Tackling Occlusion in Siamese Tracking with Structured Dropouts

Deepak K. Gupta, Efstratios Gavves and Arnold Smeulders
QUVA Lab and Informatics Institute
University of Amsterdam, The Netherlands

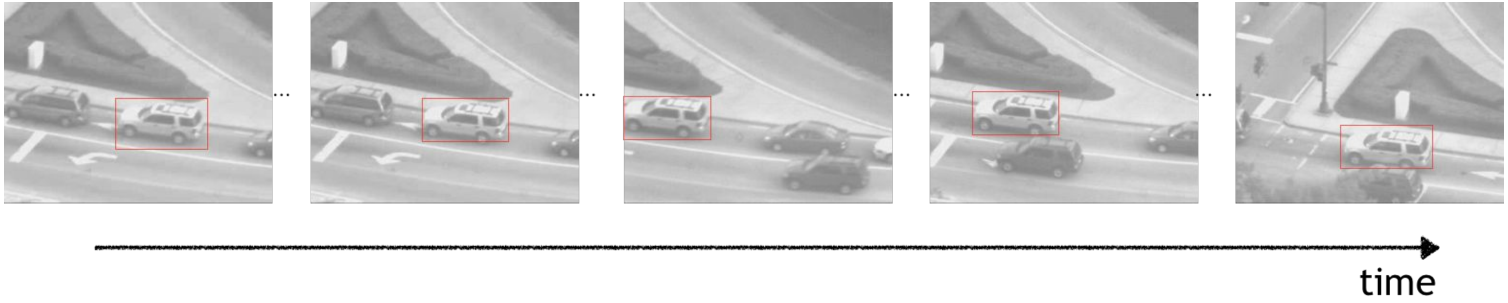
Paper Id: 1590

UNIVERSITY OF AMSTERDAM

QUVA Deep Vision Lab

Visual Object Tracking

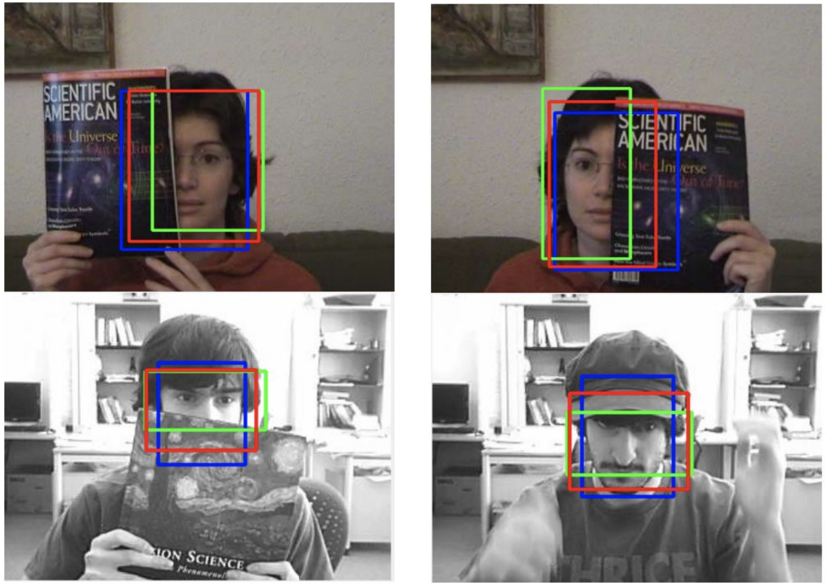
Object Tracking refers to predicting the location of target throughout a video sequence based on a ground-truth provided in only the first frame.



Example video frames encountered in tracking

Several open challenges in tracking include handling occlusion, target rotations, large illumination variations, etc. and we are working towards bringing trackers closer to reality.

Occlusion in tracking

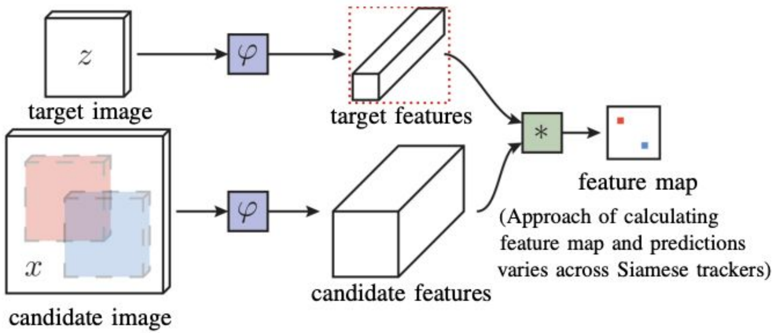


Example images from OTB100 dataset showing ground-truth (blue), predictions from SiamRPN++ (green) and SD-SiamRPN (red).

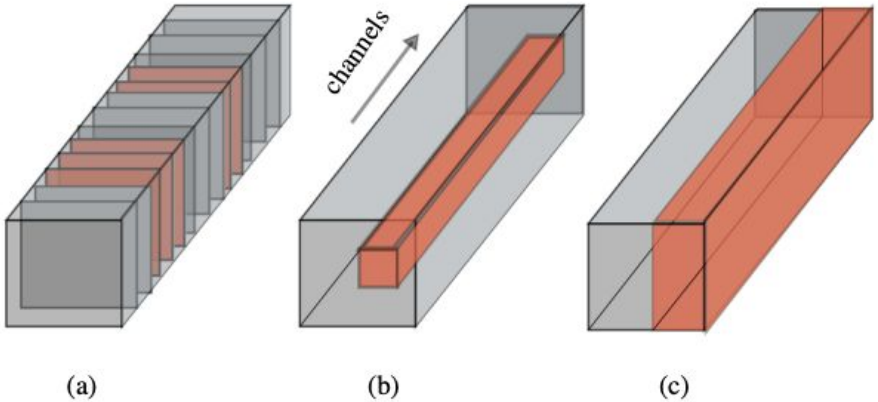
Occlusion cannot be learnt since it does not have a representation - occlusion happens when the representation/information is missing in part of an image.

Can architectural modifications in the tracking model help to tackle occlusion?

Tackling occlusion with structured dropouts



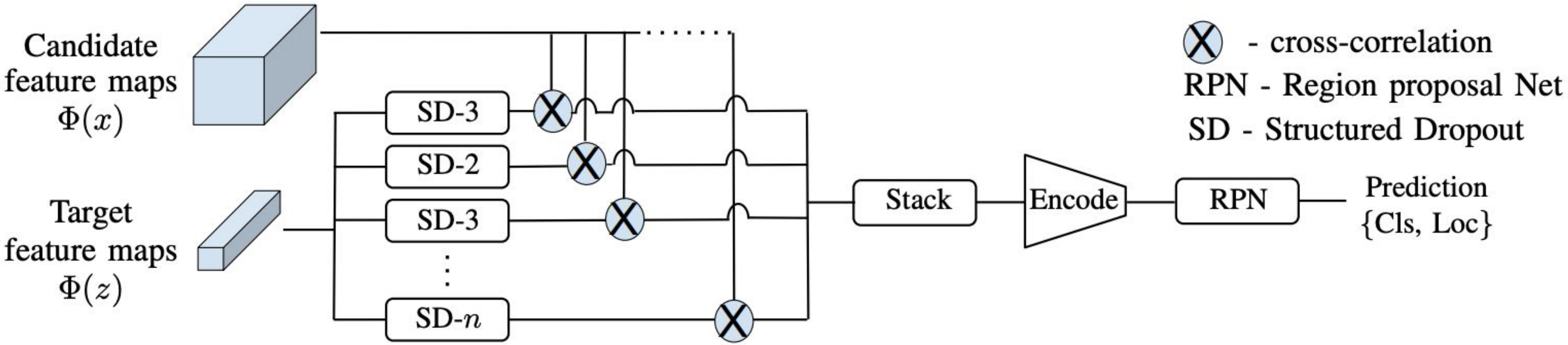
General structure of Siamese trackers



Structured dropouts for sampling against occlusion;
(a) channel (b) segment and (c) slice dropouts.

- Segment - involves dropping randomly a fraction of feature channels from the template feature map.
- Segment - drops a random patch from the feature map across all channels.
- Slice - predefined set of slices are dropped from one of the 4 sides of edges of the feature map - non stochastic.

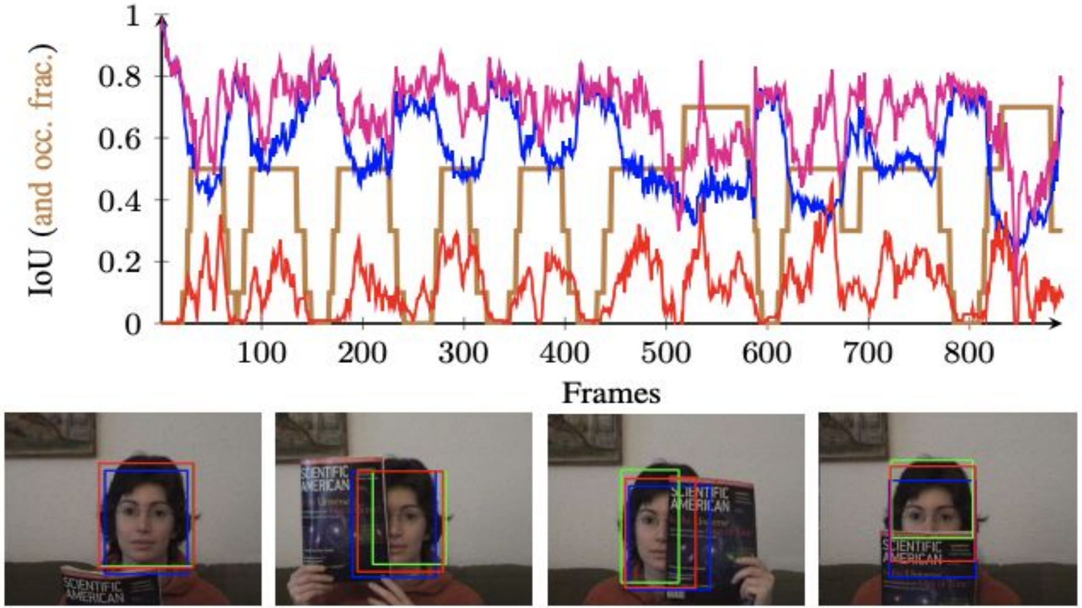
Tracker with structured dropouts



Schematic representation of SiamRPN architecture equipped with structured dropouts.

Tackling occlusion with structured dropouts

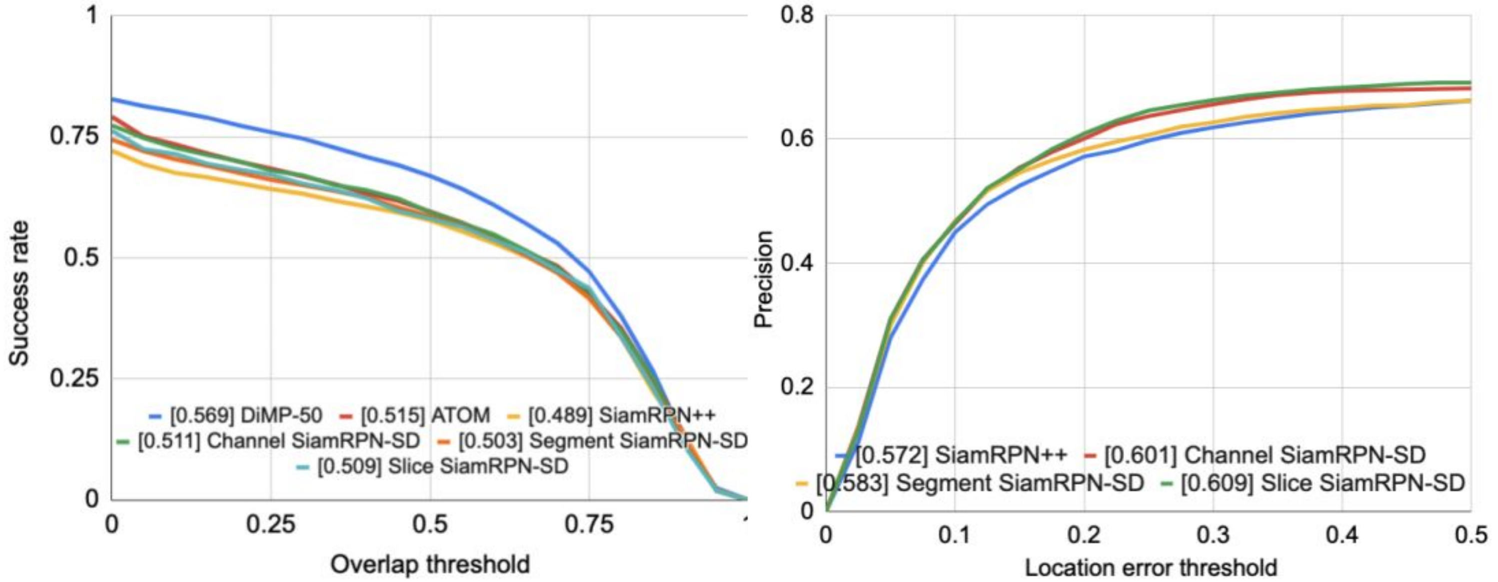
Structured dropouts help in improved localization of the target under occlusion



Top: IoU for predictions without (blue) and with (magenta) structured dropouts, and improvement obtained with structured dropouts (red). Occlusion fraction is denoted by brown. Bottom: examples showing occlusion.

Experiments

Improved performance scores with structured dropouts



Performance plots of OPE for success rate and precision on LaSOT dataset.

Experiments

Approach	OTB2015 Pr	OTB2015 Acc	VOT2018 EAO	VOT2018 Acc
SiamRPN++	0.890	0.683	0.414	0.600
DiMP-50 [5]	-	0.684	0.440	0.597
UPDT [33]	-	0.702	0.378	0.536
ATOM [34]	-	0.669	0.401	0.590
SiamRPN-MC	0.876	0.681	0.417	0.599
exp-SiamRPN-SD-channel	0.908	0.695	0.416	0.591
SiamRPN-SD-channel	0.912	0.702	0.421	0.601
SiamRPN-SD-segment	0.896	0.698	0.410	0.588
SiamRPN-SD-slice	0.914	0.701	0.418	0.598

Precision and accuracy scores on OTB2015 and VOT2018 datasets for SiamRPN++ with and without different dropouts. Results for other SOTA trackers as well are shown.

Conclusions

- Structured dropouts show a promising direction towards tackling occlusion in Siamese trackers.
- For most datasets, adding structured dropouts helped to make trackers more robust in presence of occlusions.
- Further research towards elegant sampling of the dropouts could help in further improvement of tracker performance.