Deep Meta Learning for Real-Time Target-Aware Visual Tracking

Supplementary Material

In this supplementary material, we attach a figure to show additional examples of how the meta-learner can be beneficial in the task of visual tracking, and we provide more success plots on LaSOT dataset by different attributes, and we attach a video file to show more tracking results on the OTB dataset.

1. Effectiveness of Meta-Learner

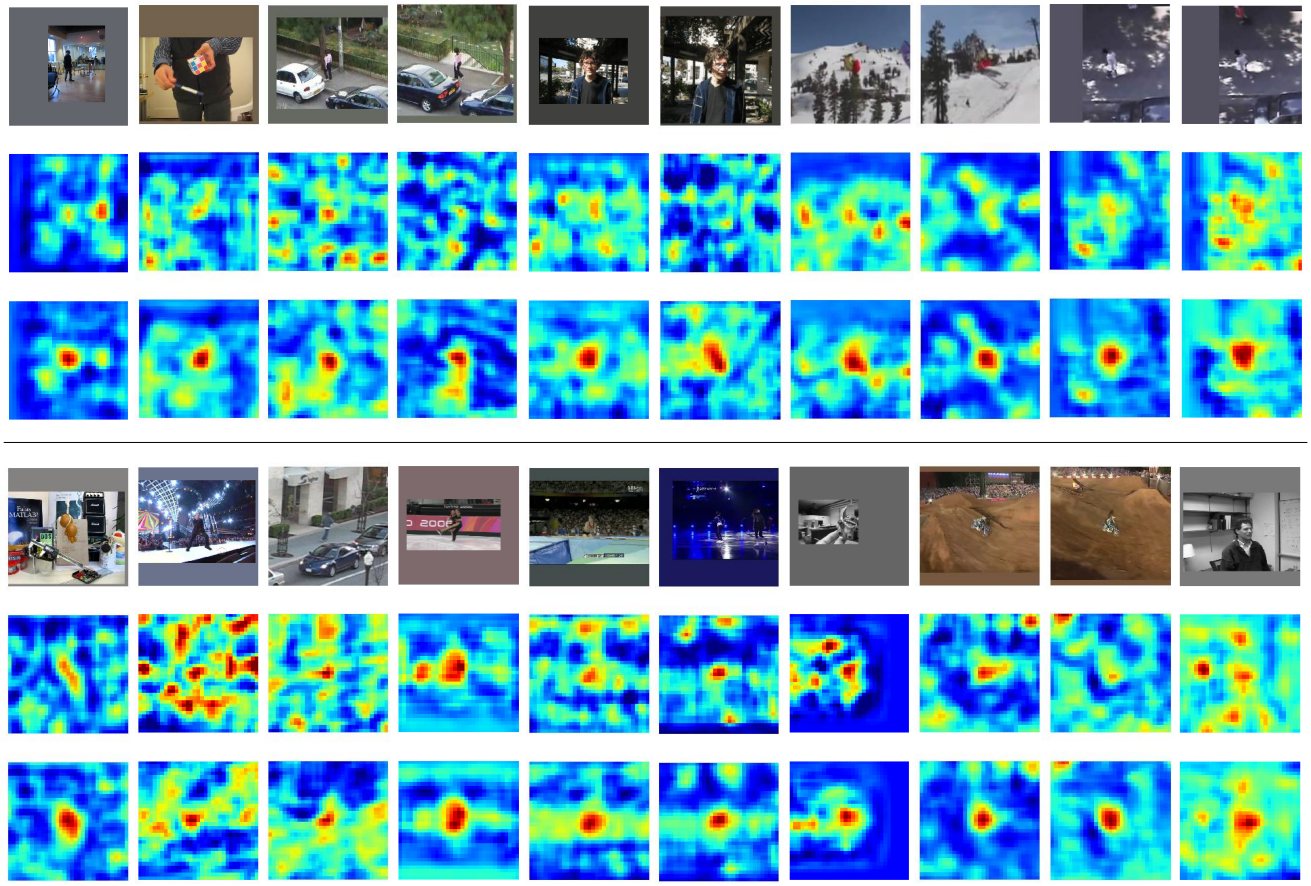
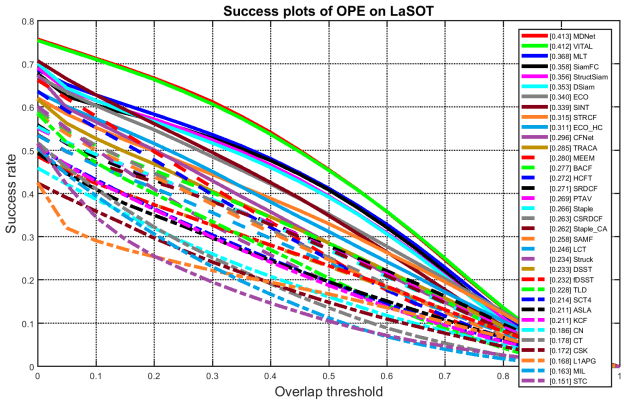
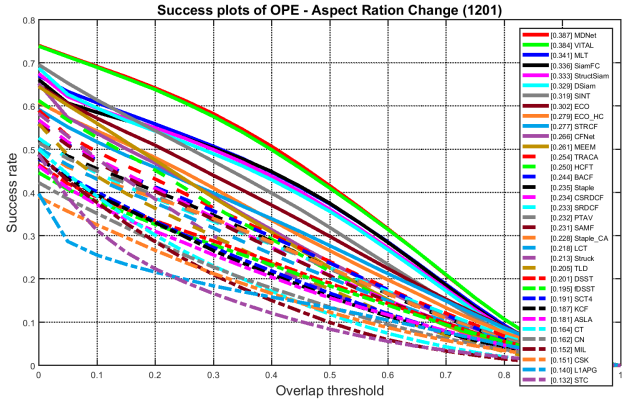


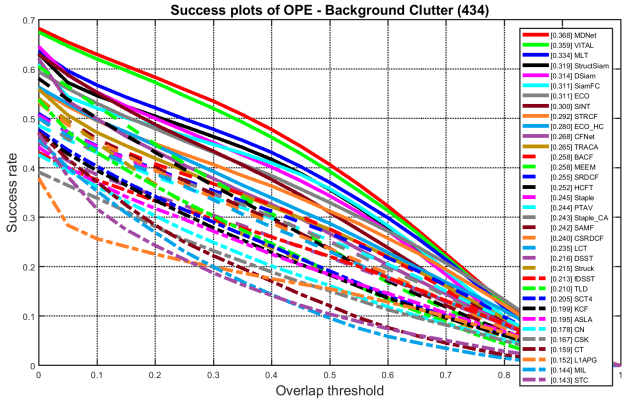
Figure A: **Visualization for the effect of the target-specific feature space.** This shows some example image patches z (1st and 4th row) with the changes in response maps $\hat{y}$ ***before*** (2nd and 5th row) and ***after*** (3rd and 6th row) applying our adaptive weights w^{target} generated by our meta-learner. Extended from Fig. 6 of the original paper

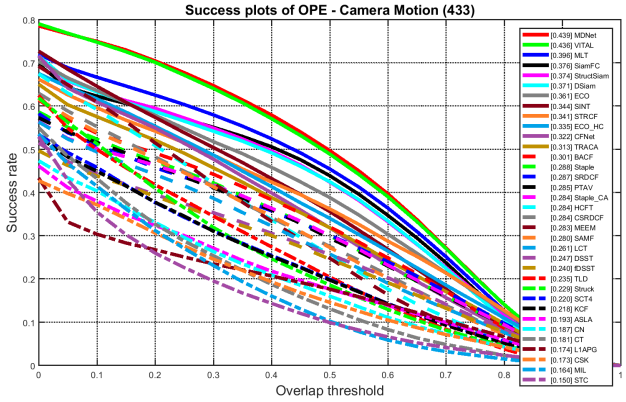
Fig. A shows that the adaptive target-specific weights w^{target} generated by our meta-learner are effective and beneficial for visual tracking. The context images z are shown with the target object fixed at the center of the image, and the response maps *without* and *with* the adaptive weights w^{target} are shown. The response maps show that the target-specific weights help the tracker adapt to various target appearance changes and localize the target, and are also effective in avoiding false positives by suppressing the responses from distractors in the background.

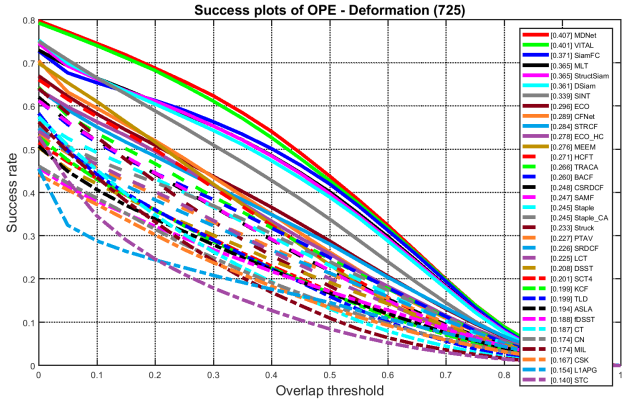
2. Additional Success Plots for LaSOT Dataset

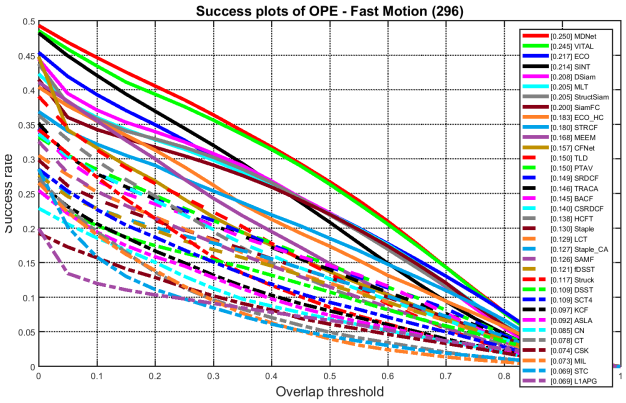


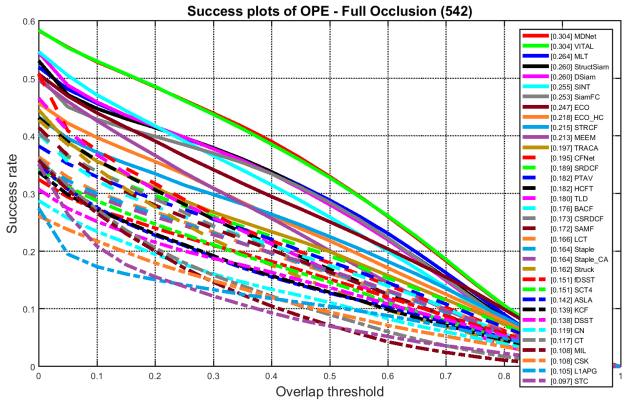












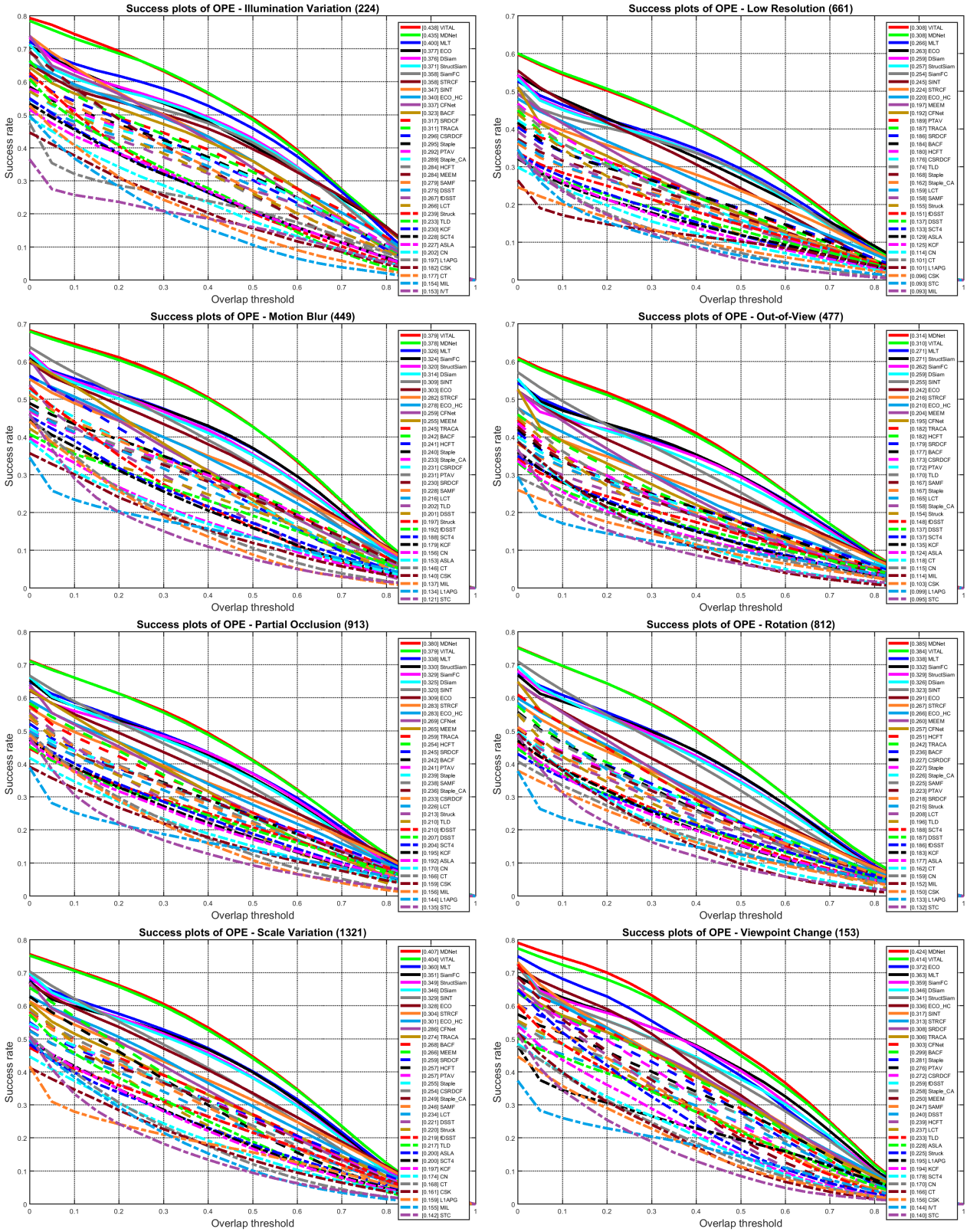
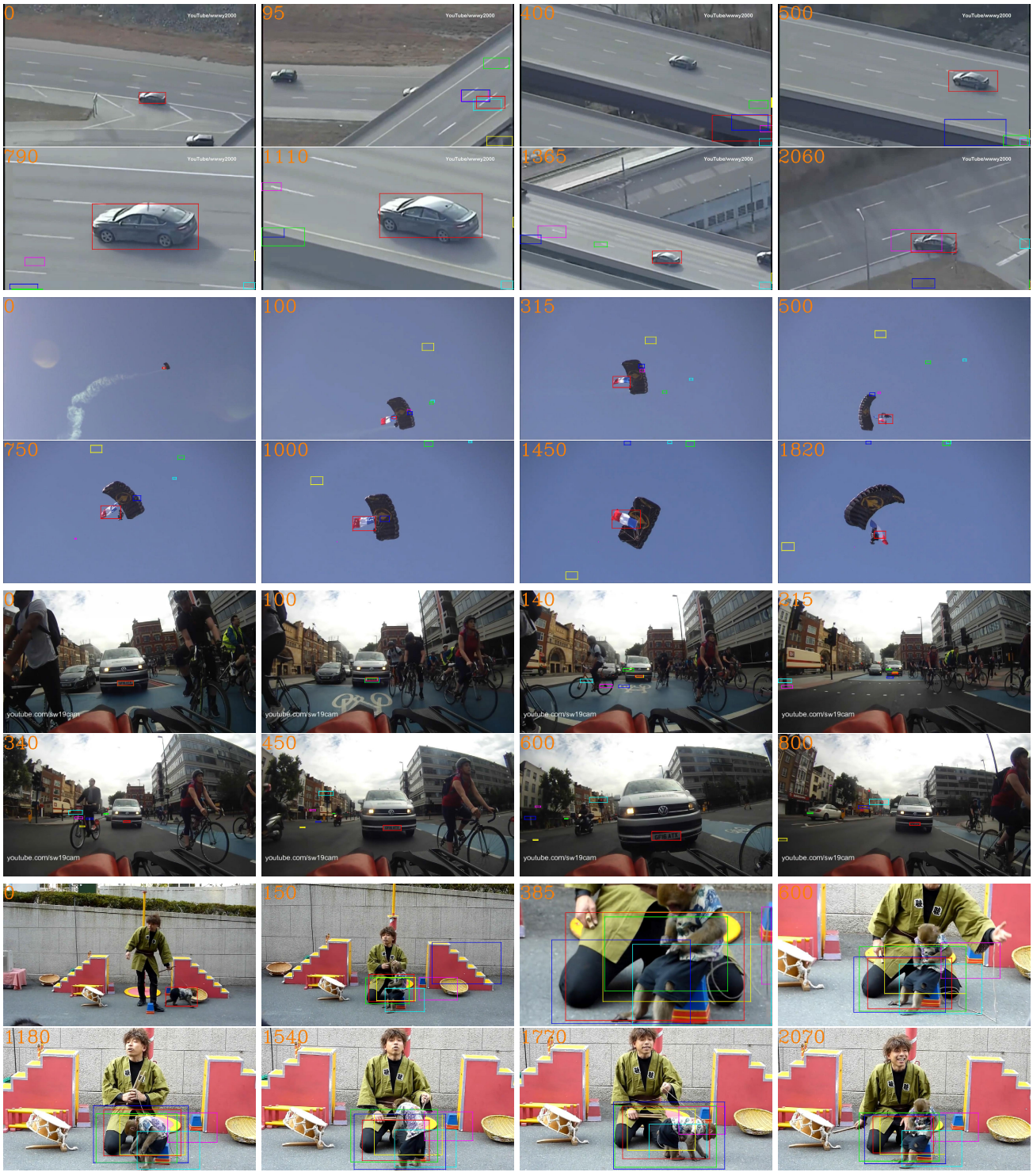


Figure B: **Additional success plots for LaSOT dataset.** Best viewed zoomed in on a high resolution display.

MLT shows robustness to background clutter, camera motion and illumination variation attributes of the LaSOT dataset.

3. Qualitative Results for LaSOT Dataset

The results of **MLT**, **SiamFC**, **StructSiam**, **ECO-HC**, **Staple$_{CA}$**, **SRDCF** are shown in the bounding boxes with respective colors. Frame numbers are indicated in orange on the top left corner. Best viewed zoomed in on a high resolution display.


0
95
400
500
290
1110
1365
2060
0
100
315
500
750
1000
1450
1820
0
100
140
215
340
450
600
800
0
160
385
600
1180
1540
1770
2070

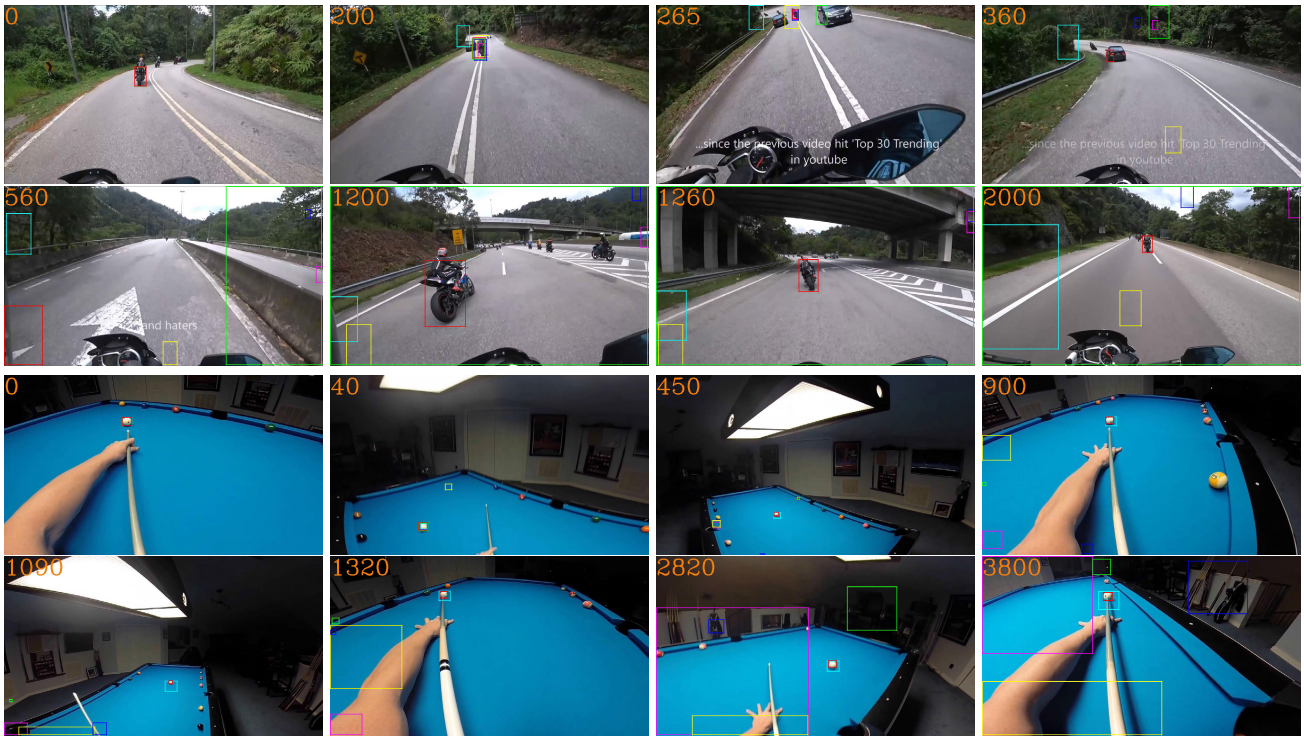


Figure C: **Qualitative Results on LaSOT dataset.** Shown for *boat-19*, *bus-1*, *bus-2*, *car-18*, *car-20*, *flag-1*, *licenseplate-4*, *monkey-17*, *motorcycle-4*, *pool-19* sequences.

4. Preliminary Results on TrackingNet Dataset

We show preliminary results of our method evaluated on the **test set** of TrackingNet dataset, which consists of 511 sequences. AUC of the OPE success plots are shown as performance measure. More comparisons are available in TrackingNet paper. [1]

	MLT	SiamFC	CFNet	ECO-HC	STAPLE$_{\text{CA}}$	BACF	SRDCF	SAMF	ASLA	DSST
TrackingNetTest	**0.581**	0.571	0.578	0.541	0.529	0.523	0.521	0.504	0.478	0.464

Table 1: **Quantitative results on TrackingNetTest dataset.** MLT denotes the proposed algorithm. The proposed algorithm shows similar performance rank as in LaSOT dataset experiments.

5. Ablation Study of Individual Components

We show the results of the ablation experiments that analyzes the impact of each component. The result shows that while all components contribute to performance improvements, meta-kernel weights have the biggest impact.

[1] M. Muller, A. Bibi, S. Giancola, S. Alsubaihi, and B. Ghanem. "Trackingnet: A large-scale dataset and benchmark for object tracking in the wild". ECCV 2018.

Kernel Weights	Channel Attention	ℓ^2-Normalization	AUC
✓	✓	✓	0.611
	✓	✓	0.571
✓		✓	0.601
✓	✓		0.580
			0.564

Table 2: **Ablation study of individual components.** Experiments are performed on OTB-2015 dataset. Performance is denoted in AUC of OPE success plot.

6. Video Results on the OTB dataset

Please refer to the attached video file MLT_ICCV2019_ID4309_DEMO.mp4.