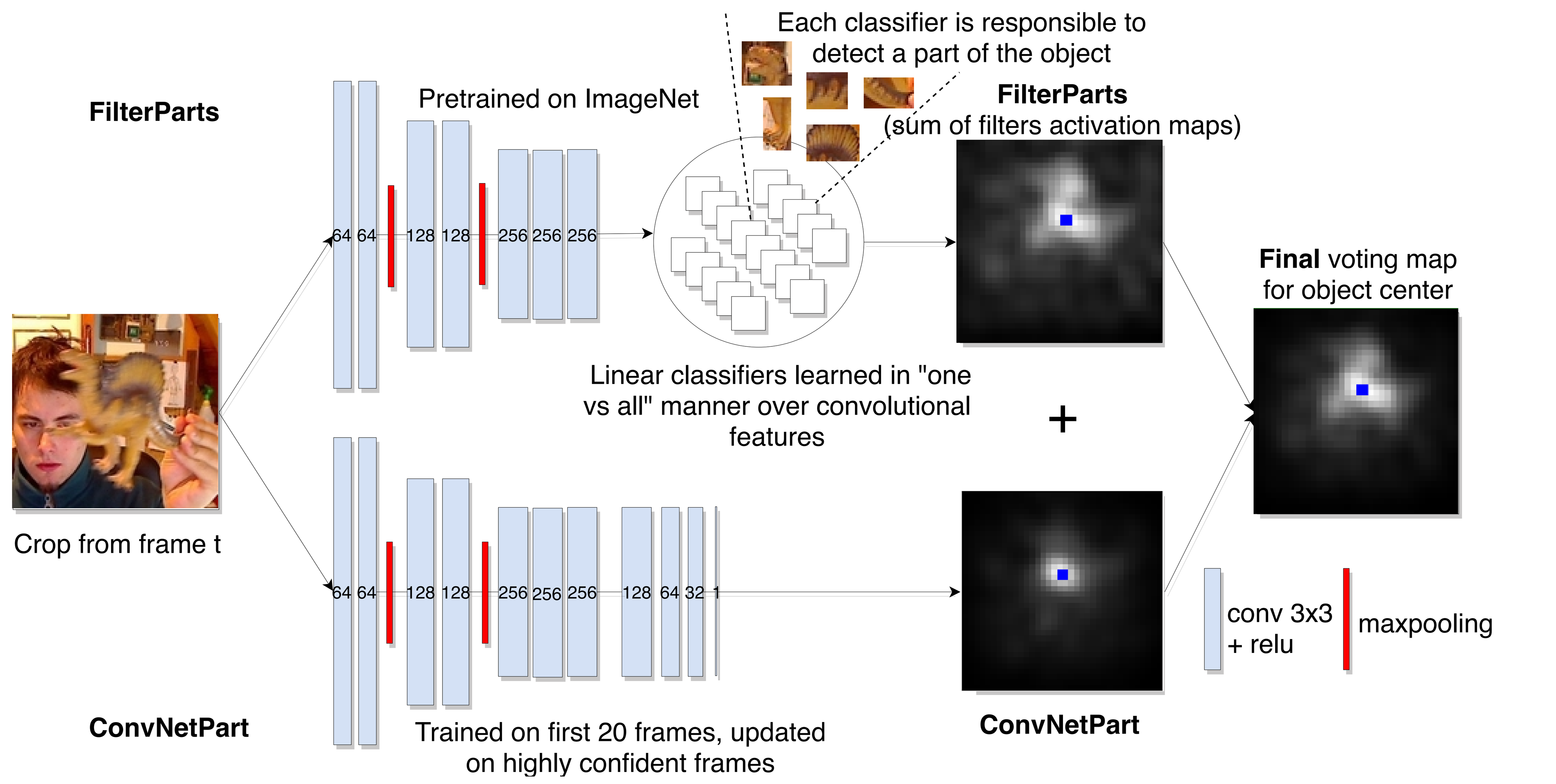
Choosing Parts:

- each **part** has a corresponding **patch** and differs from other by size, location and its reliability role
- choose only **highly discriminative linear filters** over the feature maps
- **positives** inside the voted bounding box, **negatives** outside

Promoting parts validated in time:

- each part votes for the target center and the votes are **monitored** over time
- parts that vote **consistently in agreement** with the chosen center are promoted
- roles: **candidate**, **reliable** and **gold**

3. Society of Tracking Parts (STP)



4. One sample versus all

Efficient multi-class filter learning:

- Weighted least squares needs **n matrix inversions** ($\min \frac{1}{n}\|\mathbf{W}^{\frac{1}{2}}(\mathbf{D}\mathbf{c}_i - \mathbf{y}_i)\|^2 + \lambda \mathbf{c}_i^\top \mathbf{c}_i$)
- **Theoretical novelty**: the weighted solution has the same direction with the simple one, and we computed the exact ratio
- **Our Theorem**: Given $y_- = 0, y_+ = 1, w_- = 1, w_+ = w_i = n-1, n$ the number of classifiers and c_i the solution to least square: $c_i\_weighted = \frac{n}{1+(n-1)\mathbf{d}_i^\top \mathbf{c}_i} * c_i$

Optimizations:

- **Single matrix inversion** for all parts
- **Matrix Inversion Lemma** to rewrite the equation and invert 200 times smaller matrix

(Robust SOTA) Results

Tracker	VOT17 Dataset [1]		
	EAO	$R\downarrow$	$A\uparrow$
STP (ours)	**0.309**	**0.76**	0.44
CFWCR [2]	0.303	1.2	0.48
ECO [3]	0.28	1.13	0.48
CCOT [4]	0.267	1.31	0.49

Metrics:

- **EAO**: sota on VOT17; 3rd score on VOT16
- **R**: STP outperforms the published methods in terms of **robustness** on both VOT17 and VOT16 [5] by a large margin.
- We **strongly outperform the others** on occlusion (identified by VOT17/16 as the **most difficult case**)

We used the exact **same set of parameters** on all videos from both VOT17 and VOT16.

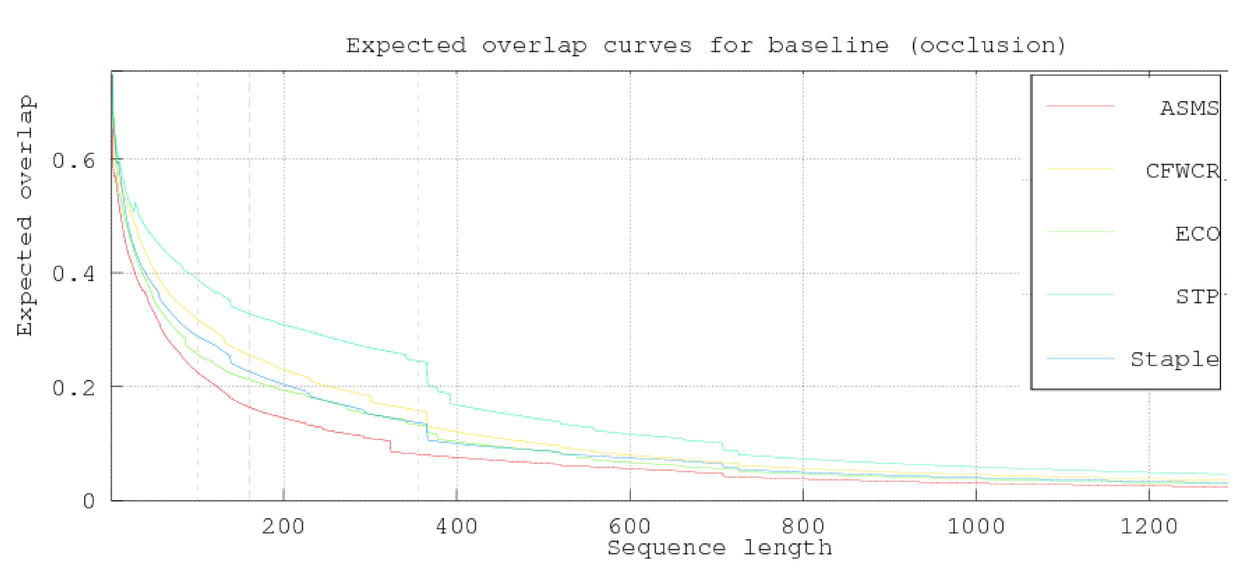
Ablation study

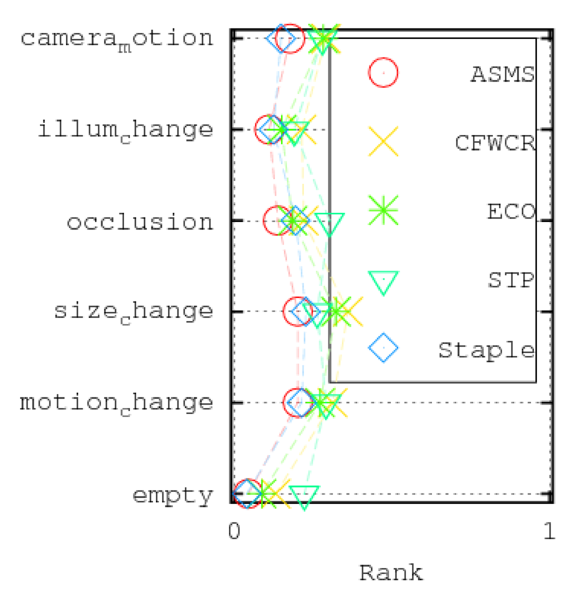
Version	VOT17 Dataset [1]		
	EAO	$R\downarrow$	$A\uparrow$
FilterParts	0.25	0.99	0.42
ConvNetPart	0.20	2.09	0.43
Combined	**0.309**	**0.765**	**0.44**
One role	0.262	0.99	0.44
Reliability roles	**0.309**	**0.765**	**0.44**
No update	0.28	0.95	0.43
Full update	0.284	0.92	**0.44**
HCFs update	**0.309**	**0.765**	**0.44**

Elements that significantly improved our final performance:

- Combining the **two pathways** (one robust and one that easily adapt to changes)
- Validating parts in time based on **co-occurrence** signal between parts
- ConvNetPart update only on **HCFs** (also a form of co-occurrence between the two pathways)

Occlusion and other difficult cases





References

[1] Kristan et al. VOT17. In *ICCV Workshops*, 2017.

[2] He et al. CFWCR. In *ICCV Workshops*, 2017.

[3] Danelljan et al. ECO. In *CVPR*, 2017.

[4] Danelljan et al. CCOT. In *ECCV*, 2016.

[5] Kristan et al. VOT16. In *ECCV Workshops*, 2016.