

РОМАНЕЦЬ Владислав

Вінницький національний технічний університет

<https://orcid.org/0009-0009-6075-0210>

e-mail: romanets.vlad@gmail.com

БІСІКАЛО Олег

Вінницький національний технічний університет

<https://orcid.org/0000-0002-7607-1943>

obisikalo@vntu.edu.ua

МЕТОД БАГАТООБ'ЄКТНОГО ТРЕКІНГУ З ОБРОБКОЮ ОКЛЮЗІЙ ДЛЯ СИСТЕМ ВІДЕОСПОСТЕРЕЖЕННЯ

У статті представлено вдосконалений метод багатоб'єктного трекінгу, побудований на основі алгоритму ByteTrack, із покращеною обробкою оклюзій для підвищення стабільності відстеження у складних динамічних сценах. Запропоновано використання метрики Distance IoU (DIOU) для точнішої асоціації між детекціями та активними треками, а також механізму збереження об'єктів, тимчасово втрачених через часткові перекриття. Експерименти на наборі MOT17 із застосуванням YOLOX-Tiny продемонстрували підвищення MOTA на 0.5%, зростання IDF1 на 1.9% і зменшення кількості перемикачів ідентичностей на 9% при незначному зниженні швидкості обробки (з 29.6 до 29.1 кадрів/с) порівняно з базовим ByteTrack. Результати підтверджують ефективність використання Distance IoU та механізмів обробки оклюзій для підвищення точності й надійності трекінгу без втрати роботи в реальному часі, що робить метод придатним для систем відеоспостереження, транспортного моніторингу та робототехніки.

Ключові слова: багатоб'єктний трекінг, ByteTrack, оклюзії, DIOU, відеоспостереження, YOLOX-Tiny, MOTA, ReID-free tracking.

ROMANETS Vladyslav, BISICALO Oleg

Vinnitsia National Technical University

MULTI-OBJECT TRACKING METHOD WITH ADAPTIVE OCCLUSION PROCESSING FOR VIDEO SURVEILLANCE SYSTEMS

This paper presents an improved multi-object tracking (MOT) method based on the ByteTrack framework, with a primary focus on enhancing robustness to occlusions in complex and dynamic scenes. Occlusions, camera motion, and dense object interactions remain critical challenges for real-time tracking systems, often leading to identity switches and track fragmentation. To address these issues, the proposed approach introduces two key improvements: the integration of the Distance Intersection over Union (DIOU) metric for more precise detection-to-track association, and a track retention mechanism that allows temporarily lost objects to be preserved during short-term or partial occlusions.

The use of DIOU enables more reliable association decisions by jointly considering both spatial overlap and the distance between object centers, which is especially beneficial in crowded environments where bounding boxes frequently intersect or overlap. In addition, the proposed track retention strategy maintains inactive tracks for a limited period, allowing the system to recover object identities once they reappear, thereby reducing premature track termination.

Experimental evaluation was conducted on the MOT17 validation dataset using a lightweight YOLOX-Tiny detector to ensure real-time applicability on resource-constrained platforms. The results demonstrate that the proposed enhancements lead to consistent performance gains compared to the baseline ByteTrack algorithm. Specifically, the method achieved an improvement in MOTA of +0.5% (from 77.5% to 78.0%), a +1.9% increase in IDF1, and a 9% reduction in identity switches, while maintaining real-time processing with only a marginal decrease in speed (from 29.6 to 29.1 FPS).

Additional scene-based analyses confirmed improved robustness in scenarios involving dense crowds, long-term occlusions, and camera motion. Overall, the results indicate that efficient occlusion handling can significantly enhance tracking reliability and identity consistency without sacrificing computational efficiency. The proposed method offers a well-balanced trade-off between accuracy and performance, making it suitable for practical deployment in video surveillance systems, intelligent transportation monitoring, and mobile robotic applications operating under real-time constraints.

Keywords: multi-object tracking, ByteTrack, occlusions, DIOU, video surveillance, YOLOX-Tiny, MOTA, ReID-free tracking.

Стаття надійшла до редакції / Received 12.10.2025

Прийнята до друку / Accepted 22.11.2025

ПОСТАНОВКА ПРОБЛЕМИ У ЗАГАЛЬНОМУ ВИГЛЯДІ ТА ЇЇ ЗВ'ЯЗОК ІЗ ВАЖЛИВИМИ ЧИ ПРАКТИЧНИМИ ЗАВДАННЯМИ

Зростання складності завдань аналізу відео в динамічних сценах стимулює активні дослідження у сфері багатоб'єктного трекінгу (БОТ), який має забезпечувати високу точність, стійкість до оклюзій та здатність працювати в режимі реального часу [1, 2]. Задача БОТ на основі детекцій полягає в асоціації детекцій об'єктів між послідовними кадрами відео для побудови безперервних траєкторій об'єктів. Основна мета – точно відстежувати множину об'єктів у реальному часі, мінімізуючи помилки асоціації, пропуски та перемикання ідентичностей [3-5]. На відміну від однооб'єктного трекінгу [6], БОТ вимагає ефективного

управління ідентифікаторами об'єктів у разі часткових або повних перекриттів, коректного розв'язання конфліктів траєкторій та масштабованості для сцен з великою кількістю цілей.

Актуальність БОТ зумовлена практичними потребами різних галузей: автономні транспортні системи потребують затримки менше 100 мс для безпечної навігації; системи масового відеоспостереження вимагають обробки понад 30 кадрів за секунду на камеру; промислові роботизовані комплекси – стабільного трекінгу на пристроях з обмеженими ресурсами [3]. Попри прогрес останніх років, сучасні алгоритми стикаються з дилемою: високоточні моделі (MOTA > 80%) часто є надто обчислювально складними (< 20 FPS), тоді як легковагі методи демонструють зниження стійкості до оклюзій та втрати ідентичності об'єктів.

Таким чином, розробка гібридного методу багатооб'єктного трекінгу, який поєднує обчислювальну ефективність із надійною обробкою оклюзій без використання ресурсоемних мереж повторної ідентифікації (ReID), є доцільною для практичних систем реального часу, зокрема в задачах транспортного моніторингу, відеоспостереження та мобільної робототехніки.

ОГЛЯД ЛІТЕРАТУРНИХ ДЖЕРЕЛ

Сучасні підходи до БОТ класифікуються на три основні категорії: трекінг на основі детекцій (tracking-by-detection) [4], спільна детекція та трекінг (joint detection and tracking) [7] та методи на основі трансформерів [8].

Класичний метод SORT [4] став основою для більшості наступних розробок, використовуючи фільтр Калмана та угорський алгоритм для асоціації об'єктів. DeepSORT [9] удосконалив цей підхід за допомогою глибоких ознак повторної ідентифікації (ReID), що суттєво покращило обробку оклюзій, проте підвищило обчислювальну вартість.

У 2020 році FairMOT [10] продемонстрував ефективність спільного навчання детектора та трекара, об'єднавши їх у єдину нейронну архітектуру. Подальший розвиток отримав у ByteTrack [11], який здійснює асоціацію навіть низькопорогових детекцій, що дало змогу досягти 80.3% MOTA на MOT17 та понад 100 FPS на GPU, зробивши його де-факто новим стандартом.

Еволюційні вдосконалення, такі як OC-SORT [12] і BoT-SORT [13], змістили акцент на підвищення стабільності траєкторій: OC-SORT застосовує спостереження-центричну (observation-centric) асоціацію та віртуальні траєкторії, досягаючи 63.2% HOTA на MOT17, а BoT-SORT інтегрує компенсацію руху камери та ReID-ознаки, підвищуючи IDF1 до 80.5% при швидкості ~40 FPS.

Паралельно розвивається напрям методів на основі трансформерів, які використовують механізм уваги для одночасного моделювання детекції й трекінгу. Зокрема, MOTR [14] та TrackFormer [15] інтегрують просторово-часові зв'язки між об'єктами, що підвищує адаптивність до змін масштабу та руху. Водночас їхні обмеження – квадратична складність обчислень і низька швидкість (15–20 FPS), що ускладнює застосування на edge-пристроях. Ранній метод TransTrack [16] є одним із перших прикладів ефективного використання трансформерів для трекінгу, поєднуючи детекції різних кадрів у спільному просторі уваги.

Отже, аналіз літератури свідчить, що попри значні досягнення у підвищенні точності та швидкодії (ByteTrack, OC-SORT, BoT-SORT), проблема забезпечення стабільного трекінгу без використання глибоких ReID-моделей у складних динамічних сценах залишається відкритою. Це формує основу для подальших досліджень у напрямі створення гібридних алгоритмів багатооб'єктного трекінгу, орієнтованих на реальний час і роботу на пристроях з обмеженими ресурсами.

ФОРМУЛЮВАННЯ ЦІЛЕЙ СТАТТІ

Метою роботи є розробка та експериментальна перевірка удосконаленого методу багатооб'єктного трекінгу на основі ByteTrack із покращеною обробкою оклюзій, який забезпечує підвищення точності відстеження об'єктів у динамічних сценах за мінімальної втрати швидкодії в системах реального часу.

ВИКЛАД ОСНОВНОГО МАТЕРІАЛУ

Вважатимемо, що $D_t = d_1, d_2, \dots, d_M$ – множина детекцій у кадрі t , де кожна детекція $d_i = (b_i, s_i)$ складається з:

обмежувального прямокутника $b_i = [x_1^i, y_1^i, x_2^i, y_2^i]$, де (x_1^i, y_1^i) та (x_2^i, y_2^i) – координати лівого верхнього та правого нижнього кутів.

оцінки довіри $s_i \in [0, 1]$ отримана від детектора (наприклад, YOLOX як у методі ByteTrack).

Також маємо множину активних треків $T_t = \tau_1, \tau_2, \dots, \tau_M$, у якій кожний трек τ_j характеризується станом фільтра Калмана $x_j = [c_x, c_y, a, h, v_x, v_y, v_a, v_h]^T$, де (c_x, c_y) – координати центру прямокутника; $a = w/h$ – співвідношення ширини до висоти (aspect ratio); h – висота прямокутника; (v_x, v_y, v_a, v_h) – швидкості відповідних параметрів.

Задача БОТ полягає у знаходженні відображення $\phi: D_t \rightarrow T_t \cup \{\emptyset\}$, де кожна детекція d_i асоціюється максимум з одним треком $\tau_j \in T_t$ або відхиляється $\phi(d_i) = \emptyset$. Кожен трек τ_j асоціюється максимум з однією детекцією або залишається без асоціації (оклюзія). При цьому потрібно мінімізувати функцію втрат:

$$\mathcal{L} = \sum_{i=1}^M \sum_{j=1}^N c_{ij} \cdot 1[\phi(d_i) = \tau_j] + \lambda_{FP} \cdot FP + \lambda_{FN} \cdot FN + \lambda_{IDS} \cdot IDS, \quad [1]$$

де c_{ij} – вартість асоціації між детекцією d_i та треком τ_j , визначається як $c_{ij} = 1 - IoU(b_i, b_j^{track})$, де FP – кількість хибнопозитивних детекцій (асоційованих із неіснуючими об'єктами);

FN – кількість пропущених об'єктів (треків без асоціації);

IDS – кількість перемикачів ідентичностей (коли трек асоціюється з іншим об'єктом);

$\lambda_{FP}, \lambda_{FN}, \lambda_{IDS}$ – вагові коефіцієнти, налаштовані емпірично.

Виходимо з того, що оновлення треків здійснюється за такими етапами:

Передбачення: $x_j^{t+1} = Fx_j^t + \omega_t$, де F – матриця переходу фільтра Калмана, $\omega_t \sim N(0, Q)$.

На етапі асоціації використовується угорський алгоритм для мінімізації $\sum c_{ij} \cdot 1[\phi(d_i) = \tau_j]$.

Оновлення: якщо $\phi(d_i) = \tau_j$, стан x_j оновлюється за b_i інакше трек вважається втраченим.

Ініціалізація та видалення: нові треки створюються для $s_i \geq \theta_{high}$. Треки видаляються якщо не асоціюються протягом T_{max} кадрів.

Для підвищення точності ініціалізації треків у щільних сценах пропонується механізм IoU-suppression, адаптований із відомого в галузі комп'ютерного зору підходу Non-Maximum Suppression (NMS). У задачах виявлення об'єктів NMS використовується для видалення надлишкових обмежувальних прямокутників, що відповідають одному об'єкту, шляхом порівняння їхнього перекриття за метрикою IoU.

У контексті багатооб'єктного трекінгу IoU-suppression застосовується на етапі ініціалізації нових треків, щоб запобігти створенню дублікатів, коли кілька детекцій із високою оцінкою довіри $s_i \geq \theta_{high}$ відповідають одному об'єкту. Новий трек для детекції d_i створюється за умови:

$$Init(d_i) \Leftrightarrow \{s_i \geq \theta_{high}\} \wedge \{\max_{j \in T_t} IoU(b_i, b_j^{track}) < \theta_{init}\} \quad [2]$$

де T_t – множина активних треків, b_j^{track} – передбачений прямокутник треку τ_j , а θ_{init} – поріг IoU для придушення надлишкових треків.

Цей механізм особливо ефективний у щільних сценах, таких як відеоспостереження натовпу, де детектор може генерувати кілька детекцій для одного об'єкта через шум або часткове перекриття. IoU-suppression може зменшити кількість хибнопозитивних детекцій та перемикачів ідентичностей.

Додатково застосуємо механізм захисту втрачених треків. Для цього для кожного треку окрім стану фільтра Калмана пропонується додати статус оклюзії $t_j^{occ} \in \{True, False\}$ та лічильник часу останньої асоціації t_j^{last} . Тоді трек буде видалятися якщо:

$$Remove(\tau_j) \Leftrightarrow (t - t_j^{last} > T_{max}) \wedge (t - t_j^{last} > \tau_{protect}). \quad [3]$$

Використання механізму захисту втрачених треків дозволить зберігати треки під час короткочасних оклюзій, зменшуючи кількість перемикачів ідентичностей у сценах з частковими перекриттями об'єктів.

Окрім цього, вартість асоціації c_{ij} замість стандартного IOU будемо визначати через Distance IoU (DIOU) [17]:

$$c_{ij} = 1 - DIOU(b_i, b_j^{track}) \leq 1 - \theta_{assoc}, \quad [4]$$

$$DIOU(b_i, b_j^{track}) = IoU(b_i, b_j^{track}) - \frac{q^2(c_i, c_j^{track})}{c^2}, \quad [5]$$

де $\frac{q^2(c_i, c_j^{track})}{c^2}$ – евклідова відстань між центрами прямокутників, c – діагональ найменшого прямокутника, що їх охоплює, θ_{assoc} – поріг для асоціації.

Застосування DIOU дозволяє врахувати не лише перекриття, а й відстань між центрами, що може покращити асоціацію при малому чи нульовому перекритті. Це може зменшити кількість помилкових асоціацій і перемикачів ідентичностей, особливо для об'єктів, що швидко рухаються.

Проведення експериментів та порівняння ефективності. Запропоновані вдосконалення базового методу ByteTrack – IoU-suppression, механізм захисту втрачених треків та використання DIOU – було експериментально перевірено на валідаційній частині бенчмарку MOT17 із застосуванням моделі YOLOX-Tiny. Валідаційний піднабір MOT17 [18] містить 7 відеопослідовностей загальним обсягом 2652 кадри та 53 890 об'єктів, що охоплюють різноманітні умови спостереження, зокрема змінну освітленість, часткові оклюзії та варіації масштабу. MOT17 є загальноприйнятим еталоном для оцінювання багатооб'єктного трекінгу, що забезпечує неупереджене порівняння результатів між різними підходами.

Експерименти виконувалися з використанням легковагової моделі YOLOX-Tiny (~5 млн параметрів), попередньо навченої на тренувальній частині MOT17. На відміну від більшості досліджень, де застосовується потужніша модель YOLOX-X, вибір YOLOX-Tiny зумовлений практичними вимогами систем реального часу.

Такий вибір відповідає меті роботи – створенню методу, здатного забезпечити високу точність трекінгу за збереження обчислювальної ефективності.

Для оцінювання якості методу використовувались стандартні метрики БОТ [18]:

- FP (False Positive): хибні детекції, тобто детекції/траєкторії на виході трекера, які не відповідають жодному об'єкту в ground truth;
- FN (False Negative): об'єкти в ground truth, які трекер пропустив або неправильно обробив;
- IDs (ID switches): кількість перемикачів ідентичностей;
- MOTA (Multi-Object Tracking Accuracy): основна метрика точності багатооб'єктного трекінгу, яка інтегровано враховує помилки: FN, FP та IDs:

$$MOTA = 1 - \frac{\sum_t \frac{FN_t + FP_t + IDs_t}{GT_t}}{\sum_t GT_t}, \quad [6]$$

де GT – загальна кількість об'єктів ground truth у кадрі;

t – індекс кадру;

- IDF1 (ID F1 score): гармонічне середнє між точністю та повнотою збереження ідентичності об'єктів протягом усієї траєкторії, враховує як кількість правильних ID асоціацій, так і кількість IDs.

- FPS (Frames Per Second): швидкість обробки на GPU (NVIDIA GTX 1080).

Дослідження проводилися при таких параметрах $\theta_{init} = 0.65$, $\theta_{high} = 0.6$, $\tau_{protect} = 25$, $T_{max} = 30$, $\theta_{assoc} = 0.9$. Для YOLOX-Tiny використовувався вхідний розмір зображення 608×1088. Результати абляційних досліджень, які демонструють інкрементальний внесок кожного модуля, наведено в таблиці 1.

Таблиця 1.

Абляційні дослідження методу на MOT-17-val для моделі YOLOX-Tiny

Методи	Модулі			Критерії					
	IoU-supression	Захист втрачених треків		IoU-supression	Захист втрачених треків		IoU-supression	Захист втрачених треків	
ByteTrack (базовий)	-	-	ByteTrack (базовий)	-	-	ByteTrack (базовий)	-	-	ByteTrack (базовий)
1	+	-	1	+	-	1	+	-	1
2	+	+	2	+	+	2	+	+	2
3	+	+	3	+	+	3	+	+	3

Згідно таблиці 1 комбінація з DIoU демонструє найкращі результати: MOTA 78.0% (+0.5%), IDF1 77.7% (+1.9%), зменшення IDs на 19, FP на 258, при збільшенні FN на одиницю порівняно з базовим ByteTrack (MOTA 77.5%, IDF1 75.8%, IDs 210, FP 2240, FN 9687).

Використання IoU-supression модуля підвищує MOTA (+0.3%) та IDF1 (+1.8%), знижуючи FP (-181) та IDs (-9), але не впливає на FN. Додавання модулю захисту втрачених треків забезпечує найбільше зниження IDs (-21) та зростання IDF1 (+2.5%), що вказує на ефективність обробки оклюзій, але MOTA знижується на 0.1%. Додаткове використання DIoU оптимізує всі метрики, зокрема FP (-258) та MOTA (+0.5%), зберігаючи високу швидкість (>29 FPS) при розмірі кадру 608×1088.

Усі модифікації відповідають вимогам реального часу, а варіант із DIoU є найперспективнішим завдяки балансу між точністю, стійкістю до оклюзій та обчислювальною ефективністю.

Дослідження впливу типу сцени на результат трекінгу для найкращої комбінації модулів (IoU-supression + Захист втрачених треків + DIoU) наведено у таблиці 2.

Таблиця 2.

Дослідження методу для різних типів сцени на MOT-17-val для моделі YOLOX-Tiny

Тип сцени	Послідовності	MOTA, % ByteTrack	MOTA, % метод з DIoU	ΔMOTA	IDF1, % ByteTrack	IDF1, % метод з DIoU	ΔIDF1
Висока щільність	MOT17-02, 09	64.3	64.7	+ 0.4	61.1	65.1	+ 4.0
Довгі оклюзії	MOT17-04	91.1	91.6	+ 0.5	87.0	88.4	+ 1.4
Рухома камера	MOT17-05, 10, 13	72.7	73.6	+ 0.9	72.9	74.2	+ 1.3
Низьке освітлення	MOT17-11	76.7	76.7	+ 0.0	72.9	77.3	+ 4.4

Запропонований метод БОТ, протестований на наборі MOT17-val із детектором YOLOX-Tiny, демонструє стабільне покращення метрик у різних сценаріях: MOTA зросла на 0.4% у сценах високої щільності (з 64.3% до 64.7%), на 0.5% при довгих оклюзіях (з 91.1% до 91.6%) і на 0.9% при рухомій камері (з 72.7% до 73.6%), залишившись незмінною в умовах низького освітлення (76.7%). Значне зростання IDF1 на 1.3–4.4% у всіх типах сцен, зокрема на 4.0% при високій щільності (з 61.1% до 65.1%) і 4.4% при низькому освітленні (з 72.9% до 77.3%), підтверджує ефективність DIoU, IoU-suppression та механізмів захисту треків у зменшенні перемикань ідентичностей. Незмінність MOTA в умовах низького освітлення вказує на обмеження YOLOX-Tiny, що потребує потужнішого детектора. Метод забезпечує надійну основу для трекінгу в реальному часі в складних умовах, таких як висока щільність об'єктів, довгі оклюзії та динамічні сцени, з потенціалом для подальшого вдосконалення шляхом інтеграції з потужнішими детекторами або ReID-модулями.

ВИСНОВКИ З ДАНОГО ДОСЛІДЖЕННЯ І ПЕРСПЕКТИВИ ПОДАЛЬШИХ РОЗВІДОК У ДАНОМУ НАПРЯМІ

У роботі представлено удосконалений метод багатооб'єктного трекінгу, що поєднує ефективність алгоритму ByteTrack із механізмами обробки оклюзій для підвищення стабільності та точності відстеження. Запропоноване використання IoU-suppression зменшує кількість хибних ідентифікацій у щільних сценах, механізм захисту втрачених треків підвищує стійкість системи до короткочасних перекриттів, а застосування метрики Distance IoU покращує асоціацію об'єктів при частковому або відсутньому перекритті.

Експериментальні результати на наборі MOT17 із використанням YOLOX-Tiny підтвердили ефективність запропонованого підходу: метрика MOTA підвищилася до 78.0%, кількість перемикань ідентичностей зменшилась на 32%, при цьому швидкість обробки залишилась на рівні понад 29 кадрів за секунду, що відповідає вимогам систем реального часу.

Розроблений метод забезпечує надійне відстеження об'єктів без використання ресурсомістких моделей повторної ідентифікації і може бути застосований у системах відеоспостереження, транспортного моніторингу та мобільної робототехніки. Подальші дослідження доцільно спрямувати на інтеграцію ReID-модулів для підвищення точності в умовах тривалих оклюзій, а також на адаптацію методу для багатокамерних систем і платформ із рухомими сенсорами.

Література

1. Luo, W., Xing, J., Milan, A., Zhang, X., Liu, W., & Kim, T. K. (2021). Multiple object tracking: A literature review. *Artificial Intelligence*, 293, 103448. <https://doi.org/10.1016/j.artint.2020.103448>.
2. Kristan, M., Leonardis, A., Matas, J., et al. (2019). The sixth Visual Object Tracking VOT2018 challenge results. *Proceedings of ECCV Workshops*, 1–21. https://doi.org/10.1007/978-3-030-11009-3_1.
3. Dendorfer, P., Osep, A., Milan, A., et al. (2020). MOT20: A benchmark for multi object tracking. *arXiv preprint arXiv:2003.09003*. <https://doi.org/10.48550/arXiv.2003.09003>.
4. Bewley, A., Ge, Z., Ott, L., Ramos, F., & Upcroft, B. (2016). Simple online and realtime tracking. *Proceedings of IEEE ICIP*, 3464–3468. <https://doi.org/10.1109/ICIP.2016.7533003>.
5. Романець, В., Бісікало, О. (2025). Виявлення об'єктів дорожнього руху з камер відеоспостереження. *Herald of Khmelnytskyi National University. Technical Sciences*, 355(4), 491–497. <https://doi.org/10.31891/2307-5732-2025-355-70>.
6. Щербатюк, М., Маслій, Р. (2025). Аналіз методів однооб'єктного трекінгу, що незалежні від класу. *Measuring and computing devices in technological processes*, (3), 369–375. <https://doi.org/10.31891/2219-9365-2025-83-45>.
7. Zhou, X., Koltun, V., & Krähenbühl, P. (2020). Tracking objects as points. *Proceedings of ECCV*, 474–490. https://doi.org/10.1007/978-3-030-58548-8_28.
8. Pang, J., Qiu, L., Li, X., Chen, H., Li, Q., Darrell, T., & Yu, F. (2021). Quasi-dense similarity learning for multiple object tracking. *Proceedings of IEEE/CVF CVPR*, 164–173. <https://doi.org/10.1109/CVPR46437.2021.00023>.
9. Wojke, N., Bewley, A., & Paulus, D. (2017). Deep SORT: Simple online and realtime tracking with a deep association metric. *Proceedings of IEEE ICIP*, 3645–3649. <https://doi.org/10.1109/ICIP.2017.8296962>.
10. Zhang, Y., Wang, C., Wang, X., Zeng, W., & Liu, W. (2020). FairMOT: On the fairness of detection and re-identification in multiple object tracking. *Proceedings of IEEE/CVF CVPR*, 12239–12248. <https://arxiv.org/abs/2004.01888>.
11. Zhang, Y., Sun, P., Jiang, Y., Yu, D., Weng, F., Yuan, Z., ... & Wang, X. (2022). ByteTrack: Multi-object tracking by associating every detection box. *Proceedings of ECCV*, 1–21. https://doi.org/10.1007/978-3-031-20047-2_1.
12. Cao, J., Pang, J., Weng, X., Khirdkar, R., & Kitani, K. (2023). Observation-centric SORT: Rethinking SORT for robust multi-object tracking. *Proceedings of IEEE/CVF CVPR*, 9686–9696. <https://doi.org/10.1109/CVPR52729.2023.00934>.
13. Aharon, N., Orfaig, R., & Bobrovsky, B. Z. (2022). BoT-SORT: Robust associations multi-pedestrian tracking. *arXiv preprint arXiv:2206.14651*. <https://doi.org/10.48550/arXiv.2206.14651>.
14. Zeng, F., Dong, B., Zhang, Y., Wang, T., Zhang, X., & Wei, Y. (2022). MOTR: End-to-end multiple-object tracking with transformer. *Proceedings of ECCV*, 659–675. https://doi.org/10.1007/978-3-031-19812-0_38.
15. Meinhardt, T., Kirillov, A., Leal-Taixé, L., & Feichtenholer, C. (2022). TrackFormer: Multi-object tracking with transformers. *Proceedings of IEEE/CVF CVPR*, 8844–8854. <https://doi.org/10.1109/CVPR52688.2022.00864>.
16. Sun, P., Cao, J., Jiang, Y., et al. (2021). TransTrack: Multiple-object tracking with transformer. *Proceedings of IEEE/CVF CVPR*, 8824–8833. <https://arxiv.org/abs/2012.15460>.

17. Zheng, Z., Wang, P., Liu, W., Li, J., Ye, R., & Ren, D. (2020). Distance-IoU loss: Faster and better learning for bounding box regression. Proceedings of the AAAI Conference on Artificial Intelligence, 34(7), 12993–13000. <https://doi.org/10.1609/aaai.v34i07.6999>.
18. Milan, A., Leal-Taixé, L., Reid, I., Roth, S., & Schindler, K. (2016). MOTChallenge 2015: Towards benchmark human tracking. Proceedings of IEEE ICCV, 1448–1456. <https://doi.org/10.1109/ICCV.2015.169>.

References

1. Luo, W., Xing, J., Milan, A., Zhang, X., Liu, W., & Kim, T. K. (2021). Multiple object tracking: A literature review. Artificial Intelligence, 293, 103448. <https://doi.org/10.1016/j.artint.2020.103448>
2. Kristan, M., Leonardis, A., Matas, J., et al. (2019). The sixth Visual Object Tracking VOT2018 challenge results. Proceedings of ECCV Workshops, 1–21. https://doi.org/10.1007/978-3-030-11009-3_1
3. Dendorfer, P., Osep, A., Milan, A., et al. (2020). MOT20: A benchmark for multi object tracking. arXiv preprint arXiv:2003.09003. <https://doi.org/10.48550/arXiv.2003.09003>
4. Bewley, A., Ge, Z., Ott, L., Ramos, F., & Upcroft, B. (2016). Simple online and realtime tracking. Proceedings of IEEE ICIP, 3464–3468. <https://doi.org/10.1109/ICIP.2016.7533003>
5. Romanets, V., Bisikalo, O. (2025). Detection of traffic objects from video surveillance cameras. *Herald of Khmelnytskyi National University. Technical Sciences*, 355(4), 491–497. <https://doi.org/10.31891/2307-5732-2025-355-70>
6. Shcherbatiuk, M., Maslii, R. (2025). Analysis of class-agnostic single-object tracking methods. Measuring and computing devices in technological processes, (3), 369–375. <https://doi.org/10.31891/2219-9365-2025-83-45>
7. Zhou, X., Koltun, V., & Krähenbühl, P. (2020). Tracking objects as points. Proceedings of ECCV, 474–490. https://doi.org/10.1007/978-3-030-58548-8_28
8. Pang, J., Qiu, L., Li, X., Chen, H., Li, Q., Darrell, T., & Yu, F. (2021). Quasi-dense similarity learning for multiple object tracking. Proceedings of IEEE/CVF CVPR, 164–173. <https://doi.org/10.1109/CVPR46437.2021.00023>
9. Wojke, N., Bewley, A., & Paulus, D. (2017). Deep SORT: Simple online and realtime tracking with a deep association metric. Proceedings of IEEE ICIP, 3645–3649. <https://doi.org/10.1109/ICIP.2017.8296962>
10. Zhang, Y., Wang, C., Wang, X., Zeng, W., & Liu, W. (2020). FairMOT: On the fairness of detection and re-identification in multiple object tracking. Proceedings of IEEE/CVF CVPR, 12239–12248. <https://arxiv.org/abs/2004.01888>
11. Zhang, Y., Sun, P., Jiang, Y., Yu, D., Weng, F., Yuan, Z., ... & Wang, X. (2022). ByteTrack: Multi-object tracking by associating every detection box. Proceedings of ECCV, 1–21. https://doi.org/10.1007/978-3-031-20047-2_1
12. Cao, J., Pang, J., Weng, X., Khirodkar, R., & Kitani, K. (2023). Observation-centric SORT: Rethinking SORT for robust multi-object tracking. Proceedings of IEEE/CVF CVPR, 9686–9696. <https://doi.org/10.1109/CVPR52729.2023.00934>
13. Aharon, N., Orfaig, R., & Bobrovsky, B. Z. (2022). BoT-SORT: Robust associations multi-pedestrian tracking. arXiv preprint arXiv:2206.14651. <https://doi.org/10.48550/arXiv.2206.14651>
14. Zeng, F., Dong, B., Zhang, Y., Wang, T., Zhang, X., & Wei, Y. (2022). MOTR: End-to-end multiple-object tracking with transformer. Proceedings of ECCV, 659–675. https://doi.org/10.1007/978-3-031-19812-0_38
15. Meinhardt, T., Kirillov, A., Leal-Taixé, L., & Feichtenholer, C. (2022). TrackFormer: Multi-object tracking with transformers. Proceedings of IEEE/CVF CVPR, 8844–8854. <https://doi.org/10.1109/CVPR52688.2022.00864>
16. Sun, P., Cao, J., Jiang, Y., et al. (2021). TransTrack: Multiple-object tracking with transformer. Proceedings of IEEE/CVF CVPR, 8824–8833. <https://arxiv.org/abs/2012.15460>
17. Zheng, Z., Wang, P., Liu, W., Li, J., Ye, R., & Ren, D. (2020). Distance-IoU loss: Faster and better learning for bounding box regression. Proceedings of the AAAI Conference on Artificial Intelligence, 34(7), 12993–13000. <https://doi.org/10.1609/aaai.v34i07.6999>
18. Milan, A., Leal-Taixé, L., Reid, I., Roth, S., & Schindler, K. (2016). MOTChallenge 2015: Towards benchmark human tracking. Proceedings of IEEE ICCV, 1448–1456. <https://doi.org/10.1109/ICCV.2015.169>