

УДК 330.4:004.93'1:629.73-519

5.2.2. Математические, статистические и инструментальные методы в экономике

**СРАВНИТЕЛЬНЫЙ АНАЛИЗ
ЭКОНОМИЧЕСКОЙ ЭФФЕКТИВНОСТИ
АРХИТЕКТУР СЕМЕЙСТВА YOLO
ИСПОЛЪЗУЕМЫХ ПРИ ОБНАРУЖЕНИИ
ЧЕЛОВЕКА НА АЭРОФОТОСНИМКАХ С
БПЛА**

Вострокнутов Александр Евгеньевич
к.э.н., доцент кафедры системного анализа и
обработки информации
РИНЦ SPIN-код: 2237-4408
ORCID 0000-0002-8948-427X
Scopus Author ID: 57218528421
f_dop@mail.ru
*ФГБОУ ВО «Кубанский государственный
аграрный университет имени И. Т. Трубилина»,
Краснодар, Россия*

Самойлик София Михайловна
Магистрант
РИНЦ SPIN-код: 1644-2810
*Кубанский государственный аграрный
университет имени И.Т. Трубилина, Краснодар,
Россия*

Статья посвящена сравнительному анализу экономической эффективности архитектур семейства YOLO (версий v3, v4, v5, v8, v11 и новейшей v26) при решении задачи автоматического обнаружения человека на аэрофотоснимках, получаемых с беспилотных летательных аппаратов в ходе поисково-спасательных операций. Актуальность работы обусловлена высокой стоимостью летного часа БПЛА и прямым влиянием алгоритмов компьютерного зрения на время поиска. В статье разработана авторская квалиметрическая модель оценки версий YOLO по четырем критериям с весовыми коэффициентами (точность обнаружения сверхмелких объектов, скорость обработки в реальном времени, энергоэффективность и аппаратные требования, устойчивость к сложным ракурсам и перекрытиям). На основе интегрального бала каждой версии рассчитано требуемое время полета и стоимость операции. Результаты показывают, что YOLOv26, набравшая максимальный балл при оценке, позволяет выполнить операцию за 1 час, тогда как YOLOv5 требует 1,9 часа. Экономия при использовании версии YOLOv26 составляет 47% на одну летную смену, что делает YOLOv26 наиболее рациональным выбором с экономической точки зрения

УДК 330.4:004.93'1:629.73-519

5.2.2. Mathematical, statistical and instrumental methods in economics

**COMPARATIVE ANALYSIS OF THE COST-
EFFECTIVENESS OF THE YOLO FAMILY OF
ARCHITECTURES USED FOR HUMAN
DETECTION IN AERIAL PHOTOGRAPHS
FROM UAVS**

Vostroknutov Alexander Evgenievich
Candidate of Economic Sciences, Associate
Professor of the Department of Systems Analysis and
Information Processing
RSCI SPIN-code: 2237-4408
ORCID 0000-0002-8948-427X
Scopus Author ID: 57218528421
f_dop@mail.ru
*Kuban State Agrarian University named after I. T.
Trubilin, Krasnodar, Russia*

Samoylik Sofia Mikhailovna
Master's student
RSCI SPIN-code: 1644-2810
*Kuban State Agrarian University named after I.T.
Trubilin, Krasnodar, Russia*

The article is devoted to a comparative analysis of the cost-effectiveness of the YOLO family of architectures (versions v3, v4, v5, v8, v11 and the latest v26) in solving the problem of automatic human detection on aerial photographs obtained from unmanned aerial vehicles during search and rescue operations. The relevance of the work is due to the high cost of the UAV flight hour and the direct influence of computer vision algorithms on the search time. The article develops the author's qualimetric model for evaluating YOLO versions according to four criteria with weighting coefficients (accuracy of detecting ultra-small objects, real-time processing speed, energy efficiency and hardware requirements, resistance to complex angles and overlaps). Based on the integral score of each version, the required flight time and cost of the operation are calculated. The results show that YOLOv26, which scored the highest score in the evaluation, allows the operation to be performed in 1 hour, while YOLOv5 requires 1.9 hours. Savings when using the YOLOv26 version amount to 47% per flight shift, which makes the YOLOv26 the most rational choice from an economic point of view

Ключевые слова: ЭКОНОМИЧЕСКАЯ ЭФФЕКТИВНОСТЬ, БПЛА, ОБНАРУЖЕНИЕ ЧЕЛОВЕКА, СПАСАТЕЛЬНАЯ ОПЕРАЦИЯ, АЭРОФОТОСНИМКИ, КОМПЬЮТЕРНОЕ ЗРЕНИЕ, СРАВНИТЕЛЬНЫЙ АНАЛИЗ

Keywords: ECONOMIC EFFICIENCY, UAV, HUMAN DETECTION, RESCUE OPERATION, AERIAL PHOTOGRAPHS, COMPUTER VISION, COMPARATIVE ANALYSIS

<http://dx.doi.org/10.21515/1990-4665-220-005>

ВВЕДЕНИЕ

Активное внедрение беспилотных летательных аппаратов открыло новые возможности в сфере поисково-спасательных операций. Соответственно научные и прикладные акценты в этой сфере сменились на быстрое и точное автоматическое обнаружение людей на видео и фото, которые получают с бортовых камер дрона. Это обуславливается экономической эффективностью, лежащей в основе процесса поисково-спасательной операции. Один час использования беспилотного летательного аппарата может обходиться от 5,8 до 7,1 тыс. руб. Следовательно, алгоритмы обнаружения и их точность прямо пропорционально влияют на экономическую эффективность данного процесса. Применение алгоритмов компьютерного зрения отличается от классического видеонаблюдения, к которому мы так привыкли и имеет ряд специфик:

1. Изменчивость объектов на снимке. Съемка ведется либо строго сверху или под острым углом, что сильно искажает контуры человеческого тела [16].
2. Визуальный шум. Плотная застройка, лес, тени от каждого объекта, вода и размытие изображения при движении создают много визуального шума.
3. Сверхмалый масштаб объектов. Большая высота полетов (от 50 метров и выше) обуславливает сверхмалую площадь изображения человека в кадре (около 15x15 пикселей) [12].

4. Аппаратные ограничения. Для автономного реагирования в режиме реального времени алгоритм должен выполняться бесперебойно на бортовом компьютере дрона, у которого есть лимит по грузоподъемности, энергопотреблению, мощности бортового компьютера и времени автономной работы [13].

Учитывая данную специфику, выбор оптимального метода для задач обнаружения человека на аэрофотоснимках сводится к балансу высокой чувствительности к сверхмелким объектам (метрика mAP) и скоростью обработки (FPS)[14].

Основной целью исследования является провести сравнительный анализ экономической эффективности архитектур семейства YOLO (версии v3, v4, v5, v8, v11, v26), используемых для обнаружения человека на аэрофотоснимках с БПЛА в ходе поисково-спасательных операций и определить архитектуру, обеспечивающую минимальные затраты при заданном уровне надежности обнаружения.

Задачи исследования:

- выполнить ретроспективный обзор эволюции методов обнаружения объектов на аэрофотоснимках, включая классические алгоритмы и современные одностадийные детекторы, и визуальные трансформеры;
- разработать квалиметрическую матрицу критериев оценки версий YOLO с учетом специфики поисково-спасательных операций с БПЛА;
- присвоить весовые коэффициенты критериям на основе их значимости для экономической эффективности операции;
- провести количественную оценку времени полета и стоимости спасательной операции для каждой версии;
- сравнить полученные затраты и определить архитектуру с наилучшими экономическими показателями.

МАТЕРИАЛЫ И МЕТОДЫ

В последние 20 лет методы обнаружения объектов пережили стремительную эволюцию. До массового внедрения нейросетей (до 2012 года) применялись алгоритмы, которые основывались на ручном извлечении признаков, то есть дата-инженеры писали математические правила вручную, чтобы описать, как выглядит человек. Самыми популярными методами, например, является HOG (Гистограммы направленных градиентов) в связке с классификатором SVM (метод опорных векторов)[1]. Суть алгоритма состоит в том, что он разбивает изображение на ячейки и вычисляет направления резких перепадов яркости, в результате чего получается абстрактный силуэт контуров объекта, затем по картинке скользит «окно поиска». Если градиенты внутри окна совпали с заложенным вручную математическим шаблоном человека, то классификатор SVM подает сигнал об обнаружении. Сегодня эти методы считаются устаревшими, хотя они и требуют небольшое количество ресурсов, они не справляются со сложными фонами и изменениями ракурса, выдавая большое количество ошибок.

Позже появились двухстадийные детекторы, которые совершили революцию в точности. Этот метод имитирует работу человеческого глаза: сначала боковым зрением фиксируется что-то интересное, а потом осуществляется фокусировка взгляда, чтобы рассмотреть. Метод делится на две стадии:

1. Стадия RPN (Region Proposal Network). Подсеть сканирует карту признаков изображения и генерирует «якори» (прямоугольные рамки), как бы отмечая объекты, которые могут быть похожи на человека.

2. Стадия RoI Pooling и Классификация. «Якори» приводятся к единому размеру и прогоняются через нейросеть-классификатор, которая определяет, что именно находится в кадре и с точностью до пикселя выравнивает границы рамки.

Этот метод обладает высокой точностью на микрообъектах, даже если изображение 10x10 пикселей, но крайне медлителен, для обработки тысячи гипотез необходимы мощные видеокарты. Как говорилось ранее, дроны имеют свое аппаратное ограничение вычислительной мощности, например, микрокомпьютер Jetson Nano, который может обрабатывать только 1-4 кадра в секунду [13]. С такой задержкой дрон не сможет лететь на автопилоте и облетать препятствия в воздухе.

Еще позже появились ранние одностадийные детекторы (SSD) [3]. Этот метод предполагает сокращение описанного ранее метода двухстадийных детекторов до одной стадии. Нейросеть должна была искать координаты и определять класс объекта за один подход. Изображение проходит сквозь нейросети, которые сжимают его и на разных этапах сжатия SSD делает предсказания. На изображении с высоким разрешением осуществляется поиск мелких объектов, а на изображении с низким разрешением крупных объектов на весь кадр. Этот метод оказался неподходящим для большой высоты полета. На больших изображениях нейросети видят мелкие детали (линии, углы), не понимая человек это или нет, а на глубоких слоях, где накапливается контекст, маленькое изображение человека стирается из-за сильного сжатия изображения. Таким образом, данные о наличии на изображении человека теряют свою правдивость и точность [3].

С 2020 года популярность набирают визуальные трансформеры Vision Transformers (DETR [4], RT-DETR [5]). На такой архитектуре работают современные текстовые ИИ, например, ChatGPT. Вместо чтения изображения по пикселям, трансформер разрезает изображения на «патчи», затем включается механизм Self-Attention (Самовнимание) [4]. Алгоритм автоматически вычисляет связь каждого «патча» со всеми остальными одновременно, вне зависимости от их удаленности друг от друга. Трансформер хорошо справляется с перекрытиями, например, если

человек стоит под крышей и видна только его часть (часть плеча, ноги), он свяжет воедино крышу, тень, плечо и ноги, распознав человека. Но механизм имеет квадратичную вычислительную сложность [5]. Чем выше разрешение снимка (у дронов камеры снимают в 4К разрешении), тем быстрее заполняется оперативная память, поэтому технология затрудняет массовое применение для легких дронов.

Носителями отраслевого стандарта и вершиной (на данный момент) эволюции одностадийных детекторов является семейство YOLO (You Only Look Once) [14]. Самые современные его версии v8, v9, v10, v11 совмещают в себе скорость SSD, но не пропускают даже мелкие детали. YOLO делит изображение на сетку и та ячейка этой сетки, которая попадает в центр искомого объекта становится ответственной за его обнаружение и сразу выдает координаты рамки. Для того, чтобы видеть мелкие объекты используется пирамида признаков, внутри сети информация движется в двух направлениях. Сначала алгоритм берет информацию о контексте из глубоких слоев, затем перемещает ее обратно в слой, где изображение в высоком разрешении [6, 7]. Получается, что алгоритм видит мелкого человека и одновременно определяет, что это точно он. Ранее рамка, в которую должен был помещаться объект была четко определена, но человек редко статичен, он может бежать, размахивать руками, ногами, лежать на траве и уже не будет подходить под заданные параметры человека. Семейство YOLO же предсказывает центр объекта и расстояние до его краев во все стороны без четких рамок и параметров объекта, что позволяет существенно повысить точность обнаружения [8, 16]. Также существенным плюсом является оптимизация веса моделей (3-6 Мб), они легко помещаются в кэш мобильных процессоров дрона [14].

Таким образом, наиболее перспективной группой методов обнаружения людей на видео и фото, которые получают с бортовых камер дрона является семейство YOLO, включающее шесть версий.

РЕЗУЛЬТАТЫ РАБОТЫ И ИХ ОБСУЖДЕНИЕ

Для проведения дальнейшего сравнительного анализа необходимо ввести критерии оценки методов обнаружения человека на аэрофотоснимках с БПЛА. Так как в качестве базового метода было выбрано YOLO, будем сравнивать его самые значимые версии между собой. Это версии YOLOv3[6], YOLOv4[7], YOLOv5, YOLOv8 [8], YOLOv11[15], YOLOv26[10].

Для выбора рациональной версии метода было предложено использовать такие критерии, как архитектура извлечения микропризнаков, скорость и отказ от постобработки, энергоэффективность и аппаратные требования, устойчивость к сложным ракурсам и перекрытиям. Именно эти критерии, на наш взгляд, оказывают существенное влияние на эффективность процесса поисково-спасательных операций с применением БПЛА.

Критерий «Архитектура извлечения микропризнаков» означает способность нейросети распознать объект размером меньше, чем 32x32 пикселей без потери их признаков после сжатия изображения [12].

Критерий «Скорость и отказ от постобработки» отображает время задержки, которое необходимо для математического анализа одного кадра (1 кадр/с) [9].

Критерий «Энергоэффективность и аппаратные требования» обозначает объем потребляемой оперативной памяти (VRAM), вычислительную сложность (FLOPs), и потребление энергии [13].

Критерий «Устойчивость к сложным ракурсам и перекрытиям» означает способность алгоритма находить человека при съемке сверху, под острым углом и частично перекрытого человека [16].

Для каждого критерия был введен коэффициент важности, отражающий степень влияния каждого критерия на эффективность процесса. По степени важности первым идет критерий «Точность обнаружения сверхмелких объектов», так как рабочая высота полета у дрона от 50 метров над землей. Если алгоритм не видит объекты с такой высоты, то тогда разработка всей системы обнаружения не имеет смысла. Следующим по важности является критерий «Скорость обработки в реальном времени». Так как дрон летит со скоростью от 30 км/ч и для того, чтобы при обнаружении человека система автопилота могла взять его на сопровождение или затормозить дрон, задержка должна быть минимальной. Третьим по важности критерием является «Энергоэффективность и аппаратные требования». Каждый дрон имеет аппаратные ограничения и алгоритм должен корректно работать в условиях ограниченной мощности микрокомпьютера. Четвертый по важности критерий – «Устойчивость к сложным ракурсам и перекрытиям». Алгоритм должен быть способен распознавать человека с разных ракурсов.

В таблице 1 приведены критерии оценки алгоритмов YOLO, а в таблице 2 приведены характеристики ключевых версий YOLO, которые помогут оценить алгоритмы.

Таблица 1 – Предложенные критерии и шкалы оценки алгоритмов YOLO

№ п/п	Наименование критерия	Коэффициент важности	Шкала оценки
1	Точность обнаружения сверхмелких объектов	0,35	Отлично – 2 балл Частично – 1 балл Плохо – 0 баллов
2	Скорость обработки в реальном времени	0,3	Отлично – 2 балл Частично – 1 балл Плохо – 0 баллов
3	Энергоэффективность и аппаратные требования	0,2	Отлично – 2 балл Частично – 1 балл Плохо – 0 баллов
4	Устойчивость к сложным ракурсам и перекрытиям	0,15	Отлично – 2 балл Частично – 1 балл Плохо – 0 баллов
Итого		1	

Как видно из таблицы 2, семейство YOLO развивалось постепенно. Сначала оно претерпело инфраструктурные изменения: ушли от тяжелого C-фреймворка Darknet [6, 7], который существенно усложнял развертывание на бортовых компьютерах дронов, к PyTorch и меньшему размеру (Nano и Small), что позволило оптимизировать аппаратную нагрузку. Далее произошли архитектурные изменения. Отказались от «якорей» [8], которые подгоняли форму человека под конкретные

параметры, и внедрили блоки пространственного внимания, позволившие искать микроскопические изображения людей в сложной среде и со сложных ракурсов. Завершили эволюцию аппаратные изменения, которые значительно ускорили работу, свели задержки к нулю и сделали нейронную сеть еще более внимательной к деталям [9, 10].

Таблица 2 – Характеристики ключевых версий YOLO

№ п/п	Наименование версии	Достоинства	Недостатки
1	YOLOv3 (2018 г.)	Пирамида признаков для поиска объектов разных размеров (FPN). Стабильная детекция в реальном времени	Маленькие изображения людей теряются в глубоких слоях нейросети. Жесткие «якори», которые пропускают человека при съемке сверху. Сложна для внедрения. С-фреймворк Darknet тяжело развернуть на компьютерах дронов
2	YOLOv4 (2020 г.)	Улучшенная резкость, лучше видит людей с высоты, математическая оптимизация	Тяжелые функции активации, которые перегревают процессор. Сохранение С-фреймворк Darknet. Жесткие «якори», которые пропускают человека при съемке сверху
3	YOLOv5 (2020 г.)	Переход на PyTorch, возможность конвертации в (TensorRT, TFLite). Версии Nano и Small (до 10 Мб, высокий FPS), экономия батареи	Жесткие «якори», которые пропускают человека при съемке сверху. Для качественного обнаружения с большой высоты требует дополнительного программного решения
4	YOLOv8	Отказ от «якорей»,	Сложный модуль DFL

	(2023 г.)	поиск объекта основан на математическом центре. Единая экосистема для обучения на пользовательских датасетах	(Distribution Focal Loss) перегревает дешевые платы БПЛА, вызывает задержки
5	YOLOv11 (2024 г.)	Механизм пространственного самовнимания (C2PSA) помогает понять контекст (отличает мелкого человека от других объектов). Модель на 22% легче (меньше параметров)	Сложный модуль DFL (Distribution Focal Loss) перегревает дешевые платы БПЛА, вызывает задержки. Требуется много оперативной памяти
6	YOLOv26 (2026 г.)	Сеть выдает ответ с первого прохода. Удаление DFL, скорость работы повысилась на 43%. Внедрение функций потерь (ProgLoss, STAL), которые отслеживают пропуск сверхмелких объектов	Смена логики требует переписывания интеграции в автопилоты. Функция STAL предъявляет серьезные требования к качеству и попиксельной точности ручной разметки обучающих датасетов

В таблице 3 представлены результаты оценки ключевых версий YOLO. По итогам оценки лидирующую позицию заняла YOLOv26 (2026 г.) [10], она представляет собой наиболее подходящую архитектуру, которую наиболее рационально использовать для задач детекции людей с помощью БПЛА.

Таблица 3 – Результаты оценки ключевых версий YOLO

Версия архитектуры	Точность обнаружения сверхмелких объектов	Скорость обработки в реальном времени	Энергоэффективность и аппаратные требования	Устойчивость к сложным ракурсам и перекрытиям	Взвешенный балл
YOLOv3 (2018 г.)	0	0	0	0	0
YOLOv4 (2020 г.)	1	1	0	0	0,65
YOLOv5 (2020 г.)	1	1	2	0	1,05
YOLOv8 (2023 г.)	1	1	1	2	1,15
YOLOv11 (2024 г.)	2	1	1	2	1,50
YOLOv26 (2026 г.)	2	2	2	2	2

Для получения максимальной пользы от YOLOv26 при анализе 4К изображений в сочетании с данной архитектурой используются технологии SANI (Slicing Aided Hyper Inference)[11], которые разрезают изображения на перекрывающиеся квадраты и прогоняют каждый из них через модель YOLOv26, а потом соединяют эти квадраты воедино, что значительно повышает точность результата без необходимости значительно сжимать снимок до стандартных размеров. Еще сочетают YOLOv26 с тепловизорами (Thermal TIR), так как обычные камеры могут плохо работать в условиях тумана и темного времени суток, необходимо просто до обучить модель на инфракрасных датасетах [14].

Проведенный сравнительный анализ использования методов семейства YOLO при распознавании человека в видеопотоке и фотоснимках, получаемых с БПЛА, в процессе поисково-спасательной операции позволяет получить данные о сравнительной экономической эффективности применения рассматриваемых версий архитектур. В среднем 1 час эксплуатации дрона составляет 6500 руб. Если сделать

допущение, что необходимое время поиска заданной площади обратно пропорционально интегральному баллу по результатам оценки систем (таблица 3). Тогда при идеальном алгоритме (балл будет равен 2,0) использование дрона в спасательной операции будет занимать один час полета. Следовательно, время использования дрона можно определить, как:

$$T = T_0 \frac{2,00}{Q}, \quad (1)$$

где T – время полета дрона в поисково-спасательной операции, час;

T_0 – время полета дрона при использовании идеального алгоритма, час;

Q – интегральная оценка версии алгоритма, балл.

Полученные результаты сравнительной эффективности версий архитектур семейства YOLO приведены в таблице 4.

Таблица 4 – Результаты сравнительной эффективности версий архитектур семейства YOLO

Версия архитектуры	Требуемое время полета, час	Затраты на использование дрона, руб.	Экономия, %
YOLOv3	Операция невозможна	-	-
YOLOv4	3,08	20020,00	-47
YOLOv5	1,90	12350,00	База
YOLOv8	1,74	11310,00	Экономия ~8
YOLOv11	1,33	8645,00	Экономия ~30
YOLOv26	1,00	6500,00	Экономия ~47

Использование YOLOv26 вместо YOLOv5, как самой распространенной из «легких» версий снижает стоимость спасательной операции примерно на 47% на одну летную смену. Таким образом, выбор более современной архитектуры детекции является не только техническим улучшением, но и прямым экономическим фактором, позволяющим либо сэкономить бюджет, либо при тех же затратах многократно увеличить площадь поиска или количества вылетов. Для поисково-спасательных работ, где каждая минута на счету, использование YOLOv26 является наиболее рациональным с экономической точки зрения.

ЗАКЛЮЧЕНИЕ

Проведенный сравнительный анализ показал, что архитектура YOLOv26 превосходит предыдущие версии по всем четырем рассматриваемым критериям и обеспечивает максимальную точность обнаружения сверхмелких объектов, нулевую программную задержку за счет отказа от постобработки NMS и удаления модуля DFL, высокую энергоэффективность и устойчивость к сложным ракурсам. Интегральный балл архитектуры YOLOv26 составил 2,0 против 1,05 у YOLOv5. При стоимости часа полета 6500 руб. и допущении обратной пропорциональности времени интегральному баллу, YOLOv26 позволяет выполнить операцию за 1 час с затратами 6500 руб., тогда как YOLOv5 требует 1,9 часа, что составляет более 12 тыс. руб. Таким образом, замена YOLOv5 на YOLOv26 снижает стоимость одной спасательной операции на 47% без потери качества обнаружения. Следовательно, выбор более современной архитектуры детекции является не только технически, но и прямым экономическим фактором, позволяющим либо сэкономить бюджет, либо при тех же затратах многократно увеличить площадь поиска.

СПИСОК ЛИТЕРАТУРЫ

1. Dalal N., Triggs B. Histograms of oriented gradients for human detection // 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05). – IEEE, 2005. – Vol. 1. – P. 886-893.
2. Ren S., He K., Girshick R., Sun J. Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks // IEEE Transactions on Pattern Analysis and Machine Intelligence. – 2017. – Vol. 39. – № 6. – P. 1137-1149.
3. Liu W. et al. SSD: Single Shot MultiBox Detector // European Conference on Computer Vision (ECCV). – Springer, Cham, 2016. – P. 21-37.
4. Carion N. et al. End-to-End Object Detection with Transformers // European Conference on Computer Vision (ECCV). – Springer, Cham, 2020. – P. 213-229.
5. Lv W. et al. DETRs Beat YOLOs on Real-time Object Detection // Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). – 2024. – P. 5121-5130.
6. Redmon J., Farhadi A. YOLOv3: An Incremental Improvement // arXiv preprint arXiv:1804.02767. – 2018.
7. Bochkovskiy A., Wang C. Y., Liao H. Y. M. YOLOv4: Optimal Speed and Accuracy of Object Detection // arXiv preprint arXiv:2004.10934. – 2020
8. Jocher G. et al. Ultralytics YOLO (v5, v8, v11) [Электронный ресурс] // GitHub Repository. – 2020-2024. – URL: <https://github.com/ultralytics/ultralytics> (дата обращения: 22.04.2026).
9. Wang A. et al. YOLOv10: Real-Time End-to-End Object Detection // arXiv preprint arXiv:2405.14458. – 2024.
10. YOLOv26 Object Detection Format [Электронный ресурс] // Labelformat Docs. – 2025. – URL: <https://labelformat.com/formats/object-detection/yolov26/> (дата обращения: 22.04.2026). (Подтверждение характеристик новейшего Edge-флагмана YOLOv26: End-to-End NMS-Free архитектуры и хирургического удаления слоя DFL).
11. Akyon F. C., Altinuc S. O., Temizel A. Slicing aided hyper inference and fine-tuning for small object detection (SAHI) // 2022 IEEE International Conference on Image Processing (ICIP). – IEEE, 2022. – P. 966-970. (Обоснование программного метода нарезки 4K-изображений для поиска сверхмелких целей).
12. Zhu P. et al. Detection and Tracking Meet Drones Challenge // IEEE Transactions on Pattern Analysis and Machine Intelligence. – 2021. – Vol. 44. – № 11. – P. 7380-7399. (Масштабный анализ датасета VisDrone и выявление проблем искажения при съемке строго сверху / nadir view).
13. Wu X. et al. Deep learning for unmanned aerial vehicle-based object detection and tracking: A survey // IEEE Geoscience and Remote Sensing Magazine. – 2021. – Vol. 10. – № 1. – P. 91-124. (Анализ проблемы энергоэффективности БПЛА, концепции Edge Computing и аппаратных ограничений).
14. Сацюк А.В., Белый Р.В., Ищенко А.Е. Оценка эффективности алгоритмов YOLO для обнаружения объектов в реальном времени во встраиваемых системах беспилотных транспортных средств // Сборник научных трудов Донецкого института железнодорожного транспорта. – 2024. – № 4 (75). – С. 73-82.
15. Бу С., Абламейко С.В. RLD-YOLO: новый метод обнаружения объектов на изображениях беспилотных летательных аппаратов с использованием модели нейронной сети YOLOv11 // Журнал Белорусского государственного университета. Математика. Информатика. – 2024. – № 3. – С. 14-26.

16. Смирнов П. И., Иванов А. С. Методы компенсации геометрических искажений при автоматическом распознавании малых объектов на аэрофотоснимках с БПЛА // Информационные технологии и вычислительные системы. – 2023. – № 2. – С. 78-86.

References

1. Dalal N., Triggs B. Histograms of oriented gradients for human detection // 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05). – IEEE, 2005. – Vol. 1. – P. 886-893.
2. Ren S., He K., Girshick R., Sun J. Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks // IEEE Transactions on Pattern Analysis and Machine Intelligence. – 2017. – Vol. 39. – № 6. – P. 1137-1149. (Pervoistochnik dvuhstadijnyh setej, analiz seti generacii gipotez RPN).
3. Liu W. et al. SSD: Single Shot MultiBox Detector // European Conference on Computer Vision (ECCV). – Springer, Cham, 2016. – P. 21-37. (Obosnovanie rannih odnostadijnyh setej i problemy «semanticheskogo razryva» pri szhatii).
4. Carion N. et al. End-to-End Object Detection with Transformers // European Conference on Computer Vision (ECCV). – Springer, Cham, 2020. – P. 213-229. (Fundamental'naja rabota o vnedrenii vizual'nyh transformerov DETR i mehanizma samovnimanija).
5. Lv W. et al. DETRs Beat YOLOs on Real-time Object Detection // Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). – 2024. – P. 5121-5130. (Issledovanie oblegchennoj arhitektury RT-DETR i analiz ejo resursoemkosti).
6. Redmon J., Farhadi A. YOLOv3: An Incremental Improvement // arXiv preprint arXiv:1804.02767. – 2018. (Pervoistochnik primenenija bazovoj piramidy priznakov FPN).
7. Bochkovskiy A., Wang C. Y., Liao H. Y. M. YOLOv4: Optimal Speed and Accuracy of Object Detection // arXiv preprint arXiv:2004.10934. – 2020. (Vnedrenie PANet i metodov peredachi kontekstnyh priznakov «snizu-vverh»).
8. Jocher G. et al. Ultralytics YOLO (v5, v8, v11) [Elektronnyj resurs] // GitHub Repository. – 2020-2024. – URL: <https://github.com/ultralytics/ultralytics> (data obrashhenija: 22.04.2026). (Oficial'naja dokumentacija: perehod na Anchor-free v v8 i vnedrenie blokov prostranstvennogo vnimanija C2PSA v v11).
9. Wang A. et al. YOLOv10: Real-Time End-to-End Object Detection // arXiv preprint arXiv:2405.14458. – 2024. (Kljuchevaja nauchnaja rabota issledovatelej Universiteta Cinhua, matematicheski obosnovavshaja polnyj otkaz ot algoritma postobrabotki NMS).
10. YOLOv26 Object Detection Format [Elektronnyj resurs] // Labelformat Docs. – 2025. – URL: <https://labelformat.com/formats/object-detection/yolov26/> (data obrashhenija: 22.04.2026). (Podtverzhdenie harakteristik novejshego Edge-flagmana YOLOv26: End-to-End NMS-Free arhitektury i hirurgicheskogo udalenija sloja DFL).
11. Akyon F. C., Altinuc S. O., Temizel A. Slicing aided hyper inference and fine-tuning for small object detection (SAHI) // 2022 IEEE International Conference on Image Processing (ICIP). – IEEE, 2022. – P. 966-970. (Obosnovanie programmnogo metoda narezki 4K-izobrazhenij dlja poiska sverhmelkih celej).
12. Zhu P. et al. Detection and Tracking Meet Drones Challenge // IEEE Transactions on Pattern Analysis and Machine Intelligence. – 2021. – Vol. 44. – № 11. – P. 7380-7399. (Masshtabnyj analiz dataseta VisDrone i vyjavlenie problem iskazhenija pri s#emke strogo sverhu / nadir view).
13. Wu X. et al. Deep learning for unmanned aerial vehicle-based object detection and tracking: A survey // IEEE Geoscience and Remote Sensing Magazine. – 2021. – Vol. 10.

– № 1. – P. 91-124. (Analiz problemy jenergojeffektivnosti BPLA, koncepcii Edge Computing i apparatnyh ogranicenij).

14. Sacjuk A.V., Belyj R.V., Ishhenko A.E. Ocenka jeffektivnosti algoritmov YOLO dlja obnaruzhenija ob#ektov v real'nom vremeni vo vstraivaemyh sistemah bespilotnyh transportnyh sredstv // Sbornik nauchnyh trudov Doneckogo instituta zheleznodorozhnogo transporta. – 2024. – № 4 (75). – S. 73-82.

15. Vu S., Ablamejko S.V. RLD-YOLO: novyj metod obnaruzhenija ob#ektov na izobrazhenijah bespilotnyh letatel'nyh apparatov s ispol'zovaniem modeli nejronnoj seti YOLOv11 // Zhurnal Belorusskogo gosudarstvennogo universiteta. Matematika. Informatika. – 2024. – № 3. – S. 14-26.

16. Smirnov P. I., Ivanov A. S. Metody kompensacii geometricheskih iskazhenij pri avtomaticheskom raspoznavanii malyh ob#ektov na ajerofotosnimkah s BPLA // Informacionnye tehnologii i vychislitel'nye sistemy. – 2023. – № 2. – S. 78-86.