Lecture Notes of TICMI
Vol. 27, 2026, 1–15

Does Re-ID Really Help in Multi-Object Tracking?

Tomasz Stanczyk[a,b*], and Francois Bremond[a,b]

[a] *Inria, France*; [b] *Université Côte d'Azur, France*

Multi-object tracking (MOT) is a critical problem in computer vision with applications in video surveillance, behavior analysis, and subject monitoring. In the tracking-by-detection paradigm, detections are associated across video frames to form tracklets using cues such as motion, position, appearance. Appearance-based association relies on re-identification (re-ID) algorithms, which we claim to yield limited improvements in terms of MOT. This study investigates the impact of re-ID on the tracking performance of a MOT algorithm within a popular yet diversified MOT17 dataset. We evaluate the tracker with and without re-ID modules, considering various appearance thresholds and additional constraints aiming to improve the re-ID impact. Our findings indicate that while re-ID can provide marginal improvements in specific scenarios, it often fails to significantly enhance overall performance and can even degrade it under certain conditions. We conclude that re-ID, although beneficial in specific instances, is not a universally robust solution for improving MOT performance.

Keywords and phrases: multi-object tracking, person re-identification, association cues

1 Introduction

Multi-object tracking (MOT) aims to associate objects across consecutive video frames. It plays a crucial role in applications such as video surveillance, subject monitoring, and behavior analysis. In the tracking-by-detection paradigm, detections are obtained per frame and linked into tracklets using cues like motion, position, and appearance. The appearance cue involves comparing bounding boxes visually, typically via re-identification (re-ID) algorithms [5, 3].

Trackers like BoT-SORT [1] and Deep OC-SORT [10] incorporate re-ID into the association process. As shown in Tab. 1, re-ID often yields only marginal improvements. This raises the question of whether re-ID meaningfully enhances MOT performance or if the gains stem from tuning and stochasticity. Intuitively, appearance-based matching seems helpful, echoing how humans track people visually.

This work investigates the impact of re-ID on MOT performance using BoT-SORT [1] evaluated on all validation sequences from the MOT17 dataset [11]. We test (1) no re-ID, (2) a re-ID fine-tuned on MOT17, and (3) a generic re-ID trained on large-scale external data. We further vary the appearance thresholds (Sec. 3.4) and introduce constraints on re-ID usage, such as maximal occlusion and minimum bounding box size.

To examine the effect of ideal detections, we also evaluate BoT-SORT using ground truth bounding boxes, again comparing the three re-ID setups. Additionally, we analyze correct and incorrect inter-frame matches based on re-ID appearance distances (Sec. 3.2), with distribution plots to visualize the match quality.

*Corresponding author. Email: tomasz.stanczyk@inria.fr

ISSN: 1512-0511 print

Table 1: Table extracts from BoT-SORT [1] and Deep OC-SORT [10] works demonstrating performance differences with and without using a re-ID.

MOT algorithm setting	HOTA	IDF1
BoT-SORT, no re-ID	69.11	81.53
BoT-SORT, with re-ID	69.17	82.07
Deep OC-SORT baseline, no re-ID	68.13	79.52
Deep OC-SORT baseline, with re-ID	68.59	80.18
Deep OC-SORT baseline, with re-ID and dynamic appearance	68.65	80.45

Because MOT17 sequences vary in difficulty and scene characteristics, we report results per sequence rather than only in aggregate. This helps assess whether re-ID brings global benefits or only helps in specific scenarios.

In summary, our contributions are: (I) evaluating whether re-ID helps when integrated into an MOT tracker; (II) identifying conditions under which re-ID is most beneficial.

The paper proceeds as follows: related works (Sec. 2); methodology and setup (Sec. 3); experiments and analysis (Sec. 4); general discussion (Sec. 5); and conclusions (Sec. 6).

2 Related works

Tracking-by-detection links detection bounding boxes across video frames. A widely used algorithm, ByteTrack [16], detects objects using YOLOX [4] (pre-trained on the target dataset) and associates them using IoU-based matching between detections and Kalman Filter [6] predictions. BoT-SORT [1] extends ByteTrack by adding camera motion compensation and a re-ID module for appearance-based association. OC-SORT [2] modifies the Kalman Filter trajectory using detection-informed visual motion estimation during occlusions. Deep OC-SORT [10] builds on this by incorporating re-ID features, camera motion compensation, and adaptive weighting.

Tab. 1 shows performance gains from adding re-ID in BoT-SORT and Deep OC-SORT are modest. In this study, we focus on BoT-SORT to explore when and how re-ID improves MOT performance. We evaluate two re-ID systems: Fast-reid [5] and ISR [3]. Fast-reid is a general-purpose neural re-ID framework used in BoT-SORT, where it is fine-tuned (SBS-50 model) on datasets including MOT17. ISR is a SwinTransformer-based [8, 14] model trained on large-scale data, designed for high generalizability.

We also briefly mention the GHOST tracker [13], which integrates re-ID with on-the-fly domain adaptation via batch normalization over same-frame detections during training and inference. However, the pre-trained model is unavailable, and the method is not readily transferable to other trackers, making integration difficult.

3 Methodology

Our goal is to assess whether re-ID genuinely improves association in MOT. We begin by outlining the dataset used in our experiments. To evaluate re-ID performance independently of other tracking components, we introduce a custom assessment script. We then describe the re-ID architectures and detections used for tracking and evaluation on MOT. Finally, we explain how re-ID is usedwithin the tracking process, including our modifications and applied constraints.

3.1 Dataset

We evaluate the MOT algorithm under various re-ID settings using the validation set of the MOT17 dataset [11], which contains the second halves of seven distinct training sequences. These sequences vary in pedestrian density, lighting, camera viewpoint and motion, resolution, and frame rate, making the dataset highly diverse. As shown in Sec. 4, re-ID can improve tracking in some scenarios but may degrade it in others.

3.2 Feature extraction and re-ID assessment

We develop an assessment script to analyze re-ID performance in a tracking context. Unlike standard re-ID evaluation—which focuses on retrieving matching identities from a gallery—tracking requires matching detections across consecutive frames. A common approach is to compute cosine distances between normalized re-ID features and associate detection pairs with the lowest distances, provided they fall below a similarity threshold.

Our script simulates this behavior by using ground truth detections to extract bounding boxes and compute re-ID features for each frame. For each box at frame t, we compute cosine distances to boxes from frame $t-1$ and earlier unmatched boxes from $t-k$ ($k>1$). Matches are made based on minimal distance, with a maximum allowed threshold of 0.5—consistent with most of our full tracking evaluations.

The output includes per-sequence counts and distance distributions of correct and incorrect matches. We also generate corresponding plots for individual sequences and the overall dataset.

Note that this evaluation is not used to produce our main tracking results, but serves as an auxiliary analysis to understand re-ID matching behavior.

3.3 Re-ID and detections considered

To evaluate the impact of re-ID on tracking performance, we use the BoT-SORT [1] algorithm. As a baseline, we consider the variant without re-ID, equivalent to ByteTrack [16] with camera motion compensation. We then evaluate several re-ID models.

First, we use Fast-reid [5] with the SBS S50 architecture, trained on MOT17 [11], as in the original BoT-SORT. For comparison, we also include two Fast-reid models trained on external re-ID datasets: Market-1501 [17] and MSMT17 [15]. Additionally, we test ISR [3], a SwinTransformer-based [8] model trained on a large and diverse corpus (thus not specifically on MOT17 as Fast-reid), reaching remarkable performance on the independent re-ID problem.

All models are evaluated within BoT-SORT using the YOLOX detector provided by the ByteTrack [16] authors, trained on MOT17. To isolate the effect of re-ID from detection quality, we also run evaluations using ground truth detections from MOT17 [11].

3.4 The use of re-ID

We evaluate BoT-SORT [1] using its standard re-ID pipeline with added constraints. Each tracklet maintains an appearance feature vector, updated after each detection match via Exponential Moving Average (EMA). Re-ID features are extracted per bounding box, and BoT-SORT computes both IoU and cosine (appearance) distances for each tracklet-detection pair. Matches with cosine distance above the appearance threshold or IoU distance above the proximity threshold are discarded. The final cost matrix takes the lower of the two distances and is passed to the Hungarian algorithm [7] for bipartite matching. For details, see [1].

(a) (b) (c) (d)

Figure 1: An example of the re-ID (Fast-reid [5]) being confused. At frame t (a), the bounding box of the considered person (b) was associated with the bounding box of the other person (d) at t-1 instead of the same person (c) at t-1. Match was based on the minimum re-ID cosine distance score. Such cases might pose difficulties also to human observers. Source: MOT17 [11].

We argue that re-ID is not always beneficial in MOT: occlusions, small bounding boxes, or poor lighting can yield misleading features. Fig. 1 shows an example where occlusion led to incorrect matching—difficult even for a human to resolve.

To mitigate such issues, we introduce two constraints: (1) maximal occlusion and (2) minimal bounding box size, aiming to exclude unreliable crops from re-ID matching. For occlusion, we filter out boxes heavily covered by others (identified via lower bottom y-coordinates). We use thresholds of 0.5 and 0.75, excluding more occluded boxes. We also apply a minimal in-frame visibility threshold of 0.8 to ensure that most of the subject is within the image.

For bounding box size, we test four pixel area thresholds (4000, 6000, 8000, 10000), excluding smaller boxes likely lacking detail. These filters aim to ensure only meaningful crops contribute to re-ID features.

Lastly, we test three appearance similarity thresholds (0.1, 0.2, 0.5) for cosine distance-based matching. The default value of 0.5 is originally used by BoT-SORT [1] with Fast-reid [5] trained on MOT17 [11].

4 Experiments

We conduct a series of experiments to answer the central question: *Does re-ID really help in Multi-Object Tracking?* Following the methodology in Sec. 3, each subsection addresses a specific aspect:

- How effective is re-ID for matching bounding boxes across consecutive frames? (Sec. 4.1)
- Can correct vs. incorrect matches be clearly distinguished? (Sec. 4.2)
- How much can re-ID improve MOT with perfect detections? (Sec. 4.3)
- How much does a pre-trained re-ID help with real (imperfect) detections? (Sec. 4.4)
- Can a pre-trained re-ID be more effective with usage constraints? (Sec. 4.5)
- How much does a generic re-ID help with imperfect detections? (Sec. 4.6)
- Can a generic re-ID be more effective with usage constraints? (Sec. 4.7)

Table 2: Re-ID assessment matches based on Fast-reid [5], all bounding boxes included.

All bounding boxes				
Sequence	Total matches	Correct matches	Incorrect matches	Accuracy
MOT17-02	18559	18264	295	0.9841
MOT17-04	47515	47473	42	0.9991
MOT17-05	6911	6489	422	0.9389
MOT17-09	5319	5272	47	0.9912
MOT17-10	12820	12535	285	0.9778
MOT17-11	9419	9380	39	0.9959
MOT17-13	11620	11454	166	0.9857
OVERALL	112163	110867	1296	0.9884

Table 3: Re-ID assessment matches based on Fast-reid [5], maximum bounding box overlap 0.75.

Max overlap=0.75				
Sequence	Total matches	Correct matches	Incorrect matches	Accuracy
MOT17-02	9330	9254	76	0.9919
MOT17-04	39398	39340	58	0.9985
MOT17-05	4738	4635	103	0.9783
MOT17-09	3721	3691	30	0.9919
MOT17-10	10507	10295	212	0.9798
MOT17-11	7490	7448	42	0.9944
MOT17-13	9955	9828	127	0.9872
OVERALL	85139	84491	648	0.9924

Table 4: Re-ID assessment matches based on Fast-reid [5], maximum bounding box overlap 0.5.

Max overlap=0.5				
Sequence	Total matches	Correct matches	Incorrect matches	Accuracy
MOT17-02	7211	7182	29	0.996
MOT17-04	28740	28710	30	0.999
MOT17-05	3616	3541	75	0.9793
MOT17-09	3134	3116	18	0.9943
MOT17-10	9082	8911	171	0.9812
MOT17-11	6077	6059	18	0.997
MOT17-13	7945	7861	84	0.9894
OVERALL	65805	65380	425	0.9935

Since aggregate results may mask sequence-level variation, we report performance per sequence to highlight when re-ID is beneficial. Due to space constraints, we include distribution plots for all sequences combined in the main paper and present single sequence-focused plots in the supplementary material.

To evaluate tracking performance, we report HOTA [9] and IDF1 [12], focusing on association quality. MOTA is omitted as it primarily reflects detection quality. Due to space constraints, IDF1 results are available in the supplementary material.

4.1 Re-ID assessment matches in numbers

We run the re-ID assessment from Sec. 3.2 using Fast-reid [5] on all MOT17 sequences, with ground truth detections to isolate re-ID behavior. Re-ID usage constraints are applied as described in Sec. 3.4.

Tabs. 2 to 6 report the number of correct and incorrect matches per sequence and overall, under different constraint settings. As constraints tighten—e.g., lower maximum occlusion or higher minimum box size—both the total number of matches and incorrect matches decrease significantly.

Importantly, the effect varies across sequences, reflecting differences in characteristics such as crowd density and bounding box scale. By reducing incorrect matches, re-ID can have a more positive impact on tracking performance.

Table 5: Re-ID assessment matches based on Fast-reid [5], minumum bounding box size 4000 pixels.

Min size=4000				
Sequence	Total matches	Correct matches	Incorrect matches	Accuracy
MOT17-02	7958	7896	62	0.9922
MOT17-04	47403	47362	41	0.9991
MOT17-05	4958	4801	157	0.9683
MOT17-09	5319	5272	47	0.9912
MOT17-10	7247	7166	81	0.9888
MOT17-11	8474	8450	24	0.9972
MOT17-13	4118	4073	45	0.9891
OVERALL	85477	85020	457	0.9947

Table 6: Re-ID assessment matches based on Fast-reid [5], minumum bounding box size 6000 pixels.

Min size=6000				
Sequence	Total matches	Correct matches	Incorrect matches	Accuracy
MOT17-02	4420	4385	35	0.9921
MOT17-04	44377	44335	42	0.9991
MOT17-05	4518	4414	104	0.977
MOT17-09	5319	5272	47	0.9912
MOT17-10	5129	5111	18	0.9965
MOT17-11	7901	7884	17	0.9978
MOT17-13	2643	2612	31	0.9883
OVERALL	74307	74013	294	0.996

4.2 Re-ID assessment matches in distribution plots

Fig. 2 presents cosine distance distributions of correct (green) and incorrect (red) matches from our re-ID assessment across all MOT17 sequences. The effect of applying bounding box constraints is clearly visible. Note that frequency scales vary between plots.

Selecting an appropriate cosine distance threshold is critical for re-ID-based matching. A low threshold may reject true matches; a high one may accept false ones. The plots reveal that setting this threshold is non-trivial: trade-offs exist between rejecting incorrect matches and preserving correct ones. Even a small number of incorrect matches can degrade tracking performance, while rejecting valid ones may limit re-ID's benefit.

As shown later in Secs. 4.4 and 4.6, changes in threshold do not always align with expectations compared to the unconstrained distributions in Fig. 2(a)–(b). We also include all the distribution plots for a selected single sequence, MOT17-02 and place them in the supplementary material.

4.3 Tracker runs with re-ID and ground truth detections

We run BoT-SORT using ground truth detections and evaluate performance with both Fast-reid [5] and ISR [3], as shown in Tab. 7. Three cosine distance thresholds are tested for re-ID matching: 0.1, 0.2, and 0.5—the latter corresponding to BoT-SORT's original setting (0.25 threshold and computed re-ID distance divided by 2). For reference, we also report results without re-ID.

As seen in Tab. 7, re-ID adds little value when detections are near-perfect and IoU with motion compensation already performs well. In some cases, re-ID even slightly worsens performance—both per sequence and overall.

For instance, in MOT17-05, frequent scene entry/exit makes IoU matching difficult, while large boxes benefit re-ID. In contrast, MOT17-13 features small, overlapping boxes (e.g., crowded bus scenes), making it harder for re-ID to extract meaningful features, leading to degraded performance.

We do not apply re-ID usage constraints (e.g., max overlap or min box size) in this setup, as ground truth boxes are assumed to be of high quality, and tracking performance is already strong.

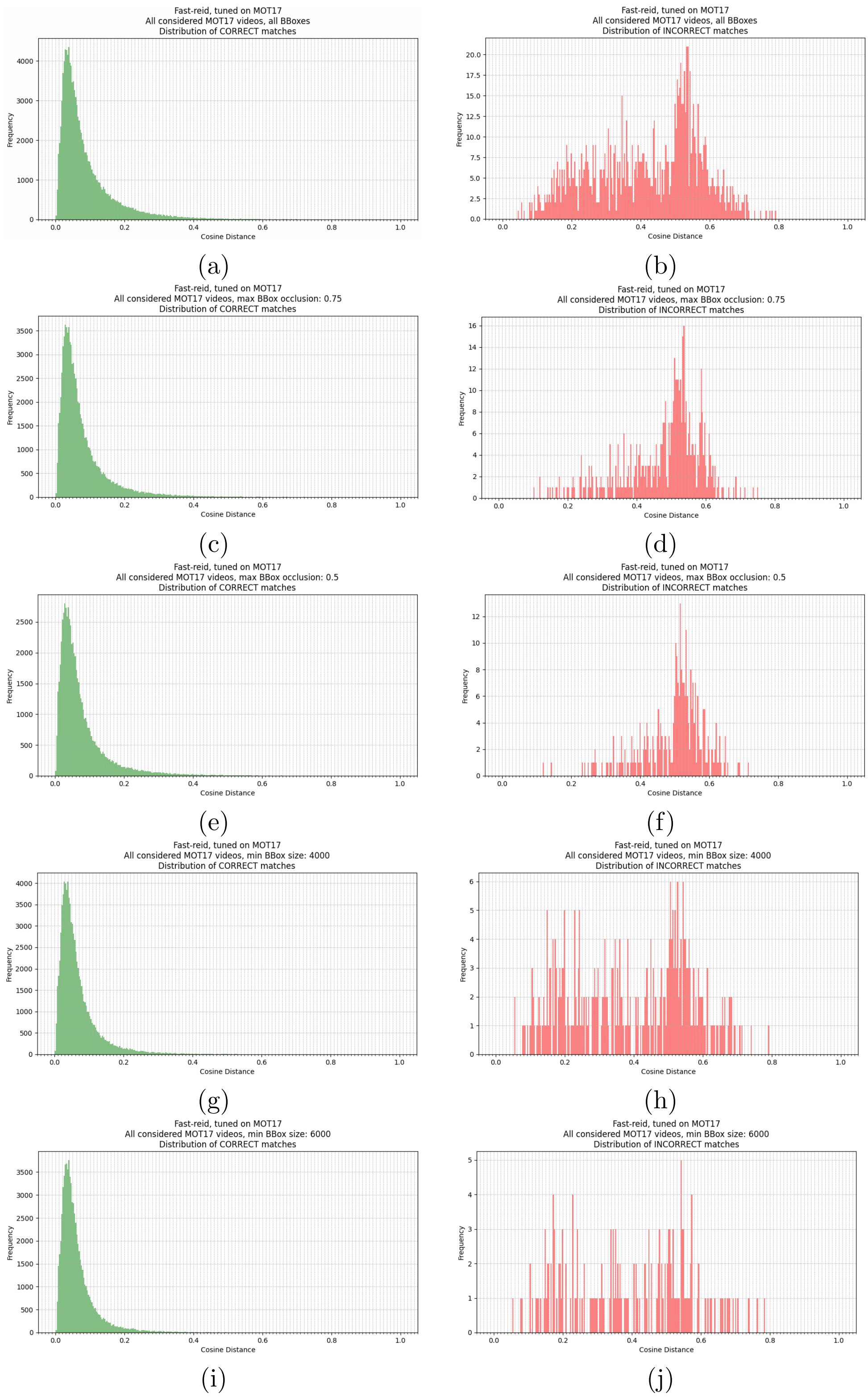

Figure 2: Distribution plots of correct and incorrect matches for all considered sequences based on the re-ID assessment with Fast-reid [5]. Note different frequency scales.

Table 7: HOTA scores of BoT-SORT [1] runs with ground truth detections and different re-IDs.

Sequence	No re-ID	Fast-reid app_th=0.1	Fast-reid app_th=0.2	Fast-reid app_th=0.5	ISR app_th=0.1	ISR app_th=0.2	ISR app_th=0.5
MOT17-02	99.131	99.131	99.131	99.131	99.131	99.131	99.131
MOT17-04	99.913	99.913	99.913	99.913	99.913	99.913	99.913
MOT17-05	89.734	89.873	90.89	88.619	90.842	89.528	89.56
MOT17-09	94.055	94.055	94.055	94.055	94.04	94.04	94.04
MOT17-10	98.968	98.968	98.968	96.928	98.933	98.192	98.098
MOT17-11	99.167	99.167	99.167	99.167	99.167	99.167	99.167
MOT17-13	95.897	95.897	96.953	96.953	96.377	96.793	96.953
COMBINED	98.458	98.466	98.585	98.231	98.544	98.411	98.412

Table 8: HOTA scores of BoT-SORT with YOLOX and Fast-reid variants.

Sequence	No re-ID	Market-1501 app_th=0.5	MSMT17 app_th=0.5	MOT17 app_th=0.1	MOT17 app_th=0.2	MOT17 app_th=0.5
MOT17-02	47.131	48.158	46.491	47.11	49.304	48.919
MOT17-04	78.976	78.125	77.607	78.574	79.046	78.433
MOT17-05	60.078	59.601	61.433	60.063	61.469	61.664
MOT17-09	67.941	66.624	65.853	67.94	65.878	66.653
MOT17-10	57.204	56.464	57.519	57.207	59.565	58.506
MOT17-11	66.697	66.519	66.481	66.699	66.699	66.698
MOT17-13	69.833	68.423	69.033	69.822	69.791	68.42
COMBINED	68.428	67.878	67.588	68.231	68.951	68.466

4.4 Tracker runs with Fast-reid and YOLOX detections

We run BoT-SORT with YOLOX detections and different Fast-reid [5] variants. SBS-50 models pre-trained on Market-1501 [17], MSMT17 [15], and MOT17 [11] are evaluated. For Market-1501/MSMT17 models, we use the BoT-SORT default threshold of 0.5. For the MOT17-pretrained model, we test thresholds of 0.1, 0.2, and 0.5. A no re-ID baseline is also included. Results are shown in Tab. 8.

When using re-ID trained on external datasets (Market-1501, MSMT17), MOT performance generally decreases, though minor gains appear on select sequences. The MOT17-pretrained model brings slight improvements overall. Varying the cosine threshold affects results inconsistently—some sequences benefit from re-ID, while others show degraded performance.

Overall, re-ID yields only modest gains compared to the no re-ID baseline. Sequence characteristics (e.g., crowd density, box size as mentioned in the previous section) strongly influence the usefulness of re-ID-based matching.

4.5 Tracker runs with Fast-reid conditioned and YOLOX detections

We run BoT-SORT with YOLOX detections and Fast-reid pre-trained on MOT17, applying our re-ID usage constraints as variables. We include: (i) the base version without constraints, (ii) variants with max overlap or min box size thresholds, and (iii) the no re-ID baseline. The cosine distance threshold is fixed at 0.5, following the original BoT-SORT [1]. Results are shown in Tab. 9.

Performance changes vary across sequences. While filtering out occluded bounding boxes might seem beneficial, it often negates re-ID's contribution, leading to results close to the no re-ID baseline. Likely, too many boxes are excluded, and association defaults to IoU-based matching.

Table 9: HOTA scores of BoT-SORT with YOLOX and Fast-reid (threshold 0.5) under our constraints.

Sequence	No re-ID	Primary	Max overlap 0.75	Max overlap 0.5	Min size 4000	Min size 6000	Min size 8000	Min size 10000
MOT17-02	47.131	48.919	47.066	47.114	49.953	50.134	50.189	50.142
MOT17-04	78.976	78.433	78.976	78.976	78.434	78.771	78.131	78.966
MOT17-05	60.078	61.664	60.063	60.073	61.726	61.726	61.728	61.734
MOT17-09	67.941	66.653	67.948	67.94	66.651	66.645	66.651	66.651
MOT17-10	57.204	58.506	58.688	57.185	57.575	59.557	59.561	57.158
MOT17-11	66.697	66.698	66.698	66.7	66.696	66.699	66.694	66.694
MOT17-13	69.833	68.42	69.831	69.835	68.42	68.424	69.807	69.82
COMBINED	68.428	68.466	68.559	68.423	68.513	68.9	68.668	68.852

Table 10: HOTA scores of BoT-SORT with YOLOX detections and ISR variants.

Sequence	No re-ID	app_th=0.1	app_th=0.2	app_th=0.5
MOT17-02	47.131	47.767	47.29	48.101
MOT17-04	78.976	77.9	77.875	76.774
MOT17-05	60.078	60.078	60.078	60.284
MOT17-09	67.941	59.667	62.992	66.748
MOT17-10	57.204	54.614	54.665	57.117
MOT17-11	66.697	66.697	66.697	66.697
MOT17-13	69.833	69.686	69.48	68.621
COMBINED	68.428	67.322	67.403	67.345

For minimal bounding box size, some sequences benefit (e.g., MOT17-05, with frequent scene entry/exit), while others suffer (e.g., MOT17-13, with many small boxes limiting feature quality). Overall, the most helpful constraint is a 6000-pixel minimum box size. Still, the performance differences remain modest.

4.6 Tracker runs with ISR and YOLOX detections

We run BoT-SORT with YOLOX detections and the generic ISR [3] re-ID model, using cosine distance thresholds of 0.1, 0.2, and 0.5. For comparison, we also include the variant without re-ID. Results are shown in Tab. 10.

ISR yields slight improvements in some cases (e.g., MOT17-02), but also causes visible drops (e.g., MOT17-04). While MOT17-02 includes occlusions, it features many large bounding boxes favorable for re-ID. MOT17-04, in contrast, presents small, occluded groups under low lighting—conditions under which ISR struggles.

Despite strong standalone performance on re-ID benchmarks, ISR struggles in the cases present in MOT, where occlusions, small detections, and poor lighting limit its utility.

4.7 Tracker runs with ISR conditioned and YOLOX detections

We run BoT-SORT with YOLOX detections and the generic ISR re-ID model, applying our bounding box constraints to examine their effect on tracking. We include the primary variant (no constraints) and the variant without re-ID, all using the cosine distance threshold of 0.5 as in the original BoT-SORT [1]. Results are shown in Tab. 11.

The max-overlap constraint appears to fully negate the influence of ISR, suggesting IoU matching alone was sufficient—especially given ISR's limited effectiveness in MOT scenarios.

Table 11: HOTA scores of BoT-SORT with YOLOX and ISR (threshold 0.5) under our constraints.

Sequence	No re-ID	Primary	Max overlap 0.75	Max overlap 0.5	Min size 4000	Min size 6000
MOT17-02	47.131	48.101	47.131	47.131	47.607	47.821
MOT17-04	78.976	76.774	78.976	78.976	76.774	76.774
MOT17-05	60.078	60.284	60.078	60.078	60.332	60.2
MOT17-09	67.941	66.748	67.941	67.941	66.748	66.748
MOT17-10	57.204	57.117	57.204	57.204	56.18	56.19
MOT17-11	66.697	66.697	66.697	66.697	66.697	66.697
MOT17-13	69.833	68.621	69.833	69.833	68.621	68.633
COMBINED	68.428	67.345	68.428	68.428	67.177	67.202

The minimal bounding box size constraint leads to mixed outcomes: slight gains (e.g., MOT17-02), no change (MOT17-04), or drops (MOT17-13). Overall, ISR tends to degrade or match the no re-ID baseline, even with constraints applied.

Since ISR benefits less from the bounding box size constraint than Fast-reid (Sec. 4.5), we limit evaluation to 4000 and 6000 pixel thresholds.

5 Discussion

This section summarizes the insights from Sec. 4. As shown, re-ID is not a universally reliable association cue for MOT. Its impact depends heavily on detection quality: when detections are strong, re-ID may offer slight gains, but often adds little beyond what intersection over union (IoU) can already achieve. When detections are poor, re-ID may introduce incorrect associations and harm performance.

Performance improves slightly when re-ID is trained on the target dataset, but its benefit remains limited compared to other cues. Moreover, success depends not only on the re-ID model, but also on proper tuning of related MOT parameters—such as appearance thresholds, matching thresholds, and high-confidence filters [1, 16]. Using a generic or out-of-domain re-ID without tuning typically leads to performance drops.

The effect of re-ID also varies across sequences. In diverse datasets like MOT17, it may help in some scenarios (e.g., people entering/exiting the scene), especially when subjects are large enough for reliable feature extraction. In others, where bounding boxes are small or occluded, re-ID often misleads. In cases with high-quality detections and clear visibility, re-ID becomes redundant.

Overall, while re-ID can assist tracking in some settings, its gains are modest, and its success depends on careful, case-specific tuning. As such, it is not a consistently robust cue for multi-object tracking.

6 Conclusion

We evaluated the impact of integrating re-identification (re-ID) into multi-object tracking using BoT-SORT on the MOT17 dataset. Through extensive experiments with different re-ID models, thresholds, and bounding box constraints, we found that while re-ID can help in specific scenarios, its overall contribution to tracking performance is limited. Beneficial use requires careful tuning and dataset-specific training. Due to variability across conditions and the risk of performance degradation, re-ID should be applied selectively. In many cases, simpler cues like intersection over union may prove more effective.

Acknowledgements

This work has been supported by the French government, through the 3IA Cote d'Azur Investments in the project managed by the National Research Agency (ANR) with the reference number ANR-23-IACL-0001.

This work was granted access to the HPC resources of IDRIS under the allocation 2025-AD011014370 made by GENCI.

References

[1] Nir Aharon, Roy Orfaig, and Ben-Zion Bobrovsky. Bot-sort: Robust associations multi-pedestrian tracking. *arXiv preprint arXiv:2206.14651*, 2022.

[2] Jinkun Cao, Jiangmiao Pang, Xinshuo Weng, Rawal Khirodkar, and Kris Kitani. Observation-centric sort: Rethinking sort for robust multi-object tracking. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 9686–9696, 2023.

[3] Zhaopeng Dou, Zhongdao Wang, Yali Li, and Shengjin Wang. Identity-seeking self-supervised representation learning for generalizable person re-identification. In *Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV)*, pages 15847–15858, October 2023.

[4] Zheng Ge, Songtao Liu, Feng Wang, Zeming Li, and Jian Sun. Yolox: Exceeding yolo series in 2021. *arXiv preprint arXiv:2107.08430*, 2021.

[5] Lingxiao He, Xingyu Liao, Wu Liu, Xinchen Liu, Peng Cheng, and Tao Mei. Fastreid: A pytorch toolbox for general instance re-identification. *arXiv preprint arXiv:2006.02631*, 2020.

[6] R. E. Kalman. A New Approach to Linear Filtering and Prediction Problems. *Journal of Basic Engineering*, 82(1):35–45, 03 1960.

[7] H. W. Kuhn. The hungarian method for the assignment problem. *Naval Research Logistics Quarterly*, 2(1-2):83–97, 1955.

[8] Ze Liu, Yutong Lin, Yue Cao, Han Hu, Yixuan Wei, Zheng Zhang, Stephen Lin, and Baining Guo. Swin transformer: Hierarchical vision transformer using shifted windows. In *Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV)*, 2021.

[9] Jonathon Luiten, Aljosa Osep, Patrick Dendorfer, Philip Torr, Andreas Geiger, Laura Leal-Taixé, and Bastian Leibe. Hota: A higher order metric for evaluating multi-object tracking. *International Journal of Computer Vision*, pages 1–31, 2020.

[10] Gerard Maggiolino, Adnan Ahmad, Jinkun Cao, and Kris Kitani. Deep oc-sort: Multi-pedestrian tracking by adaptive re-identification. *arXiv preprint arXiv:2302.11813*, 2023.

[11] A. Milan, L. Leal-Taixé, I. Reid, S. Roth, and K. Schindler. MOT16: A benchmark for multi-object tracking. *arXiv:1603.00831 [cs]*, March 2016. arXiv: 1603.00831.

[12] Ergys Ristani, Francesco Solera, Roger Zou, Rita Cucchiara, and Carlo Tomasi. Performance measures and a data set for multi-target, multi-camera tracking. In Gang Hua and Hervé Jégou, editors, *Computer Vision – ECCV 2016 Workshops*, pages 17–35, Cham, 2016. Springer International Publishing.

[13] Jenny Seidenschwarz, Guillem Brasó, Ismail Elezi, and Laura Leal-Taixé. Simple cues lead to a strong multi-object tracker. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2023.

[14] Ashish Vaswani, Noam Shazeer, Niki Parmar, Jakob Uszkoreit, Llion Jones, Aidan N Gomez, Ł ukasz Kaiser, and Illia Polosukhin. Attention is all you need. In I. Guyon, U. Von Luxburg, S. Bengio, H. Wallach, R. Fergus, S. Vishwanathan, and R. Garnett, editors, *Advances in Neural Information Processing Systems*, volume 30. Curran Associates, Inc., 2017.

[15] Longhui Wei, Shiliang Zhang, Wen Gao, and Qi Tian. Person transfer gan to bridge domain gap for person re-identification, 2018.

[16] Yifu Zhang, Peize Sun, Yi Jiang, Dongdong Yu, Fucheng Weng, Zehuan Yuan, Ping Luo, Wenyu Liu, and Xinggang Wang. Bytetrack: Multi-object tracking by associating every detection box. In *Proceedings of the European Conference on Computer Vision (ECCV)*, 2022.

[17] Liang Zheng, Liyue Shen, Lu Tian, Shengjin Wang, Jingdong Wang, and Qi Tian. Scalable person re-identification: A benchmark. In *2015 IEEE International Conference on Computer Vision (ICCV)*, pages 1116–1124, 2015.

Appendices

This supplementary material contains the following appendinces, skipped in the main paper due to the space limit:

- IDF1 scores of the tracker performance in all the cases examined (5 tables).
- Re-ID assessment distribution plots of correct and incorrect matches for a single video sequence: MOT17-02 [11] (10 plots).

Table 12: IDF1 scores of BoT-SORT with ground-truth detections and different re-ID models.

Sequence	No re-ID	Fast-reid app_th=0.1	Fast-reid app_th=0.2	Fast-reid app_th=0.5	ISR app_th=0.1	ISR app_th=0.2	ISR app_th=0.5
MOT17-02	98.931	98.931	98.931	98.931	98.931	98.931	98.931
MOT17-04	99.957	99.957	99.957	99.957	99.957	99.957	99.957
MOT17-05	86.547	86.637	87.598	85.285	87.568	86.456	86.486
MOT17-09	92.161	92.161	92.161	92.161	92.161	92.161	92.161
MOT17-10	99.383	99.383	99.383	97.667	99.383	98.563	98.546
MOT17-11	99.456	99.456	99.456	99.456	99.456	99.456	99.456
MOT17-13	95.406	95.406	96.074	96.074	96.074	96.074	96.074
COMBINED	98.151	98.156	98.255	97.923	98.253	98.094	98.094

Table 13: IDF1 scores of BoT-SORT with YOLOX and different FastReID variants.

Sequence	No re-ID	Market-1501 app_th=0.5	MSMT17 app_th=0.5	MOT17 app_th=0.1	MOT17 app_th=0.2	MOT17 app_th=0.5
MOT17-02	56.968	57.454	55.47	56.917	60	58.608
MOT17-04	91.021	89.996	89.152	90.527	90.864	90.299
MOT17-05	75.124	74.146	78.124	75.136	77.969	77.804
MOT17-09	79.985	79.106	78.05	79.985	78.832	79.091
MOT17-10	76.157	76.166	78.368	76.157	81.087	78.745
MOT17-11	77.326	76.967	77.057	77.326	77.326	77.326
MOT17-13	89.533	87.44	87.898	89.499	89.431	87.496
COMBINED	80.92	80.249	79.956	80.678	81.984	81.121

Table 14: IDF1 scores of BoT-SORT with YOLOX detections and FastReID (appearance threshold 0.5) using our constraints.

Sequence	No re-ID	Primary	Max overlap 0.75	Max overlap 0.5	Min size 4000	Min size 6000	Min size 8000	Min size 10000
MOT17-02	56.968	58.608	56.908	56.93	60.351	60.133	60.173	60.143
MOT17-04	91.021	90.299	91.021	91.023	90.301	90.788	89.626	90.931
MOT17-05	75.124	77.804	75.136	75.136	77.856	77.913	77.913	77.913
MOT17-09	79.985	79.091	79.985	79.985	79.091	79.091	79.091	79.091
MOT17-10	76.157	78.745	78.836	76.12	76.525	79.221	79.166	76.115
MOT17-11	77.326	77.326	77.326	77.326	77.326	77.326	77.326	77.326
MOT17-13	89.533	87.496	89.499	89.499	87.496	87.496	89.499	89.499
COMBINED	80.92	81.121	81.193	80.91	81.192	81.674	81.243	81.532

Table 15: IDF1 scores of BoT-SORT with YOLOX detections and ISR variants.

Sequence	No re-ID	app_th=0.1	app_th=0.2	app_th=0.5
MOT17-02	56.968	57.264	58.558	58.993
MOT17-04	91.021	89.806	89.477	87.59
MOT17-05	75.124	75.124	75.124	75.219
MOT17-09	79.985	71.286	75.362	79.315
MOT17-10	76.157	73.311	73.311	78.081
MOT17-11	77.326	77.326	77.326	77.326
MOT17-13	89.533	89.362	89.226	87.7
COMBINED	80.92	79.633	79.898	79.723

Table 16: DF1 scores of BoT-SORT with YOLOX detections and ISR (appearance threshold 0.5, with constraints).

Sequence	No re-ID	Primary	Max overlap 0.75	Max overlap 0.5	Min size 4000	Min size 6000
MOT17-02	56.968	58.993	56.968	56.968	58.111	57.934
MOT17-04	91.021	87.59	91.021	91.021	87.59	87.59
MOT17-05	75.124	75.219	75.124	75.124	75.239	75.04
MOT17-09	79.985	79.315	79.985	79.985	79.315	79.315
MOT17-10	76.157	78.081	76.157	76.157	75.849	75.867
MOT17-11	77.326	77.326	77.326	77.326	77.326	77.326
MOT17-13	89.533	87.7	89.533	89.533	87.7	87.7
COMBINED	80.92	79.723	80.92	80.92	79.335	79.295

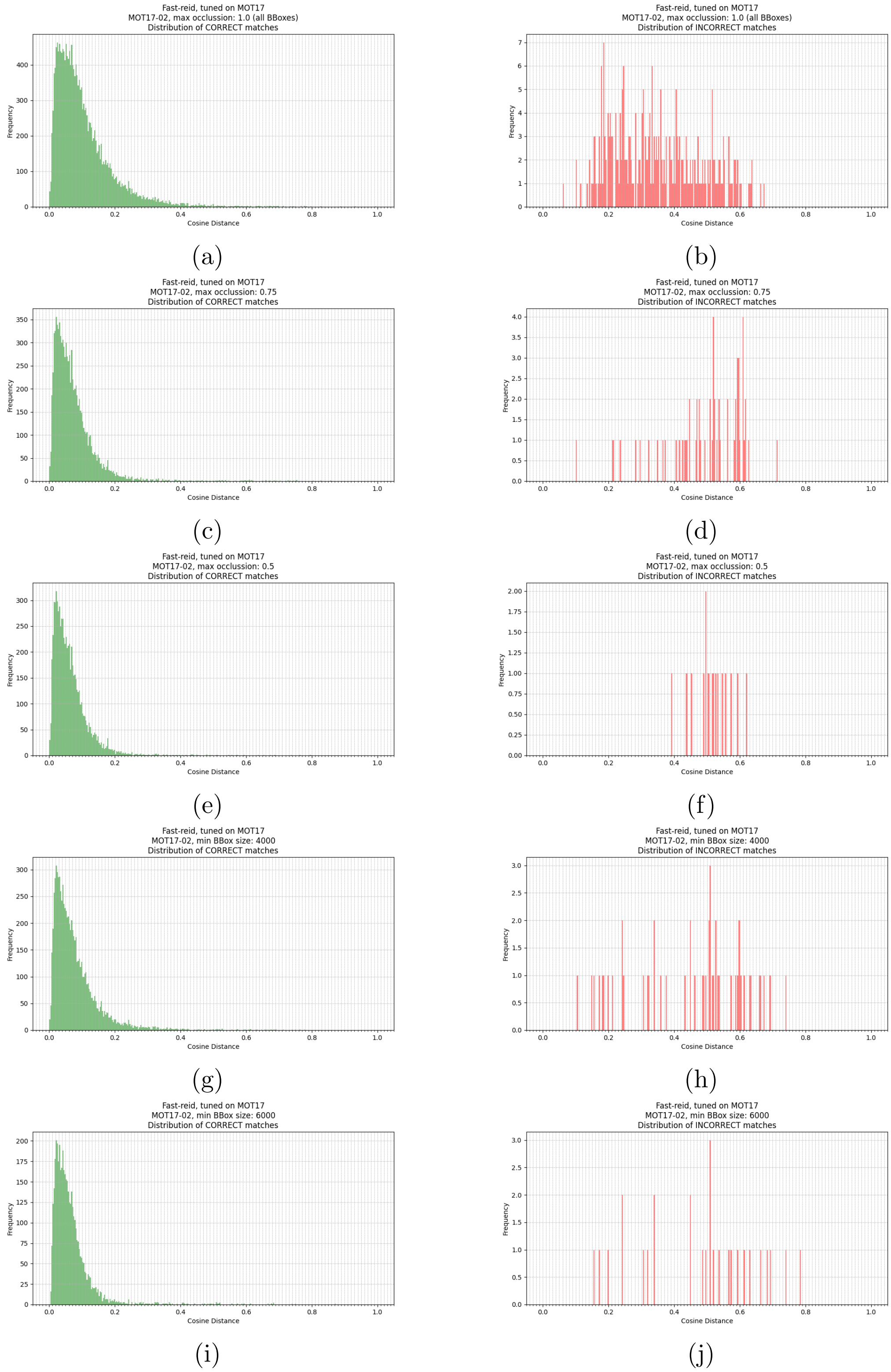

Figure 3: Distribution plots of correct and incorrect matches for the MOT17-02 sequence based on the re-ID assessment with Fast-reid [5]. Note different frequency scales.