Table R1: Comparison results of our method and the uniform resizing baseline with its crop hyper-parameters varied (*i.e.*, setting c^w and c^h using two different choices as in Sec. 3.1, varying the search-image context factor, and varying the template-image size to alleviate the mismatch of the object size between the search and temple images). The experiments are based on OSTrack-256 and tested on LaSOT. The context factor and size of search/template images are displayed as `context factor/size`. The choices of c^w, c^h: ❶ $c^w = b^w, c^h = b^h$ and ❷$c^w = c^h = (b^w + b^h)/2$

	Original	c^w, c^h	Object Size Mismatched				Object Size Matched				Ours
Search	4/256	4/256	4.5/256	5/256	5.5/256	6/256	4.5/256	5/256	5.5/256	6/256	5/256
Template	2/128	2/128	2/128	2/128	2/128	2/128	2/114	2/102	2/93	2/85	2/128
c^w, c^h	❶	❷	❶	❶	❶	❶	❶	❶	❶	❶	❶
AUC	69.1	68.5	69.0	69.5	68.2	67.5	69.1	68.8	68.0	67.4	**70.2**

Table R2: Ablation on the effect of applying our non-uniform resizing on the template and search images separately.

#	Non-uniform Resizing		LaSOT	
	Template	Search	AUC	P
①	✔		64.8	71.0
②		✔	**70.2**	**76.2**
③	✔	✔	69.6	75.5

Table R3: Further validation of applying our non-uniform resizing to both the performance-oriented version (backbone: Swin Transformer-Base) and speed-oriented version (backbone: Swin Transformer-Tiny) of SwinTrack [15]. Note that we use the v1 version of SwinTrack as the baseline because only the code of v1 version is publicly available.

Type	Trackers	Size	LaSOT	
			AUC	P
Speed-oriented	SwinTrack-T-Zoom	224	**68.5**	**72.9**
	SwinTrack-T	224	66.7	70.6
Performance-oriented	SwinTrack-B-Zoom	224	**70.5**	**75.4**
	SwinTrack-B	224	69.6	74.1
	SwinTrack-B-384	384	70.2	75.3

Table R4: Ratios of challenging scenarios appearing in the different tracking datasets. Our method can achieve significant performance gains under challenging scenarios with drastic movement (*e.g.*, Fast Motion (FM), Viewpoint Change (VC)) or significant drift after some period of time (*e.g.*, Full Occlusion (FOC), Out-of-View (OV)) owning to the enlarged visual field. Note that VC is not labeled in TrackingNet.

Datasets	Ratios of Challenging Scenarios				AUC Gains (OSTrack)
	Significant Drift		Drastic Motion		
	FOC	OV	FM	VC	
LaSOT	42.1%	37.1%	18.9%	11.8%	+1.1%
LaSOT$_{ext}$	62.7%	21.3%	58.7%	39.3%	+3.1%
TNL2K	13.9%	33.4%	24.0%	44.3%	+2.2%
TrackingNet	4.7%	4.0%	10.0%	-	+0.1%