Supplementary Material of

Tracking-by-Trackers with a Distilled and Reinforced Model

A Methodology

A.1 MDP Auxiliary Functions

The function $\psi : \mathcal{A} \times \mathbb{R}^4 \rightarrow \mathbb{R}^4$ used to obtain the bounding box b_t given a_t and the previous bounding box b_{t-1} is defined such that

$$\psi(a_t, b_{t-1}) = \begin{cases} x_t = x_{t-1} + \Delta x_t \cdot w_{t-1} \\ y_t = y_{t-1} + \Delta y_t \cdot h_{t-1} \\ w_t = w_{t-1} + \Delta w_t \cdot w_{t-1} \\ h_t = h_{t-1} + \Delta h_t \cdot h_{t-1} \end{cases} \tag{9}$$

. The function $\phi : \mathbb{R}^4 \times \mathbb{R}^4 \rightarrow \mathcal{A}$ employed to obtain the expert action $a_t^{(\mathbf{t})}$ given the teacher bounding boxes $b_t^{(\mathbf{t})}, b_{t-1}^{(\mathbf{t})}$ is defined as

$$\phi(b_t^{(\mathbf{t})}, b_{t-1}^{(\mathbf{t})}) = \begin{cases} \Delta x_t^{(\mathbf{t})} = (x_t^{(\mathbf{t})} - x_{t-1}^{(\mathbf{t})})/w_{t-1}^{(\mathbf{t})} \\ \Delta y_t^{(\mathbf{t})} = (y_t^{(\mathbf{t})} - y_{t-1}^{(\mathbf{t})})/h_{t-1}^{(\mathbf{t})} \\ \Delta w_t^{(\mathbf{t})} = (w_t^{(\mathbf{t})} - w_{t-1}^{(\mathbf{t})})/w_{t-1}^{(\mathbf{t})} \\ \Delta h_t^{(\mathbf{t})} = (h_t^{(\mathbf{t})} - h_{t-1}^{(\mathbf{t})})/h_{t-1}^{(\mathbf{t})} \end{cases} \tag{10}$$

.

A.2 Curriculum Learning Strategy

A curriculum learning strategy [69] is designed to further facilitate and improve the student's learning. After terminating each E_j, a success counter C_j for M_j is increased if $\mathbf{s}$ performs better than $\mathbf{t}$ in that interaction, i.e. if the first cumulative reward, received up to $\widehat{T}_j$, is greater or equal to the one obtained by the second. In formal terms, C_j is updated if the following condition holds

$$\sum_{i=1}^{\widehat{T}_j} r_i \geq \sum_{i=1}^{\widehat{T}_j} r_i^{(\mathbf{t})}. \tag{11}$$

The counter update is done by testing students that interact with M_j by exploiting $\mathbf{s}_\pi(s_t \mid \theta')$. The terminal video index $1 \geq \widehat{T}_j \geq T_j$ is successively increased during the training procedure by a central process which checks if $\frac{C_j}{E_j} \geq \tau$. After each update of $\widehat{T}_j$, C_j is reset to zero. By this setup, we ensure that, at every increase of $\widehat{T}_j$, students face a simpler learning problem where they are likely to succeed and in a shorter time, since they have already developed a tracking policy that, up to $\widehat{T}_j - 1$, is at least good as the one of $\mathbf{t}$. We found $\tau = 0.25$ to work well in practice.

B Experimental Setup

Benchmarks and Performance Measures. In this subsection we offer more details about the benchmark datasets and the relative performance measures employed to validate our methodology.

GOT-10k Test Set. The GOT-10k [71] test set is composed of 180 videos. Target objects belong to 84 different classes, and 32 forms of object motion are present. An interesting note is that, except for the class *person*, there is no overlap between the object classes in the training and test splits. For the *person* class, there is no overlap in the type of motion. The evaluation protocol proposed by the authors is the one-pass evaluation (OPE) [74], while the metrics used are the average overlap (AO) and the success rates (SR) with overlap thresholds 0.50 and 0.75.

OTB-100. The OTB-100 [74] benchmark is a set of 100 challenging videos and it is widely used in the tracking literature. The standard evaluation procedure for this dataset is the OPE method while the Area Under the Curve (AUC) of the success and precision plot, referred as success score (SS) and precision scores (PS) respectively, are utilized to quantify trackers' performance.

UAV123. The UAV123 benchmark [72] proposes 123 videos that are inherently different from traditional visual tracking benchmarks like OTB and VOT, since it offers sequences acquired form low-altitude UAVs. To evaluate trackers, the standard OTB methodology [74] is exploited.

LaSOT. A performance evaluation was also performed on the test set of LaSOT benchmark [73]. This dataset is composed of 280 videos, with a total of more than 650k frames and an average sequence length of 2500 frames, that is higher than the lengths of the videos contained in the aforementioned benchmarks. The same methodology and metrics used for the OTB [74] experiments are employed.

VOT2019. The VOT benchmarks are datasets used in the annual VOT tracking competitions. These sets change year by year, introducing increasingly challenging tracking scenarios. We evaluated our trackers on the set of the VOT2019 challenge [75], which provides 60 highly challenging videos. Within the framework used by the VOT committee, trackers are evaluated based on Expected Average Overlap (EAO), Accuracy (A) and Robustness (R) [76]. Differently from the OPE, the VOT evaluation protocol presents the automatic re-initialization of the tracker when the IoU between its estimated bounding box and the ground-truth becomes zero.

C Additional Results

C.1 Impact of Transfer Set

We evaluated how performance change considering other sources of video data. By respecting the idea that unbiased demonstrations of the teachers should be

Table 6: Teacher-based statistics of the LaSOT transfer set.

Teachers	$\beta = 0.5$		
	# traj	AO	$\lvert\mathcal{D}\rvert$
$\mathbf{T}_{\mathtt{K}}$	16	0.835	80
$\mathbf{T}_{\mathtt{M}}$	32	0.830	160
$\mathbf{T}_{\mathtt{E}}$	44	0.817	220
$\mathbf{T}_{\mathtt{S}}$	87	0.852	435
$\mathbf{T}_{\mathtt{P}}$	106	0.856	530

Table 7: Performance of the proposed trackers considering the training set of LaSOT as transfer set.

	Tracker	GOT-10k			UAV123		LaSOT		OTB-100	
		AO	$SR_{0.50}$	$SR_{0.75}$	SS	PS	SS	PS	SS	PS
$\beta = 0.5$	TRAS	0.242	0.252	0.086	0.329	0.437	0.222	0.166	0.254	0.337
	TRAST	0.475	0.552	0.248	0.553	0.746	0.463	0.432	0.577	0.760
	TRASFUST	0.468	0.529	0.221	0.594	0.803	0.470	0.452	0.666	0.885

employed, we used the training set of the LaSOT benchmark [73]. This dataset is smaller than the training set of GOT-10k and contains 1120 videos with approximately 2.83M frames. After filtering the trajectories, we obtained the transfer set $\mathcal{D}$ which specification are given in Table 6.

The results are shown in Table 7. The amount of training samples is lower than the amount obtained by filtering the GOT-10k transfer set with $\beta = 0.8$, and the proposed trackers present a behaviour that reflects the loss of data (as seen in Table 4). This experiment suggests that the quantity of data has more impact than the quality of data.

C.2 Success and Precision Plots on OTB-100

In Figures 9 and 10 the success plots and precision plots for different sequence categories of the OTB-100 benchmark are presented.

C.3 Video

At this link `https://youtu.be/uKtQgPk3nCU`, we provide a video showing the tracking abilities of our proposed trackers. For each video, the predictions of TRAS, TRAST and TRASFUST are shown. For TRAST and TRASFUST, we report also the tracker which prediction was chosen as output proposes (with the term "CONTROLLING" next to the tracker's name).

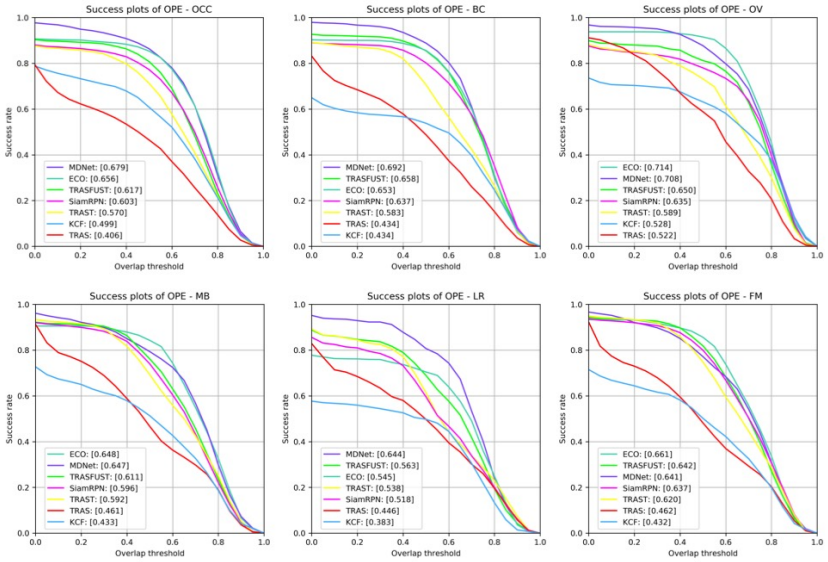

Fig. 9: Success plots on OTB-100 presenting the performance of the proposed trackers and the teachers on tracking situations with: occlusion (OCC); background clutter (BC); out of view (OV); motion blur (MB); low resolution (LR); fast motion (FM).

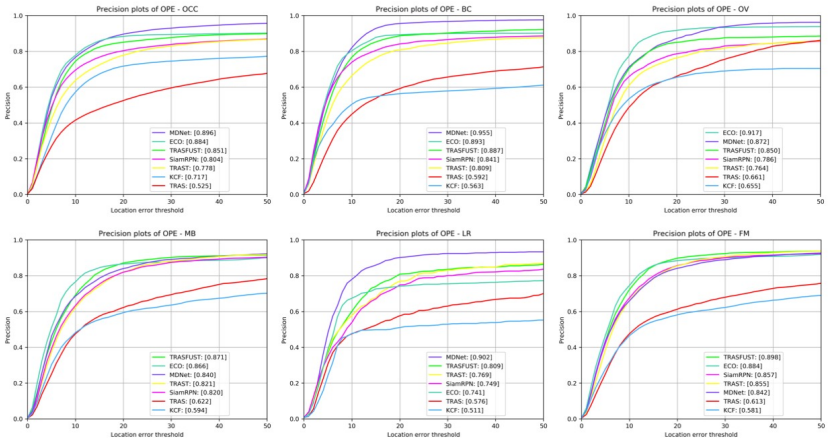

Fig. 10: Precision plots on OTB-100 presenting the performance of the proposed trackers and the teachers on tracking situations with: occlusion (OCC); background clutter (BC); out of view (OV); motion blur (MB); low resolution (LR); fast motion (FM).