

北京航空航天大学学报
*Journal of Beijing University of Aeronautics and Astronautics*
ISSN 1001-5965,CN 11-2625/V

# 《北京航空航天大学学报》网络首发论文

题目：OH-Track:一种具有自适应遮挡感知能力的鲁棒多目标跟踪算法
作者：曾可，余旺盛，秦先祥，龙思宇，白龙飞，杨少波
DOI：10.13700/j.bh.1001-5965.2026.0542
收稿日期：2026-07-31
网络首发日期：2026-11-17
引用格式：曾可，余旺盛，秦先祥，龙思宇，白龙飞，杨少波．OH-Track:一种具有自适应遮挡感知能力的鲁棒多目标跟踪算法[J/OL]．北京航空航天大学学报．https://doi.org/10.13700/j.bh.1001-5965.2026.0542



**网络首发**：在编辑部工作流程中，稿件从录用到出版要经历录用定稿、排版定稿、整期汇编定稿等阶段。录用定稿指内容已经确定，且通过同行评议、主编终审同意刊用的稿件。排版定稿指录用定稿按照期刊特定版式（包括网络呈现版式）排版后的稿件，可暂不确定出版年、卷、期和页码。整期汇编定稿指出版年、卷、期、页码均已确定的印刷或数字出版的整期汇编稿件。录用定稿网络首发稿件内容必须符合《出版管理条例》和《期刊出版管理规定》的有关规定；学术研究成果具有创新性、科学性和先进性，符合编辑部对刊文的录用要求，不存在学术不端行为及其他侵权行为；稿件内容应基本符合国家有关书刊编辑、出版的技术标准，正确使用和统一规范语言文字、符号、数字、外文字母、法定计量单位及地图标注等。为确保录用定稿网络首发的严肃性，录用定稿一经发布，不得修改论文题目、作者、机构名称和学术内容，只可基于编辑规范进行少量文字的修改。

**出版确认**：纸质期刊编辑部通过与《中国学术期刊（光盘版）》电子杂志社有限公司签约，在《中国学术期刊（网络版）》出版传播平台上创办与纸质期刊内容一致的网络版，以单篇或整期出版形式，在印刷出版之前刊发论文的录用定稿、排版定稿、整期汇编定稿。因为《中国学术期刊（网络版）》是国家新闻出版广电总局批准的网络连续型出版物（ISSN 2096-4188，CN 11-6037/Z），所以签约期刊的网络版上网络首发论文视为正式出版。

# OH-Track:一种具有自适应遮挡感知能力的鲁棒多目标跟踪算法

曾可[1]，余旺盛[2,✉]，秦先祥[2]，龙思宇[1]，白龙飞[1]，杨少波[2]

（1. 空军工程大学 研究生院，陕西省西安市 710051；2. 空军工程大学 信息与导航学院，陕西省西安市 710077）
*通信作者 余旺盛，E-mail：xing_fu_yu@sina.com

**摘要** 针对多目标跟踪(Multi-Object Tracking, MOT)在遮挡场景下面临的轨迹漂移和身份频繁切换等严峻挑战，以SFTrack 为基本框架，提出一种处理遮挡场景的多目标跟踪算法(Occlusion Handling-Track, OH-Track)。首先，设计鲁棒遮挡感知模块(Robust Occlusion-Aware Module, ROAM)，通过几何前后景判定与运动方向一致性过滤，为每个目标生成连续的遮挡系数。其次，提出遮挡感知自适应卡尔曼滤波(Occlusion Aware Adaptive-Kalman Filtering, OAA-KF)，利用瞬时加速度动态调节过程噪声，并融合检测置信度与遮挡系数自适应缩放观测噪声，使滤波器在遮挡和检测退化时自主调整预测与观测的信任权重。在关联阶段，第一级引入遮挡感知偏移(Occlusion-Aware Offset, OAO)模块，基于遮挡系数修正 IoU 匹配代价，抑制被遮挡轨迹的错误匹配；第二级采用遮挡感知代价补偿(Occlusion-Aware Cost Compensation, OACC)机制，主动压低遮挡轨迹的匹配代价，补偿外观特征带来的距离虚高，减少漏匹配与轨迹断裂。在 VisDrone2019 和 DanceTrack 两个遮挡密集型数据集上的实验表明，OH-Track 在 MOTA 和 IDF1 上性能均取得显著提升。相较于基准 SFTrack，在 VisDrone2019 上，MOTA，IDF1 分别提升 0.20%和 0.35%。在 DanceTrack 上，采用两种检测器进行评估：使用低性能检测器 YOLO11s 时，MOTA，IDF1 分别提升 0.41%与 2.61%；使用经微调的高性能 YOLOX 检测器时，MOTA，IDF1 分别提升 0.10%和 2.99%。以上结果充分验证了所提方法在保持身份一致性和应对遮挡方面的有效性。

**关键词** 多目标跟踪，遮挡处理，卡尔曼滤波，TbD，数据关联，身份保持
中图分类号 TP391.41 文献标志码 A DOI：10.13700/j.bh.1001-5965.2026.0542

# OH-Track: A Robust Multi-Object Tracking Algorithm with Adaptive Occlusion Awareness

ZENG Ke[1], YU Wangsheng[2,✉], QIN Xianxiang[2], LONG Siyu[1], BAI Longfei[1], YANG Shaobo[2]

(1. Graduate School, Air Force Engineering University, Xi'an 710051, Shaanxi, China; 2. School of Information and Navigation, Air Force Engineering University, Xi'an 710077, Shaanxi, China)

*E-mail: xing_fu_yu@sina.com

**Abstract** To address the severe challenges of trajectory drift and frequent identity switches in multi-object tracking (MOT) under occlusion scenarios, this paper proposes an occlusion handling multi-object tracking algorithm, termed Occlusion Handling-Track (OH-Track), based on the SFTrack framework. First, a Robust Occlusion-Aware Module (ROAM) is designed to generate continuous occlusion coefficients for each target through geometric foreground-background discrimination and motion-direction consistency filtering. Second, an Occlusion Aware Adaptive-Kalman Filtering (OAA-KF) method is proposed, which dynamically adjusts the process noise using instantaneous acceleration and adaptively scales the observation noise by integrating detection confidence and occlusion coefficients, enabling the filter to autonomously adjust the trust weights between prediction and observation under occlusion and detection degradation. In the association stage, an Occlusion-Aware Offset (OAO) module is introduced at the first stage to correct the IoU matching cost based on occlusion coefficients

收稿日期：2026-07-31
基金项目：陕西省自然科学基础研究计划项目（2025JC-YBMS-255、2024JC-YBQN-0699）
Fund：Shaanxi Provincial Natural Science Basic Research Program (2025JC-YBMS-255 ， 2024JC-YBQN-0699)
网络首发时间：2026-11-17 17:05:47 网络首发地址：https://link.cnki.net/urlid/11.2625.V.20261117.1653.001

and suppress erroneous matching of occluded trajectories. At the second stage, an Occlusion-Aware Cost Compensation (OACC) mechanism is adopted to proactively reduce the matching cost of occluded trajectories, compensate for the inflated distance caused by appearance features, and reduce missed matching and trajectory fragmentation. Experiments on two occlusion-intensive datasets, VisDrone2019 and DanceTrack, show that OH-Track achieves significant improvements in both MOTA and IDF1. Compared with the baseline SFTrack, MOTA and IDF1 are improved by 0.20% and 0.35%, respectively, on VisDrone2019. On DanceTrack, when evaluated with two detectors, the improvements are as follows: with the low-performance detector YOLO11s, MOTA and IDF1 are improved by 0.41% and 2.61%, respectively; with the fine-tuned high-performance YOLOX detector, MOTA and IDF1 are improved by 0.10% and 2.99%, respectively. These results fully verify the effectiveness of the proposed method in maintaining identity consistency and handling occlusion.
**Key words** multi-object tracking, occlusion handling, Kalman filtering, TbD, data association, identity preservation

多目标跟踪旨在从视频序列中恢复所有感兴趣目标的运动轨迹，其核心目标是为每帧中检测到的对象分派并保持独一无二的身份标识(ID)，从而构建起目标完整的运动轨迹[1]。如今，基于检测的跟踪(Tracking-by-Detection, TbD)范式占据主导，如何跨帧可靠地关联检测结果已成为制约跟踪性能的关键瓶颈。位置关联作为该类方法的骨架，通过位置预测器[2]，即使用运动模型预测目标在下一帧的位置，并利用空间一致性度量[3]与匈牙利匹配算法[4]完成数据关联，因其计算高效且对密集场景适应性强而被广泛采用。典型的 SORT(Simple Online and RealtimeTracking)[5]算法以卡尔曼滤波器(Kalman Filter, KF)作为位置预测器，以马氏距离或交并比(IoU)作为相似度度量，再通过匈牙利算法求解最优匹配。这一架构虽在常规场景中表现稳健，却对遮挡等复杂干扰极为敏感。

遮挡给位置关联带来的冲击主要体现在三个层面。其一，同类目标发生相互遮挡时，检测器难以区分前景与背景，极易输出截断的边界框或直接漏检被遮挡目标，导致输入的观测本身已包含严重误差。其二，离散且线性的卡尔曼滤波器对检测质量高度依赖：当低质量检测框频繁出现(如遮挡导致的只包含部分身体的边框)，KF 的更新步骤会将这些有偏观测盲目融合进状态向量，使预测轨迹逐渐偏离真实运动，进而引发轨迹漂移甚至身份跳变。其三，在关联阶段，不准确的检测与不稳定的预测会共同造成位置代价矩阵的混淆，即矩阵无法准确反映检测与轨迹之间的关联，使得匈牙利算法极易做出错误指派，最终表现为频繁的 ID 切换[6]。

针对上述困境，本文以 SFTrack[7]为基准，提出 OH-Track，主要贡献如下：首先提出鲁棒遮挡感知模块(ROAM)，其通过比较两框的底部纵坐标与面积判定前后景关系，再引入帧间速度向量的余弦相似度滤除同向同行目标的干扰；仅当几何上构成前后景且运动方向不一致时，才认定存在真实遮挡；其次提出遮挡感知自适应卡尔曼滤波，通过双通道机制动态调节过程噪声 Q 与观测噪声 R，使滤波器能够在遮挡与检测退化时自主调整对预测和观测的信任偏向。最后引入 OAO 模块和提出 OACC 模块，分别用于 SFTrack 的一级和二级关联，OAO 模块将原始 IoU 相似度与由遮挡系数导出的不可信度进行加权融合，系统性提高对被遮挡轨迹的整体关联代价，优先保障未遮挡轨迹的关联准确性，而 SFTrack 在第二级关联时引入了颜色直方图相似度和缩放图像的均方误差，但在遮挡发生时，两者匹配的代价通常会提高，而 OACC 模块强制降低被遮挡轨迹的匹配阈值，使该轨迹能够成功匹配检测框。

## 1 相关工作

运动模型是 TbD 范式下跟踪器的核心组件之一，其作用是利用历史状态信息预测目标在当前帧的可能位置，该预测结果随后作为检测-轨迹匹配的重要依据。卡尔曼滤波(Kalman Filter, KF)因其在线性高斯假设下的最优估计特性，自 1960 年提出以来便被广泛应用于目标跟踪领域[8]。然而，现实世界中的目标运动往往是非线性的，标准卡尔曼滤波的线性假设在目标发生长时间遮挡或运动方向剧烈变化时会产生显著的估计误差。

在基于卡尔曼滤波的经典跟踪器中，SORT(Simple Online and Realtime Tracking)是 TbD 范式下最具影响力的早期工作之一，由 Bewley 等人于 2016 年提出[5]。SORT 采用极简框架，其核心包括卡尔曼滤波用于运动状态估计，以及基于检测框交并比(IoU)的匈牙利算法用于数据关联。SORT 证明了仅

依赖运动信息的轻量级方案即可在实时条件下取得具有竞争力的性能，但其在目标长时间遮挡时身份切换严重。为克服这一不足，DeepSORT 由 Wojke 等人于 2017 年提出，在 SORT 基础上集成了深度外观特征，通过预训练 CNN 提取检测框描述子，并将马氏距离(运动)与余弦距离(外观)加权融合进行关联，有效减少了身份切换次数[9]。ByteTrack 由 Zhang 等人于 2022 年提出，其创新在于关联几乎每一个检测框的策略，即先关联高置信度检测框，再利用低置信度检测框恢复被遮挡目标，从而减少漏检和碎片化轨迹[10]。OC-SORT(Observation-Centric SORT)由 Cao 等人于 2023 年提出，重新思考了卡尔曼滤波在无观测更新时的误差累积问题，通过引入以观测为中心的虚拟轨迹修正机制，在遮挡和非线性运动场景下显著提升了鲁棒性[11]。StrongSORT 对 DeepSORT 进行了系统性的模块化改进，并提出了 AFLink(无外观全局关联)和 GSI(高斯平滑插补)两个辅助算法，提供了强劲且公平的基线[12]。

除标准卡尔曼滤波外，研究者还发展了多种非线性滤波变体以应对更复杂的运动模式。扩展卡尔曼滤波(Extended Kalman Filter, EKF)通过对非线性系统进行一阶泰勒展开实现线性化。有研究者提出了自适应扩展卡尔曼滤波(Adaptive EKF)，通过动态调整过程噪声来平衡运动模型与传感器观测之间的信任，从而提升非线性机动目标的跟踪精度[13]；此外，Center Track-EKF 将 EKF 集成到 Center Track 框架中，用于改善多目标跟踪中的运动预测[14]。无迹卡尔曼滤波(Unscented Kalman Filter, UKF)则采用确定性采样的 Sigma 点传递均值和协方差，避免了 EKF 的线性化误差。在 SORT 框架基础上，有工作构建了多无迹卡尔曼滤波跟踪器，实验表明其在密集同向目标场景下优于 KF 和 EKF[15]；LUKF-Track 将 UKF 与李群流形相结合，提出轻量级且与外观无关的方法，在 MOT17、MOT20 和 DanceTrack 等具有高度非线性运动的数据集上取得了领先结果[16]。

然而，这些方法仍会受到遮挡导致的负面影响，尤其在目标密集和遮挡频率高的场景下，该问题更为明显。在本文中，我们直接分析遮挡的严重程度，以减轻因遮挡引起的边界框偏移和 ID 跳变。

## 2　本文算法

本文各个核心模块并非相互独立，遮挡感知自适应卡尔曼滤波（OAA-KF）、遮挡感知便宜模块（OAO）以及遮挡感知补偿机制（OACC）三个模块均依赖鲁棒遮挡感知模块（ROAM）输出的遮挡系数。首先，鲁棒遮挡感知模块（ROAM）处于最前端，其核心作用是量化每个目标当前帧的遮挡严重程度。该系数是后续所有自适应操作的先决条件，只有获取此刻遮挡状态后，滤波器与关联策略才能做出有针对性的调整。其次，遮挡感知自适应卡尔曼滤波（OAA-KF）以 ROAM 输出遮挡系数、检测置信度以及加速度幅值为依据，动态调节观测噪声与过程噪声。该步骤在数据关联之前完成，因为滤波器的预测状态直接决定了后续匹配中检测框与预测轨迹的代价矩阵。最后，两级关联优化策略（OAO+OACC）将 ROAM 提供的遮挡系数与 OAA-KF 优化后的预测结果共同作为输入：一级关联利用遮挡系数主动压低被遮挡轨迹的 IoU 匹配分以抑制误匹配，二级关联则利用遮挡系数补偿因外观退化而虚高的匹配代价。可见，ROAM 为后续模块提供遮挡感知基础条件，OAA-KF 为关联阶段的模块提供可靠的位置先验，各大模块协同优化最终的跟踪结果。

### 2.1 鲁棒遮挡感知模块

本文提出鲁棒遮挡感知模块(ROAM)，在逐帧跟踪过程中为每个目标生成一个连续遮挡系数$o_i \in [0,1]$。该系数表示目标当前被其他目标遮挡的最大面积比例，用于在数据关联和状态更新中对被遮挡目标进行特殊处理，减少因遮挡造成的身份切换和轨迹中断。

当目标经历未被遮挡→被遮挡→遮挡消失的状态变化时，其检测框往往在消失前后呈现显著跳变：遮挡发生前，检测框紧贴目标真实轮廓；经遮挡期间纯运动模型预测后，重新出现的检测框与预测位置之间可能产生较大残差。这种运动状态的突变若不加区分地输入卡尔曼滤波更新，会导致状态估计出现剧烈抖动甚至发散。因此，本文通过三个方面判断是否遮挡，首先是前景优先，遮挡者通常位于被遮挡者前方，反映在图像上表现为更靠近图像底部；其次是空间交叠约束，遮挡必须伴随检测框之间的显著交叠；最后是运动差异性，当两目标运动方向高度一致(如结伴行人)时，即使交叠也往往不是真正的相互遮挡，只有在运动方向差异较大(如舞蹈视频中不同目标时常相反运动并造成遮挡)

时，判断遮挡发生并计算被遮挡目标的遮挡系数。具体如图 1 所示。

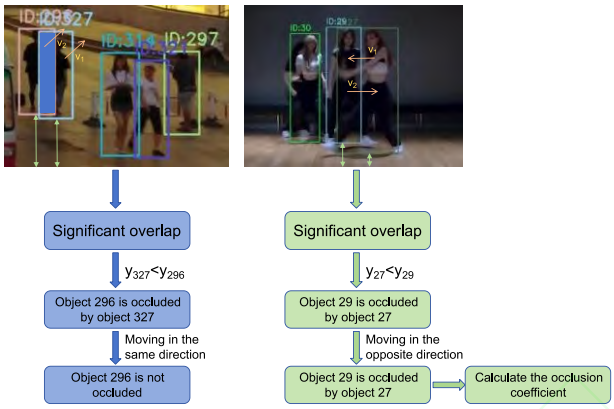
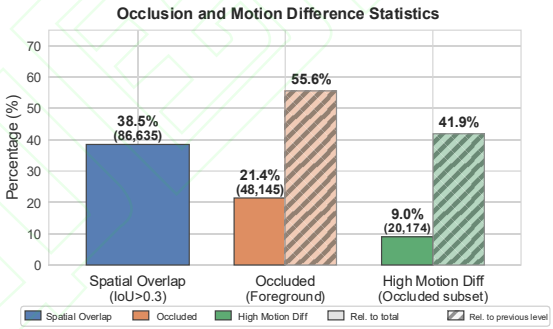


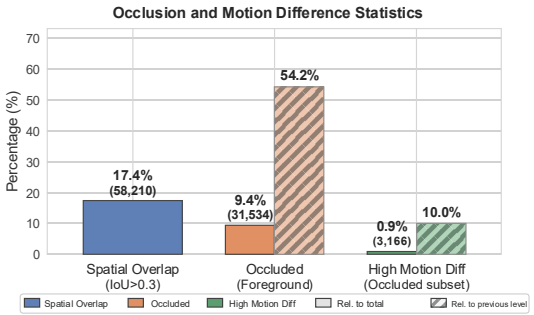

图 1 遮挡系数计算流程
Fig.1 Calculation process of the occlusion coefficient

我们在 DanceTrack 与 VisDrone2019-MOT 数据集上对遮挡判断的三个方面进行统计，统计结果如图 2 所示。

其中(a)为 DanceTrack 数据集的统计结果，结果表明，该数据集中，有大量目标都存在与其他目标发生空间交叠的情况，占所有目标数量的 38.5%，在发生交叠的目标中，又有约 55.6%的目标判定为被遮挡，本文主要处理从未遮挡到遮挡再到脱离遮挡这一情形，主要通过运动差异性判断该情形，因此还统计出运动差异性过大的目标占被遮挡目标的 41.9%，说明本文算法对该数据集有强大的适配性；(b)为 VisDrone2019-MOT 数据集的统计结果，相对来说，被遮挡目标的占比较低于 DanceTrack，占所有目标的 9.4%，也存在不少遮挡情况，但运动差异性大的运动目标较少，因此算法在该数据集上的性能提升较弱。




(a)DanceTrack 数据集统计结果




(b)VisDrone2019-MOT 统计结果

图 2 数据集统计数据结果
Fig.2 Statistical results of the datasets

ROAM 计算过程如下：

假设第 $t$ 帧存在 $N$ 个目标，其边界框与身份标识构成集合 $\left\{(id_i^t, b_i^t)\right\}$，其中 $b_i^t=(x_{i,1}^t, y_{i,1}^t, x_{i,2}^t, y_{i,2}^t)$。为每一目标维护其历史中心坐标，目标$i$在当前帧的中心点定义为：

$$\mathbf{c_i^t}=(\frac{x_{i,1}^t+x_{i,2}^t}{2},\frac{y_{i,1}^t+y_{i,2}^t}{2}) \tag{1}$$

若目标$i$在上一帧存在且被成功关联，则其速度向量为：

$$\mathbf{v_i^t}=\mathbf{c_i^t}-\mathbf{c_i^{t-1}} \tag{2}$$

否则 $\mathbf{v_i^t}$=0。两目标$i$与$j$之间的运动相似度定义为归一化余弦相似度：

$$S(\mathbf{v}_i^t,\mathbf{v_j^t})=\begin{cases}\frac{1}{2}\left(\frac{\mathbf{v_i^t}\mathbf{v_j^t}}{\left\|\mathbf{v_i^t}\right\|\left\|\mathbf{v_j^t}\right\|}+1\right), \left\|\mathbf{v_i^t}\right\|>\varepsilon, \left\|\mathbf{v_j^t}\right\|>\varepsilon \\ 1, Otherwise\end{cases} \tag{3}$$

其中 $\varepsilon=10^{-6}$ 防止除零。运动相似度越高，表示二者运动方向越一致。

对于任意两边界框框 $b_i^t$ 与 $b_j^t$，首先计算标准交并比 $IoU(b_i^t, b_j^t)$。同时，通过几何先验判断目标$j$是否为目标$i$的前景遮挡物，计算过程如下：

$$O_j \prec O_i \Leftrightarrow y_{j,2}^t > y_{i,2}^t \tag{4}$$

其中 $O_j \prec O_i$ 表示目标 $i$ 被目标 $j$ 遮挡，目标 $i$ 被目标 $j$ 遮挡的面积比例定义为重叠面积占自身面积之比：

$$\omega_{i,j}=\frac{b_i^t \cap b_j^t}{b_i^t} \tag{5}$$

其中，$b_i^t$ 与 $b_j^t$ 分别为目标 $i$ 与目标 $j$ 的边界框所占面积。综合以上要素，目标 $i$ 的遮挡分数由所有满足目标的目标 $j$ 的最大重叠比例给出：

$$\tilde{\omega}_{i,j}=\begin{cases}\omega_{i,j}, IoU\left(b_i^t, b_j^t\right)>\tau, O_j \prec O_i, S(v_i^t, v_j^t)<\tau_m \\ 0, Otherwise\end{cases} \tag{6}$$

$$o_i=\min\left(1, \max\left(\tilde{\omega}_{i,j}\right)\right) \tag{7}$$

其中 $\tau$ 与 $\tau_m$ 分别为 IoU、运动相似度阈值，$\tilde{\omega}_{i,j}$ 为遮挡分数，表示目标 $i$ 与目标 $j$ 遮挡的严重程度，$o_i$ 为优化后的遮挡分数，限制 $\tilde{\omega}_{i,j}$ 不大于 1。最后，当前帧所有目标的历史中心更新为$c_i^t$，供下一帧使用。

ROAM 输出的遮挡系数 $\tilde{\omega}_{i,j}$ 本质上是对目标被遮挡面积比例的瞬时估计，其准确性受检测框精度和运动估计质量的双重影响，不可避免存在误判风险。若该系数直接被后续模块不加防护地使用，将可能引发两类负面影响：（1）对滤波环节的影响——若遮挡系数被高估（将未遮挡目标误判为遮挡），OAA-KF 会过度放大观测噪声，使滤波器对高质量检测反应迟钝，导致轨迹偏移加剧；反之若 $\sigma_i$ 被低估（真实遮挡被判别为未遮挡），滤波器则会盲目信任受污染的观测框，同样导致轨迹偏移。（2）对关联环节的影响——遮挡系数被高估会使一级关联中被遮挡轨迹的 IoU 相似度被过度压低，导致真实匹配被错误抑制；$\sigma_i$ 低估则会使二级关联中本应放大的代价未能得到有效补偿，致使遮挡轨迹因外观退化而漏匹配。

因此为判断目标是否被真实遮挡，从前景优先，空间交叠以及运动差异性三个方面着手实现，解决目标从未被遮挡→被遮挡→遮挡消失状态变化导致的轨迹偏移与断裂。首先，ROAM 自身的前端验

证通过同时施加空间交叠阈值 $\tau$ 与运动方向一致性阈值 $\tau_m$，仅在几何与运动双重证据支撑下才认定遮挡发生，从源头减少了因随机交叠或结伴同行引起的误触发。

### 2.2 遮挡感知自适应卡尔曼滤波(Occlusion-Aware Adaptive Kalman Filtering, OAA-KF)

在标准卡尔曼滤波多目标跟踪框架中，过程噪声协方差$Q$与观测噪声协方差$R$通常仅根据目标高度静态缩放，无法动态响应目标的运动状态变化与观测质量波动。当目标发生急转、加速等机动时，恒定速度模型的预测偏差急剧增大，但固定的$Q$迫使滤波器仍过度信赖失准的预测，导致正确的观测被错误排除，轨迹随之中断，并在后续重识别时产生身份切换(ID Switch)。与此同时，固定的$R$对低置信度或受遮挡污染的检测与高质量检测赋予同等更新权重，易使轨迹状态被噪声观测污染而偏离真实目标，进一步增加错误关联与身份混淆的风险。此外，固定的噪声参数使得新息协方差矩阵无法反映当前时刻的实际不确定性，造成门控阈值系统性失配，从而大幅推高错误关联率。上述问题使得非自适应卡尔曼滤波器在复杂场景下身份切换次数显著增加，为此，本文引入一种根据瞬时加速度与检测可信度动态调整噪声协方差的自适应机制，以有效抑制 IDS。

#### 2.2.1 加速度驱动过程噪声自适应

目标运动建模为恒定速度模型，其基础过程噪声协方差 $Q$ 由目标高度$h$按 DeepSORT 标准生成：

$$Q = diag(\sigma_p^2, \sigma_p^2, \sigma_a^2, \sigma_p^2, \sigma_v^2, \sigma_v^2, \sigma_{\dot{a}}^2, \sigma_v^2) \tag{8}$$

其中 $\sigma_p^2 = \frac{h}{20}$ 为位置分量 $(x, y, h)$ 的过程噪声标准差，$\sigma_a = 10^{-2}$ 为宽高比$a$的过程噪声标准差，$\sigma_v = \frac{h}{160}$ 为速度分量$(\dot{x}, \dot{y}, \dot{h})$的过程噪声标准差，$\sigma_{\dot{a}} = 10^{-5}$ 为宽高比变化速度$\dot{a}$的过程噪声标准差。

本文利用当前帧与上一帧的速度差计算瞬时加速度幅值 $a$ (像素/帧 ²)：

$$a = \left\| v_t - v_{t-1} \right\| \tag{9}$$

定义过程噪声放大因子：

$$q = 1 + \alpha_Q \times \tanh(a) \tag{10}$$

其中 $\alpha_Q$ 为最大额外放大倍数。最后计算实际过程噪声协方差：

$$\tilde{Q} = q \times Q \tag{11}$$

该设计使$\tilde{Q}$随加速度幅值平滑增长，当 $a = 0$ 时 $q = 1$，退化为基础噪声；当 $a \gg 1$ 时 $q \approx 1 + \alpha_Q$，预测不确定性显著增加。由此，滤波器在后续更新中自动加大对当前观测的依赖程度，从而快速补偿机动引起的预测偏差。

#### 2.2.2 融合置信度与遮挡的观测噪声自适应

检测结果的可靠性受检测器置信度 $c \in [0,1]$ 和目标遮挡程度 $o \in [0,1]$(0 无遮挡，1完全遮挡)共同影响。常规处理往往忽略遮挡或单独利用置信度，难以在遮挡初期、检测框严重变形时仍保持跟踪稳定性。本文提出将二者同时纳入观测噪声的调节中。

首先定义不可靠性因子：

$$\alpha_{conf} = 1 - c, \alpha_{occ} = o_i \tag{12}$$

其中$c$为检测框置信度，$o_i$为上节计算得到的遮挡分数，表示当前目标$i$被遮挡目标遮挡的严重程度，然后分别计算由置信度低和遮挡严重引起的噪声放大倍数：

$$r_{conf} = 1 + \left( R_{\max}^{conf} - 1 \right) \cdot \alpha_{conf} \tag{13}$$

$$r_{occ} = 1 + \left( R_{\max}^{occ} - 1 \right) \cdot \alpha_{occ} \tag{14}$$

其中 $R_{\max}^{conf}$ 和 $R_{\max}^{occ}$ 为最大放大倍数，最终的观测噪声放大因子采用最保守策略，即取两者最大值：

$$r_{scale} = \max(r_{conf}, r_{occ}) \tag{15}$$

更新步的观测噪声协方差由基础观测噪声$R$缩放得到：

$$\tilde{R} = r_{scale} \cdot R, R = diag(\sigma_p^2, \sigma_p^2, \sigma_{a,obs}^2, \sigma_p^2) \tag{16}$$

其中$\sigma_p^2 = \frac{h}{20}$为位置分量$(x, y, h)$的观测噪声标准差，$\sigma_{a,obs} = 10^{-1}$为宽高比观测噪声标准差。当$c = 1$且$o = 0$时，$r_{scale} = 1$，观测噪声不变；置信度下降或遮挡加剧时，$r_{scale}$增大，观测在状态更新中的权重被压低，滤波器转而更信任模型预测，有效抑制低质量观测对轨迹的干扰。

### 2.3 两级关联优化策略

跟踪关联阶段，本文采用 SFTrack 为基准框架，引入遮挡自适应卡尔曼滤波的同时，优化该算法的一级和二级关联策略。SFTrack 以 ByteTrack 为基线框架，通过运动补偿机制，传统手工特征完成低置信度检测关联以及优化跟踪初始化策略三个方法协同实现对多目标的稳定跟踪。

其中，运动补偿机制基于稀疏光流与特征点提取计算仿射变换矩阵，约束目标边界框宽高比保持不变，有效修正无人机，运动相机等剧烈、不规则运动带来的边界框失真问题，提升卡尔曼滤波的运动预测精度。第一次关联仅使用 IoU 作为匹配依据，第二级关联采用传统手工特征完成低置信度检测关联：用颜色直方图相似度与缩放图像均方误差(MSE)作为外观匹配依据，结合交并比(IoU)完成剩余轨迹与低置信度检测的二次关联，在低分辨率、遮挡场景下仍能保持稳定匹配。最后优化跟踪初始化策略打破现有方法仅从高置信度检测启动跟踪的局限，同时利用高、低置信度检测初始化新轨迹，并通过同类高置信度检测的相似度阈值过滤低置信度误报，在降低目标漏检率的同时控制误检数量。

需要指出的是，关联阶段的有效性高度依赖于前两步的输出：一方面，ROAM 提供的遮挡系数直接用于修正代价矩阵；另一方面，OAA-KF 输出的优化预测框决定了后续代价矩阵计算是否可靠。

#### 2.3.1 第一级关联优化策略

第一级匹配在原始 IoU 相似度的基础上引入了鲁棒遮挡感知模块(ROAM)输出的轨迹遮挡系数，对相似度矩阵进行自适应修正。该方法来自于 OA-SORT 中的 OAO 模块。由于遮挡会导致检测框畸变、前景背景混淆，传统仅用 IoU 计算位置关联代价时，会出现代价混淆：被遮挡目标的检测框不准，IoU 无法真实反映轨迹与检测的匹配度；相似目标、重叠目标易出现错误匹配，直接引发 ID 切换、ID 丢失；匈牙利算法按最小代价分配时，会把遮挡目标匹配给错误轨迹。因此 OAO 模块基于遮挡系数，动态修正 IoU 匹配代价，降低遮挡物体的错误匹配，其计算过程如下：

$$S_{i,j} = \tau_o (1 - o_i) + (1 - \tau_o) \cdot IoU(x_i, b_j) \tag{17}$$

其中$IoU(x_i, b_j)$表示轨迹$i$与检测$j$的交并比。ROAM 输出各轨迹的遮挡系数向量$o \in [0,1]$，$o_i$为轨迹$i$的瞬时遮挡程度(0 表示无遮挡，1 表示完全遮挡)。$\tau_o \in [0,1]$,定义为平衡因子。所有$S_{i,j}$被截断至$[0,1]$以保证相似度有界。最终该级关联使用的代价矩阵为:

$$C_{i,j} = 1 - S_{i,j} \tag{18}$$

当轨迹处于明显遮挡时，即使其预测框与某个检测框的 IoU 仍较高，$(1 - o_i)$的压低作用使相似度大幅下降，代价增大，从而在第一级关联中主动抑制遮挡轨迹的潜在误匹配，有效减少身份切换。

#### 2.3.2 第二级关联代价补偿策略

在目标被部分或完全遮挡时，其外观表征发生显著变化：遮挡物引入的干扰像素会污染目标的颜色直方图统计，破坏外观模板的特征分布，导致匹配距离异常增大。此时，若直接采用原始外观距离进行关联，即使是真实的目标-检测配对，其代价也可能因外观污染而虚高，从而被错误拒配。

针对这一问题，本文在第二级关联中采用与第一级相反的代价调控策略—遮挡感知代价补偿机制(OACC)，该机制主动压低遮挡轨迹的匹配代价，以补偿外观线索的退化。同样记轨迹$i$的遮挡系数为$o_i \in [0,1]$，其值越大表示目标受遮挡越严重。在计算融合距离时，对匹配代价施加以下缩放：

$$s_i = 1 + \gamma \cdot o_i \tag{19}$$

$$\tilde{D}_{i,j} = \frac{D_{i,j}}{s_i} \tag{20}$$

其中$\gamma$为惩罚因子，控制缩放程度；$s_i$为轨迹$i$的缩放系数；$D_{i,j}$为轨迹$i$与检测$j$的原始相似度；$\tilde{D}_{i,j}$为缩放后的代价。

第二级匹配依赖颜色直方图与外观模板等表观线索。遮挡引发的特征污染使得这些线索的判别力下降，相同身份间的表观距离被系统性高估。通过压低几何分量的权重，实质上放宽了对遮挡轨迹的关联阈值，以几何先验的让渡换取表观线索的容错空间，从而将那些因遮挡时外观干扰而距离增大、但确属同一目标的检测重新纳入可匹配范围，有效抑制了遮挡场景下的漏匹配与轨迹断裂。

### 2.4 算法流程

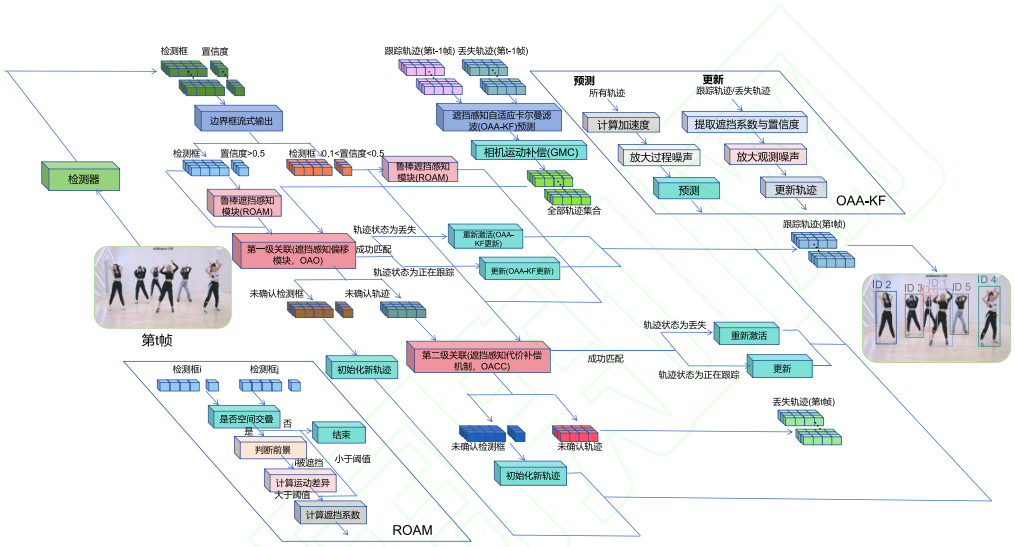



图 3 算法流程

Fig.3 Algorithm flow

算法流程如图 3 所示。首先将视频序列的每一帧输入检测器，实时完成目标边界框检测，输出带置信度的检测框；并将检测结果送入本文的跟踪算法。该算法首先将检测框分为低置信度检测框与高置信度检测框，其中置信度小于 0.5 且大于 0.1 被归类为低置信度检测框，置信度大于 0.5 被归类为高置信度检测框，随后两类检测框被分别送入鲁棒遮挡感知模块(ROAM)，并输出所有检测框的遮挡系数。与此同时，对上一帧中所有已有跟踪轨迹通过遮挡感知自适应卡尔曼滤波(OAA-KF)预测位置，并与检测框协同执行多级关联匹配：首先，在第一级关联中，构建高置信度检测框与所有预测跟踪轨迹的 IoU 代价矩阵，在该关联中，代价矩阵会被遮挡感知偏移模块(OAO)优化，降低遮挡发生时的错误匹配数，成功关联的轨迹根据前一帧状态（丢失状态与正在跟踪状态）进行激活或更新；第二级关联将未匹配轨迹与低置信度检测框融合 IoU、颜色直方图相似度、缩放图像均方误差计算代价矩阵再次匹配，在该关联中，使用遮挡感知代价补偿机制(OACC)抑制遮挡场景下的漏匹配与轨迹断裂，同样将成功关联的轨迹进行重新激活或更新，最终返回有效的跟踪结果并将轨迹送入下一帧，未被匹配的检测框（未确认检测框）会初始化新一个跟踪轨迹并送入下一帧，而未匹配的上一帧跟踪轨迹（未确认轨迹）被送入丢失轨迹集合中。

## 3 实验结果与分析

### 3.1 实验环境和数据集

为充分验证所提方法在处理遮挡场景下的有效性，本文选用 VisDrone2019 与 DanceTrack 两个公

开基准数据集进行实验。二者分别从无人机航拍与密集人群交互两个维度，为遮挡条件下的跟踪挑战提供了良好的测试平台。数据集示例如图 4 所示。

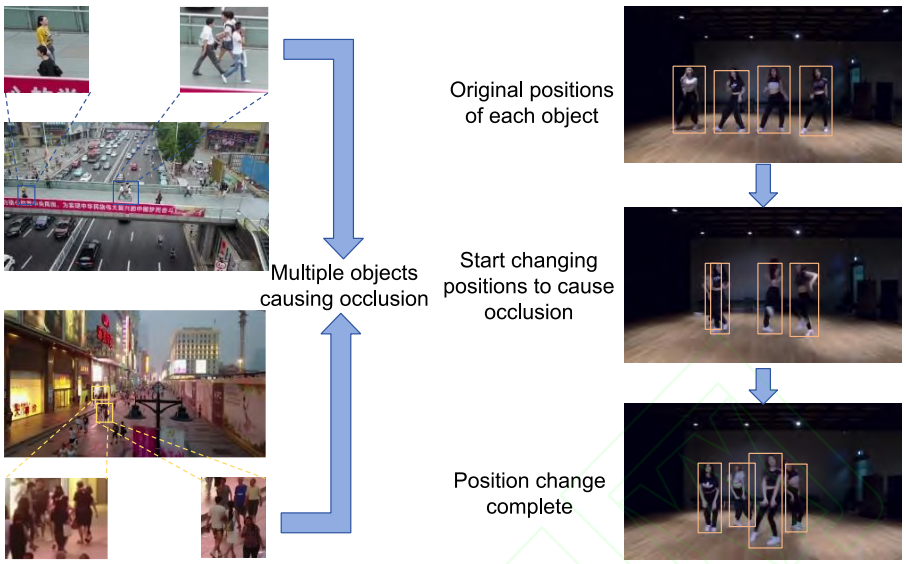



图 4　数据集示例图
Fig.4　Sample images from the datasets

VisDrone2019 数据集由天津大学机器学习与数据挖掘实验室 AISKYEYE 团队收集并发布[17]，是面向无人机视觉任务的大规模基准数据集。该数据集包含 288 个视频片段，共 261,908 帧图像，以及 10,209 张静态图像，由搭载不同型号摄像头的多种无人机平台在 14 个中国城市的不同场景、不同天气与光照条件下采集完成。数据集提供了超过 260 万个目标边界框的人工标注，涵盖行人、人、自行车、小汽车、货车、卡车、三轮车、带篷三轮车、公交车和摩托车等 10 个目标类别。尤为重要的是，VisDrone2019 的标注中明确包含了场景可见性、目标类别及遮挡等属性信息，这使其成为研究航拍视角下遮挡问题的理想平台。事实上，已有大量研究指出航拍图像中的目标普遍存在遮挡与重叠问题，并基于 VisDrone2019 数据集开展遮挡条件下的目标检测与跟踪研究。此外，该数据集中的目标尺度变化显著、分布密集，进一步加剧了遮挡场景下的跟踪难度。

DanceTrack 数据集由 Sun 等人于 2022 年提出[18]，是一个专门针对多目标跟踪中外观判别力不足问题而构建的大规模多人跟踪基准。该数据集包含 100 个视频序列，其核心设计理念在于：现有跟踪数据集中的目标大多具有可区分的外观特征，重识别(Re-ID)模型足以建立目标关联，但多目标跟踪方法应当在目标外观缺乏足够判别力时依然有效。DanceTrack 中的数据具有两个显著特征：外观高度相似，目标人物穿着极为相似甚至难以区分的服装；运动模式多样，目标处于复杂的运动与交互模式中。这些舞蹈视频中的人物频繁发生相对位置交换与严重遮挡。由于外观线索被极大削弱，跟踪算法被迫更多地依赖运动分析来完成目标关联，这使得 DanceTrack 成为检验跟踪方法在严重遮挡与外观歧义条件下鲁棒性的极具挑战性的基准。

综上所述，VisDrone2019 提供了无人机航拍视角下包含丰富遮挡标注的大规模多样化数据，而 DanceTrack 则聚焦于外观相似与复杂运动导致的密集遮挡场景。两个数据集从不同视角覆盖了遮挡问题的核心挑战，为本文所提方法的充分验证提供了坚实的数据基础。最后，实验环境配备 RTX4090 显卡。所有仿真都是基于 PyTorch 框架使用 Python 实现的，具体版本详情见表 1。

表1 实验环境

Table 1 Experimental environment

| ExperimentalEnvironment | Version |
|---|---|
| OperatingSystem | Windows10Pro |
| GPU | NVIDIAGeForceRTX4090(24G) |
| ProgrammingLanguage | Python3.9.21 |
| CUDA | CUDA11.3 |
| DeepLearningFramework | Pytorch1.12.0 |

## 3.2 定性分析

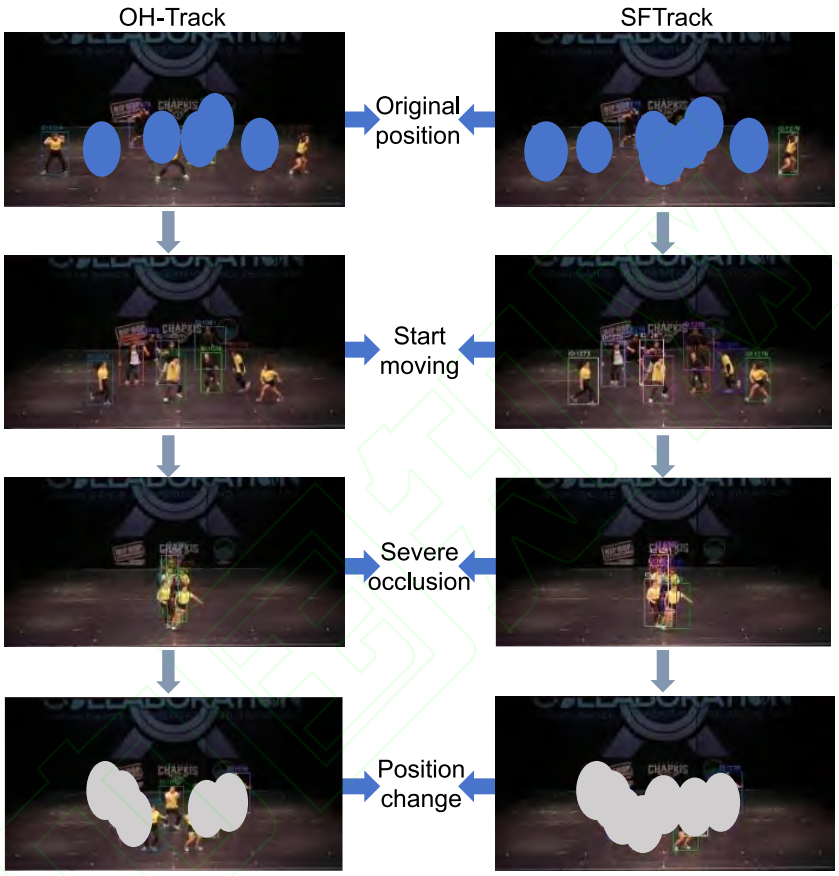



图 5 DanceTrack 的跟踪结果可视化，其中椭圆阴影为 ID 跳变的目标

Fig.5 Visualization of tracking results on DanceTrack, where targets with ID switches are marked by elliptical shading.




图 6 VisDrone2019 的跟踪结果可视化
Fig.6　Visualization of tracking results on VisDrone2019

为直观展示跟踪效果，本文从 DanceTrack 和 VisDrone2019 数据集中各选取一个序列的部分帧进行可视化，结果如图 5、6 所示。以图 5 为例，该序列共包含 9 个目标，选取的 4 帧图像分别对应目标初始位置、开始移动、形成严重遮挡以及目标位置改变四种典型情况。从图中可以观察到，本文算法发生身份跳变的目标仅有 5 个，而基准算法中出现身份跳变的目标达到 7 个，表明本文算法在发生遮挡时具有更强的身份保持能力。VisDrone2019 数据集可视化如图 6 所示，该序列包含大量密集小目标，该序列中的行人会出现大量遮挡，如图中放大的行人目标，SFTrack 在遮挡发生时都发生了身份跳变，比如 ID: 645 跳变为 683，ID: 632 跳变为 691，而本文算法仍能保持相应目标的 ID。

### 3.3 定量分析

#### 3.3.1 消融实验分析

表2 VisDrone2019-MOT数据集消融实验结果

Table 2 Ablation study results on the VisDrone2019-MOT dataset

| Method | MOTA | IDF1 | IDS | FN | FP | FPS |
|---|---|---|---|---|---|---|
| ①:SFTrack | 38.18% | 53.60% | 1151 | 193200 | 11899 | 9.65 |
| ②:①+OAA-KF | 38.18% | 53.75% | 1113 | 193300 | 11880 | 7.46 |
| ③:②+OAO | 38.31% | 53.76% | 1040 | 192885 | 11949 | 6.63 |
| ④:③+OACC | 38.38% | 53.95% | 1011 | 192579 | 12027 | 6.19 |
| ⑤:①+OACC | 38.22% | 53.71% | 1127 | 193014 | 12031 | 7.75 |
| ⑥:①+OAO | 38.30% | 53.73% | 1062 | 192874 | 11942 | 6.66 |

如表 2 所示，本文以 SFTrack 为基本框架，渐进式融合了 OAA-KF，OAO，OACC 等方法，并在 VisDrone2019-MOT 的测试集上进行跟踪，检测器采用的是 YOLOv11n，并在 VisDrone2019-DET 检测数据集上进行 200 轮的微调。方法②将 SFTrack 的标准卡尔曼滤波替换为遮挡感知自适应卡尔曼滤波，略微提高了 IDF1，并减少了 IDS，增强了跟踪算法的身份保持能力。其次，方法③在方法②的第一级匹配引入遮挡感知偏移模块，该模块基于遮挡系数，动态修正 IoU 匹配代价，降低遮挡物体的错误匹配，使 IDS 减少了 73，MOTA 提高了 0.13%。然后在方法③的第二级匹配引入遮挡感知代价补偿机制，将那些因遮挡时外观干扰而距离增大、但确属同一目标的检测重新纳入可匹配范围，有效抑制了遮挡场景下的漏匹配与轨迹断裂。最后将 OAO 模块与 OACC 机制分别引入基准模型 SFTrack，都得到了一定的性能提升，其中 OAO 模块的性能提升更大，达到 38.30%的 MOTA 以及 53.73%的 IDF1。

随着各个模块的逐步集成，渐进式提升了跟踪性能，但也不可避免地降低跟踪效率，除了基准算法 SFTrack，其余消融模型都引入了遮挡系数前置模块，由于遮挡系数的计算流程较为复杂，导致整体效率偏低。整体上看，OACC 机制对效率的影响极低，在消融模块中 FPS 最高，而 OAO 模块对效率的影响较大，主要归因于第一级关联的边界框较多，每个边界框在 OAO 模块中进行代价矩阵的计算。

表3 Danctrack数据集消融实验结果

Table 3 Ablation study results on the DanceTrack dataset

| Method | MOTA(%) | IDF1(%) | IDS | FN | FP | FPS |
|---|---|---|---|---|---|---|
| ①:SFTrack | 71.84/90.87 | 45.32/58.30 | 1853/2645 | 60668/15295 | 872/2596 | 13.30/5.13 |
| ②:①+OAA-KF | 71.94/91.01 | 46.14/59.25 | 1828/2511 | 60429/15126 | 922/2511 | 13.29/4.71 |
| ③:②+OAO | 72.19/90.97 | 47.77/61.29 | 1673/2499 | 60010/15080 | 925/2743 | 13.10/4.22 |
| ④:③+OACC | 72.25/90.97 | 47.93/61.29 | 1677/2499 | 59871/15080 | 927/2743 | 12.89/4.00 |
| ⑤:①+OAO | 72.11/90.98 | 46.67/60.80 | 1679/2486 | 60230/15088 | 892/2733 | 13.01/4.04 |
| ⑥:①+OACC | 71.86/90.87 | 45.40/58.30 | 1779/2645 | 60582/15295 | 924/2596 | 13.27/4.20 |

本文在 DanceTrack 数据集上进行了相同的消融实验与评估，检测器采用分别采用检测能力弱的 YOLOv11s 官方权重，类别只取行人；以及在 DanceTrack 上微调过的 YOLOX，具有极强的检测性能。如表 3 所示每个指标前者为检测能力弱的官方 YOLOv11s，后者为 DanceTrack 微调的强大检测器 YOLOX。首先是官方权重的 YOLOv11s，跟踪评估时只取行人类别，OAKF 使基线模型提高了 0.1%的 MOTA 和 0.82%的 IDF1，降低了 IDS，OAO 模块进一步提升了 0.25%的 MOTA 与 1.63%的 IDF1，IDS 降低了 155，OACC 也实现了各指标的提升。对于 DanceTrack 微调后的强大检测器 YOLOX，它的 MOTA 与 IDF1 普遍更高，消融实验上也实现了算法性能的渐进提升，但是由于检测

能力强大，低置信度的二级匹配极少，这也导致优化二级匹配的 OACC 机制的对算法的跟踪性能没有提升，不仅如此，在单独引入 OACC 和 OAO 时，弱检测器中 OAO 对性能提升的主要贡献在 MOTA，IDF1 提升较少，但在强检测器的情况下 IDF1 提升较大。而 OACC 仅在弱检测器时实现了跟踪性能的提升。

因此当检测器性能较低时，进入二次匹配的低置信度检测框较多，处于二次匹配阶段的 OACC 机制能够带来相应提升；当检测器性能已经很高时，低置信度检测框较少，很难影响到全局的指标计算，导致 OACC 的增益微弱，但仍能保持性能不下降；因此，OACC 更适合作为检测质量不足时的增强模块，或与弱检测器配合使用。

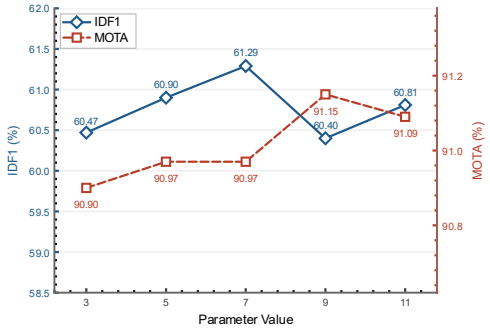
最后从实验结果可见，本文 OH-Track 在 VisDrone2019 与 DanceTrack 两个数据集上呈现差异化的增益形态：在 VisDrone2019 上，ID 切换次数相对基准 SFTrack 降低 140 次，但 MOTA 和 IDF1 仅提升 0.20 和 0.35 个百分点；而在 DanceTrack 上，使用低质量检测器的情况下 IDF1 相对提升达 2.61 个百分点，MOTA 提升 0.41 个百分点，若使用高质量检测器 IDF1 提升 2.99 个百分点，MOTA 提升提升 0.1 个百分点，显然本文算法在 DanceTrack 上表现更佳。

这是由于 VisDrone2019 为无人机航拍视角，目标尺度极小且分布极为密集，普通检测器对小目标的召回能力本就受限，压缩了 MOTA 的提升空间。与此同时，该数据集目标被遮挡的占比低于 DanceTrack 数据集，导致主要处理遮挡问题的 OH-Track 在 VisDrone2019 数据集上比较受限。

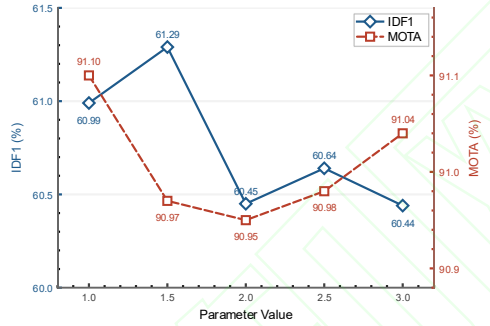
其次，DanceTrack 聚焦于地面舞蹈场景，其显著特征为目标运动剧烈和遮挡发生特别频繁。本文提出的遮挡感知自适应卡尔曼滤波（OAA-KF）通过瞬时加速度动态放大过程噪声，恰好应对了舞蹈中频繁的急停、转向和跳跃等非线性机动，有效避免了轨迹在快速运动中被错误预测，同时遮挡系数主导的动态放大观测噪声有效抑制了低质量检测框对卡尔曼滤波的干扰，二者协同运作，显著提高在 DanceTrack 上的跟踪精度。

与 VisDrone2019 的消融实验相同，DanceTrack 上的每帧处理速度也随模块递增而变慢，另外还可以看出高质量检测器极大拉低了 FPS，FPS 从 12—13 降至 4—5，说明了检测器的检测速度直接关系到跟踪效率的高低。

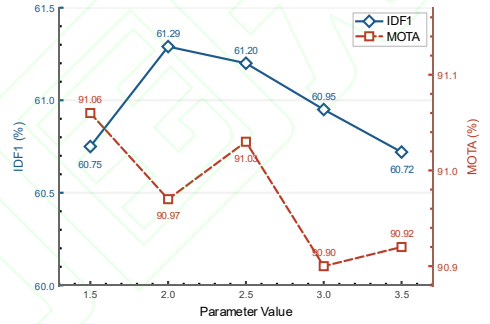
### 3.3.2 参数设置实验




(a) 过程噪声最大放大倍数$\alpha_Q$设置实验结果




(b) 置信度引起的观测噪声.
最大放大倍数实验结果




(c) 遮挡系数引起的观测噪声最大放大倍数实验结果

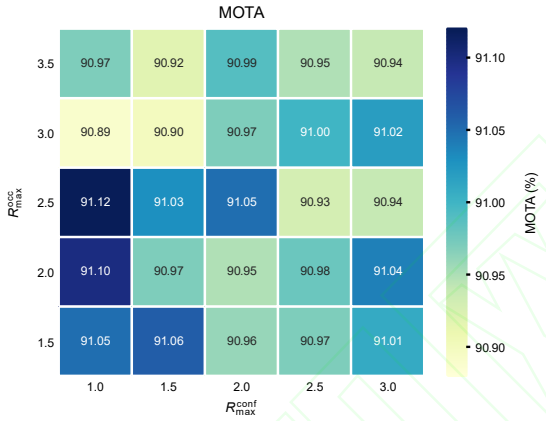
图 7 放大倍数设置实验结果
Fig.7 Experimental results for different magnification settings

本文在 DanceTrack 数据集进行参数设置实验，并使用 DanceTrack 上微调过的 YOLOX 作为检测器，具体实验如图 7 所示。(a)为不同$\alpha_Q$值的实验结果折线图，具体将参数$R_{max}^{conf}$与参数$R_{max}^{occ}$分别被随机固定为 1.5 与 2.0，并以 2 为步长取不同的$\alpha_Q$值进行实验，实验结果表明，综合 IDF1 与 MOTA 的结果，取$\alpha_Q = 7.0$为最佳，MOTA 达到最大值 61.29%，IDF1 仅略低于最大值 91.15%；(b)为不同$R_{max}^{conf}$值的实验结果折线图，参数$R_{max}^{occ}$依然被随机固定为 2.0，参数$\alpha_Q$为最佳值 7.0，并以 0.5 为步长取不同$R_{max}^{conf}$值进行实验，实验结果表明，综合 IDF1 与 MOTA 的结果，取$R_{max}^{conf} = 1.5$为最佳，MOTA 最大且远高于其他取值，IDF1 也较为优秀；最后(c)为不同$R_{max}^{occ}$值的实验结果折线图，参数$\alpha_Q$与参数$R_{max}^{conf}$分别取最佳值 7.0 与 1.5，并以 0.5 为步长取不同$R_{max}^{occ}$值进行实验，实验结果表明，综合 IDF1 与 MOTA 的结果，$R_{max}^{occ} = 2.0$与$R_{max}^{occ} = 2.0$的实验结果均表现优秀，但 MOTA 的指标综合性更强，因此取 MOTA 最大的$R_{max}^{occ} = 2.0$。

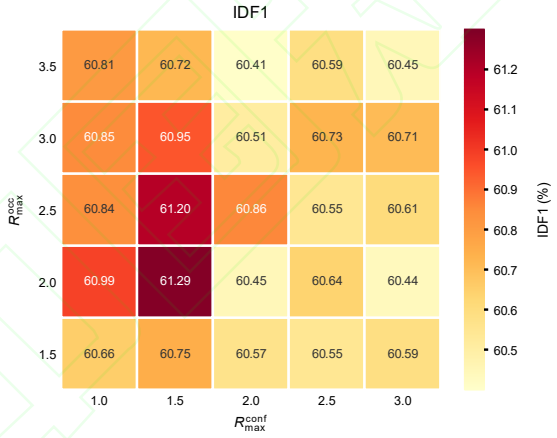
综合三组实验结果，最优参数组合确定为$\alpha_Q$=7.0、$R_{max}^{conf}$=1.5、$R_{max}^{occ}$=2.0，该组合在平衡身份保持

与检测精度上表现最佳，取得了最高的 IDF1 和稳健的 MOTA；虽然$\alpha_Q$=9.0 或$R_{max}^{conf}$=1.0 能带来极轻微的 MOTA 提升，但会导致 IDF1 显著下降，而$R_{max}^{occ}$的 MOTA 呈明显的倒 U 型曲线，超过 2.0 后性能逐步衰减，同时所有最优值均未触及搜索边界，说明该组合是真实的极值点而非截断值，因此直接采用该组参数。

最后，我们对$R_{max}^{conf}$、$R_{max}^{occ}$进行了正交实验并绘制 MOTA 分数与 IDF1 分数的热力图，以此说明这两个参数的交互效应。具体实验结果如图所示。




(a)MOTA 分数热力图




(b)IDF1 分数热力图

图 8 $R_{max}^{conf}$、$R_{max}^{occ}$正交热力图
Fig.8　Orthogonal heatmaps of $R_{\max}$and $R_{\max}$

在本组正交参数范围内，MOTA 波动较小，范围为 90.89%-91.12%；IDF1 波动相对更明显，范围为 60.41%-61.29%，最高值出现在$R_{max}^{conf}$=1.5、$R_{max}^{occ}$=2.0，为 61.29%，其次为$R_{max}^{conf}$=1.5、$R_{max}^{occ}$=2.5 的 61.20%。整体来看，$R_{max}^{conf}$=1.5 附近更有利于 IDF1，而$R_{max}^{conf}$=1.0 时 MOTA 略优；$R_{max}^{occ}$取 2.0，2.5 两个指标的综合表现均不错。

另外，在以上实验中，空间交叠阈值$\tau$、运动差异性阈值$\tau_m$、平衡因子$\tau_o$以及惩罚因子$\gamma$分别默认为 0.3、0.6、0.15 以及 1.5，其中$\tau_o = 0.15$参考 OA-SORT 的设置结果。因此本文在 DanceTrack 上对空间交叠阈值$\tau$、运动差异性阈值$\tau_m$分别分析，由于 OACC 机制在高性能检测器中提升微小，因此在检测结果较差的 VisDrone2019 对惩罚因子$\gamma$进行分析，具体如表 4 所示：

表4. 其它参数设置实验结果

Table 4 Experimental results for other parameter settings

(a) 空间交叠阈值$\tau$设置实验结果

| $\tau$ | 0.2 | 0.3 | 0.4 | 0.5 | 0.6 |
|---|---|---|---|---|---|
| MOTA | 90.97% | 90.97% | 90.98% | 90.97% | 90.97% |
| IDF1 | 61.29% | 61.29% | 61.42% | 61.34% | 61.34% |

(b) 运动差异性阈值$\tau_m$设置实验结果

| $\tau_m$ | 0.2 | 0.4 | 0.6 | 0.8 | 1.0 |
|---|---|---|---|---|---|
| MOTA | 90.97% | 90.93% | 90.98% | 91.01% | 91.13% |
| IDF1 | 61.29% | 61.28% | 61.42% | 61.39% | 61.00% |

(c) 惩罚因子$\gamma$设置实验结果

| $\gamma$ | 1.5 | 2.5 | 3.5 | 4.5 | 5.5 |
|---|---|---|---|---|---|
| MOTA | 38.38% | 38.40% | 38.41% | 38.41% | 38.42% |
| IDF1 | 53.95% | 53.93% | 53.90% | 53.90% | 53.85% |

整体来看，三组参数对性能的影响都较小，说明模型稳定性较好；空间交叠阈值$\tau$在 0.4 时 MOTA 和 IDF1 同时达到最高，此外在阈值小于 0.3 或大于 0.5 时指标变化微弱；而运动差异性阈值$\tau_m$在 1.0 时 MOTA 最高但 IDF1 最低，$\tau_m$=0.6 时 IDF1 最高，身份保持能力最强，同时 MOTA 不低；惩罚因子 γ 的影响非常微弱，γ 增大时 MOTA 略升、IDF1 略降，γ=3.5 或 4.5 较均衡。

### 3.3.3 对比实验分析

表5 VisDrone2019数据集对比实验结果

Table 5 Comparative experimental results on the VisDrone2019 dataset

| Method | MOTA | IDF1 | IDS | FN | FP |
|---|---|---|---|---|---|
| BoT-SORT[19] | 25.98% | 46.02% | 1185 | 193083 | 49932 |
| ByteTrack[20] | 24.60% | 40.27% | 2296 | 200432 | 46043 |
| BoostTrack[21] | 24.57% | 40.80% | 1626 | 213968 | 33281 |
| JdeTrack[22] | 23.87% | 39.00% | 2790 | 230431 | 17960 |
| OC-SORT[11] | 34.31% | 41.36% | 7078 | 189417 | 22686 |
| Hybrid-SORT[3] | 34.27% | 43.91% | 4027 | 199322 | 15990 |
| SFTrack | 38.18% | 53.60% | 1151 | 193200 | 11899 |
| Ours | 38.38% | 53.95% | 1011 | 192579 | 12027 |

从表 5 中可以看出，早期方法如 BoT-SORT、ByteTrack 等在无人机视角下整体精度较低，MOTA 普遍在 25%左右，且 FP 和 FN 均处于较高水平。其中 JdeTrack 虽然 FN 较低仅 17960，但 FP 高达 230431，导致 MOTA 仅为 23.87%，说明其在复杂场景中容易产生大量误检。OC-SORT 和 Hybrid-SORT 在 MOTA 上提升至约 34%，显著改善了检测与关联的整体精度，但 ID 切换次数急剧增加，主要是因为为引入相机运动补偿模块，致使 IDF1 偏低，反映出身份保持能力不足。

SFTrack 取得了此前最优的 38.18%MOTA 和 53.60%IDF1，同时将 IDS 控制在 1542 次，在检测精度和身份一致性之间实现了较好的平衡。本文方法在此基础上进一步提升了所有核心指标：MOTA 达到 38.38%，IDF1 达到 53.95%，均为所有对比方法中的最高值。具体而言，相比于 SFTrack，所提方法在 MOTA 上提升了 0.20 个百分点，IDF1 提升了 0.35 个百分点；同时，ID 切换次数从 1542 下降至 1011，降低约 34.4%，假正例从 198192 减少到 192579。假负例虽由 10861 略升至 12027，但综合假正例和 ID 切换的显著减少，整体跟踪准确率仍然取得领先，证明了本文算法在抑制误检与保持轨迹身份一致性方面的突出优势。

表6 DanceTrack数据集对比实验结果

Table 6　Comparative experimental results on the DanceTrack dataset

| Method | MOTA(%) | IDF1(%) | IDS | FN | FP |
|---|---|---|---|---|---|
| Boostrack | 60.39/86.99 | 29.02/45.70 | 2680/2556 | 80315/25988 | 6197/739 |
| Bytetrack | 67.99/90.79 | 36.35/56.03 | 2295/2690 | 62983/15937 | 6803/2106 |
| BoT-SORT | 69.96/90.49 | 37.18/59.26 | 2010/2948 | 59841/15115 | 5780/3358 |
| Jdetrack | 62.63/90.81 | 34.52/56.28 | 2397/2643 | 77830/16011 | 3906/2041 |
| OC-SORT | 72.63/87.63 | 38.29/53.61 | 3509/2751 | 56436/24027 | 1689/1062 |
| Hybrid-SORT | 68.44/89.21 | 45.05/64.74 | 1770/1731 | 68445/21250 | 839/1310 |
| SFTrack | 71.84/90.87 | 45.32/58.30 | 1853/2645 | 60668/15295 | 872/2596 |
| Ours | 72.25/90.97 | 47.93/61.29 | 1677/2499 | 59871/15080 | 927/2743 |

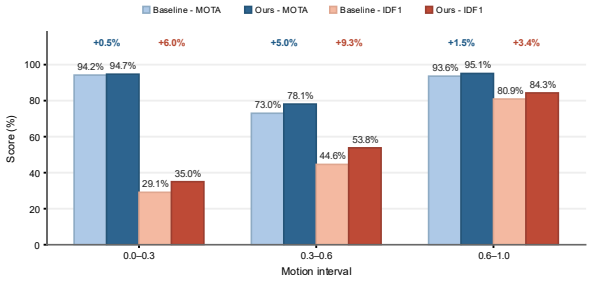
从表 6 中可以看出，检测器质量对跟踪性能具有决定性影响。当采用高质量检测器时，所有方法的 MOTA 均大幅提升至 87%以上，FN 和 FP 显著下降；而低质量检测器下的跟踪更能反映关联算法自身的鲁棒性。

在低质量检测器条件下，本文方法取得了 72.25%的 MOTA，与最高的 OC-SORT(72.63%)基本持平，但 IDF1 达到 47.93%，领先第二名 SFTrack2.61 个百分点，同时 IDS 仅为 1677，为所有方法中最低。相比之下，OC-SORT 虽然 MOTA 略高，但其 IDS 高达 3509，IDF1 仅为 38.29%，表明身份保持能力严重不足；SFTrack 的 IDF1(45.32%)与 IDS(1853)已属较优，但本文方法在身份一致性上的优势更为突出。这充分说明本文算法对低质量检测带来的漏检和误检具有极强的鲁棒性，能够有效抑制身份切换。

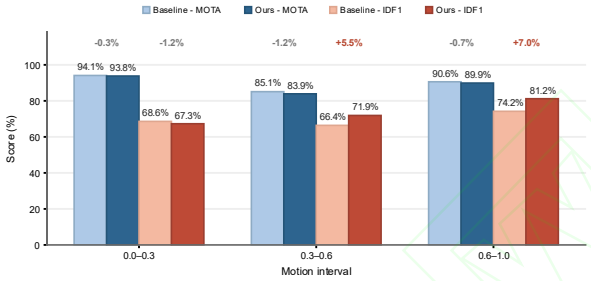
在高质量检测器条件下，本文方法取得了最高的 MOTA，达到 90.97%，较 SFTrack 和 BoT-SORT 有微弱但稳定的领先。IDF1 达到 61.29%，略低于 Hybrid-SORT 的 64.74%，但仍保持在第一梯队，且显著优于 SFTrack58.30%和 OC-SORT 的 53.61%。在 IDS 方面，本文方法的 2499 次虽高于 Hybrid-SORT 的 1731 次，但考虑到 MOTA 的领先以及高质量检测下更丰富的轨迹恢复，这一表现仍具有竞争力。总体而言，结合高 MOTA 和高 IDF1，本文方法在高质量检测下实现了检测精度与身份保持的良好平衡。

### 3.3.4 分层统计结果

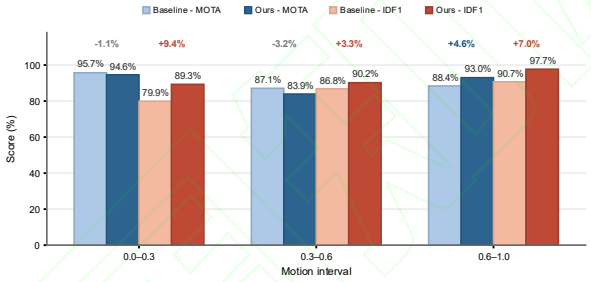
为体现遮挡自适应的有效性，本文在某个序列的每一帧的每个目标中取最大遮挡系数，若遮挡系数处于 0-0.3 区间，则该帧被视为轻度遮挡。相应地，在 0.3-0.6 区间视为中度遮挡，在 0.6-1.0 区间视为严重遮挡，本文随机选取 DanceTrack 的三个序列，分别是 dancetrack0019，dancetrack0047，dancetrack0079，实验结果如图 9 所示。




(a)dancetrack0019 分层统计实验结果




(b)dancetrack0047 分层统计实验结果




(c)dancetrack0079 分层统计实验结果

图 9 分层统计实验结果

Fig.9 Experimental results of stratified analysis

整体上看，本文算法对每个序列在不同遮挡系数区间计算的 MOTA 分数普遍提升偏低，甚至有少量降低，但 IDF1 分数总体提升较大，说明本文算法在轻度遮挡，中度遮挡，严重遮挡的情况下具有良好的身份保持能力。对于不同序列，本文算法在不同遮挡系数区间的提升有差距，在 dancetrack0019 上提升效果最好，三个区间的不同指标都得到提升，特别是 IDF1 分数提升较大。其次是 dancetrack0079，不仅 IDF1 分数提升较高，在严重遮挡时，也显著提高了 MOTA 分数。综上所述，本文算法的自适应策略是有效的，但不同序列的不同遮挡程度存在一定差异。

## 4. 结论

本文针对遮挡为多目标跟踪任务带来的诸多困境，例如 ID 跳变，轨迹断裂等问题，提出 OH-Track 多目标跟踪算法。该算法首先创新性地提出鲁棒遮挡感知模块(ROAM)，为后续工作提供有效的遮挡系数；其次将计算出的遮挡系数，目标加速度和置信度一同引入卡尔曼滤波预测阶段，三者协同调节观测噪声与过程噪声大小，实现卡尔曼滤波的精准预测，降低 ID 跳变与轨迹断裂；最后将鲁棒遮挡感知模块输出的遮挡系数同步传入关联阶段的遮挡感知偏移(OAO)模块以及遮挡感知代价补偿机制(OACC),两者分别在一级关联和二级关联优化代价矩阵，进一步提升跟踪器的性能。

本文算法使用两个遮挡发生频率高的公开数据集，分别是 VisDrone2019-MOT 与 DanceTrack，前者包含大量车辆以及人群的密集场景，存在大量遮挡；后者是专注与舞蹈视频的多目标跟踪，不同目标时常移动并变换位置，过程会造成严重遮挡。实验结果表明，本文算法在两个数据集上均实现了更

优秀的多目标跟踪性能，提升 MOTA 的同时也提升了 IDF1，具有强大的身份保持能力。甚至在 DanceTrack 使用两种检测器进行评估，分别是带有官方权重的 YOLOv11s 以及微调后的 YOLOX，无论是消融实验还是对比实验都展现出本文算法的有效性。与其他 TbD 范式的跟踪器相比，本文算法也具有强大的表现力，仅仅 OC-SORT 或 Hybrid-SORT 部分指标略高于本文算法。

## 参考文献(References)


[1] Luo W H,Xing J L,Milan A,et al.Multiple object tracking: A literature review[J].Artificial Intelligence,2021,293:103448.

[2] Du Y H,Wan J F,Zhao Y Y,et al.GIAOTracker: A comprehensive framework for MCMOT with global information and optimizing strategies in VisDrone 2021[C]//2021 IEEE/CVF International Conference on Computer Vision Workshops (ICCVW).Piscataway:IEEE Press,2021:2809-2819.

[3] Yang M Z,Han G X,Yan B,et al.Hybrid-SORT: Weak cues matter for online multi-object tracking[C]//Proceedings of the AAAI Conference on Artificial Intelligence.Washington:AAAI Press,2024,38(7):6504-6512.

[4] Kuhn H W.The Hungarian method for the assignment problem[J].Naval Research Logistics Quarterly,1955,2(1-2):83-97.

[5] Bewley A,Ge Z,Ott L,et al.Simple online and realtime tracking[C]//2016 IEEE International Conference on Image Processing (ICIP).Piscataway:IEEE Press,2016:3464-3468.

[6] Li C J,Ma J B,Shen L,et al.Occlusion-Aware SORT: Observing Occlusion for Robust Multi-Object Tracking[J].arXiv preprint,2026,abs/2603.06034.

[7] Song I P,Lee J W.SFTrack: A Robust Scale and Motion Adaptive Algorithm for Tracking Small and Fast Moving Objects[J].arXiv preprint,2024,abs/2410.20079.

[8] Kalman R E.A new approach to linear filtering and prediction problems[J].Journal of Basic Engineering,1960,82(1):35-45.

[9] Wojke N,Bewley A,Paulus D.Simple online and realtime tracking with a deep association metric[J].arXiv preprint,2017,abs/1703.07402.

[10] Zhang Y F,Sun P T,Jiang Y,et al.ByteTrack: Multi-object tracking by associating every detection box[C]//European Conference on Computer Vision (ECCV).Cham:Springer,2022,13682:1-18.

[11] Cao J K,Pang J M,Weng X S,et al.Observation-centric SORT: Rethinking SORT for robust multi-object tracking[C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR).Piscataway:IEEE Press,2023:9686-9696.

[12] Du Y H,Zhao Z M,Song Y Z,et al.StrongSORT: Make DeepSORT great again[J].IEEE Transactions on Multimedia,2023,25:8725-8737.

[13] 严春满，吴松伦，董俊松．自适应渐消有偏扩展卡尔曼滤波在目标跟踪中的应用[J]．传感技术学报，2020（2）．
Yan C M，Wu S L，Dong J S．Application of the Adaptive Fading Biased Extended Kalman Filter in Target Tracking[J]．Chinese Journal of Sensors and Actuators，2020（2）：起止页码（in Chinese）．

[14] Yang H S,Sim C B,Jung S H.CenterTrack-EKF: Improved multi-object tracking with extended Kalman filter[J].SmartMedia Journal,2024,13(5):9.

[15] 陈智超，常家云，许震俞，等．基于 SORT 框架无迹卡尔曼滤波的多目标跟踪算法[J]．制导与引信，2025，46（4）．
Chen Z C,Chang J Y,Xu Z Y,et al.Multi-Target Tracking Algorithm Based on Unscented Kalman Filter within SORT Framework[J].Guidance & Fuze,2025,46(4) (in Chinese).

[16] Wang X,et al.A multi-object tracking method with an unscented Kalman filter on a Lie group manifold (LUKF-Track)[J].Entropy,2026,28(1):103.

[17] Zhu P F,Du D W,Wen L Y,et al.VisDrone-VID2019: The Vision Meets Drone Object Detection in Video Challenge Results[C]//2019 IEEE/CVF International Conference on Computer Vision Workshops (ICCVW).Piscataway:IEEE Press,2019:144-152.

[18] Sun P T,Cao J K,Jiang Y,et al.DanceTrack: Multi-Object Tracking in Uniform Appearance and Diverse Motion[C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR).Piscataway:IEEE Press,2022:20993-21002.

[19] Aharon N,Orfaig R,Bobrovsky B Z.BoT-SORT: Robust Associations Multi-Pedestrian Tracking[J].arXiv preprint,2022,abs/2206.14651.

[20] Zhang Y F,Sun P T,Jiang Y,et al.ByteTrack: Multi-object Tracking by Associating Every Detection Box[C]//Avidan S,Brostow G,Cissé M,et al.(Eds.) Computer Vision–ECCV 2022.Cham:Springer,2022,13682:1-18.

[21] Stanojevic V D,Todorovic B T.BoostTrack: boosting the similarity measure and detection confidence for improved multiple object tracking[J].Machine Vision and Applications,2024,35:53.

[22] Erregue I,Nasrollahi K,Escalera S.YOLO11-JDE: Fast and Accurate Multi-Object Tracking with Self-Supervised Re-ID[J].arXiv preprint,2025,abs/2501.13710.