

ИНТЕРПРЕТИРУЕМОЕ УПРАВЛЕНИЕ ИНИЦИАЦИЕЙ ТРЕКОВ С ЗАРАНЕЕ ЗАДАННЫМИ ГРАНИЦАМИ ВМЕШАТЕЛЬСТВА В СИСТЕМАХ ВОСПРИЯТИЯ БЕСПИЛОТНЫХ ЛЕТАТЕЛЬНЫХ АППАРАТОВ

Трофимов Ю.В.^{1,а}, Аверкин А.Н.^{1,2,б}, Шевченко А.В.^{1,с}, Кузнецов Е.М.^{1,д},
Лебедев А.Д.^{1,е}

¹ Государственный университет «Дубна», Дубна, Россия.

² Федеральный исследовательский центр «Информатика и управление» Российской академии наук (ФИЦ ИУ РАН), Москва, Россия.

E-mail: ^а ura_trofim@bk.ru, ^б averkin2003@inbox.ru, ^с leviathan0909@gmail.com, ^д kem.22@uni-dubna.ru,

^е lad.24@uni-dubna.ru

Аннотация. В системах воздушного наблюдения единичная ошибка обнаружения может закрепиться в виде устойчивого ложного трека и повлиять на данные, передаваемые последующим компонентам. Цель работы – уменьшить число таких ошибок, заранее ограничив масштаб вмешательства и сохранив возможность проследить логику каждого решения. Предложен внешний метод кратковременной задержки эпизодов инициации треков с низкой поддержкой. Число задержек регулируется накопительным бюджетом, а для каждого эпизода сохраняются оценка поддержки, текущий баланс бюджета, причина вмешательства и конечный статус. Метод апробирован на семи видеопоследовательностях из общедоступного набора аэросъемки, содержащего малые объекты, частичные перекрытия и заметные изменения масштаба. В эксперименте сравнивались три алгоритма сопровождения и обучаемый метод на основе ансамбля решающих деревьев. Для ByteTrack задержка была применена к 2,62% эпизодов инициации, из оперативного выхода исключены 425 ложных наблюдений и 176 ложных инициаций, а изменение объединённого показателя точности и полноты осталось в пределах установленного допуска. Для OC-SORT и SORT снижение ложных наблюдений по принятому статистическому критерию не подтверждено. Проверка 29 385 начальных фрагментов видеозаписей не выявила нарушений заранее заданных ограничений. Средние дополнительные вычислительные затраты составили 0,901 миллисекунды на кадр. Результаты показывают, что кратковременная задержка может сократить число ложных наблюдений при контролируемом масштабе вмешательства.

Ключевые слова: интеллектуальные транспортные системы, беспилотные летательные аппараты, сопровождение объектов, инициация треков, ограниченное вмешательство, интерпретируемый искусственный интеллект, кратковременная задержка.

Введение

Беспилотные аппараты применяют для наблюдения за дорогами, инспекции инфраструктуры и контроля логистических объектов. Проблема с ложным

Для цитирования:

Трофимов Ю.В., Аверкин А.Н., Шевченко А.В., Кузнецов Е.М., Лебедев А.Д. Интерпретируемое управление инициацией треков с заранее заданными границами вмешательства в системах восприятия беспилотных летательных аппаратов. *Интеллектуальный транспорт*. 2026. № 3(39). С. 69–84. DOI: 10.24412/3033-6007-2026-339-69-84.

обнаружением возникает, когда алгоритм сопровождения присваивает ему идентификатор и начинает вести его как самостоятельный объект. После этого ошибочное обнаружение уже входит в состав трека, попадает в оперативный выход системы и может быть передано модулю планирования. Обычный способ уменьшить число ложных треков – это повысить порог уверенности или потребовать повторного появления объекта, однако цена такого решения меняется вместе со сценой. Высота съёмки, освещение, масштаб и плотность объектов сдвигают распределение оценок, поэтому правило, настроенное на одной видеозаписи, в другой может задержать заметно больше настоящих целей. В работе предлагается принцип, где дополнительный слой ненадолго задерживает только небольшую долю новых эпизодов, а все остальные сразу передаёт исходному алгоритму сопровождения. Если эпизод продолжает наблюдаться, но подтверждений всё ещё недостаточно, задержка снимается и дальнейшую обработку продолжает исходный алгоритм сопровождения. Накопительный бюджет вмешательств ограничивает число задержек. Его баланс задаёт верхнюю границу как для количества вмешательств, так и для возможного отличия оперативного выхода от результата исходного алгоритма. Одновременно сохраняется возможность проследить каждое решение, так как журнал хранит поддержку эпизода, состояние бюджета, причину задержки и конечный статус.

Цель исследования – выяснить, позволяет ли такое правило уменьшить число ложных наблюдений в оперативном выходе, не выходя за заранее установленный допуск по качеству. Метод проверяется с тремя алгоритмами сопровождения и сопоставляется со случайным лесом, обученным в той же схеме внешней проверки. Отдельно оцениваются соблюдение формальных ограничений, прослеживаемость решений и дополнительные вычислительные затраты.

1. Обзор существующих работ

Алгоритмы сопровождения связывают обнаружения одного и того же объекта на последовательных кадрах, формируя его траекторию. Они прогнозируют возможное положение объекта, затем сопоставляют прогноз с новыми обнаружениями. Классической основой такого прогноза служит фильтр Калмана [1]. В SORT он объединён с взаимно однозначным сопоставлением рамок на соседних кадрах [2]. Подход хорошо работает при плавном движении, но остаётся чувствительным к пропускам и ложным обнаружениям. В более новых методах устойчивость сопровождения повышается разными способами. ByteTrack повторно рассматривает обнаружения с невысокой уверенностью и тем самым восстанавливает часть прерванных траекторий [3]. OC-SORT сильнее опирается на фактически наблюдаемое движение, а StrongSORT дополняет геометрию признаками внешнего вида [4, 5].

В процедуре сопровождения применяются и обучаемые компоненты. DeepSORT различает близко расположенные объекты по признакам внешнего вида, Tracktor переносит рамку между кадрами средствами регрессии детектора, а случайный лес объединяет решения множества деревьев [6–8]. Обучаемые модели требуют размеченных данных для настройки под конкретную задачу; при переносе на другую камеру или сцену их качество может измениться, а адаптация к новым условиям может потребовать дополнитель-

ных размеченных примеров и повторного обучения. Отдельный подход к работе с неуверенными решениями используется в селективной классификации, где модель отказывается от ответа на неуверенных примерах [9]. В предлагаемой схеме вместо полного отказа используется кратковременная задержка нового эпизода. Если эпизод продолжает наблюдаться, но уверенности по-прежнему недостаточно, его дальнейшая обработка передаётся исходному алгоритму сопровождения. Фиксированный порог уверенности и требование нескольких повторных наблюдений также ограничивают раннюю передачу слабых эпизодов. При большом числе слабых обнаружений фиксированное правило может затронуть почти все новые эпизоды. Предлагаемый метод отличается способом ограничения вмешательства.

2. Постановка задачи и критерии оценки

Задача состоит в построении внешнего правила обработки новых эпизодов треков, которое позволяет уменьшить число ложных наблюдений в оперативном выходе, не допуская заранее неограниченного вмешательства в работу исходного алгоритма сопровождения. Для каждого нового эпизода правило должно определить, передать его наблюдения без задержки либо временно удержать их, причём число таких вмешательств и возможное отличие оперативного выхода от исходного результата ограничиваются заданным бюджетом. На каждом кадре детектор формирует обнаружения, а алгоритм сопровождения связывает их с существующими треками или присваивает им новые идентификаторы. Наблюдением считается рамка обнаруженного объекта на одном кадре вместе с идентификатором трека и оценкой уверенности. Непрерывная последовательность наблюдений с одним идентификатором образует эпизод трека. Если после длительного перерыва тот же идентификатор появляется снова, дополнительный слой рассматривает его как новый эпизод. Вмешательством называется временная задержка первых наблюдений такого эпизода перед их передачей последующим компонентам. В оценке различаются оперативный выход и итоговая карта. Оперативный выход содержит только наблюдения, доступные последующим компонентам в момент обработки каждого кадра. Итоговая карта формируется после завершения видеозаписи. Если задержанный эпизод подтверждён, его ранние наблюдения можно восстановить. Основная оценка проводится по оперативному выходу.

После применения дополнительного слоя все наблюдения, переданные в оперативный выход, заново сопоставляются с разметкой. Это необходимо, так как исключение или задержка одного наблюдения может изменить соответствия для остальных наблюдений. Повторное сопоставление исключает двойной учёт. Рамка связывается с размеченным объектом при отношении площади пересечения к площади объединения не менее 0,5. Исключённые области VisDrone [10] удаляются после сопоставления. Ложной инициацией считается эпизод, который в течение полного окна решения не соответствует ни одному размеченному объекту. Само окно – это время, отведённое на наблюдение нового эпизода. Для сравнения видеозаписей разной длины число ложных наблюдений в оперативном выходе пересчитывается на 100 кадров по формуле (1).

$$O_{100} = 100 \cdot \frac{N_{\text{false}}}{N_{\text{frames}}}. \quad (1)$$

Здесь N_{false} обозначает число наблюдений эпизодов ложной инициации, переданных в оперативный выход, а N_{frames} соответствует числу кадров видеозаписи.

Качество оперативного выхода оценивается мерой F_1 , которая объединяет точность и полноту. Точность показывает долю наблюдений оперативного выхода, относящихся к истинным объектам. Полнота отражает долю размеченных объектов, сохранённых системой. Сначала проверяется, остаётся ли изменение качества в пределах установленного допуска. Нижняя граница 95-процентного доверительного интервала изменения F_1 должна быть выше значения $-0,01$. Это условие записывается формулой (2).

$$H_1 : L_{95}(\Delta F_1) > -0,01. \quad (2)$$

После выполнения первого условия проверяется снижение показателя ложных наблюдений. Верхняя граница 95-процентного доверительного интервала изменения показателя O_{100} должна быть ниже нуля. Условие проверки приведено в формуле (3).

$$H_2 : U_{95}(\Delta O_{100}) < 0. \quad (3)$$

Благодаря порядку, где второе условие проверяется только после первого, успешным нельзя признать метод, который удаляет ложные наблюдения ценой заметной потери настоящих объектов. Доверительные интервалы строятся по многократным парным выборкам из семи видеопоследовательностей. Во всех расчётах используется один и тот же набор повторов. Вместе с интервалами приводится выборочное стандартное отклонение. Эпизоды, начавшиеся слишком близко к концу записи, исключаются из основной проверки, поскольку для них недоступно полное окно решения.

3. Метод интерпретируемой кратковременной задержки эпизодов инициации треков

Последовательность обработки нового эпизода показана на рисунке 1. Сначала оцениваются его поддержка и текущий баланс бюджета. При низкой поддержке и достаточном балансе первые наблюдения эпизода временно задерживаются на срок не более заданного числа кадров. После завершения этого интервала эпизоду присваивается конечный статус.

Поддержка состоит из двух частей. Абсолютная поддержка показывает, насколько оценка уверенности превышает минимальное рабочее значение. Относительная поддержка характеризует положение нового эпизода относительно эпизодов, ранее появившихся в той же видеозаписи. Рассматриваемый эпизод в историю не включается, поэтому при расчёте используются только сведения, уже доступные к текущему моменту. Итоговая поддержка определяется формулой (4).

$$u_t = (1 - \eta)u_t^{\text{abs}} + \eta u_t^{\text{rel}}. \quad (4)$$

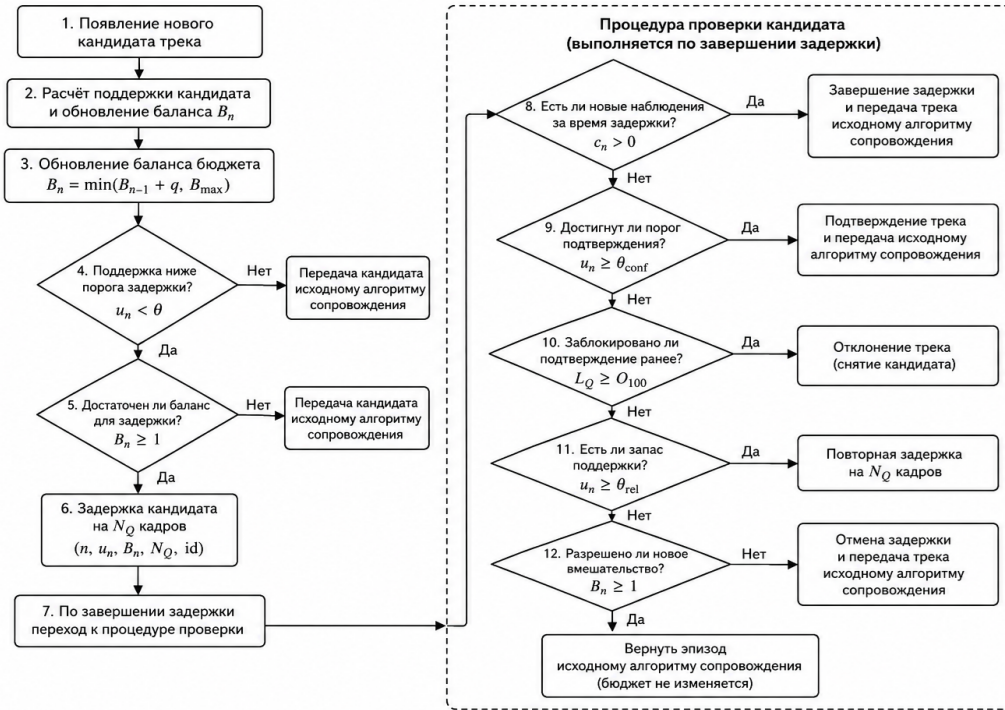


Рис. 1. Схема принятия решения при интерпретируемой кратковременной задержке эпизода инициации трека

Здесь u_t обозначает итоговую поддержку в момент времени t . Величины u_t^{abs} и u_t^{rel} соответствуют абсолютной и относительной поддержке. Параметр η определяет вклад относительной поддержки. Низкая поддержка лишь допускает задержку эпизода. Перед обработкой каждого эпизода инициации баланс пополняется на величину q_k , но не может превысить ёмкость C . Один случай задержки расходует одну единицу бюджета. Изменение баланса задаётся формулой (5).

$$\tilde{b}_k = \min(C, b_{k-1} + q_k), \quad b_k = \tilde{b}_k - a_k. \quad (5)$$

Величины b_{k-1} и b_k обозначают баланс до и после обработки эпизода. Параметр a_k равен единице, если эпизод временно задержан, и нулю в остальных случаях. Когда баланс достигает ёмкости бюджета, часть поступающего бюджета теряется. Эта потеря обозначается r_k . Суммирование изменений баланса даёт точное равенство и верхнюю границу числа вмешательств, приведённую в формуле (6).

$$N_Q(n) = b_0 + \sum_{k=1}^n q_k - \sum_{k=1}^n r_k - b_n \leq \lfloor b_0 + \sum_{k=1}^n q_k \rfloor \leq \lfloor b_0 + nq_{\max} \rfloor. \quad (6)$$

Здесь $N_Q(n)$ обозначает число случаев задержки среди первых n эпизодов, b_0 соответствует начальному балансу, а $q_{\max}$ задаёт наибольшее пополнение. Граница сохраняется при любой последовательности входных оценок. Фактическое число вмешательств может оказаться меньше, если эпизодов с низкой поддержкой недостаточно или часть бюджета теряется при заполненном накопителе. Эта граница ограничивает масштаб воздействия дополнительного слоя. Пока эпизод задержан, подтверждающие сведения накапливаются от

кадра к кадру. Повторные наблюдения увеличивают накопленную поддержку, старые наблюдения постепенно теряют вес, а пропуски уменьшают её значение. Правило накопления приведено в формуле (7).

$$H_t = \max \{0, \lambda^{\Delta t} H_{t-1} - \ln(1 - u_t + \varepsilon) - \mu(\Delta t - 1)\}, \quad P_t = 1 - \exp(-H_t). \quad (7)$$

Величина H_t обозначает накопленную поддержку. Параметр Δt соответствует числу кадров после предыдущего наблюдения, λ определяет скорость забывания, а μ задаёт штраф за пропуск. Значение P_t используется для подтверждения эпизода. Константа $\varepsilon = 10^{-6}$ предотвращает вычисление логарифма нуля. Когда накопленная поддержка достигает порога, эпизод подтверждается, а задержанные наблюдения возвращаются в итоговую карту. Если эпизод наблюдался не менее двух раз, но порог не достигнут, задержка снимается и дальнейшую обработку эпизода продолжает исходный алгоритм сопровождения. Окончательное отклонение допускается только при заранее заданном запрещающем признаке либо при отсутствии повторного наблюдения в полном окне решения. Если запись заканчивается раньше, эпизод помечается как цензурированный.

Один случай задержки временно скрывает ограниченное число наблюдений, поэтому число строк, которыми оперативный выход отличается от результата исходного алгоритма, также имеет верхнюю границу. Она определяется по формуле (8).

$$D(n) \leq L_Q N_Q(n) \leq L_Q \lfloor b_0 + \sum_{k=1}^n q_k \rfloor. \quad (8)$$

Здесь $D(n)$ обозначает число отличающихся строк, а L_Q соответствует наибольшему числу наблюдений одного эпизода, скрытых до завершения решения. Формула ограничивает возможный масштаб изменения выхода, но не определяет, являются ли задержанные наблюдения истинными или ложными. Параметры метода выбираются только по обучающим видеопоследовательностям. Сначала исключаются конфигурации, при которых изменение F_1 выходит за установленный допуск или ухудшается подтверждение истинных треков. Среди оставшихся выбирается вариант с наименьшим значением показателя ложных наблюдений. Если допустимой конфигурации нет, то дополнительная задержка не применяется и алгоритм работает в исходном режиме. Формальные свойства метода приведены в таблице 1.

Интерпретация отдельных решений является частью самого правила и не строится постфактум отдельной моделью. Для каждого нового эпизода сохраняются исходная уверенность, абсолютная и относительная поддержка, баланс бюджета до и после решения, причина немедленной передачи или задержки, накопленная поддержка и конечный статус. Все эти величины непосредственно участвуют в принятии решения. По журналу можно восстановить ход обработки эпизода и установить, почему его наблюдения были немедленно переданы в оперативный выход или задержаны, на каком основании эпизод был подтверждён либо отклонён и какой конечный статус ему присвоен. Запись также показывает, какой запас остаётся до формальной границы вмешательства. Такая интерпретация предназначена для инженерного разбора и обеспечивает прослеживаемость решения.

Таблица 1. Формальные свойства метода ограниченной задержки

Свойство	Формальная запись	Практический смысл
Ограниченное вмешательство	$N_Q(n) \leq \lfloor b_0 + \sum q_k \rfloor$	Число случаев задержки ограничено накопленным бюджетом
Ограничение изменения выхода	$D(n) \leq L_Q N_Q(n)$	Ограничено число временно скрытых наблюдений
Независимость от распределения оценок	Граница выполняется для любой последовательности оценок	Изменение оценок не нарушает установленный предел
Завершение решений	Каждый эпизод получает конечный статус	После обработки не остаётся незавершённых решений
Возврат неопределённого эпизода	Продолжающийся эпизод передаётся исходному алгоритму	Недостаточная поддержка не приравнивается к ложности
Сохранение исходного режима	При отсутствии допустимой конфигурации сохраняется исходный режим	Дополнительный слой не изменяет работу алгоритма без подтверждённой настройки

4. Организация эксперимента

Метод апробирован на семи видеопоследовательностях открытого набора VisDrone [10]. Набор содержит аэросъёмку с небольшими объектами, частичными перекрытиями и заметными изменениями масштаба. Во всех экспериментах использовались заранее сохранённые результаты детектора YOLOv8s, поэтому различия между режимами определялись правилами инициации треков. В качестве исходных алгоритмов сопровождения выбраны ByteTrack, OC-SORT и SORT. Все они получали один и тот же набор обнаружений. Всего обработано 2 846 кадров. Основные показатели рассчитывались по оперативному выходу – наблюдениям, доступным системе во время обработки записи. Эпизоды у конца видеозаписи исключались из основной проверки, если для них не хватало полного окна решения.

Для оценки использована схема последовательного исключения одной видеозаписи. На каждой итерации параметры выбирались по шести последовательностям, а седьмая применялась только для проверки, затем проверочная запись менялась. Таким образом, каждая последовательность один раз использовалась для итоговой оценки и не участвовала в соответствующей настройке метода. Случайный лес обучался заново внутри каждой итерации внешней проверки. Проверочная последовательность не использовалась ни при обучении, ни при выборе порога и настройке бюджета. Признаки нового эпизода не содержали сведений из будущих кадров. После применения каждого правила наблюдения оперативного выхода заново сопоставлялись с разметкой, что исключало искажение оценки после задержки отдельных наблюдений. Проверенные значения параметров и выбранные конфигурации для каждого алгоритма сопровождения приведены в таблице 2.

Таблица 2. Проверенные значения и выбранные параметры кратковременной задержки

Параметр	Проверенные значения	ByteTrack	OC-SORT и SORT
$q_{\max}$, предельное пополнение бюджета	0; 0,01; 0,03; 0,05	0,05 во всех проверках	0 в шести проверках; 0,05 в одной
C , ёмкость бюджета	2	2	2
T_Q , длительность задержки, кадров	1; 2	2	1
η , вес относительной поддержки	0,4; 0,6	0,6	0,4
Порог немедленной передачи	0,60; 0,65; 0,80	0,60	0,65 в шести проверках; 0,80 в одной
λ , коэффициент забывания	0,75; 0,90	0,90	0,75
μ , штраф за пропуск	0,15	0,15	0,15
τ_C , порог подтверждения	0,50; 0,60; 0,65; 0,70	0,70	0,50 в шести проверках; 0,65 в одной

Для ByteTrack режим с включённой задержкой был выбран во всех семи итерациях внешней проверки. Для OC-SORT и SORT в шести случаях из семи по результатам настройки задержка не применялась. Это означает, что на большинстве обучающих разбиений улучшение для этих алгоритмов не удовлетворяло установленным критериям. Для случайного леса проверялись глубины деревьев 6 и 10, минимальное число обучающих примеров в конечной вершине 4 и 8, а также пороги решения 0,35, 0,50 и 0,65. Модель включала 80 деревьев, дисбаланс классов учитывался сбалансированными весами. В варианте с бюджетным ограничением во всех итерациях были выбраны порог 0,65, предельное пополнение бюджета 0,05 и ёмкость 2. Измерения времени выполнения проводились на процессоре AMD Ryzen 7 7840HS. Перед измерением выполнялись два предварительных запуска, результаты которых не учитывались. Затем время выполнения оценивалось по 30 повторным измерениям. Учитывалось только время работы дополнительного правила на сохранённых выходах ByteTrack. Обнаружение объектов, исходное сопровождение, обучение случайного леса и сопоставление с разметкой в измерение не входили. На время измерений автоматическое освобождение неиспользуемой памяти и подробная запись служебных сообщений были отключены. Приведённые значения характеризуют вычислительные затраты дополнительного слоя, а не полное время работы системы восприятия.

5. Результаты

Основные показатели оперативного выхода приведены в таблице 3. Для ByteTrack кратковременная задержка уменьшила число ложных наблюдений с 46 368 до 45 943, а число ложных инициаций – с 10 031 до 9 855. Правило изменило обработку 392 из 14 967 новых эпизодов, то есть 2,62%.

Таблица 3. Основные показатели оперативного выхода

Алгоритм	Режим	F_1	Ложные наблюдения на 100 кадров	Ложные инициации на 100 кадров	Подтверждение истинных треков
ByteTrack	Исходный	0,403972	2017,96	457,03	0,6428
ByteTrack	Задержка	0,404400	1998,61	449,02	0,6418
OC-SORT	Исходный	0,392258	1130,66	164,63	0,5614
OC-SORT	Задержка	0,392276	1128,99	164,27	0,5595
SORT	Исходный	0,391946	1195,35	179,49	0,5632
SORT	Задержка	0,391942	1194,25	179,13	0,5613

Для OC-SORT и SORT изменения оказались значительно меньше. Режим с задержкой включался только в одной из семи итераций внешней проверки. В остальных случаях настройка оставляла исходный режим, поскольку на обучающих последовательностях безопасное улучшение не подтверждалось. Средние значения сами по себе не показывают, насколько результат устойчив при переходе между видеозаписями, поэтому изменения F_1 и показателя ложных наблюдений дополнительно проверялись по условиям (2) и (3). Результаты приведены в таблице 4.

Таблица 4. Проверка сохранения качества и снижения числа ложных наблюдений

Алгоритм	Показатель	Среднее изменение	Стандартное отклонение	95%-й доверительный интервал	Результат
ByteTrack	H_1, F_1	+0,000428	0,000341	[+0,000220; +0,000682]	Условие выполнено
ByteTrack	H_2, O_{100}	-19,35	10,08	[-26,047; -12,279]	Условие выполнено
OC-SORT	H_1, F_1	+0,000018	0,000047	[+0,000000; +0,000054]	Условие выполнено
OC-SORT	H_2, O_{100}	-1,67	4,43	[-5,024; 0,000]	Условие не выполнено
SORT	H_1, F_1	-0,000005	0,000012	[-0,000014; 0,000000]	Условие выполнено
SORT	H_2, O_{100}	-1,10	2,91	[-3,297; 0,000]	Условие не выполнено

Для ByteTrack выполнены оба условия: нижняя граница доверительного интервала изменения F_1 остаётся выше допустимого уровня, а верхняя граница интервала изменения показателя ложных наблюдений находится ниже нуля. Снижение числа ложных наблюдений тем самым подтверждено при сохранении качества в установленном диапазоне. Для OC-SORT и SORT условие сохранения качества также выполнено, однако верхняя граница интервала изменения показателя ложных наблюдений достигает нуля. Отдельно роль бюджетного ограничения оценивалась на сохранённых результатах ByteTrack. Кратковременная задержка сравнивалась с порогом уверенности, правилом подтверждения М из N, байесовским подтверждением и случайным лесом. Результаты приведены в таблице 5.

Таблица 5. Сравнение правил инициации на сохранённых результатах ByteTrack

Правило	Изменение F_1 и результат	Изменение O_{100} и результат	Доля вмешательства	Обучение
Порог уверенности	0,000000; условие выполнено	0,00; условие не выполнено	0%	Нет
Правило подтверждения М из N	-0,046263; условие не выполнено	-1386,84; не проверялось	Не рассчитывалась	Нет
Байесовское подтверждение	-0,187741; условие не выполнено	-1846,46; не проверялось	Не рассчитывалась	Нет
Случайный лес	-0,084383; условие не выполнено	-1084,26; не проверялось	Не рассчитывалась	Да
Случайный лес с бюджетом	-0,000346; условие выполнено	-85,83; условие выполнено	4,81%	Да
Кратковременная задержка	+0,000428; условие выполнено	-19,35; условие выполнено	2,60%	Нет

Правило подтверждения М из N, байесовское подтверждение и случайный лес без бюджета заметно уменьшили число ложных наблюдений, но одновременно снизили F_1 за пределами установленного допуска. Случайный лес с бюджетным ограничением выполнил оба условия и сильнее снизил показатель ложных наблюдений. Для его обучения требовались размеченные эпизоды, тогда как предлагаемое правило не содержит обучаемого классификатора. Средняя доля затронутых эпизодов составила 4,81%. Для кратковременной задержки среднее по итерациям внешней проверки было равно 2,60%, а доля 2,62% рассчитана по общему числу эпизодов во всех последовательностях.

Характеристики эпизодов ложной инициации оценивались на одном и том же наборе до и после применения правила. Если эпизод полностью исчезал из оперативного выхода, его длительность в оперативном выходе принималась равной нулю. Из-за длинного хвоста распределения и небольшой доли затронутых эпизодов изменения средних характеристик оказались невелики. При этом для ByteTrack из оперативного выхода были исключены 176 ложных инициаций и 425 ложных наблюдений. Распределение длительности и числа наблюдений в эпизодах до и после применения задержки приведено в таблице 6.

Формальные ограничения проверялись на каждом начальном фрагменте видеозаписей. Всего проанализировано 29 385 таких фрагментов. Результаты приведены в таблице 7.

Нарушений балансового равенства и границ, заданных формулами (6) и (8), не обнаружено. Наибольшее отношение фактического изменения выхода к бюджетному пределу составило 0,675, а к границе по числу случаев задержки в отдельных фрагментах достигало единицы. Эти результаты подтверждают соблюдение заданных ограничений. Вычислительные затраты измерялись на сохранённых выходах ByteTrack. Результаты приведены в таблице 8.

Среднее время работы кратковременной задержки составило 0,901 мс

Таблица 6. Характеристики эпизодов ложной инициации

Алгоритм и режим	Число ложных эпизодов	Средняя длительность, кадры	Медиана	P95	Среднее число наблюдений
ByteTrack без задержки	12 260	5,395	1	20	3,831
ByteTrack с задержкой	12 260	5,373	1	20	3,814
OC-SORT без задержки	5 210	6,308	3	22	5,307
OC-SORT с задержкой	5 210	6,307	3	22	5,306
SORT без задержки	5 479	6,109	3	21	5,233
SORT с задержкой	5 479	6,107	3	21	5,231

Таблица 7. Проверка границы изменения оперативного выхода

Алгоритм	Активные итерации внешней проверки	Наибольшее $D(n)$	Наибольшее отношение $D(n)/(L_Q N_Q(n))$	Отношение к бюджетной границе	Нарушения
ByteTrack	7	105	0,722	0,552	0
OC-SORT	1	27	1,000	0,675	0
SORT	1	27	1,000	0,659	0

Таблица 8. Время работы правил на сохранённых результатах ByteTrack

Правило	Число измерений	Среднее время, мс/кадр	Медиана, мс/кадр	95-й процентиль, мс/кадр
Порог уверенности	30	0,442	0,441	0,451
Правило подтверждения М из N	30	1,061	1,061	1,070
Байесовское подтверждение	30	0,625	0,623	0,642
Случайный лес	30	0,481	0,488	0,517
Случайный лес с бюджетом	30	0,552	0,556	0,592
Кратковременная задержка	30	0,901	0,899	0,925

на кадр, случайного леса с бюджетом – 0,552 мс. Предлагаемое правило хранит состояние эпизодов, временно сохраняет наблюдения и обновляет относительную поддержку. Для случайного леса использовался заранее сформированный набор признаков, поэтому сравнение отражает стоимость выполнения правил, а не полный цикл подготовки данных.

6. Обсуждение результатов и ограничения

Случайный лес с бюджетом снизил показатель ложных наблюдений сильнее предлагаемого правила и сохранил изменение F_1 в допустимом диапазоне, но затронул более высокую долю эпизодов. Если имеется актуальная размеченная выборка, соответствующая конкретной камере, высоте съёмки и типу сцены, и более высокая доля вмешательств допустима, обучаемый метод может быть предпочтительнее по величине эффекта. Кратковременная задержка рассчитана на условия, в которых обучение и регулярное обновление классификатора затруднены. Во время работы метод не использует разметку. Решение определяется поддержкой нового эпизода и текущим состоянием бюджета, а журнал сохраняет обе величины вместе с причиной конечного статуса. Каждое вмешательство можно разобрать без приближённой внешней модели, а общее число задержек и отличие оперативного выхода остаются ограниченными формулами (6) и (8). Сравнение со случайным лесом без бюджета показывает роль ограничения. Неограниченная модель исключила больше ложных наблюдений, но ухудшила F_1 за пределами установленного допуска. После введения бюджета оба условия были выполнены. Полученный результат показывает, что бюджетное ограничение может применяться как к обучаемому классификатору, так и к необучаемому правилу.

Для OC-SORT и SORT задержка применялась только в одной из семи итераций внешней проверки. В остальных случаях настройка оставляла исходный режим. Качество сохранилось, но устойчивое снижение числа ложных наблюдений не подтвердилось. Основным практический результат относится к ByteTrack. Имеющиеся данные не позволяют распространять его на все алгоритмы сопровождения.

Исследование ограничено семью последовательностями VisDrone и результатами одной модели обнаружения [10]. Эти записи содержат малые объекты, частичные перекрытия и изменения масштаба, но не охватывают всего разнообразия транспортных сцен. Переносимость полученных результатов на другие наборы данных, камеры, условия полёта и детекторы в настоящей работе не проверялась. В дальнейшем такую проверку целесообразно провести, например, на U2UData и V2U4Real [11, 12]. Формальная граница задаёт максимальный масштаб изменения оперативного выхода при любой последовательности входных оценок. Перед практическим применением метод необходимо оценить на независимых данных и проверить, как ложные треки влияют на конкретный модуль планирования движения.

Дополнительное ограничение связано с граничными сценариями, включая исчерпание бюджета вмешательств, отсутствие повторного наблюдения эпизода и неполное окно решения в конце последовательности. Эти случаи обрабатываются правилами метода, однако их отдельная количественная оценка требует более крупной независимой выборки.

С учётом указанных ограничений количественные выводы настоящей

работы относятся к исследованным последовательностям VisDrone, используемому детектору YOLOv8s и рассмотренным алгоритмам сопровождения. Перспективы дальнейших исследований включают независимую проверку на U2UData, V2U4Real и других наборах данных, применение дополнительных детекторов и камер, отдельные стресс-тесты перечисленных граничных сценариев и оценку влияния задержки на решения реального модуля планирования движения.

Заключение

Предложен интерпретируемый метод кратковременной задержки эпизодов инициации треков. Накопительный бюджет заранее ограничивает число вмешательств, а журнал решения сохраняет поддержку эпизода, текущий баланс бюджета, причину задержки и конечный статус. Балансовое равенство задаёт верхнюю границу числа задержек и возможного изменения оперативного выхода. Проверка 29 385 начальных фрагментов не выявила нарушений этих ограничений.

Основной результат получен для ByteTrack. Число ложных наблюдений в оперативном выходе уменьшилось на 425, число ложных инициаций – на 176. Правило изменило обработку 392 из 14 967 новых эпизодов, что составило 2,62%. Изменение F_1 осталось в допустимом диапазоне, а снижение числа ложных наблюдений было подтверждено статистически. Для OC-SORT и SORT качество также сохранилось, однако устойчивого снижения числа ложных наблюдений не установлено. Случайный лес с бюджетом сильнее снизил число ложных наблюдений, но требовал размеченных обучающих эпизодов и затрагивал более высокую долю эпизодов инициации. Предлагаемая задержка не использует обучаемый классификатор и сохраняет заранее заданный предел воздействия.

Исходный код, конфигурации, сценарии расчёта и поля журнала интерпретации опубликованы в открытом репозитории¹. Следующий этап – независимая проверка на других данных и камерах, а также оценка метода в составе реального модуля планирования движения.

Благодарности

Исследование выполнено в рамках государственного задания Министерства науки и высшего образования Российской Федерации, тема № 124112200072-2.

Список литературы

- [1] Kalman, R. E. A New Approach to Linear Filtering and Prediction Problems / R. E. Kalman // Journal of Basic Engineering. – 1960. – Vol. 82, No. 1. – P. 35–45. – DOI 10.1115/1.3662552.
- [2] Simple Online and Realtime Tracking / A. Bewley, Z. Ge, L. Ott [et al.] // Proceedings of the IEEE International Conference on Image Processing. – Piscataway, NJ : IEEE, 2016. – P. 3464–3468. – DOI 10.1109/ICIP.2016.7533003.

¹Defense4UAVSwarm [Электронный ресурс]. URL: <https://github.com/lebedeffson/Defense4UAVSwarm> (дата обращения: 17.07.2026).

- [3] ByteTrack: Multi-Object Tracking by Associating Every Detection Box / Y. Zhang, P. Sun, Y. Jiang [et al.] // European Conference on Computer Vision. – Cham : Springer, 2022. – P. 1–21. – DOI 10.1007/978-3-031-20047-2_1.
- [4] Observation-Centric SORT: Rethinking SORT for Robust Multi-Object Tracking / J. Cao, J. Pang, X. Weng [et al.] // Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. – Piscataway, NJ : IEEE, 2023. – P. 9686–9696. – DOI 10.1109/CVPR52729.2023.00934.
- [5] StrongSORT: Make DeepSORT Great Again / Y. Du, Z. Zhao, Y. Song [et al.] // IEEE Transactions on Multimedia. – 2023. – Vol. 25. – P. 8725–8737. – DOI 10.1109/TMM.2023.3240881.
- [6] Wojke, N. Simple Online and Realtime Tracking with a Deep Association Metric / N. Wojke, A. Bewley, D. Paulus // Proceedings of the IEEE International Conference on Image Processing. – Piscataway, NJ : IEEE, 2017. – P. 3645–3649. – DOI 10.1109/ICIP.2017.8296962.
- [7] Bergmann, P. Tracking Without Bells and Whistles / P. Bergmann, T. Meinhardt, L. Leal-Taixé // Proceedings of the IEEE/CVF International Conference on Computer Vision. – Piscataway, NJ : IEEE, 2019. – P. 941–951. – DOI 10.1109/ICCV.2019.00103.
- [8] Breiman, L. Random Forests / L. Breiman // Machine Learning. – 2001. – Vol. 45, No. 1. – P. 5–32. – DOI 10.1023/A:1010933404324.
- [9] Geifman, Y. Selective Classification for Deep Neural Networks / Y. Geifman, R. El-Yaniv // Advances in Neural Information Processing Systems. – 2017. – Vol. 30. – P. 4878–4887.
- [10] VisDrone-VDT2018: The Vision Meets Drone Video Detection and Tracking Challenge Results / P. Zhu, L. Wen, D. Du [et al.] // Computer Vision – ECCV 2018 Workshops. – Cham : Springer, 2019. – P. 496–518. – DOI 10.1007/978-3-030-11021-5_29.
- [11] U2UData: A Large-scale Cooperative Perception Dataset for Swarm UAVs Autonomous Flight / T. Feng, X. Wang, F. Han [et al.] // Proceedings of the 32nd ACM International Conference on Multimedia. – New York : ACM, 2024. – P. 7600–7608. – DOI 10.1145/3664647.3681151.
- [12] V2U4Real: A Real-world Large-scale Dataset for Vehicle-to-UAV Cooperative Perception / W. Li, H. Xiang, T. Wang [et al.] // Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. – Piscataway, NJ : IEEE, 2026. – P. 4728–4737.

INTERPRETABLE CONTROL OF TRACK INITIATION WITH PREDEFINED INTERVENTION BOUNDS IN UNMANNED AERIAL VEHICLE PERCEPTION SYSTEMS

Trofimov Yu. V.^{1, a}, Averkin A. N.^{1, 2, b}, Shevchenko A. V.^{1, c}, Kuznetsov E. M.^{1, d},
Lebedev A. D.^{1, e}

¹ Dubna State University, Dubna, Russia.

² Federal Research Center "Computer Science and Control" of the Russian Academy of Sciences, Moscow, Russia.

E-mail: ^a ura_trofim@bk.ru, ^b averkin2003@inbox.ru, ^c leviathan0909@gmail.com, ^d kem.22@uni-dubna.ru,

^e lad.24@uni-dubna.ru

Abstract. In aerial monitoring systems, a single detection error may develop into a persistent false track and affect data passed to downstream components. The study aims to reduce such errors while bounding the scale of intervention in advance and retaining a traceable account of each decision. An external method is proposed that briefly delays low-support track-initiation episodes. The number of delays is controlled by an accumulated budget, while the support estimate, current budget balance, reason for intervention, and final status are recorded for each episode. The method was evaluated on seven video sequences from a publicly available aerial-imagery dataset containing small objects, partial occlusions, and substantial scale changes. Three tracking algorithms and a supervised ensemble of decision trees were compared. For ByteTrack, the delay was applied to 2.62 percent of track-initiation episodes, removing 425 false observations and 176 false initiations from the operational output, while the change in the combined precision-recall measure remained within the predefined tolerance. For OC-SORT and SORT, the reduction in false observations was not confirmed under the adopted statistical criterion. An examination of 29,385 initial video fragments found no violations of the predefined limits. The additional processing time averaged 0.901 milliseconds per frame. The results show that a short delay can reduce the number of false observations while keeping the scale of intervention under control.

Keywords: intelligent transport systems, unmanned aerial vehicles, object tracking, track initiation, bounded intervention, interpretable artificial intelligence, short-term delay.

References

- [1] Kalman, R. E. (1960). A new approach to linear filtering and prediction problems. *Journal of Basic Engineering*, 82(1), 35–45. <https://doi.org/10.1115/1.3662552>
- [2] Bewley, A., Ge, Z., Ott, L., Ramos, F., & Upcroft, B. (2016). Simple online and realtime tracking. In *Proceedings of the IEEE International Conference on Image Processing* (pp. 3464–3468). IEEE. <https://doi.org/10.1109/ICIP.2016.7533003>
- [3] Zhang, Y., Sun, P., Jiang, Y., Yu, D., Weng, F., Yuan, Z., Luo, P., Liu, W., & Wang, X. (2022). ByteTrack: Multi-object tracking by associating every detection box. In *European Conference on Computer Vision* (pp. 1–21). Springer. https://doi.org/10.1007/978-3-031-20047-2_1

For citation:

Trofimov Yu. V., Averkin A. N., Shevchenko A. V., Kuznetsov E. M., Lebedev A. D. Interpretable Control of Track Initiation with Predefined Intervention Bounds in Unmanned Aerial Vehicle Perception Systems. *Intelligent transport*, 2026, no. 3(39), pp. 69–84. DOI: 10.24412/3033-6007-2026-339-69-84. (in Russian)

- [4] Cao, J., Pang, J., Weng, X., Khirodkar, R., & Kitani, K. (2023). Observation-centric SORT: Rethinking SORT for robust multi-object tracking. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition* (pp. 9686–9696). IEEE. <https://doi.org/10.1109/CVPR52729.2023.00934>
- [5] Du, Y., Zhao, Z., Song, Y., Zhao, Y., Su, F., Gong, T., & Meng, H. (2023). StrongSORT: Make DeepSORT great again. *IEEE Transactions on Multimedia*, 25, 8725–8737. <https://doi.org/10.1109/TMM.2023.3240881>
- [6] Wojke, N., Bewley, A., & Paulus, D. (2017). Simple online and realtime tracking with a deep association metric. In *Proceedings of the IEEE International Conference on Image Processing* (pp. 3645–3649). IEEE. <https://doi.org/10.1109/ICIP.2017.8296962>
- [7] Bergmann, P., Meinhardt, T., & Leal-Taixé, L. (2019). Tracking without bells and whistles. In *Proceedings of the IEEE/CVF International Conference on Computer Vision* (pp. 941–951). IEEE. <https://doi.org/10.1109/ICCV.2019.00103>
- [8] Breiman, L. (2001). Random forests. *Machine Learning*, 45(1), 5–32. <https://doi.org/10.1023/A:1010933404324>
- [9] Geifman, Y., & El-Yaniv, R. (2017). Selective classification for deep neural networks. In *Advances in Neural Information Processing Systems* (Vol. 30, pp. 4878–4887). Curran Associates.
- [10] Zhu, P., Wen, L., Du, D., et al. (2019). VisDrone-VDT2018: The vision meets drone video detection and tracking challenge results. In *Computer Vision – ECCV 2018 Workshops* (pp. 496–518). Springer. https://doi.org/10.1007/978-3-030-11021-5_29
- [11] Feng, T., Wang, X., Han, F., Zhang, L., & Zhu, W. (2024). U2UData: A large-scale cooperative perception dataset for swarm UAVs autonomous flight. In *Proceedings of the 32nd ACM International Conference on Multimedia* (pp. 7600–7608). ACM. <https://doi.org/10.1145/3664647.3681151>
- [12] Li, W., Xiang, H., Wang, T., Wu, S., Xia, Q., Wang, C., & Wen, C. (2026). V2U4Real: A real-world large-scale dataset for vehicle-to-UAV cooperative perception. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition* (pp. 4728–4737). IEEE.