

ИНФОРМАТИКА**INFORMATICS**

УДК 004

<https://doi.org/10.29235/1561-8323-2026-70-2-102-107>

Поступило в редакцию 24.02.2026

Received 24.02.2026

В. А. Иванюкович, Б. А. Залесский

*Объединенный институт проблем информатики Национальной академии наук Беларуси,
Минск, Республика Беларусь
ул. Сурганова, 6, 220012, Минск, Республика Беларусь*

**УСОВЕРШЕНСТВОВАННЫЙ АЛГОРИТМ СОПРОВОЖДЕНИЯ ОБЪЕКТА
ОДНОГО ИЗ НАПЕРЕД ЗАДАНЫХ ТИПОВ**

(Представлено членом-корреспондентом А. В. Тузиковым)

Аннотация. Разработан новый нейросетевой алгоритм отслеживания объекта, наблюдаемого на кадрах видеопотока. Алгоритм позволяет автоматически обнаруживать объект одного из наперед заданных типов, надежно сопровождать его в дальнейшем, быстро повторно обнаруживать объект в случае срыва сопровождения, детектировать другой объект нужного типа при исчезновении сопровождаемого объекта. Обнаружение объекта интереса на кадрах видео осуществляется с помощью нейронной сети-детектора, а сопровождение – разработанным алгоритмом, использующим нейронную сеть-трансформер.

Ключевые слова: видеопоток, обнаружение объекта, отслеживание объекта, нейронная сеть-детектор, сеть-трансформер

Для цитирования. Иванюкович, В. А. Усовершенствованный алгоритм сопровождения объекта одного из наперед заданных типов / В. А. Иванюкович, Б. А. Залесский // Доклады Национальной академии наук Беларуси. – 2026. – Т. 70, № 2. – С. 102–107. <https://doi.org/10.29235/1561-8323-2026-70-2-102-107>

Vladimir A. Ivanyukovich, Boris A. Zalesky

*United Institute of Informatics Problems of the National Academy of Sciences of Belarus, Minsk, Republic of Belarus
6, Surganov Str., 220012, Minsk, Republic of Belarus*

**IMPROVED ALGORITHM FOR TRACKING AN OBJECT OF ONE
OF THE SEVERAL PREDEFINED TYPES**

(Communicated by Corresponding Member Aleksandr V. Tuzikov)

Abstract. A new neural network algorithm for tracking objects observed in frames of video has been developed. The algorithm enables automatic detection of objects of one of the predefined types, reliable subsequent tracking, rapid redetection of the object if tracking was interrupted, and detection of a different object of the desired type if the tracked object disappears. Detection of the object of interest in video frames is performed using a neural network detector, and tracking is carried out by the developed algorithm using a neural network transformer.

Keywords: video, object detection, object tracking, neural network detector, transformer

For citation. Ivanyukovich V. A., Zalesky B. A. Improved algorithm for tracking an object of one of the several predefined types. *Doklady Natsional'noi akademii nauk Belarusi = Doklady of the National Academy of Sciences of Belarus*, 2026, vol. 70, no. 2, pp. 102–107 (in Russian). <https://doi.org/10.29235/1561-8323-2026-70-2-102-107>

Введение. Задача сопровождения объектов одного из заранее заданных типов, наблюдаемых на кадрах видеопотока, до сих пор остается одной из важнейших задач компьютерного зрения. В сообщении она понимается следующим образом: требуется в автоматическом режиме обнаружить на кадре объект одного из наперед заданных типов, захватить и сопровождать его до момен-

та исчезновения. В случае срыва сопровождения выполнить повторное обнаружение и захват сопровождаемого ранее или другого объекта этого же или другого типа. Для таких задач сопровождения одного объекта используется общепринятое сокращение – SOT (Single Object Tracking).

SOT-алгоритмам посвящено большое число публикаций [1; 2], в том числе обзорных. В обстоятельной статье [3] приведен подробный обзор современных SOT-трекеров, описан принцип их работы, а также проведен сравнительный анализ характеристик. Упомянутые трекеры можно классифицировать по принципу работы на три основных типа: дискриминационные, сиамские и трансформеры (а также комбинации первых двух типов с трансформерами) (рис. 1).

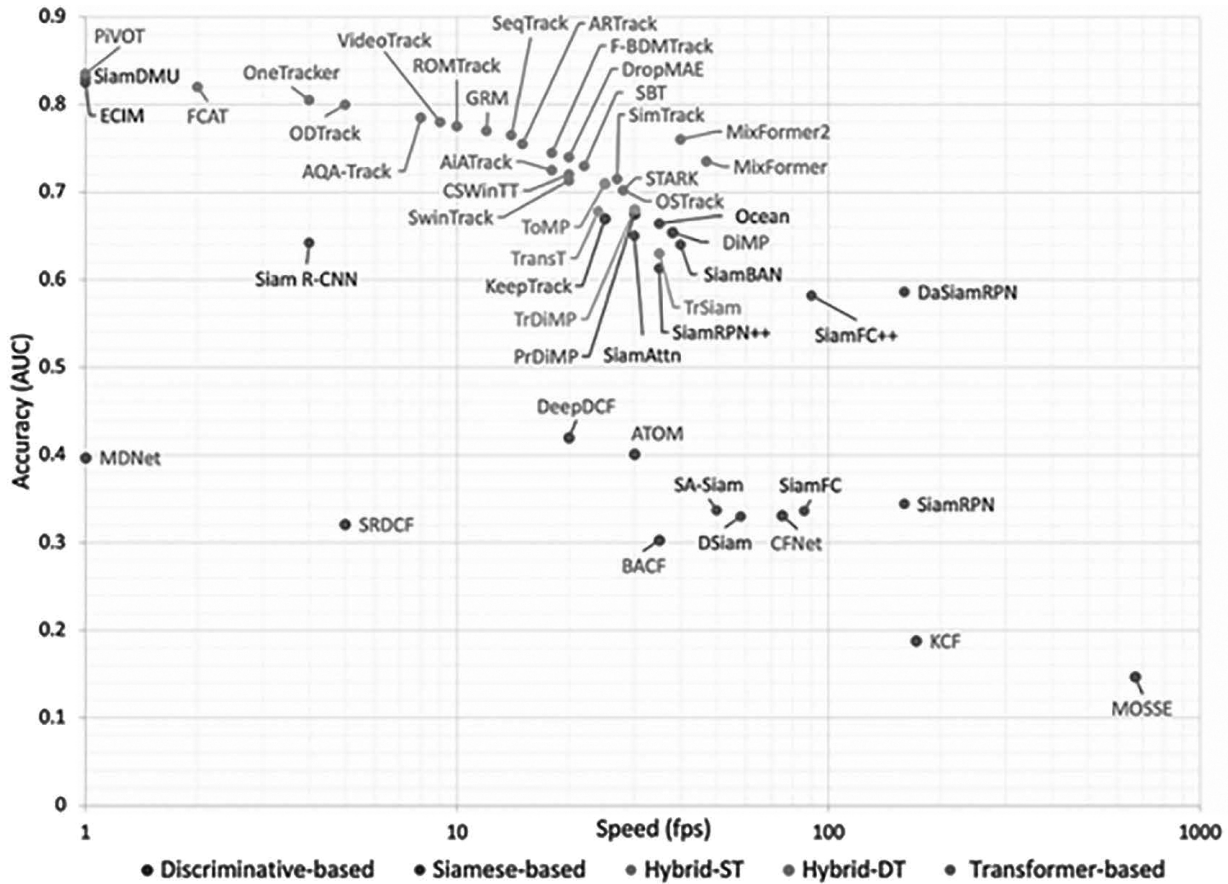


Рис. 1. Сравнительные характеристики известных в настоящее время трекеров [3]

Fig. 1. Comparative characteristics of currently known trackers [3]

Трекер, основанный на трансформерах, занимает в настоящее время лидирующие позиции [3].

Материалы и методы исследования. Предложенный нами трекер имеет смешанный тип, так как в нем используются нейросетевой детектор объектов и трансформер для нахождения особых точек (англоязычный термин «key points») на кадрах, а также разработанные нами блоки детекции и слежения, выполняющие трекинг захваченного объекта.

В [4] был представлен разработанный нами алгоритм сопровождения объекта одного из наперед заданных типов, основанный на взаимодействии нейросетевого детектора объектов и блока слежения, в котором используются полученные детектором прямоугольные рамки, ограничивающие объект (bounding boxes) и оптический поток.

Предлагаемый алгоритм основан на взаимодействии нейросетевого детектора и двух блоков слежения. Его первый блок слежения, основанный на анализе найденных детектором bounding boxes объектов, аналогичен блоку слежения разработанного ранее алгоритма, но второй блок слежения основан на анализе перемещения на кадрах видео особых точек, которые находятся нейронной сетью трансформером. Трансформер вычисляет на текущем кадре образы точек, найденных им на предыдущем, точнее, чем оптический поток.

Для автоматического обнаружения объектов могут использоваться разные детекторы объектов, выполняющиеся на вычислителе в режиме реального времени. В нашем случае был использован нейросетевой детектор YoloV10, который показал лучшие результаты на наших датасетах по сравнению с YoloV11 и YoloV12.

Для тестирования и сравнительного анализа разработанный алгоритм, как и предложенный нами ранее в [4], применялся при решении актуальной задачи автоматического обнаружения и сопровождения находящихся в воздухе коптеров, наблюдаемых видеокамерой с земной поверхности. Тестовый датасет был составлен из 17 видео с различными коптерами, снятыми разными камерами. Общее количество кадров тестового датасета 39514.

Нейросетевой алгоритм сопровождения объекта условно можно представить в виде трех блоков. Первый блок, применяемый для каждого кадра, выполняет обнаружение объекта нейросетевым детектором. Второй блок осуществляет сопровождение обнаруженного объекта с помощью нейросетевого детектора, а также нахождение трансформером особых точек на изображении объекта. Третий – выполняет сопровождение обнаруженного объекта на основе ключевых точек, найденных трансформером. В третьем блоке для сопровождения объекта используется решение, предложенное авторами.

Блок-схема на рис. 2 иллюстрирует взаимодействие блоков трекера. Переменная $I(t)$ обозначает текущий кадр.

Для подробного описания трекера используются следующие обозначения: t – номер кадра, используемый в качестве дискретного времени; I_t – кадр в момент времени t ; B – прямоугольник, ограничивающий найденный детектором объект; B_{final} – прямоугольник, ограничивающий объект, найденный трекером на текущем кадре; B_{search} – область поиска объекта на кадре с помощью детектора; m – счетчик числа кадров после неудачного обнаружения объекта детектором, на которых положение объекта предсказывается с помощью особых точек; m_{max} – максимально допустимое значение переменной m ; B_{prev} – прямоугольник, ограничивающий объект, найденный трекером на предыдущем кадре (если он был найден); B_m – прямоугольная область для обнаружения особых точек; P_0 – массив координат 2d особых точек p_0 , найденных на предыдущем кадре; P_1 – массив координат 2d особых точек p_1 , найденных на текущем кадре в области B_m ; trk_m – флаг, указывающий на сопровождение объекта ключевыми точками; trk – флаг, указывающий на сопровождение объекта в общем случае.

Приведем пошаговое описание алгоритма.

Шаг 1. Задать начальные значения $t = 0$, $m = 0$, $B_{\text{prev}} = \emptyset$, $\text{trk}_m = \text{false}$, $\text{trk} = \text{false}$.

Шаг 2. Читать текущий кадр I_t .

Шаг 3. Если $I_t = \emptyset$, STOP, иначе задать область поиска объекта на кадре детектором $B_{\text{search}} = I_t$.

Шаг 4. Выполнить обнаружение объектов на области B_{search} . Выбрать ограничивающий прямоугольник B , найденный детектором, с максимальным уровнем доверия. Если ограничивающий прямоугольник был найден, перейти на шаг 7.

Шаг 5. Если $\text{trk} = \text{false}$, перейти к шагу 16.

Шаг 6. Присвоить B_{search} область вокруг B_{prev} и провести детекцию в области B_{search} . Выбрать ограничивающий прямоугольник B , найденный детектором, с максимальным уровнем доверия. Если ограничивающий прямоугольник был найден, перейти на шаг 8, иначе перейти к шагу 10.

Шаг 7. Если $\text{trk} = \text{false}$, перейти к шагу 9.

Шаг 8. Если найденный ограничивающий прямоугольник B не пересекается с B_{prev} , перейти к шагу 10.

Шаг 9. $\text{trk}_m = \text{false}$, $\text{trk} = \text{true}$, $B_{\text{final}} = B$, $m = 0$, B_m – область вокруг B_{final} . Перейти к шагу 11.

Шаг 10. $\text{trk}_m = \text{true}$, $m = m + 1$, B_m – область вокруг B_{prev} .

Шаг 11. С помощью нейронной сети трансформера найти координаты особых точек P_1 на изображении I_t в области B_m .

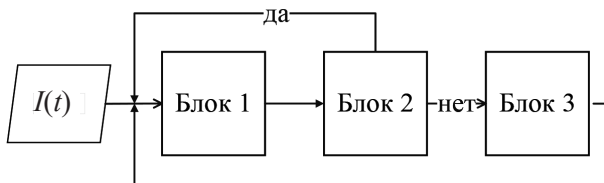


Рис. 2. Схема взаимодействия блоков трекера

Fig. 2. Diagram of tracker blocks interaction

Шаг 12. Если $trk_m = true$, $len(P_0) > 0$ и $len(P_1) > 0$, перейти к шагу 13. Иначе перейти к шагу 15.

Шаг 13. Получить B_{final} с помощью сдвига B_{prev} на вектор, равный разности среднеарифметических значений точек множеств P_1 и P_0 .

Шаг 14. Если $m > m_{max}$, то перейти к шагу 16.

Шаг 15. $P_0 = P_1$, $B_{prev} = B_{final}$. Перейти на шаг 17.

Шаг 16. $trk = false$, $trk_m = false$, $m = 0$.

Шаг 17. $t = t + 1$, вернуться на шаг 2.

На рис. 3 представлена блок-схема алгоритма.

Результаты и их обсуждение. Для тестирования алгоритма и исследования его характеристик был создан датасет, состоящий из кадров 17 видео с разными коптерами, снятых разными камерами. Тестовый датасет не использовался при обучении алгоритма. Часть датасета была создана и размечена авторами сообщения, а другая часть сформирована из видео, найденных в интернете. Весь набор данных состоит из 39514 кадров.

На кадрах созданной и размеченной авторами части датасета изображен квадрокоптер, летящий высоко в небе, в большинстве случаев малого размера (менее 20×20 при разрешении видео 1920×1080) (рис. 4).

Вторая часть составлена из восьми видео полета дронов. На одном из этих видео изображен коптер, летящий на низкой высоте на фоне плотной городской застройки, что значительно затрудняет его обнаружение и отслеживание.

Характеристики разработанного трекера сравниваются с характеристиками опубликованного нами ранее в [4] трекера и известного трекера Vutetrack [5], реализованного фирмой Ultralytics (табл. 1).

Во всех трех трекерах использовался обученный нами нейросетевой детектор YOLOv10, который, как упоминалось ранее, показал лучшие результаты на наших наборах данных. Вычислялись стандартные характеристики трекеров: Recall, Precision и F1.

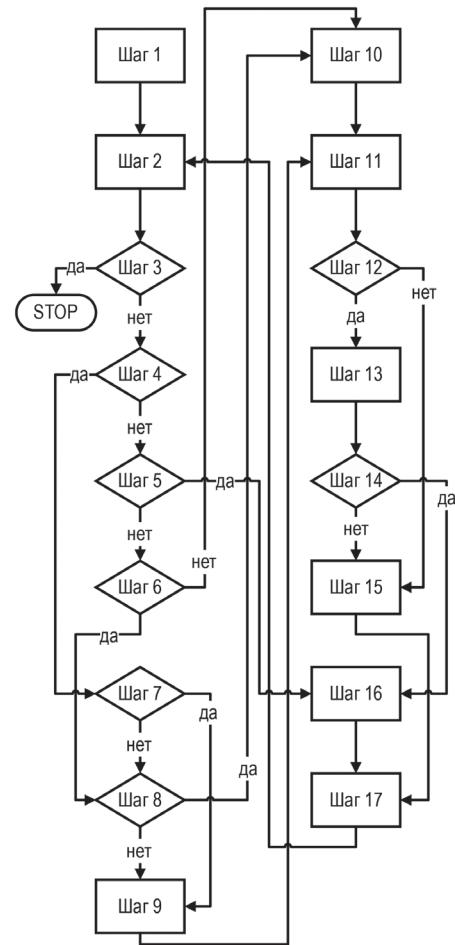


Рис. 3 Блок-схема алгоритма
Fig. 3. Algorithm flowchart



Рис. 4. Один из кадров датасета. Малым прямоугольником выделен реальный коптер, в большом для наглядности приведено его увеличенное изображение

Fig. 4. One of the frames of the dataset. The small rectangle highlights the actual drone, while the larger rectangle shows its magnified image for clarity

Далее приведены результаты тестирования алгоритмов на части датасета, созданного из видео, на котором полет малоразмерного коптера запечатлен на фоне плотной городской застройки (табл. 2).

Т а б л и ц а 1. Характеристики трекеров на всем тестовом датасете

Table 1. Characteristics of trackers on the entire test dataset

Параметр Parameter	Precision, %	Recall, %	F1, %
Детектор YOLOv10	95,69	95,85	95,77
Bytetrack	93,70	97,19	95,42
Старый трекер	95,99	97,82	96,90
Новый трекер	96,12	99,03	97,55

Т а б л и ц а 2. Характеристики трекеров на части датасета с летящим малоразмерным коптером на фоне плотной городской застройки

Table 2. Tracker characteristics for a part of the dataset with a small-sized copter flying against a dense urban background

Параметр Parameter	Precision, %	Recall, %	F1, %
Детектор YOLOv10	96,77	54,79	69,97
Bytetrack	93,35	58,99	72,30
Старый трекер	97,50	76,93	86,00
Новый трекер	95,26	86,41	90,62

Из приведенных в таблицах данных видно, что на всем датасете значение характеристик нашего трекера выше 96 % и они лучше, чем аналогичные характеристики Bytetrack: Precision – на 2,42 %, Recall – на 1,84 % и F1 – на 2,13 %. На видео с малоразмерным коптером, летящим на фоне плотной городской застройки, Precision нашего алгоритма больше, чем значение этой характеристики трекера Bytetrack, на 1,91 %, Recall – на 27,42 % и F1 – на 18,32 %.

Заключение. В работе представлен новый алгоритм сопровождения объекта, наблюдаемого на кадрах видео. Построенный алгоритм позволяет автоматически обнаружить на видео и захватить объект одного из наперед заданных типов. В представленном алгоритме для обнаружения и захвата использовался нейросетевой детектор объектов YOLOv10, хотя допустимо использование и других детекторов. Дальнейшее сопровождение объекта выполняется предложенным алгоритмом сопровождения, использующим нейронную сеть-трансформер.

Тестирование построенного алгоритма показало его преимущество перед разработанным нами ранее трекером [4]. Характеристики предложенного алгоритма Precision, Recall и F1 на тестовом датасете превысили 96 %.

В дальнейшем предполагается усовершенствовать алгоритм для реализации сопровождения нескольких объектов.

Список использованных источников

- Object Tracking. – 2024. – URL: <https://paperswithcode.com/task/object-tracking> (date of access: 05.01.2024).
- Proceedings of Conference on Computer Vision and Pattern Recognition 2025. CVPR 2025. June 11th–15th, 2025. – URL: <https://cvpr.thecvf.com/Conferences/2025> (date of access: 03.02.2026).
- Meibodi, F. A. A Deep Dive into Generic Object Tracking: A Survey / F. A. Meibodi, Sh. Alijani, H. Najjaran // arXiv preprint. – 2025, Jul. 31. – URL: <https://arxiv.org/abs/2507.23251>; <https://doi.org/10.48550/arXiv.2507.23251>
- Залесский, Б. А. Алгоритм сопровождения объекта, наблюдаемого видеокамерой / Б. А. Залесский, В. А. Иванюкович // Доклады Национальной академии наук Беларуси. – 2024. – Т. 68, № 2. – С. 105–111. <https://doi.org/10.29235/1561-8323-2024-68-2-105-111>
- ByteTrack: Multi-Object Tracking by Associating Every Detection Box / Yi Zhang, P. Sun, Yi Jiang [et al.] // Computer Vision – ECCV 2022. ECCV 2022 / eds S. Avidan, G. Brostow, M. Cissé [et al.]. – Springer, Cham., 2022. – P. 1–21. – (Lecture Notes in Computer Science; vol. 13682). https://doi.org/10.1007/978-3-031-20047-2_1

References

1. *Object Tracking* (2024). Available at: <https://paperswithcode.com/task/object-tracking> (accessed 05 January 2024).
2. *Proceedings of Conference on Computer Vision and Pattern Recognition 2025*. CVPR 2025. June 11th–15th, 2025. Available at: <https://cvpr.thecvf.com/Conferences/2025> (accessed 03 February 2026).
3. Meibodi F. A., Alijani Sh., Najjaran H. A Deep Dive into Generic Object Tracking: A Survey. *arXiv preprint*, 2025, Jul. 31. Available at: <https://arxiv.org/abs/2507.23251>; <https://doi.org/10.48550/arXiv.2507.23251>
4. Zalesky B. A., Ivanyukovich V. A. Algorithm for tracking an object observed by a video camera. *Doklady Natsional'noi akademii nauk Belarusi = Doklady of the National Academy of Sciences of Belarus*, 2024, vol. 68, no. 2, pp. 105–111 (in Russian). <https://doi.org/10.29235/1561-8323-2024-68-2-105-111>
5. Zhang Y., Sun P., Jiang Y., Yu D., Weng F., Yuan Z., Luo P., Liu W., Wang X. ByteTrack: Multi-Object Tracking by Associating Every Detection. Avidan S., Brostow G., Cissé M., Farinella G.M., Hassner T. (eds). *Computer Vision – ECCV 2022. ECCV 2022. Lecture Notes in Computer Science, vol. 13682*. Springer, Cham., 2022, pp. 1–21. https://doi.org/10.1007/978-3-031-20047-2_1

Информация об авторах

Залесский Борис Андреевич – д-р физ.-мат. наук, заведующий лабораторией. E-mail: zalesky@newman.bas-net.by.

Иванюкович Владимир Александрович – мл. науч. сотрудник. E-mail: vladimir.ivanukovich@gmail.com.

Information about the authors

Zalesky Boris A. – D. Sc. (Physics and Mathematics), Head of the Laboratory. E-mail: zalesky@newman.bas-net.by.

Ivanyukovich Vladimir A. – Junior Researcher. E-mail: vladimir.ivanukovich@gmail.com.