

УДК 004.89

## Yolo - v4 эволюция скоростной детекции объектов в компьютерном зрении

Евсеев Владислав Дмитриевич      evseevvd@student.bmstu.ru

Белов Юрий Сергеевич              yselov@bmstu.ru

КФ МГТУ им. Н.Э. Баумана, Калуга, Россия

*Статья посвящена практическому применению. Основное внимание уделено использованию алгоритма в автономных транспортных средствах, системах видеонаблюдения и промышленной автоматизации, где критически важны быстродействие и надежность. Рассмотрены ключевые инновации YOLOv4: оптимизированная CSPDarknet53, методы аугментации данных и улучшенная функция потерь, обеспечивающие работу в условиях ограниченных ресурсов.*

**Ключевые слова:** YOLOv4, детекция объектов, компьютерное зрение, реальные приложения, Edge-вычисления

В области компьютерного зрения задача детекции объектов занимает центральное место, находя применение в беспилотном транспорте, системах видеомониторинга, медицинской визуализации и промышленной автоматизации. Среди множества алгоритмов, разработанных для решения этой задачи, семейство архитектур YOLO (You Only Look Once) заслуживает особого внимания [1] благодаря уникальному сочетанию скорости обработки и точности распознавания. В 2020 г. научное сообщество познакомилось с YOLOv4 — революционной версией алгоритма, предложившей ряд инновационных решений, которые значительно превзошли показатели предыдущих итераций.

**Архитектурные инновации YOLOv4.** Основой YOLOv4 стала усовершенствованная версия Darknet53, модернизированная за счет внедрения технологии Cross Stage Partial connections (CSP). Это решение позволило достичь двух важнейших целей: существенного сокращения вычислительной нагрузки (примерно на 20 % по сравнению с оригинальной Darknet53) [2] и оптимизации градиентного потока в процессе обучения. Особого внимания заслуживает выбор функции активации — разработчики отказались от традиционных ReLU и LeakyReLU в пользу Mish, что привело к заметному росту точности распознавания.

Система агрегации признаков в YOLOv4 представляет собой тщательно продуманную комбинацию Path Aggregation Network (PANet) и Spatial Pyramid Pooling (SPP). SPP модуль особенно важен, так как он позволяет увеличить receptive field без потери разрешения, что критически важно для эффективного обнаружения объектов различных масштабов. Сохраняя проверенную временем трехмасштабную структуру детектора (аналогичную YOLOv3), разработчики обеспечили модели возможность одинаково хорошо

работать с объектами разных размеров — от крупных транспортных средств до мелких пешеходов в городской среде.

**Оптимизация производительности и точности.** Одним из наиболее значимых улучшений в YOLOv4 стала комплексная оптимизация, затронувшая как скорость работы, так и точность распознавания. Внедрение передовых методов аугментации данных, включая Mosaic и CutMix, позволило существенно улучшить качество обучения, особенно при работе с ограниченными наборами данных. Важным шагом вперед стало использование Complete Intersection over Union (CIoU) в функции потерь, что привело к более точному позиционированию bounding box.

Вычислительная эффективность модели была значительно повышена благодаря нескольким ключевым решениям. Помимо уже упомянутого сокращения параметров за счет CSPDarknet53, разработчики реализовали поддержку GPU-оптимизированных методов, включая TensorRT [3]. Это позволило достичь рекордных показателей скорости обработки при сохранении высокой точности.

**Эволюция архитектуры: современные модификации.** После первоначального релиза YOLOv4 появилось несколько специализированных модификаций, адаптированных для различных сценариев использования:

- YOLOv4-tiny — оптимизированная версия для устройств с ограниченными вычислительными ресурсами, сохраняющая приемлемую точность при значительном увеличении скорости работы;

- Scaled-YOLOv4 — масштабируемая архитектура, позволяющая гибко настраивать модель в зависимости от доступных аппаратных ресурсов и требований к точности;

- YOLOv4-csp — расширенная версия с дополнительными CSP-блоками, демонстрирующая еще более высокие показатели точности за счет усложнения архитектуры.

**Адаптация YOLOv4 для Edge-устройств.** Несмотря на впечатляющие характеристики YOLOv4 в задачах детекции объектов, ее прямое применение на edge-устройствах сталкивается с существенными ограничениями, связанными с высокими требованиями к вычислительным ресурсам и энергопотреблению [3]. Это обуславливает необходимость комплексной оптимизации архитектуры, включающей как структурные изменения модели, так и низкоуровневые аппаратные оптимизации.

Наиболее распространенным подходом к адаптации YOLOv4 для ресурсоограниченных устройств стало создание облегченной версии YOLOv4-tiny, в которой количество детектирующих голов сокращено с трех до двух, а в качестве backbone используется упрощенный вариант CSPDarknet53. Такая модификация позволяет уменьшить количество параметров модели в 4–6 раз (до 6–12 млн) при сохранении приемлемой точности детекции. Практические испытания демонстрируют, что на платформе Raspberry Pi 4 адаптированная версия обеспечивает скорость обработки 3,7 кадров в секунду по сравнению с 0,8 кадра у оригинальной модели.

**Актуальность архитектуры YOLOv4.** Данная модель компьютерного зрения сохраняет свою востребованность благодаря многолетнему опыту успешного применения в реальных условиях. В отличие от более новых, но недостаточно проверенных аналогов, она демонстрирует исключительную надежность в промышленной эксплуатации. Ключевыми преимуществами модели являются оптимизированная архитектура CSPDarknet53, обеспечивающая эффективную работу на оборудовании с ограниченными ресурсами, а также возможность масштабирования через специализированные модификации (YOLOv4-tiny, YOLOv4-csp). Поддержка научного сообщества, выражающаяся в постоянных исследованиях, регулярных обновлениях и наличии обширной учебно-методической базы, в сочетании с перечисленными характеристиками объясняет востребованность YOLOv4 в приложениях, где критически важны стабильность работы и экономическая эффективность, а не исключительно максимальные показатели точности [4].

## Литература

- [1] Конушин А.С., Куров Е.В., Вешняков А.В. Методы детекции объектов на основе сверточных нейронных сетей. *Информационные технологии и вычислительные системы*, 2021, № 2, с. 55–67.
- [2] Матвеев А.Н., Стрижов В.В. Сравнительный анализ алгоритмов детекции объектов в реальном времени. *Искусственный интеллект и принятие решений*, 2022, № 1, с. 34–48.
- [3] Bochkovskiy A., Wang C.-Y., Liao H.-Y.M. YOLOv4: optimal speed and accuracy of object detection. Arxiv:2004.10934, 2020
- [4] Redmon J., Farhadi A. YOLOv3: An incremental improvement. Arxiv preprint arXiv:1804.02767, 2018.

## Yolo - v4 evolution of real - time object detection in computer vision

Evseev Vladislav Dmitrievich

evseevvd@student.bmstu.ru

Belov Yuri Sergeevich

ysbelov@bmstu.ru

*Kaluga Branch of BMSTU, Kaluga, Russia*

The article is devoted to practical application. The main focus is on the use of the algorithm in autonomous vehicles, video surveillance systems, and industrial automation, where speed and reliability are critical. Key innovations of YOLOv4 are considered: optimized CSPDarknet53, data augmentation methods and improved loss function, ensuring operation in conditions of limited resources.

**Keywords:** YOLOv4, object detection, computer vision, real-world applications, edge computing