Online Object Tracking with Proposal Selection

Yang Hua   Karteek Alahari   Cordelia Schmid

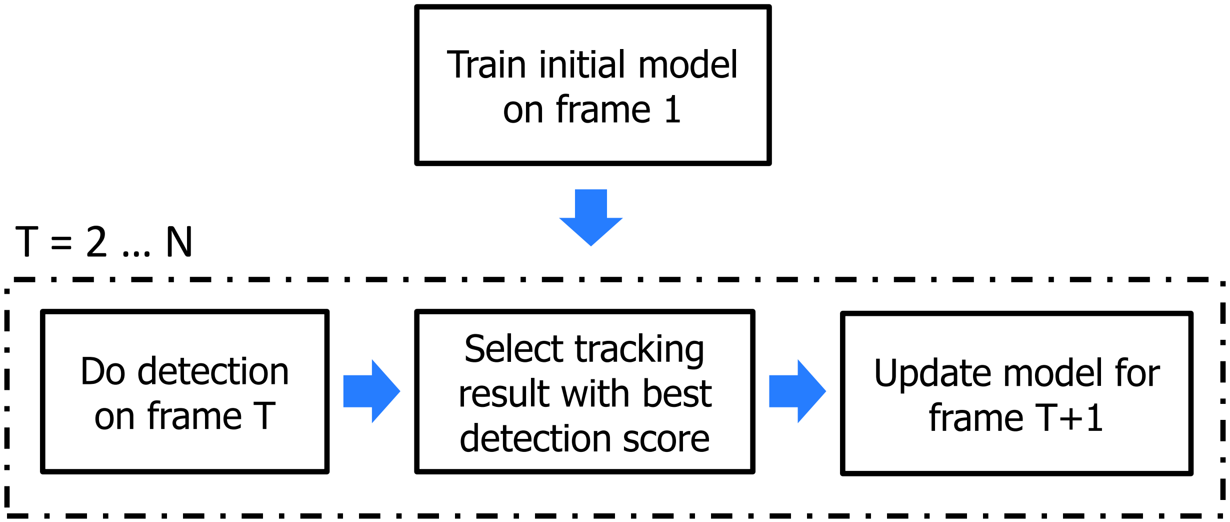
Inria Grenoble Rhône-Alpes, France

Outline

- ❑ Background
- ❑ Our approach
- ❑ Experimental results
- ❑ Summary

Background

- Tracking-by-detection paradigm has been extremely successful on diverse benchmarks [Wu et al., 2013] [Kristan et al., 2013/14] [Smeulders et al., 2014]



Y. Wu, J. Lim, and M.-H. Yang. Online object tracking: A benchmark. In CVPR, 2013.

A. W. M. Smeulders et al. Visual tracking: an experimental survey. PAMI, 2014

M. Kristan et al. The visual object tracking VOT2013/2014 challenge results. In ICCV/ECCV VOT Challenge Workshop, 2013/2014.

Background

❑ Successful tracking-by-detection methods

- Struck [Hare et al., 2011]
- PLT 13/14 [Heng et al., 2012]
- DSST [Danelljan et al., 2014]

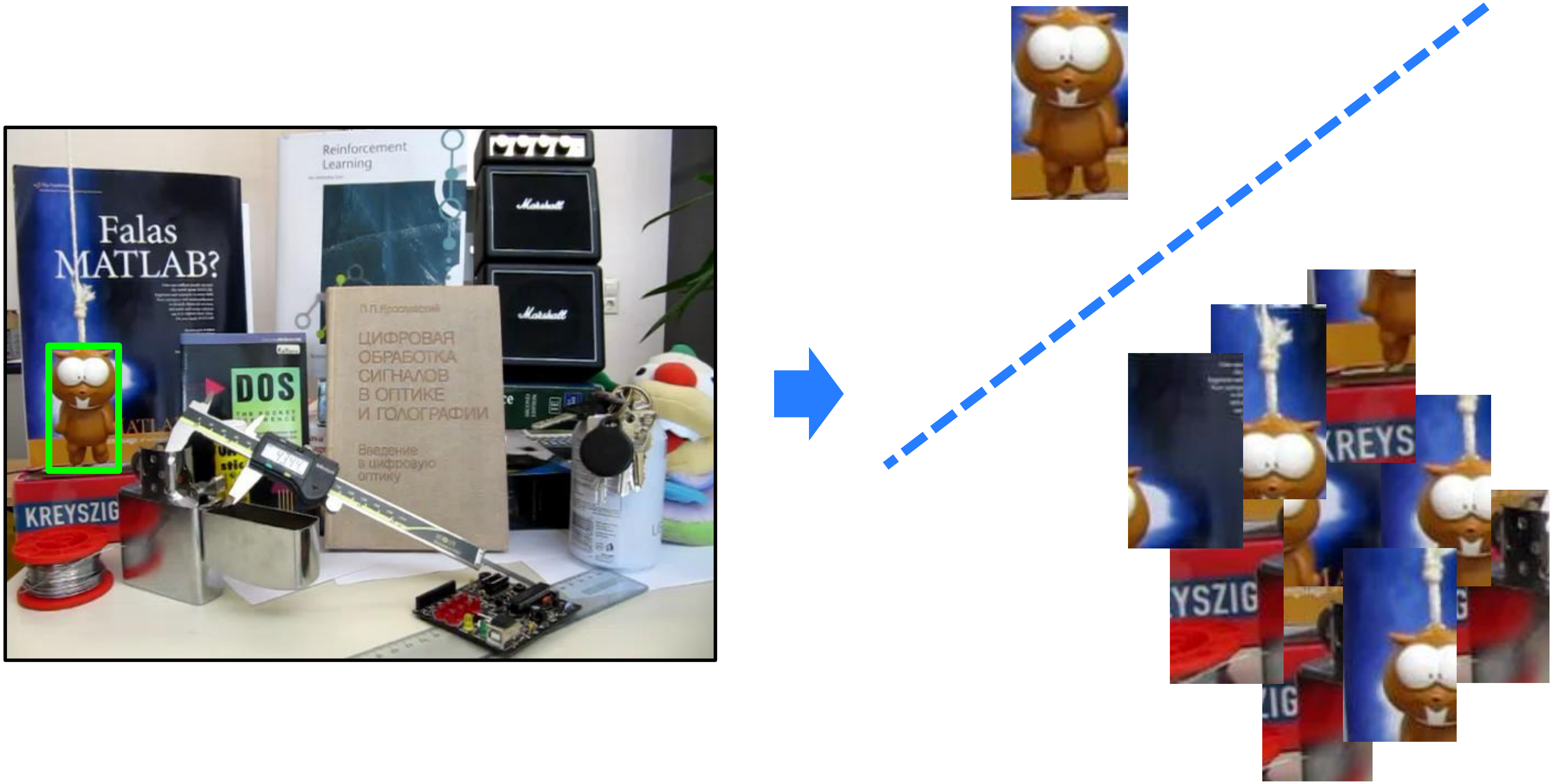
S. Hare, A. Saffari, and P. H. S. Torr. Struck: Structured output tracking with kernels. In ICCV, 2011.

C.K. Heng, Y. Sumio, M. Yuichi, T. Hajime, Shrink boost for selecting multi-lbp histogram features in object detection. In CVPR, 2012

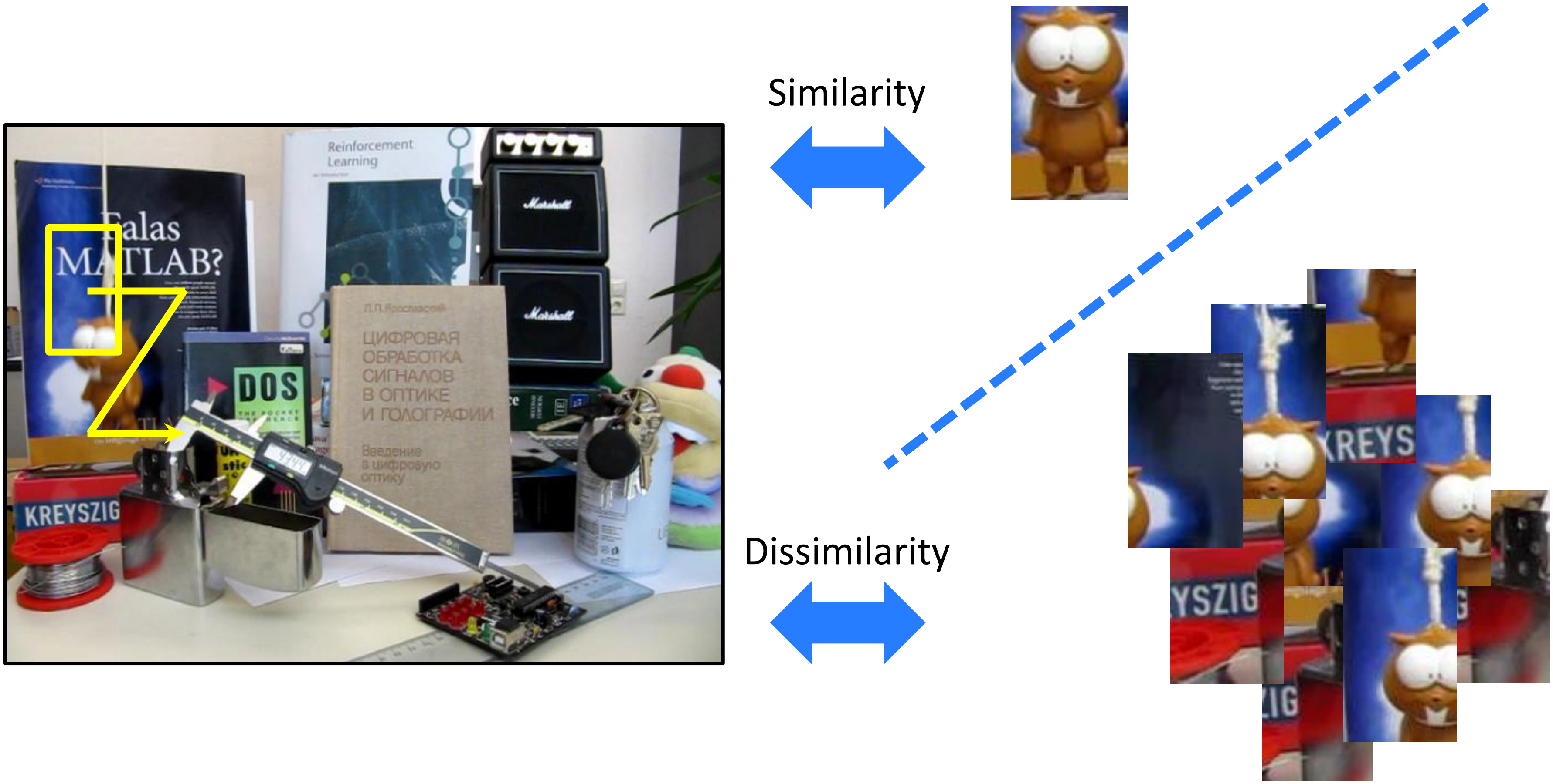
M. Danelljan, G. Hager, F. Shahbaz Khan, and M. Felsberg. Accurate scale estimation for robust visual tracking. In BMVC, 2014.

Background

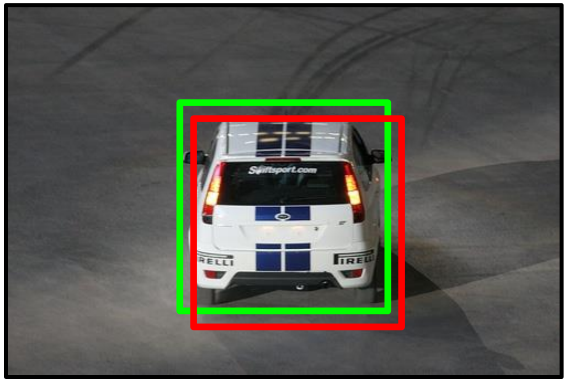
- Two key ingredients
  - Discriminative learning



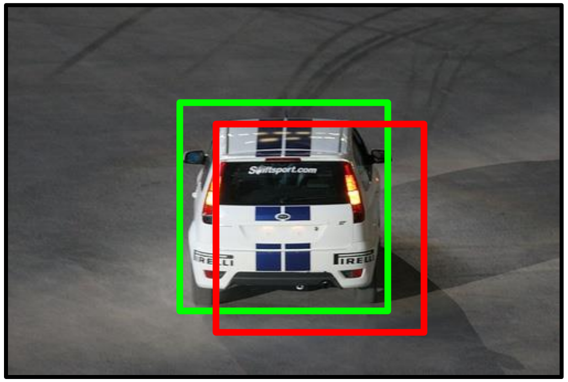
- Two key ingredients
  - Discriminative learning



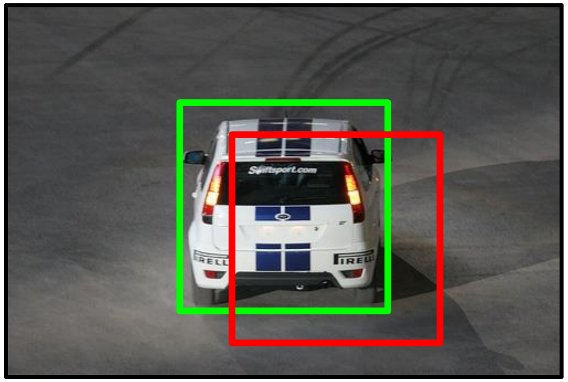
- Two key ingredients
  - Discriminative learning
  - Pixel-accurate localization



IoU = 0.9



IoU = 0.7

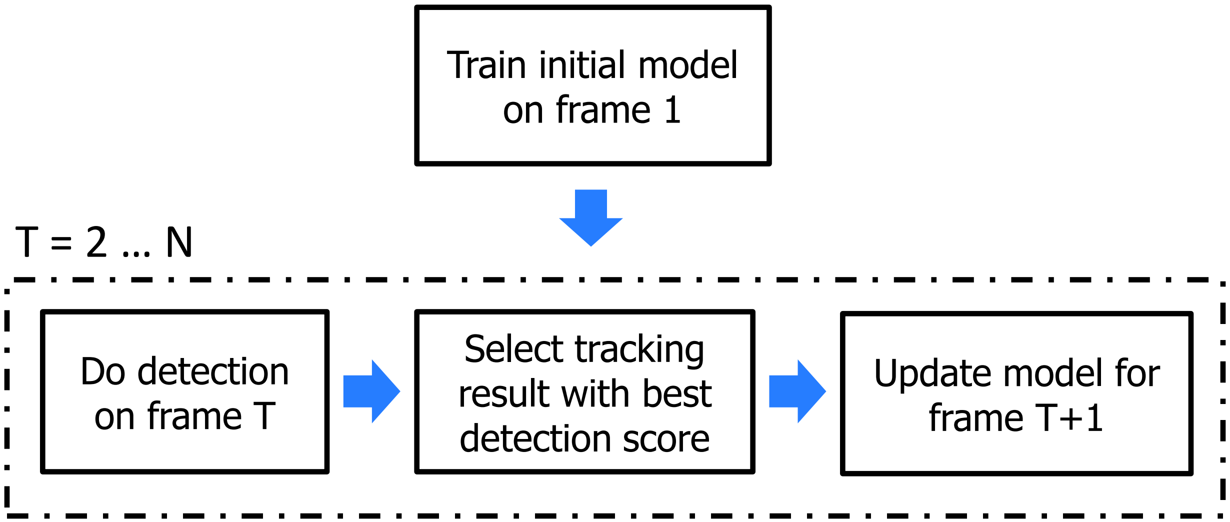


IoU = 0.5

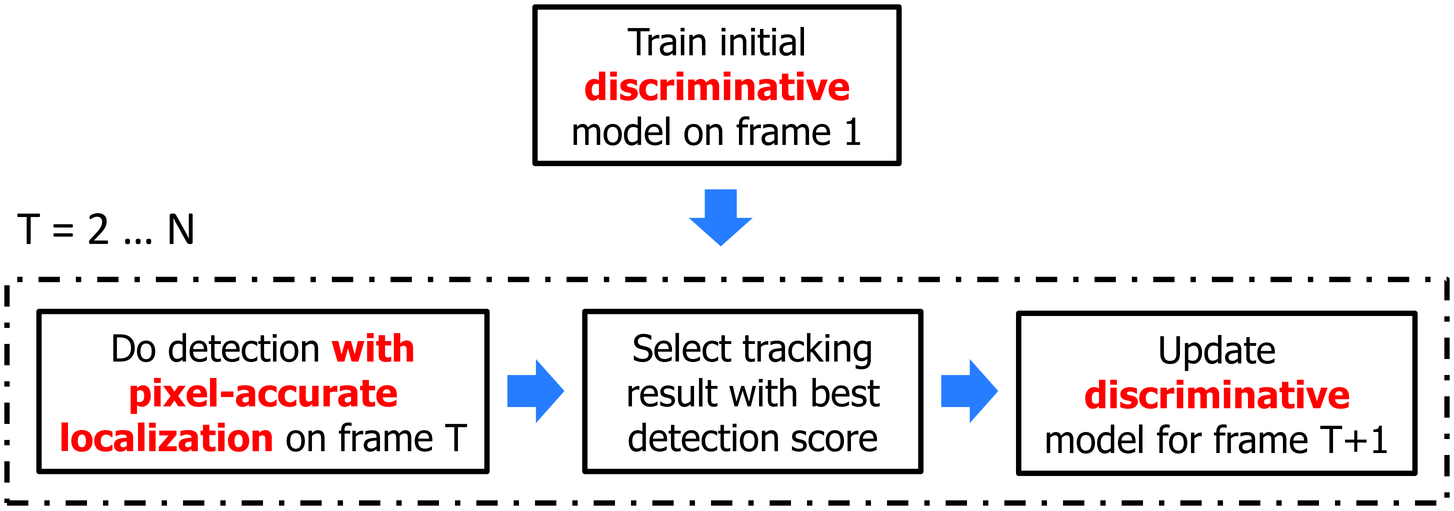
Background

❑ Two key ingredients

- Discriminative learning
- Pixel-accurate localization

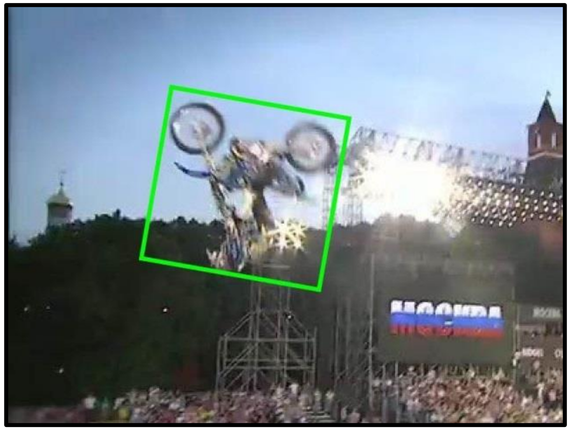


- Two key ingredients
  - Discriminative learning
  - Pixel-accurate localization

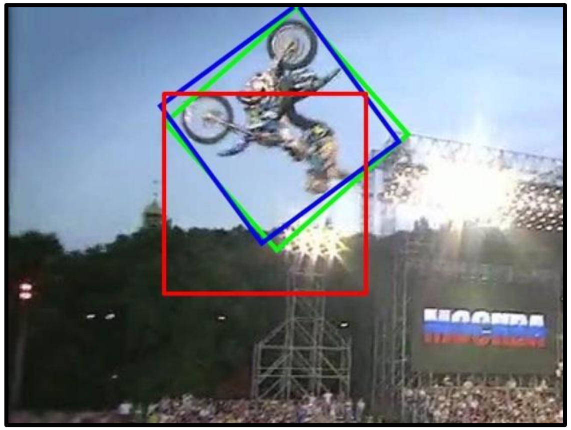


- Limitations of tracking-by-detection approaches
  - Can not handle challenging conditions where an object undergoes transformations, e.g., severe rotation

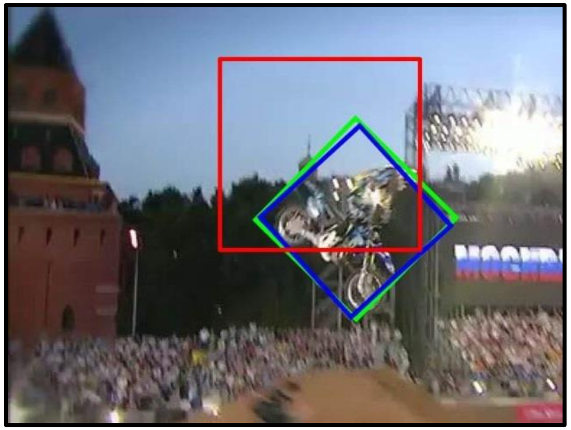
DSST Ours Groundtruth



Frame 1



Frame 10



Frame 30

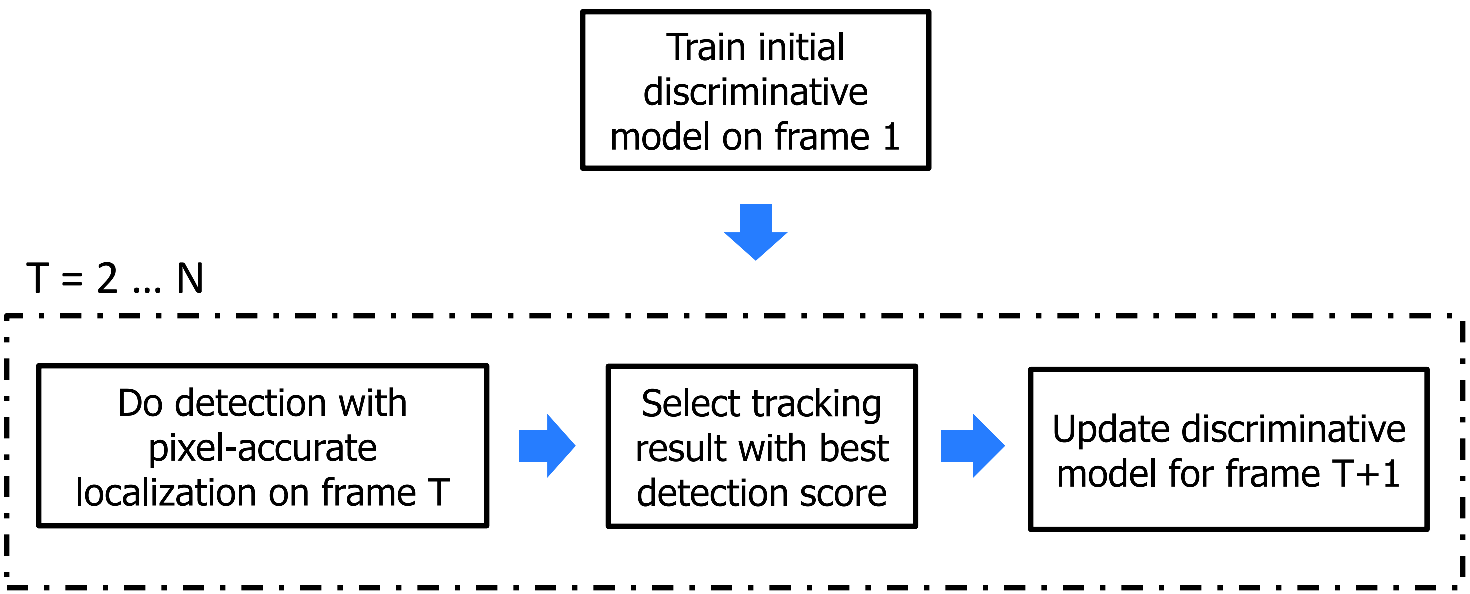
Background

❑ Limitations of tracking-by-detection approaches

- Can not handle challenging conditions where an object undergoes transformations, e.g., severe rotation
- Select tracking results based on detection score only

Our approach: Overview

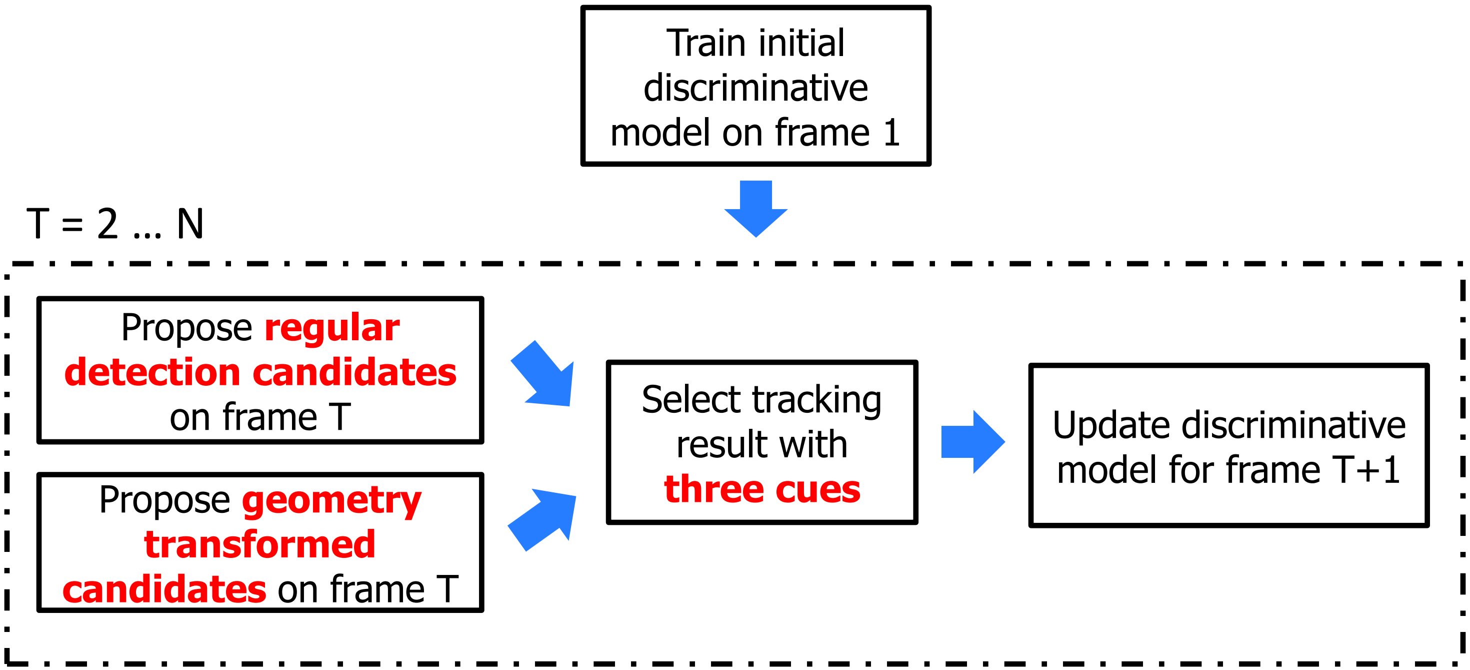
- ❑ Proposal Selection Tracker [Hua et al., 2015]



Y. Hua, K. Alahari and C. Schmid. Online Object Tracking with Proposal Selection. In ICCV, 2015

Our approach: Overview

❑ Proposal Selection Tracker [Hua et al., 2015]

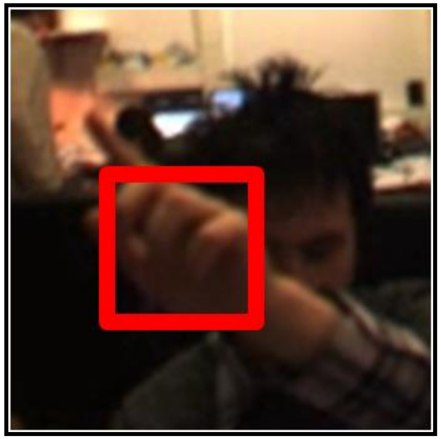


Y. Hua, K. Alahari and C. Schmid. Online Object Tracking with Proposal Selection. In ICCV, 2015

Our approach: Proposal Selection Tracker

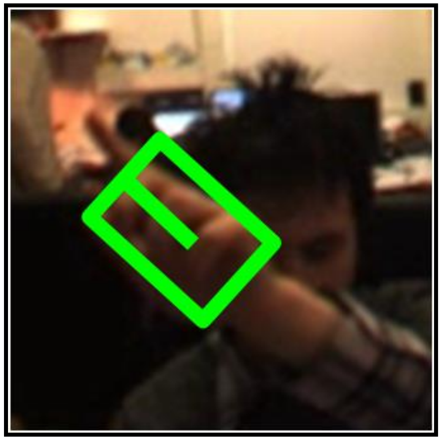
❑ Geometry proposal

- Compute frame-to-frame pixel correspondences with optical flow [Brox and Malik, 2011]
- Estimate similarity transformations with a Hough transform voting scheme



Detection proposal

Geometry proposal

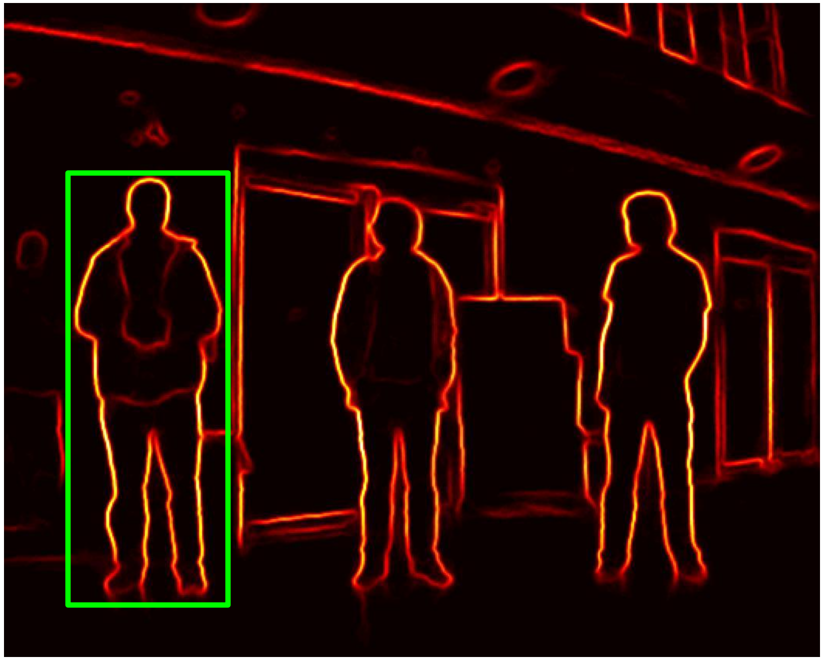


Ground truth

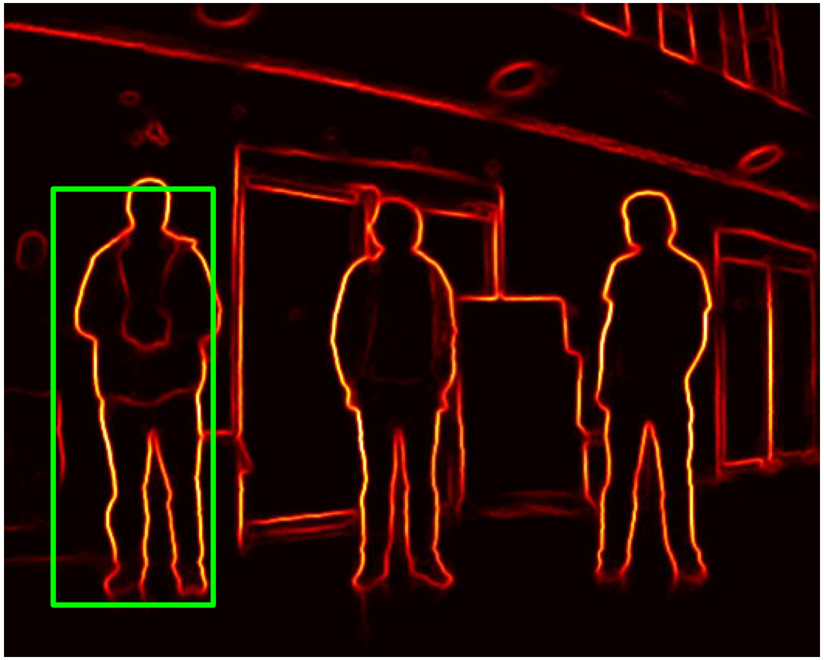
T. Brox and J. Malik. Large displacement optical flow: Descriptor matching in variational motion estimation. PAMI, 2011.

Our approach: Proposal Selection Tracker

- Multiple cues for selection
  - Detection scores
  - Edgebox score [Zitnick and Dollár, 2014], originally from edge response [Dollár and Zitnick, 2013]



High edgebox score



Low edgebox score

C. L. Zitnick and P. Dollár. Edge boxes: Locating object proposals from edges. In ECCV, 2014.
P. Dollár and C. L. Zitnick. Structured forests for fast edge detection. In ICCV, 2013.

Our approach: Proposal Selection Tracker

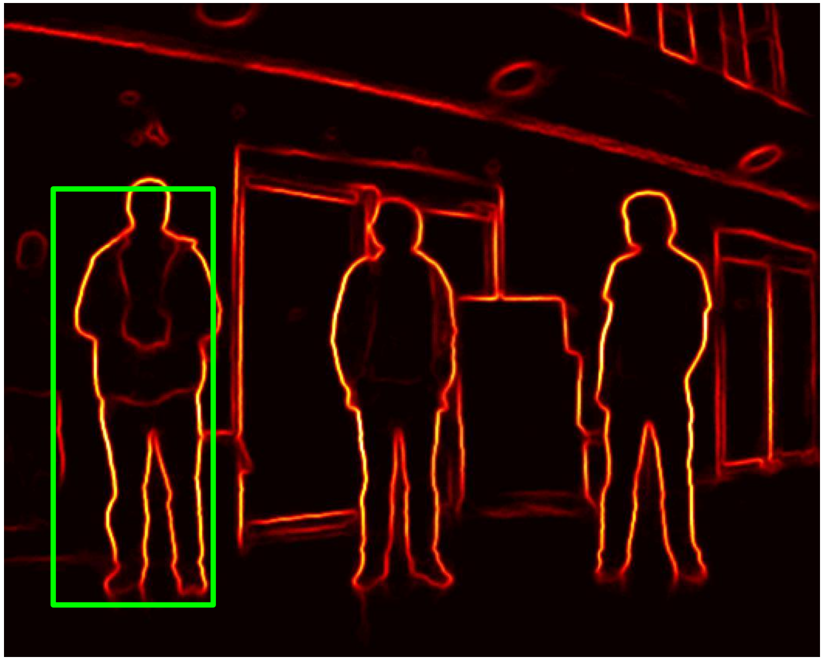
- Multiple cues for selection
  - Edgebox scores from edge responses and motion boundaries [Weinzaepfel et al., 2015] are complementary

Edge Responses Motion boundaries

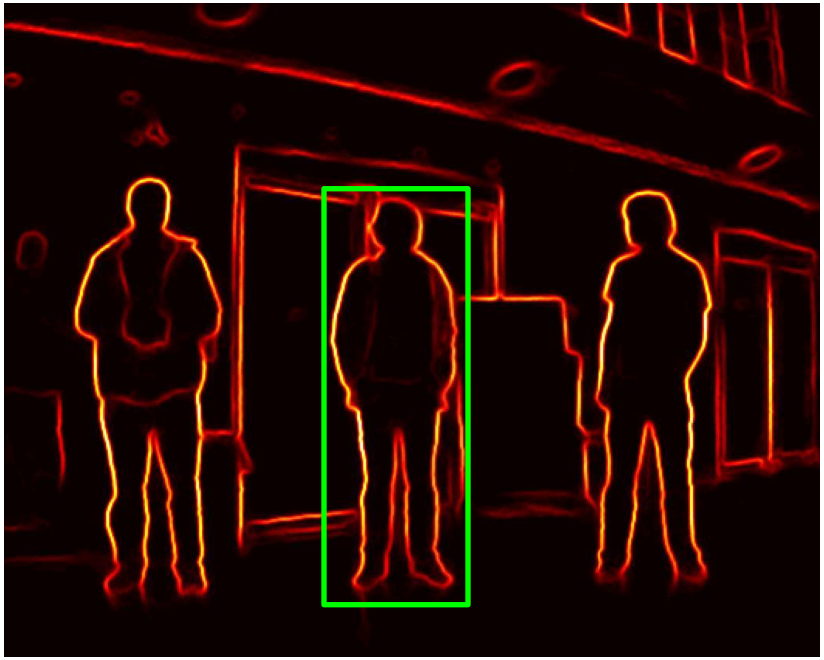
P. Weinzaepfel, J. Revaud, Z. Harchaoui, and C. Schmid. Learning to detect motion boundaries. In CVPR, 2015.

Our approach: Proposal Selection Tracker

- How to combine multiple cues?
  - When detection scores of the top candidates are very similar, we select the one with the best edgebox measure



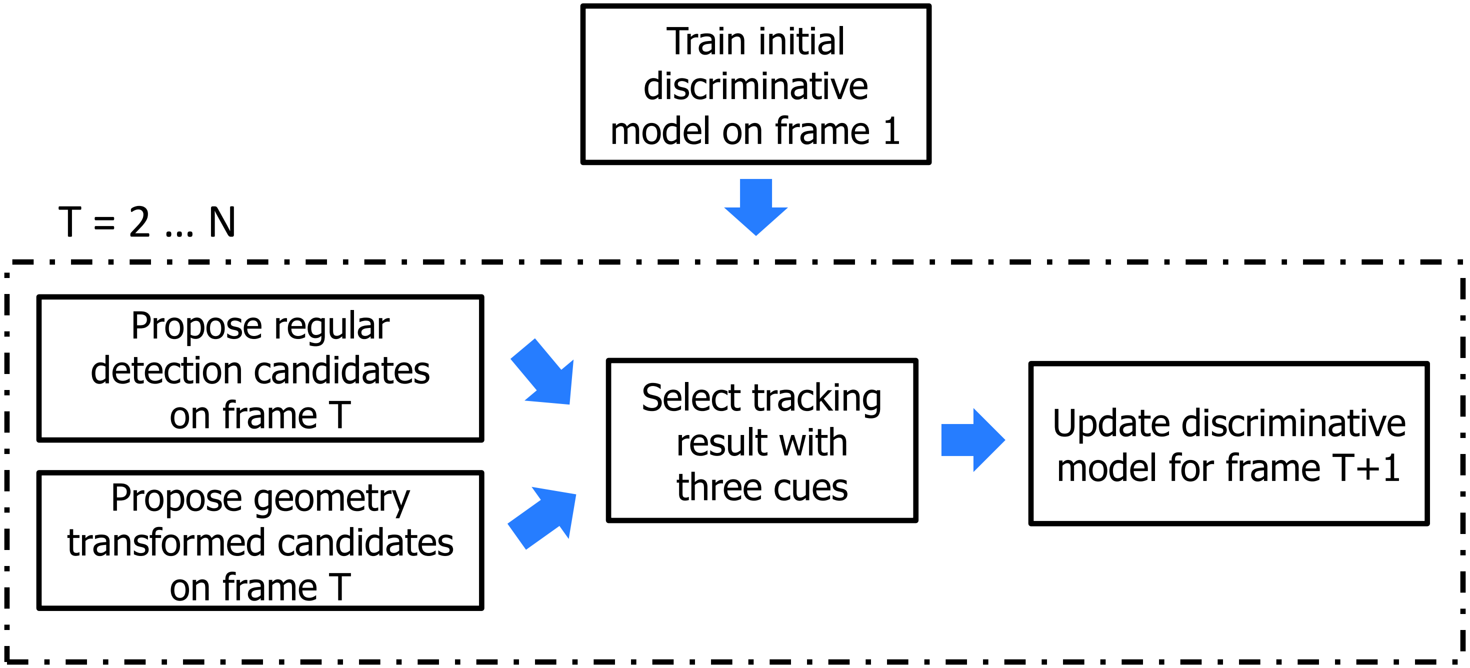
Low edgebox score, but high detection score



High edgebox score, but low detection score

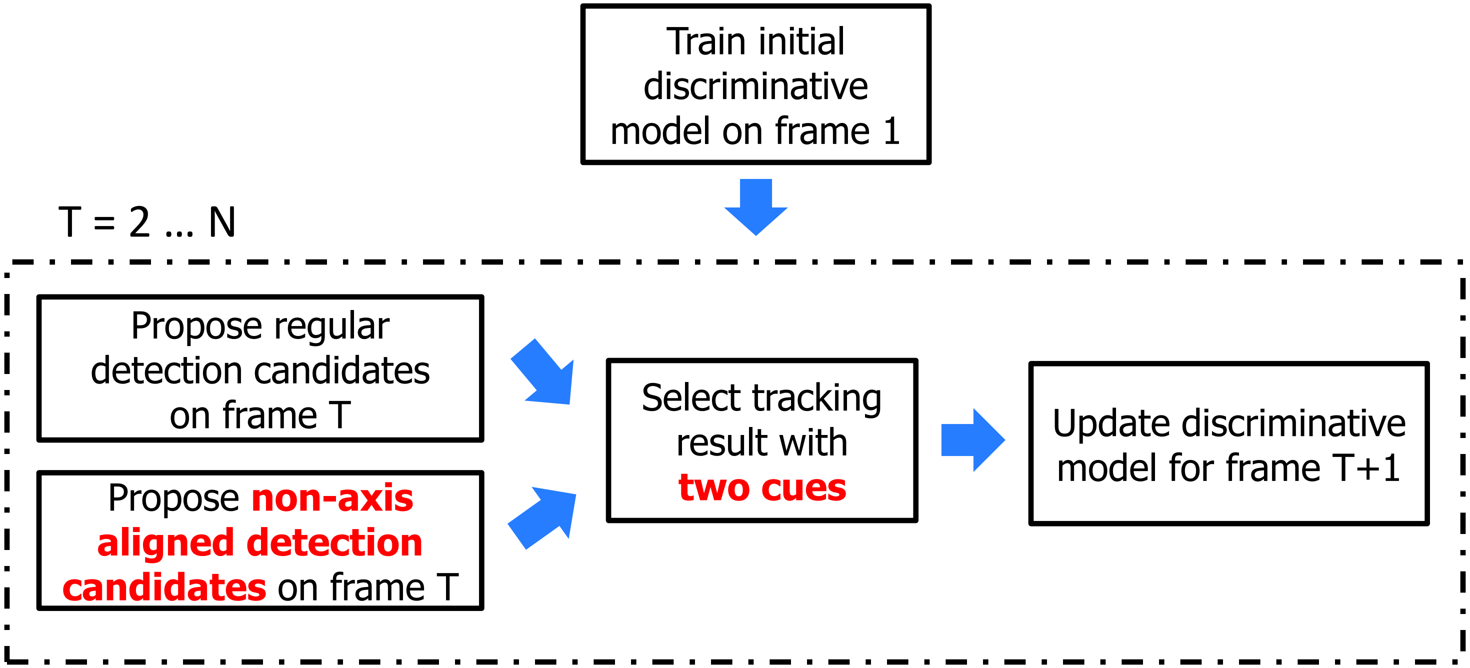
Our approach: Overview

- ❑ Proposal Selection Tracker
- ❑ Our submission



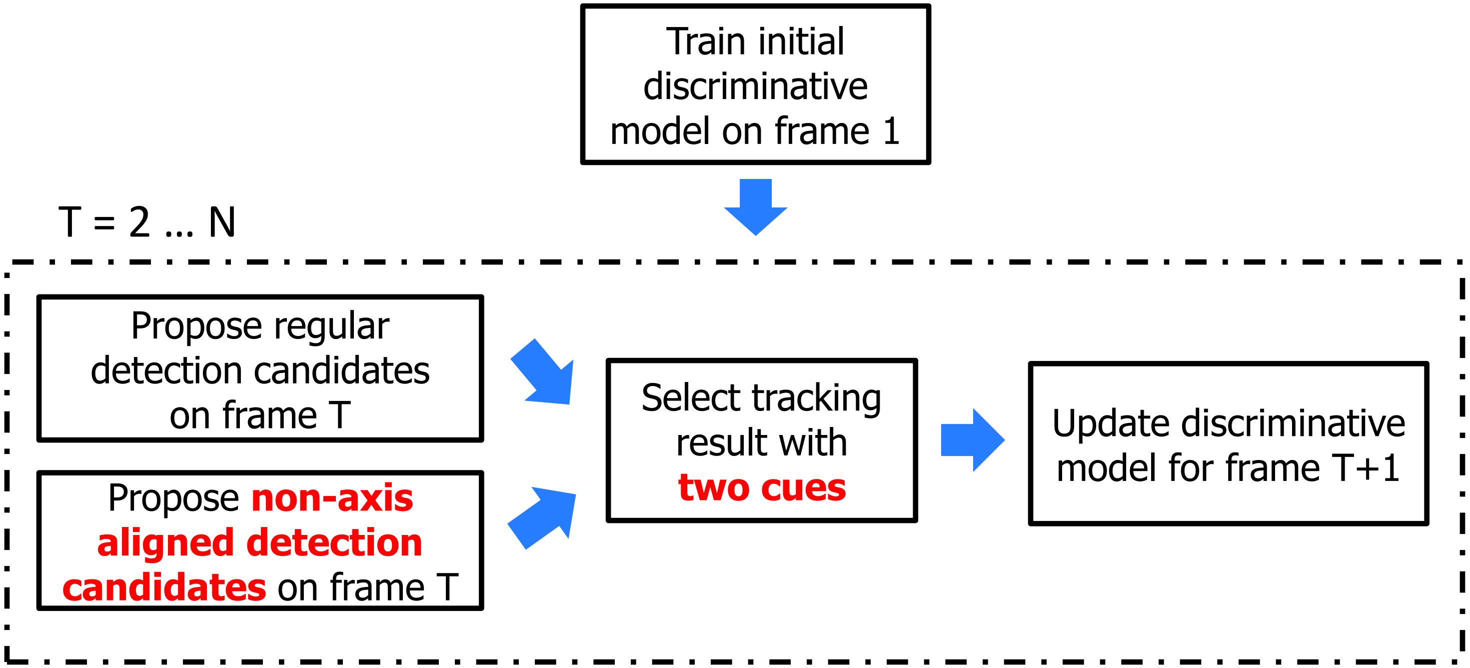
Our approach: Overview

- Proposal Selection Tracker
- Our submission
  - A simplified Proposal Selection Tracker without optical flow calculation



Our approach: Overview

- Proposal Selection Tracker
- Our submission
  - A simplified Proposal Selection Tracker without optical flow calculation
  - Same parameters for both VOT-TIR2015 and VOT2015 challenges



Our approach: Implementation details

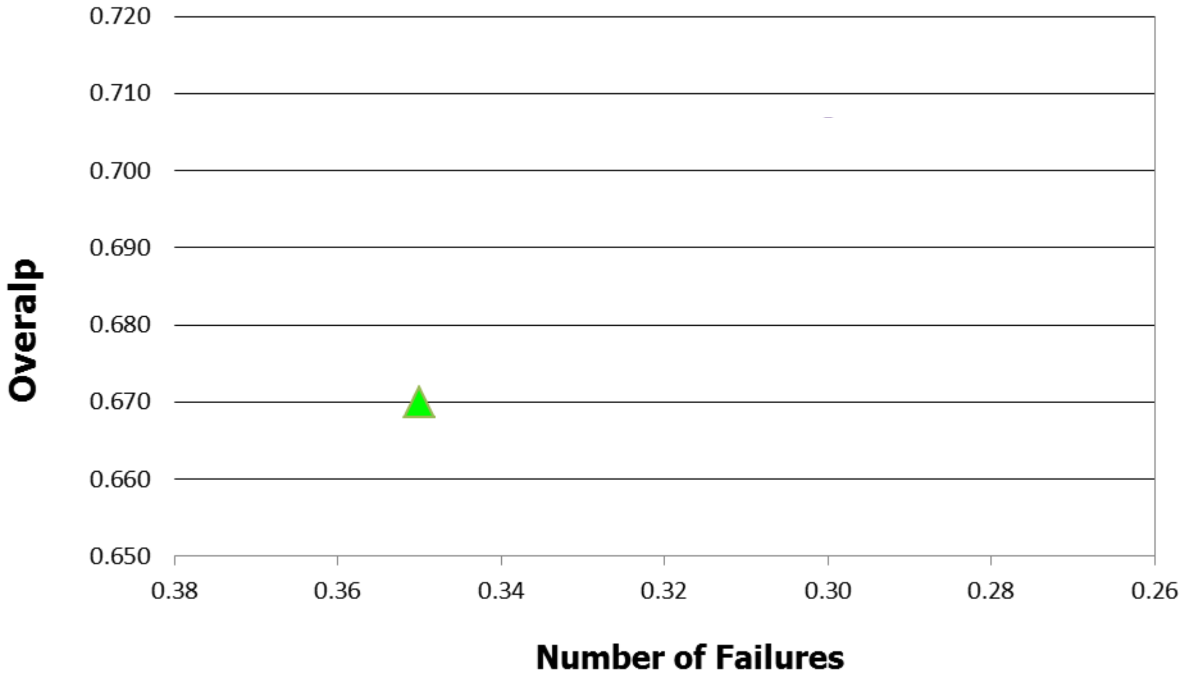
- Baseline tracking-by-detection framework
  - HOG + Linear SVM [Supancic and Ramanan, 2013]
  - Multi-scale detector with 2 pixels scanning step
- Small patch (e.g. 18x18) handling
  - Additional correlation checking
- Occlusion handling
  - Selective model updating

J.S. Supancic and D. Ramanan. Self-paced learning for long-term tracking. In CVPR, 2013.

Experimental results

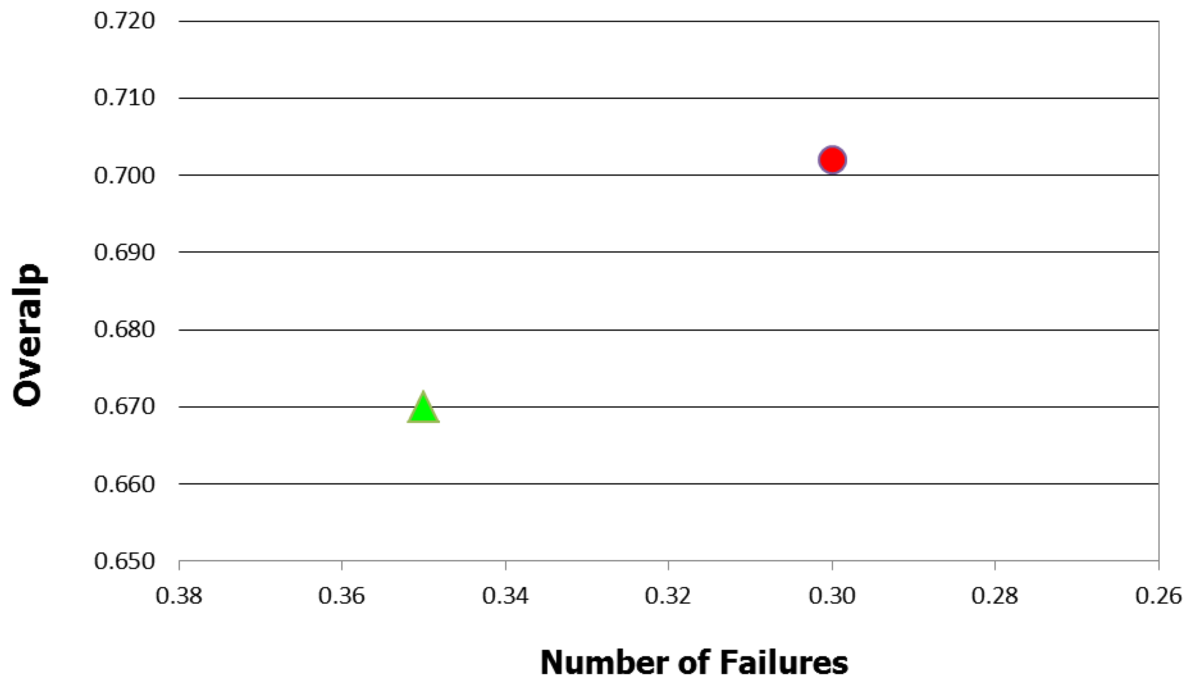
❑ VOT-TIR2015 Challenge Dataset

▲ Selection: detection score only	
Overlap	#Failures
0.670	0.35



❑ VOT-TIR2015 Challenge Dataset

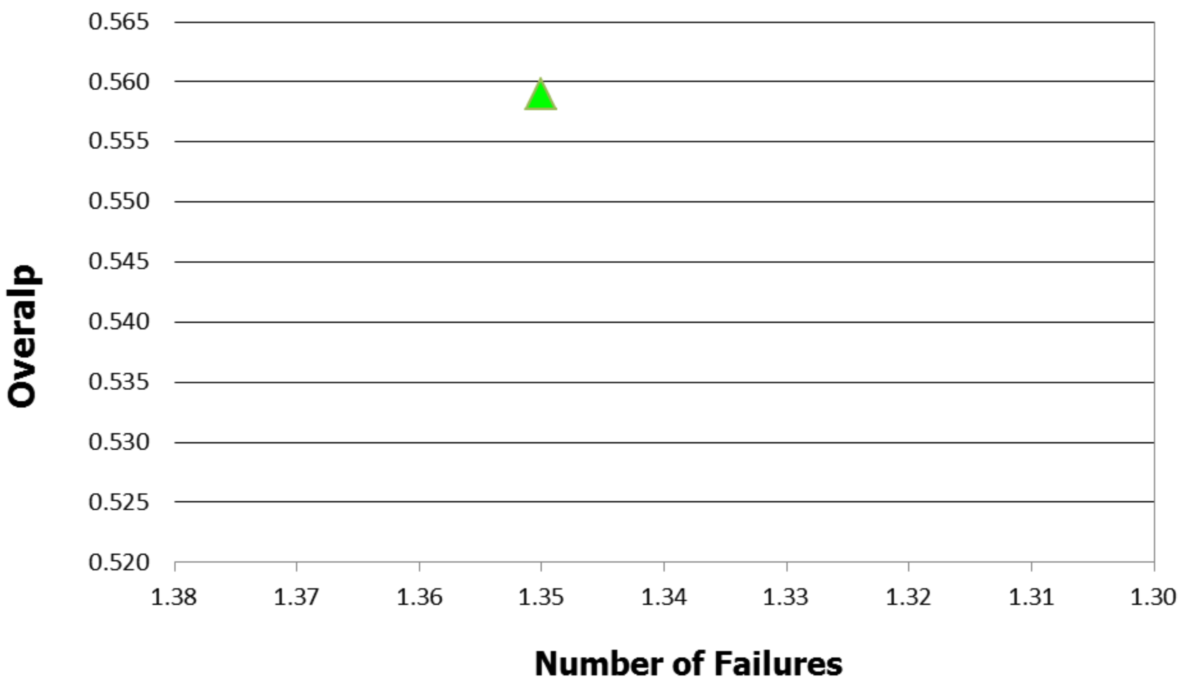
<table>
<tr><th colspan="2">Selection: detection score only</th><th colspan="2">Selection: detection + edgebox score</th></tr>
<tr><td>Overlap</td><td>#Failures</td><td>Overlap</td><td>#Failures</td></tr>
<tr><td>0.670</td><td>0.35</td><td>0.702</td><td>0.30</td></tr>
</table>



Experimental results

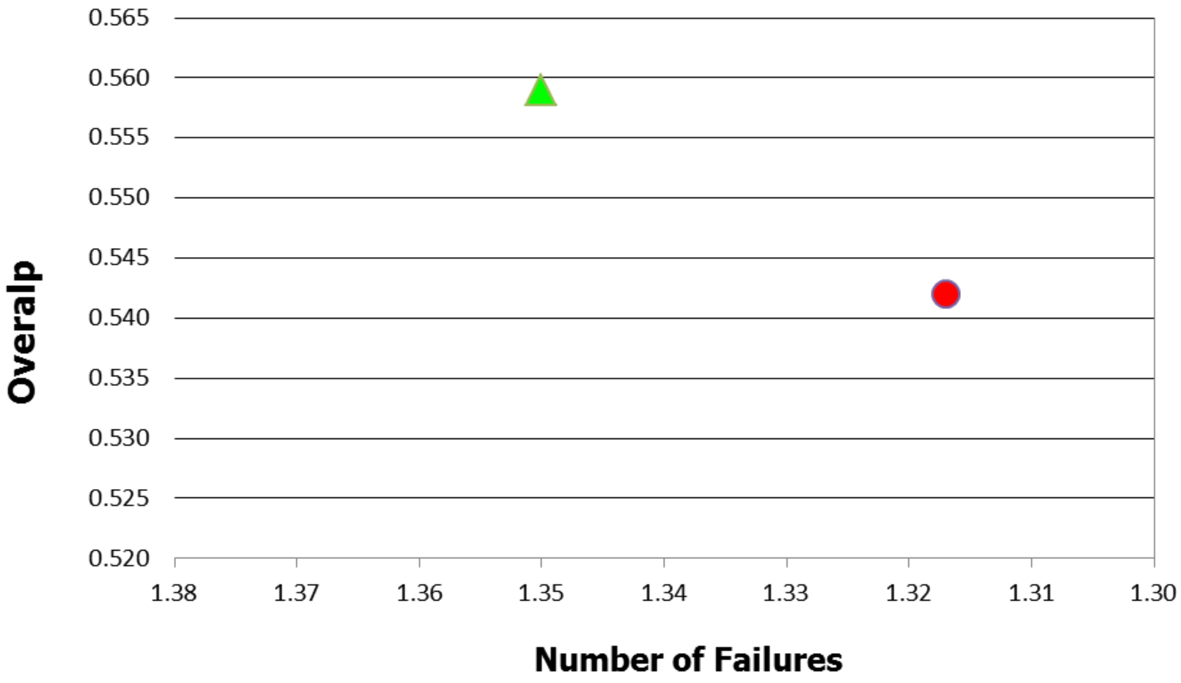
❑ VOT2015 Challenge Dataset

▲ Selection: detection score only	
Overlap	#Failures
0.559	1.35



- VOT2015 Challenge Dataset

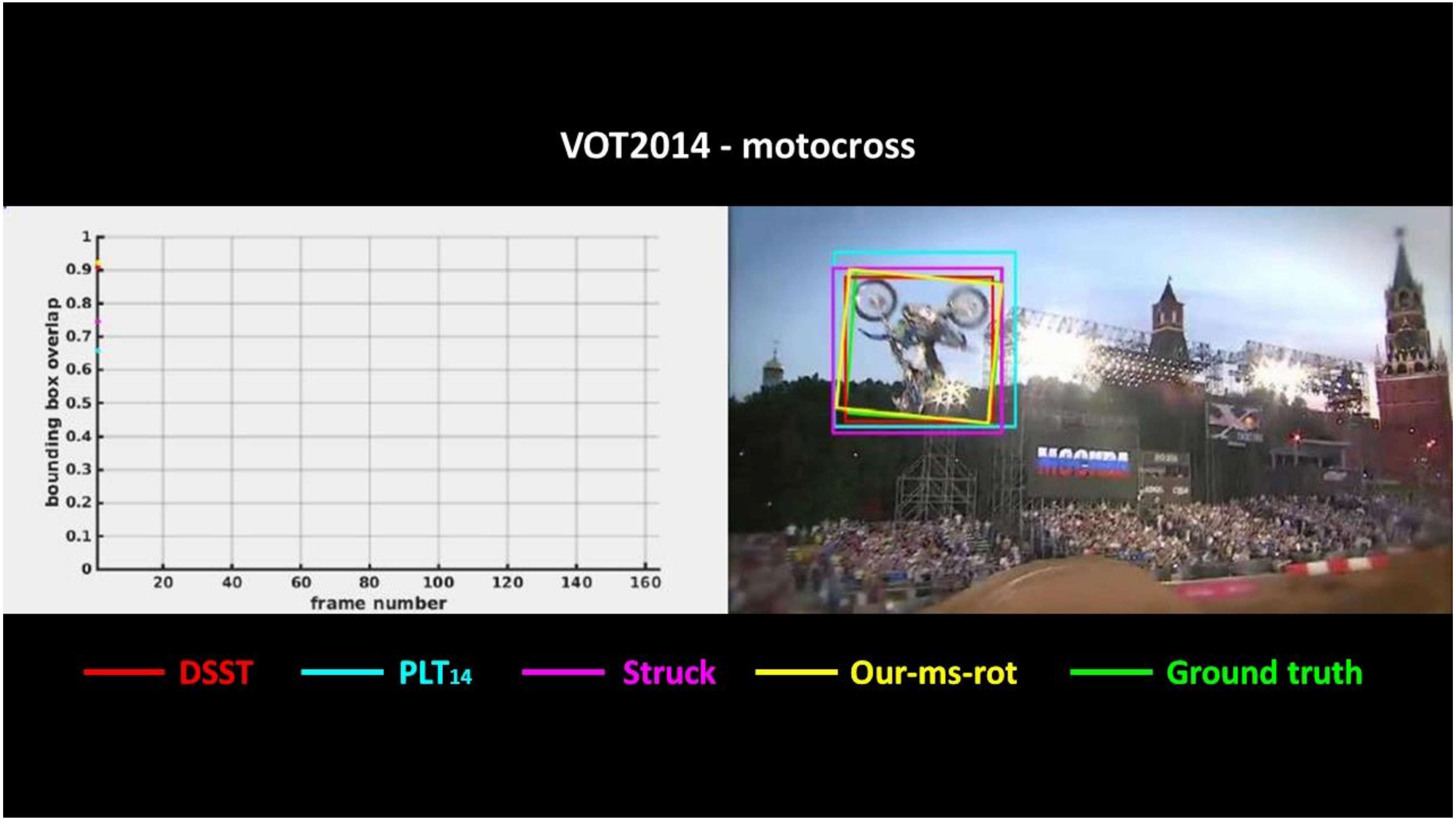
▲ Selection: detection score only		● Selection: detection + edgebox score	
Overlap	#Failures	Overlap	#Failures
0.559	1.35	0.542	**1.32**



Summary

- Extending tracking-by-detection as a general proposal and selection scheme
  - New geometry proposals
  - A novel selection scheme based on multiple cues
- Achieving good performance on VOTTIR-2015 and VOT2015 challenge datasets
- Source code is released at project page
  - http://lear.inrialpes.fr/research/pstracker/

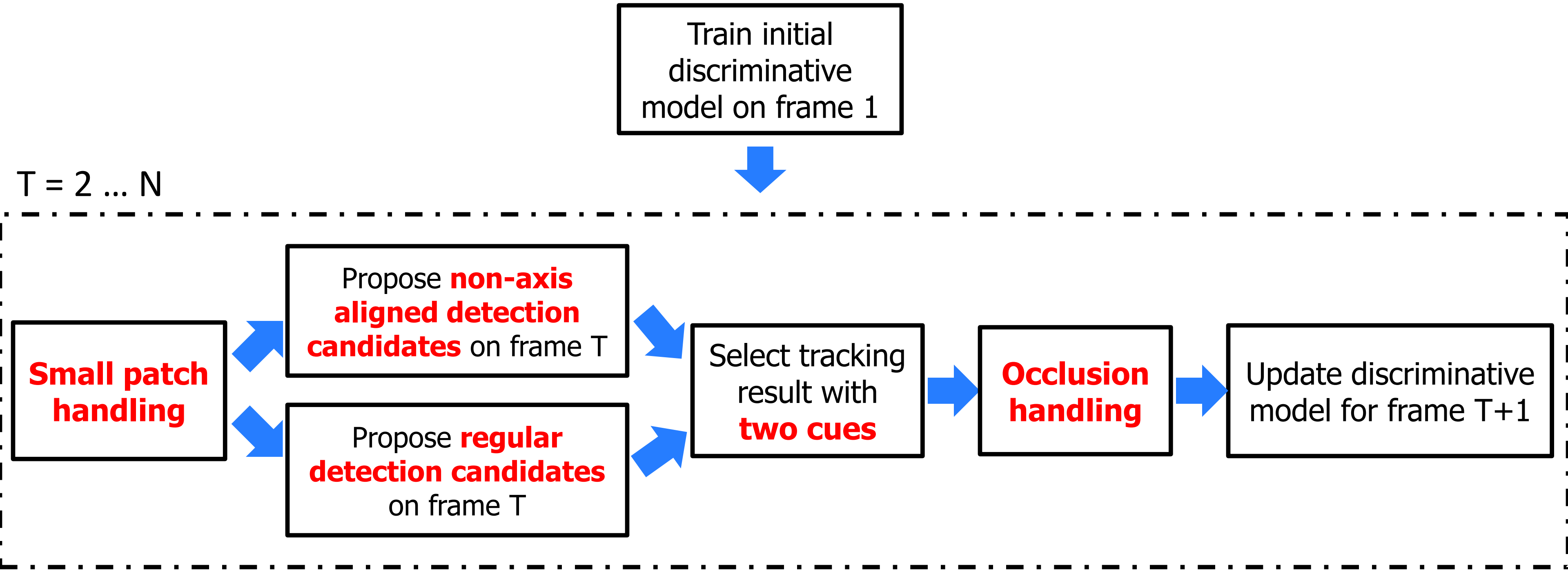
- ❑ Proposal Selection Tracker will be presented at Poster Session 3B (Tuesday, 15 Dec. 2015)



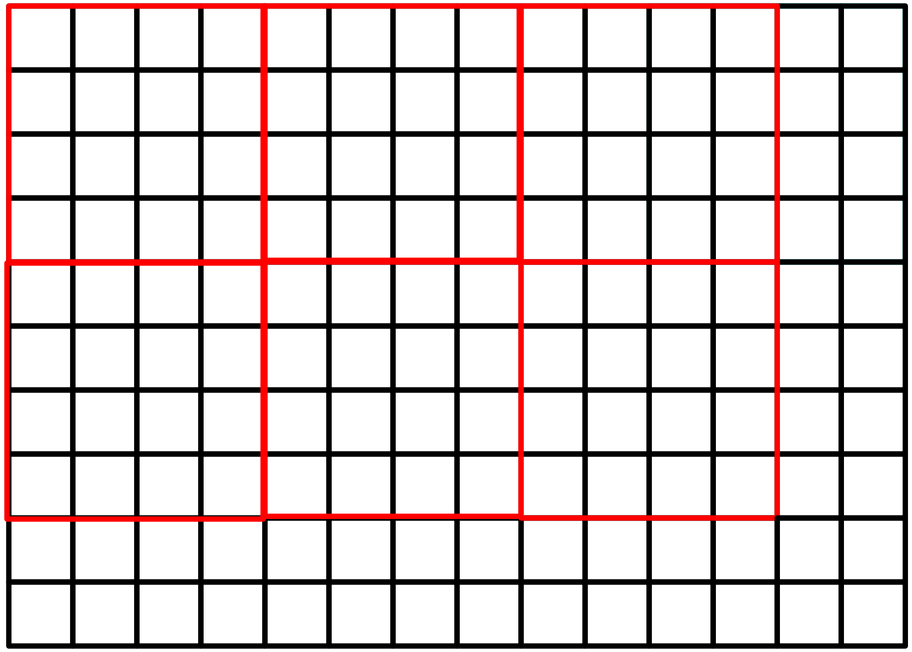
Appendix: List

Appendix: Overall framework

- sPST: overall framework

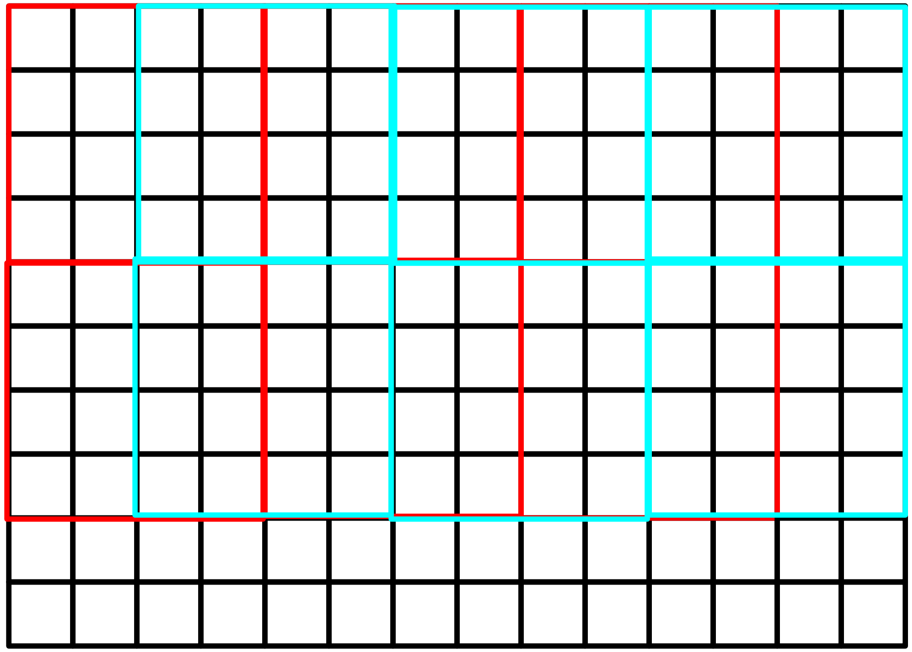


- Dense HOG



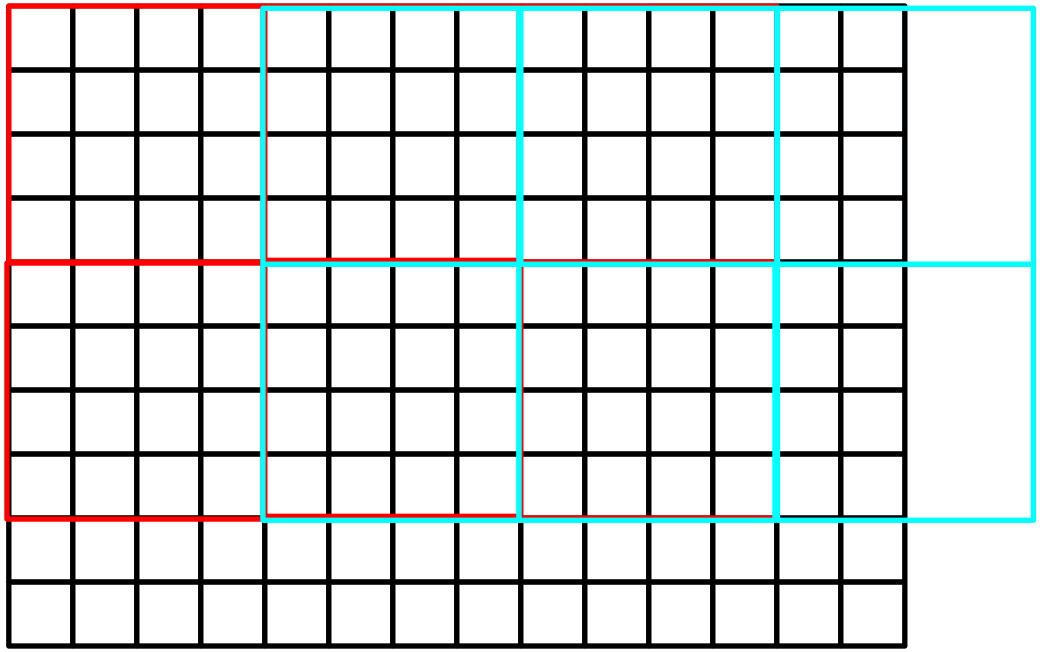
Appendix: Dense HOG

- Dense HOG



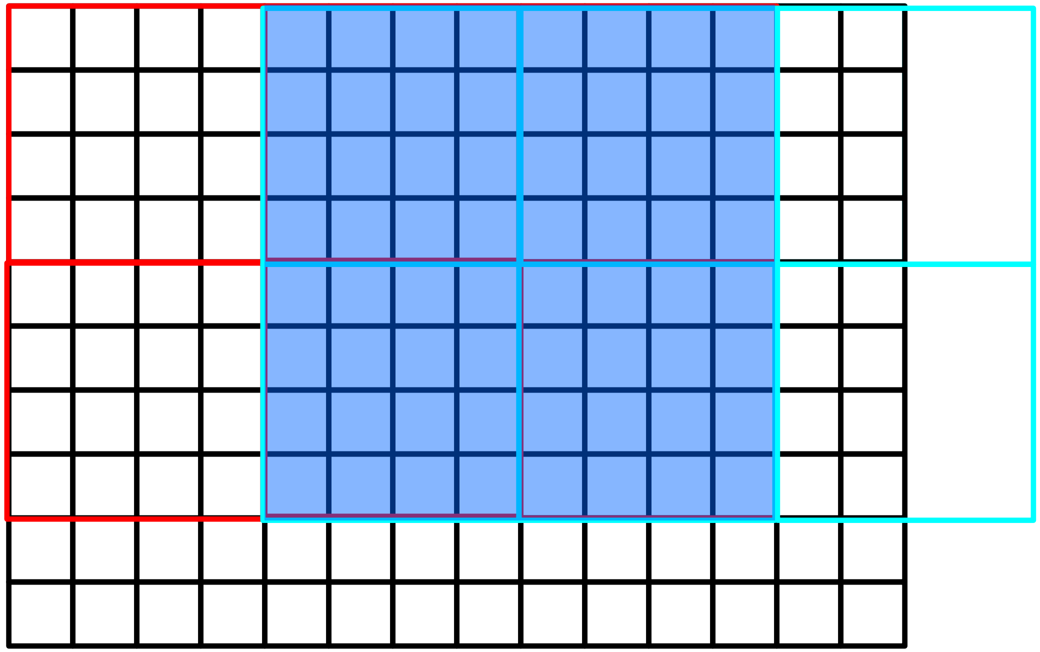
Appendix: Dense HOG

- Dense HOG



Appendix: Dense HOG

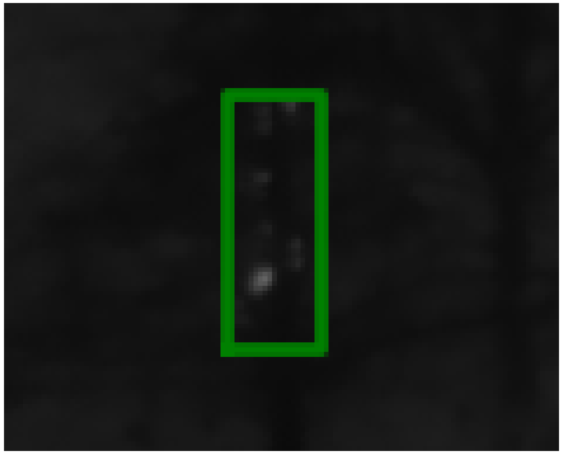
❑ Dense HOG



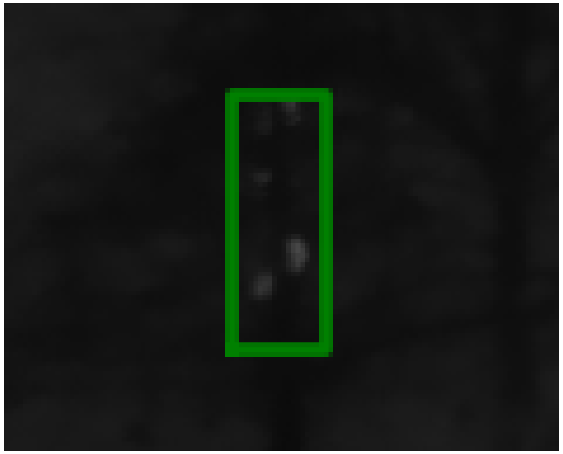
Appendix: Occlusion handling

❑ Occlusion handling

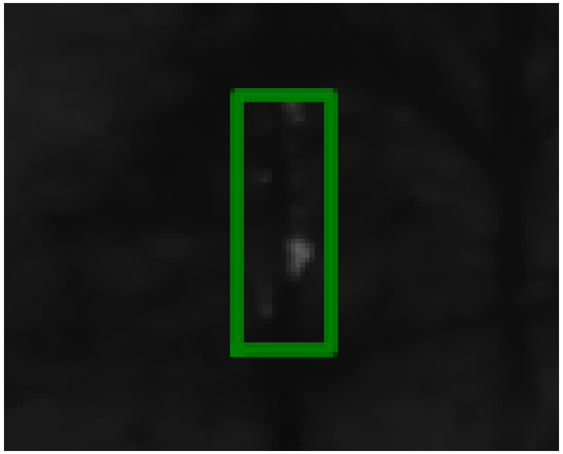
- For short-term tracking, we update model with tracking result in every frame
- Selective model updating: If tracking result in current frame is quite similar with tracking result in the previous frame, probably occlusion happens, we don't update model in current frame



Frame 246



Frame 247

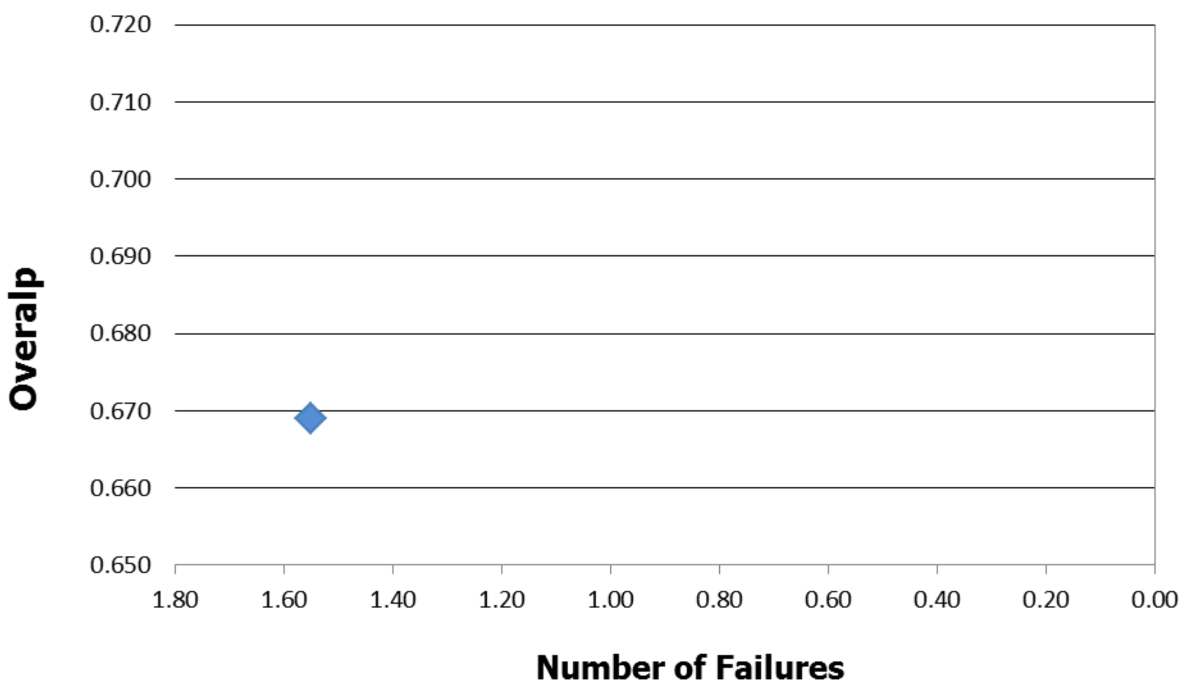


Frame 248

Experimental results

❑ VOT-TIR2015 Challenge Dataset

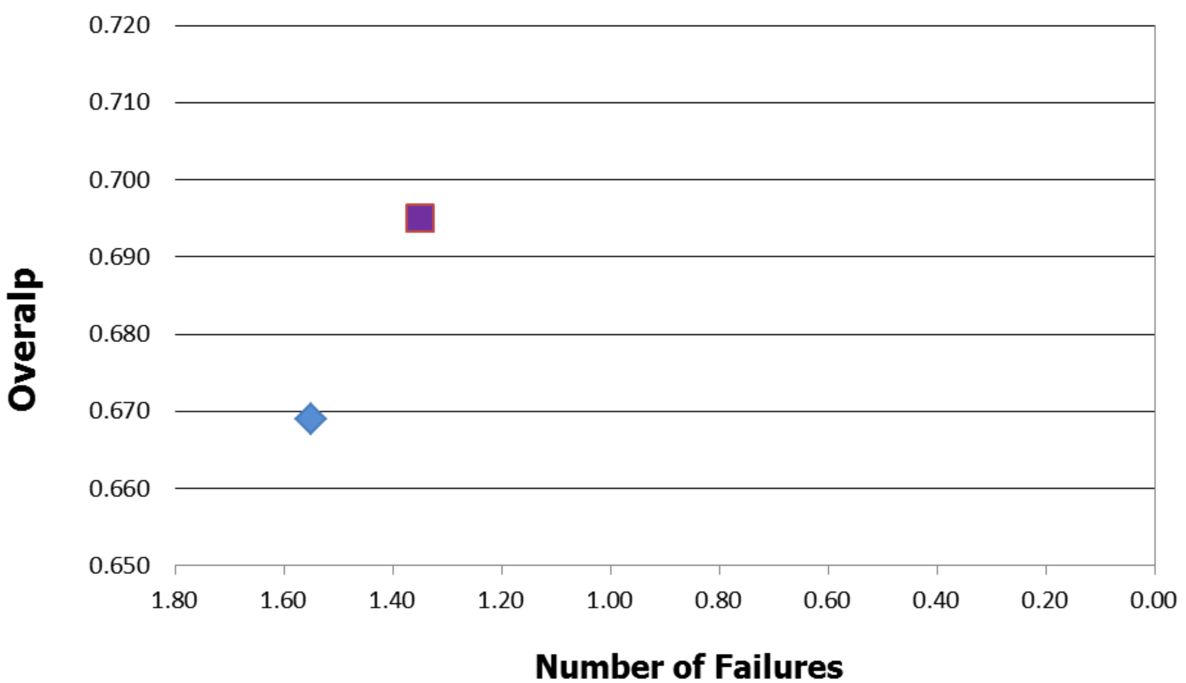
	Selection: detection score only		
	Marker	Overlap	#Failures
Baseline	◆	0.669	1.55



Experimental results

❑ VOT-TIR2015 Challenge Dataset

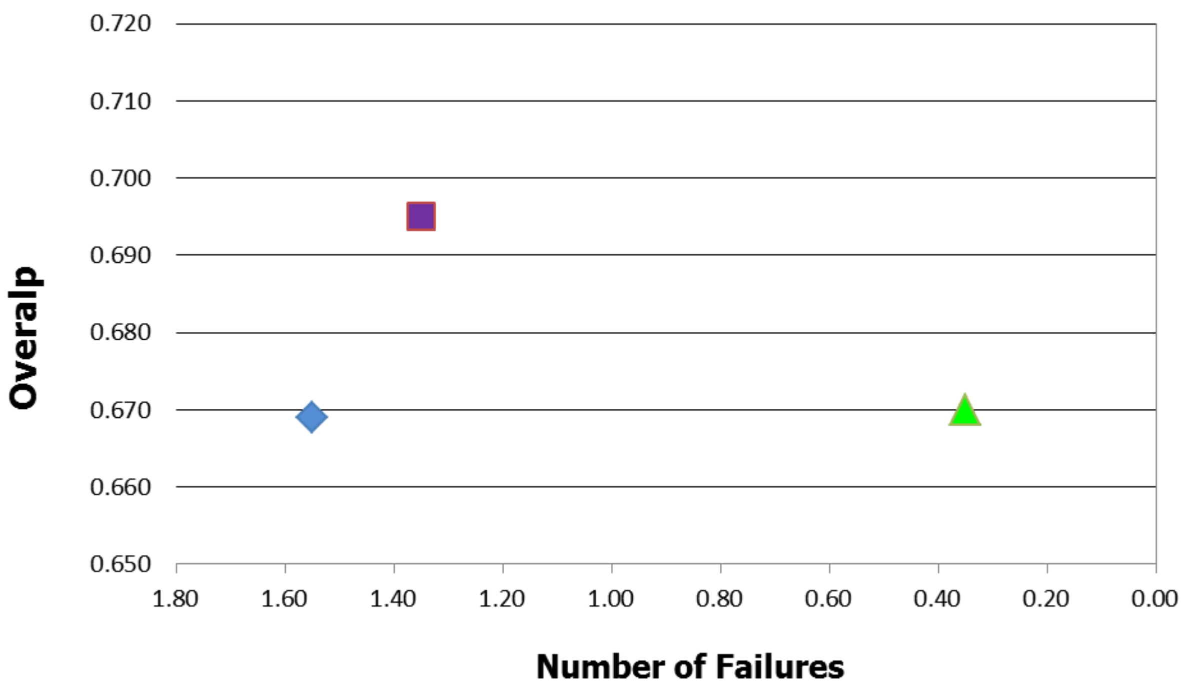
	Selection: detection score only			Selection: detection + edgebox score		
	Marker	Overlap	#Failures	Marker	Overlap	#Failures
Baseline		0.669	1.55		0.695	1.35



Experimental results

❑ VOT-TIR2015 Challenge Dataset

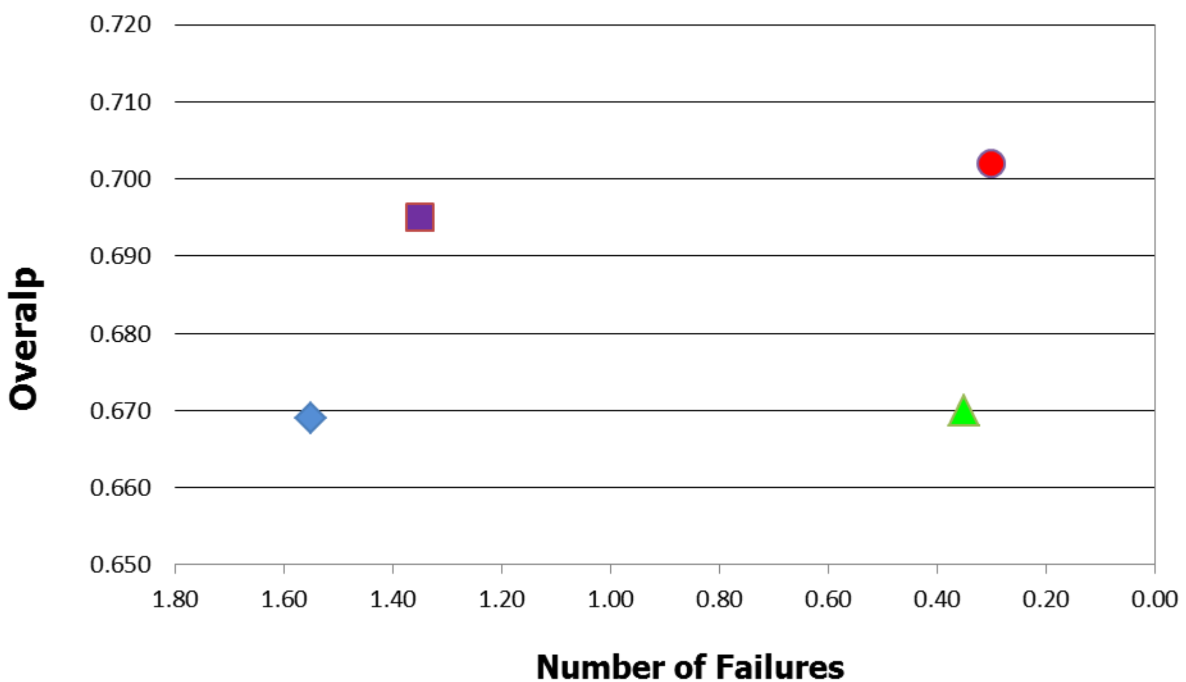
	Selection: detection score only			Selection: detection + edgebox score		
	Marker	Overlap	#Failures	Marker	Overlap	#Failures
Baseline		0.669	1.55		0.695	1.35
Submission		0.670	0.35			



Experimental results

❑ VOT-TIR2015 Challenge Dataset

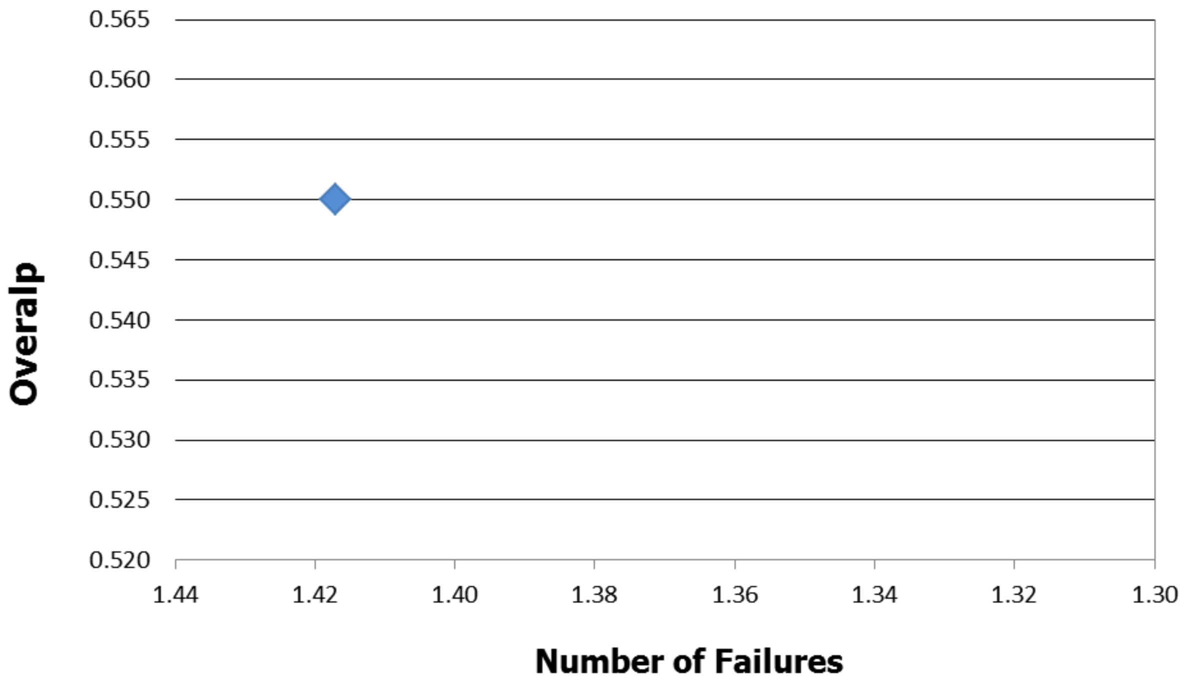
	Selection: detection score only			Selection: detection + edgebox score		
	Marker	Overlap	#Failures	Marker	Overlap	#Failures
Baseline	◆	0.669	1.55	■	0.695	1.35
Submission	▲	0.670	0.35	●	**0.702**	**0.30**



Experimental results

❑ VOT2015 Challenge Dataset

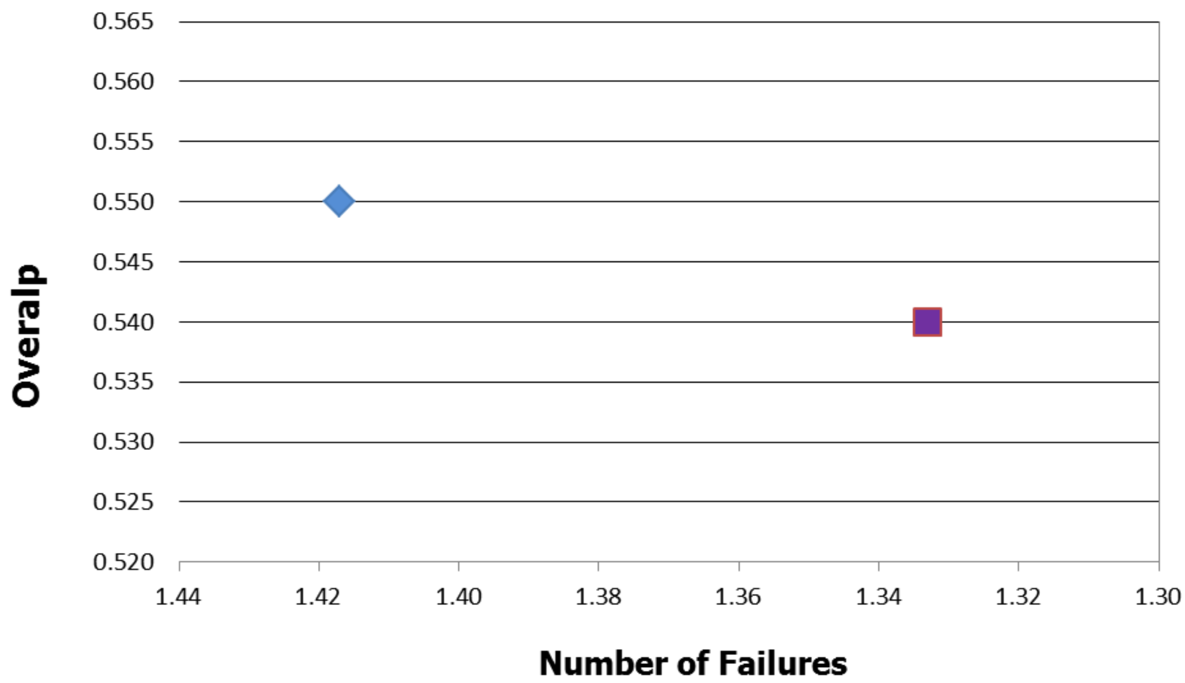
	Selection: detection score only		
	Marker	Overlap	#Failures
Baseline	◆	0.550	1.42



Experimental results

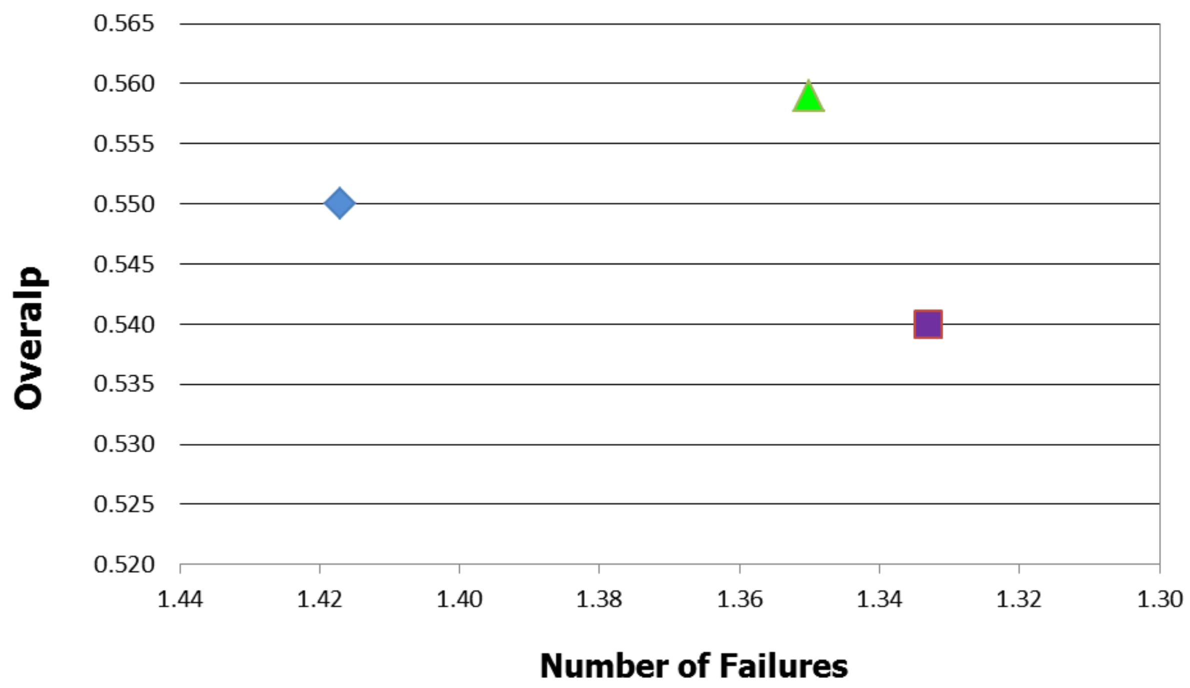
❑ VOT2015 Challenge Dataset

	Selection: detection score only			Selection: detection + edgebox score		
	Marker	Overlap	#Failures	Marker	Overlap	#Failures
Baseline	◆	0.550	1.42	■	0.540	1.33



❑ VOT2015 Challenge Dataset

	Selection: detection score only			Selection: detection + edgebox score		
	Marker	Overlap	#Failures	Marker	Overlap	#Failures
Baseline	◆	0.550	1.42	■	0.540	1.33
Submission	▲	**0.559**	1.35			



Experimental results

❑ VOT2015 Challenge Dataset

	Selection: detection score only			Selection: detection + edgebox score		
	Marker	Overlap	#Failures	Marker	Overlap	#Failures
Baseline	◆	0.550	1.42	■	0.540	1.33
Submission	▲	**0.559**	1.35	●	0.542	**1.32**