Global Filter Pruning with Self-Attention for Real-Time UAV Tracking

Mengyuan Liu[1] Yuelong Wang[1] Qiangyu Sun[2] Shuiwang Li[1,3]

[1] Guilin University of Technology, China

[2] Hubei Enshi College, China

[3] Guangxi Key Laboratory of Embedded Technology and Intelligent System, China

Introduction

Our contributions can be summarized as follows:(1)Our work provides a fresh perspective to improve efficiency and precision of UAV tracking by developing DL-based trackers with filter pruning method, which has not been well explored before. (2)We proposed a method of global filter pruning with self-attention for real-time UAV tracking, with which the proposed PS-SimaFC++ can globally compress the baseline SiamFC++ to about 60% of its original model size, meanwhile maintaining and even significantly boosting precision simultaneously.(3)We evaluate our PS-SiamFC++ on four public UAV benchmarks, i.e., UAV123@10fps, DTB70 , UAVDT and VisDrone2018 . Experimental results show that the proposed PS-SiamFC++ achieves state-of-the-art performance.

Experiments

Table 1: Comparision of average precision and speed (FPS) between PS-SiamFC++ and hand-crafted based trackers on the four benchmarks. The reported FPSs are evaluated on a single CPU. Red, blue and green respectively show the first, second and third places.

	KCF TPAMI 15	fDSST CVPR 16	Staple-CA CVPR 17	BACF ICCV 17	ECO-HC CVPR 17	MCCT-H CVPR 18	STRCF CVPR 18	ARCF-HC ICCV 19	AutoTrack CVPR 20	RACF 3DV 21	**PS-SiamFC++ Ours**
Precision	53.3	60.4	64.2	65.3	68.8	66.8	67.1	71.9	**72.3**	**75.7**	**79.1**
FPS (PC)	**655.6**	**203.6**	67.7	57.0	**88.9**	66.7	29.9	36.0	61.8	37.5	71.3
FPS (NUC)	**572.1**	**177.6**	59.4	50.1	**77.8**	58.4	26.3	31.2	54.2	32.9	62.4

Table 2: Precision and speed (FPS) (evaluated on a single GPU) comparison between PS-SiamFC++ and deep-based trackers on UAVDT [15].

	SiamR-CNN CVPR 20	D3S CVPR 20	PrDimp18 CVPR 20	KYS ECCV 20	SiamGAT CVPR 21	LightTrank CVPR 21	TransT CVPR 21	HiFT ICCV 21	SOAT ICCV 21	AutoMatch ICCV 21	**PS-SiamFC++ Ours**
Precision	66.5	72.2	73.2	79.8	76.4	80.4	**82.6**	65.2	**82.1**	**82.1**	**80.5**
FPS (GPU)	7.2	44.6	48.5	30.2	74.8	84.8	42.1	**135.3**	29.4	50.4	**291.9**

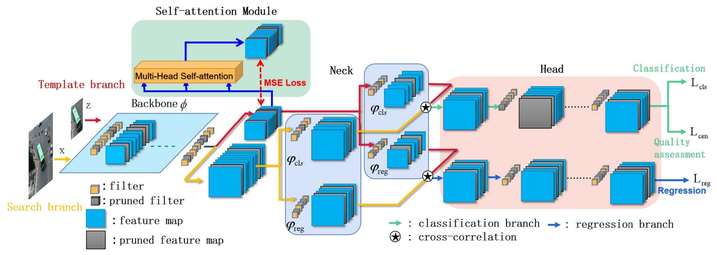
Table 3: Comparison of the proposed PS-SiamFC++ tracker with two baseline trackers in terms of model size (Parameters), precision (DP), and tracking speed (FPS) on the PC CPU.

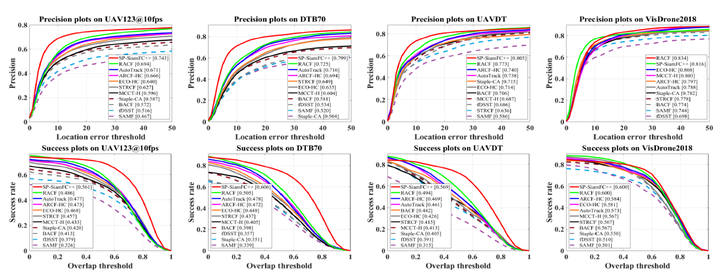
Methods	Parameters	Pruning	Self-attention	UAV123@10fps	DTB70	UAVDT	VisDrone2018	Avg. Precision	Avg. FPS (CPU)
SiamCAR [18]	8.5M			**73.7**	76.6	76.1	**80.3**	**76.7**	40.7
P-SiamCAR	5.1M	✓		70.9	68.2	72.7	74.6	71.6	**79.4**
PS-SiamCAR	5.1M	✓	✓	71.0	73.2	77.1	75.8	74.1	**79.3**
SiamFC++ [47]	9.7M			**72.8**	**80.5**	**76.2**	72.5	75.5	36.5
P-SiamFC++	5.8M	✓		71.9	**79.5**	**78.8**	**79.3**	**77.4**	**71.3**
PS-SiamFC++	5.8M	✓	✓	**74.3**	**79.9**	**80.5**	**81.6**	**79.1**	71.1

Table 4: Illustration of how the precision on DTB70 of PS-SiamFC++ varies with the global pruning ratio, with or without the self-attention module. The precisions that have been improved by the self-attention component are marked in bold.

ρ	UAV123@10fps		DTB70		UAVDT		VisDrone2018	
	w/o	w/	w/o	w/	w/o	w/	w/o	w/
0.1	70.8	**72.2**	**79.6**	79.4	**81.4**	81.3	79.6	**83.1**
0.2	71.6	**72.3**	80.0	**80.1**	76.9	**77.2**	**80.2**	77.7
0.3	71.3	**72.4**	81.0	**81.5**	**83.9**	80.3	75.6	**79.3**
0.4	71.9	**74.3**	79.5	**79.9**	78.8	**80.5**	79.3	**81.6**
0.5	**71.1**	69.2	77.6	**79.2**	77.5	**78.0**	72.7	**76.4**
0.6	68.7	**69.2**	**78.6**	77.1	76.6	**78.3**	75.2	**76.9**
0.7	**69.1**	66.0	**77.9**	74.9	77.8	**80.7**	76.7	**79.4**
0.8	65.2	**66.0**	**76.4**	70.1	74.6	**77.0**	71.8	**74.9**

Methods





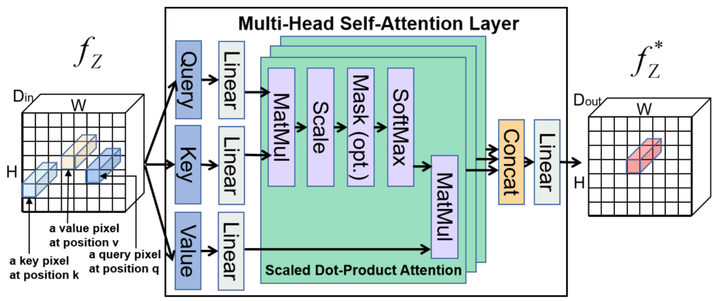
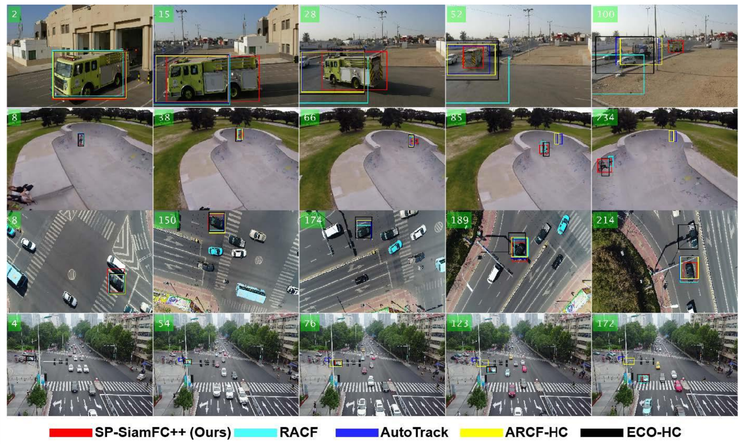


Illustration of the Multi-Head Self-Attention layer applied to the backbone output feature f_Z . Note that f_Z is encoded in a pixel-wise manner, i.e., the spatial coordinates of f_Z index the tokens, and the query, key, and value are initially the same.



Qualitative evaluation on 4 sequences from UAV123@10fps, DTB70, UAVDT and VisDrone2018 (i.e. truck1, BMX5, S0309 and uav0000294_00000_s), respectively. The results of different methods have been shown with different colors.

Conclusions

In this work, we present a method of global filter pruning with self-attention for real-time UAV tracking and achieve state-of-the-art performance on four public UAV tracking benchmarks. When using the proposed method to improve UAV tracking efficiency, experimental results reveal that the proposed method is quite effective at maintaining and even improving precision. Surprisingly, the proposed PS-SiamFC++ not only outperforms the baseline SiamFC++ in terms of efficiency (PS-SiamFC++ can run at and more than 62 FPS on a single CPU of a mini PC, i.e., Intel NUC), but it also outperforms the baseline in terms of precision on UAVDT and VisDrone2018, well combating the adverse effects of filter pruning.