

КОМПЬЮТЕРНОЕ ЗРЕНИЕ НА ОСНОВЕ НЕЙРОСЕТИ YOLO: ПРИНЦИПЫ, МЕТОДЫ И ПЕРСПЕКТИВЫ

Д.С. Давыденко¹, А.В. Акименко¹, Хассан Мохаммад Махди¹

¹ФГБОУ ВО «Воронежский государственный лесотехнический университет
имени Г.Ф. Морозова»

Аннотация. В данной статье рассматривается применение нейросетевой архитектуры YOLO (You Only Look Once) для задач компьютерного зрения. YOLO является одной из наиболее эффективных моделей для детекции объектов в реальном времени. Описываются основные принципы работы сети, методы улучшения точности и скорости, а также перспективы её развития в будущем.

Ключевые слова: компьютерное зрение, искусственный интеллект, машинное обучение

COMPUTER VISION BASED ON THE YOLO NEURAL NETWORK: PRIN- CIPLES, METHODS AND PROSPECTS

D.S. Davydenko¹, A.V. Akimenko¹, Hassan Mohammad Mehdi¹

¹Voronezh State University of Forestry and Technologies named after G.F. Morozov

Abstract. This article discusses the application of the YOLO (You Only Look Once) neural network architecture for computer vision tasks. YOLO is one of the most effective models for real-time object detection. The article describes the basic principles of the network, methods for improving accuracy and speed, as well as prospects for its future development.

Keywords: computer vision, artificial intelligence, machine learning.

Компьютерное зрение — это область искусственного интеллекта, направленная на автоматический анализ, интерпретацию и понимание визуальной информации. Оно активно развивается и находит применение во множестве отраслей, от промышленности до медицины. Основная цель компьютерного зрения заключается в том, чтобы дать машинам способность «видеть» и интерпретировать окружающий мир подобно человеку.

С появлением мощных нейросетевых моделей, таких как YOLO, компьютерное зрение вышло на новый уровень, позволяя выполнять детекцию объектов в реальном времени с высокой точностью. YOLO предлагает эффективный подход к обработке изображений, обеспечивая быстрое и точное распознавание объектов, что делает её востребованной во многих сферах.

Сегодня компьютерное зрение используется для автоматизации процессов, повышения безопасности, улучшения медицинской диагностики и создания интеллектуальных систем. В данной статье рассмотрены ключевые аспекты работы YOLO, её применение в различных отраслях, а также перспективы дальнейшего развития технологии.

Современные области применения компьютерного зрения и реальные кейсы.

1. Промышленность.

- **Контроль качества продукции:** Заводы и фабрики используют системы компьютерного зрения для обнаружения дефектов на производственных линиях. Например, Tesla применяет нейросетевые алгоритмы для автоматического контроля сборки автомобилей.

- **Роботизированные производственные линии:** Завод BMW использует компьютерное зрение для автоматического анализа собранных деталей, минимизируя риск дефектов.

2. Автономные транспортные системы.

- **Автопилот в автомобилях:** Tesla, Waymo и другие компании используют YOLO для детекции пешеходов, транспортных средств и дорожных знаков.

- **Беспилотные дроны:** Amazon Prime Air применяет компьютерное зрение для точного приземления дронов при доставке товаров.

3. Безопасность и видеонаблюдение.

- **Распознавание лиц:** Правительственные системы Китая активно применяют нейросети для мониторинга преступников в реальном времени.

- **Контроль доступа:** YOLO используется в аэропортах и на стратегических объектах для автоматизированного распознавания лиц.

4. Медицина.

- **Диагностика заболеваний:** Google DeepMind разработал систему, использующую YOLO и CNN для обнаружения рака груди на снимках маммографии.

- Распознавание аномалий в рентгеновских снимках: Системы на основе нейросетей применяются для диагностики COVID-19 по компьютерным томограммам.

5. Розничная торговля.

- Автоматизированные кассы: В Amazon Go используется компьютерное зрение для автоматического списания товаров без необходимости в кассирах.

- Анализ поведения покупателей: Walmart применяет YOLO для отслеживания передвижения клиентов в магазинах и улучшения выкладки товаров.

6. Спорт.

- Анализ движений спортсменов: Футбольные клубы используют компьютерное зрение для анализа тактики игроков и оценки их физической подготовки.

- Автоматизированные камеры: В теннисе Hawk-Eye отслеживает траектории мячей, помогая судьям принимать точные решения.

7. Сельское хозяйство.

- Мониторинг посевов: Сельскохозяйственные дроны с камерами анализируют состояние растений, обнаруживают болезни и оптимизируют процесс полива.

- Обнаружение вредителей: YOLO помогает фермерам автоматически определять вредителей на растениях, снижая потребность в химических пестицидах.

Перспективы развития компьютерного зрения связаны с улучшением архитектуры моделей, расширением сфер их применения и оптимизацией вычислительных процессов. В ближайшие годы ожидается значительный прогресс в области глубокого обучения, что позволит добиться большей точности и эффективности. Кроме того, появление квантовых вычислений может революционизировать обработку изображений, ускоряя сложные вычисления в разы. Автоматизированные системы, такие как беспилотные автомобили и медицинская диагностика, продолжат активно развиваться, а интеграция компьютерного зрения с дополненной и виртуальной реальностью откроет новые горизонты в индустрии развлечений и образования. Также возрастет роль компьютерного зрения в экологическом мониторинге, что позволит более эффективно бороться с изменением климата и контролировать состояние окружающей среды.

Особое место в общемировых тенденциях занимает развитие технологий компьютерного зрения в России. В стране наблюдается стремительный рост исследований и практических внедрений благодаря высококвалифицированным

научным кадрам, передовым исследовательским институтам и активной государственной поддержке. Российские университеты и научные центры, такие как Московский физико-технический институт, Высшая школа экономики и Санкт-Петербургский государственный университет, играют ключевую роль в разработке алгоритмов для детекции объектов, сегментации изображений и анализа видеопотоков. Государственная программа «Цифровая экономика» вместе с региональными инициативами способствует выделению финансовых средств и созданию благоприятной инфраструктуры для стартапов и инновационных компаний.

Крупные корпорации, такие как «Яндекс», «Сбер» и «ВТБ», а также молодые технологические стартапы активно инвестируют в разработку систем компьютерного зрения, применяемых в безопасности, транспортной логистике, автоматизации производства и медицинской диагностике. Внедрение этих технологий позволяет повысить качество и скорость обработки данных, снизить производственные издержки и обеспечить конкурентоспособность на международном уровне. Особенно перспективны направления, связанные с распознаванием лиц, анализом поведения в общественных местах и автоматизированным мониторингом критически важных объектов, что имеет большое значение для общественной безопасности.

В медицинской сфере отечественные разработки в области компьютерного зрения направлены на создание систем ранней диагностики, анализ изображений рентгеновских снимков и МРТ, что особенно актуально в условиях развития телемедицины. Автономные транспортные системы, включая беспилотные автомобили и дроны, также активно тестируются в различных регионах России, что свидетельствует о высоком потенциале дальнейшего развития этой технологии. В аграрном секторе применение дронов для мониторинга состояния посевов и автоматизированного обнаружения вредителей способствует повышению урожайности и снижению затрат.

Таким образом, перспективы развития компьютерного зрения в России охватывают широкий спектр направлений – от фундаментальных исследований до практических применений в различных отраслях экономики. Сочетание государственной поддержки, наличия высококвалифицированных специалистов и активной инновационной деятельности создаёт благоприятные условия для быстрого роста и внедрения отечественных технологий как на национальном, так и на глобальном уровнях.

YOLO представляет собой одношаговый метод обнаружения объектов, который преобразует задачу детекции в задачу регрессии. В отличие от двухшаговых подходов (например, R-CNN), YOLO обрабатывает изображение целиком, что значительно увеличивает скорость работы. Основные особенности YOLO:

- Разделение изображения на сетку, в каждой ячейке которой предсказывается ограничивающая рамка (bounding box), класс объекта и вероятность обнаружения.
- Использование сверточных нейросетей (CNN) для экстракции признаков.
- Совмещение локализации объектов и классификации в одном процессе, что делает сеть чрезвычайно быстрой.

Таблица 1 – Действие алгоритма

Шаг	Описание
1	Входное изображение разделяется на сетку $N \times N$.
2	Каждая ячейка отвечает за обнаружение объектов, центр которых попадает в неё.
3	YOLO предсказывает несколько ограничивающих рамок и соответствующие им вероятности принадлежности к классам.
4	Используется функция потерь, включающая ошибки классификации, локализации и предсказания ограничивающих рамок.
5	Применяется метод подавления немаксимальных значений (Non-Maximum Suppression, NMS) для удаления дублирующихся рамок.

YOLO остается одним из наиболее перспективных алгоритмов компьютерного зрения благодаря своей высокой скорости и точности. Развитие данной технологии способствует её интеграции в различные отрасли, что делает её важным инструментом в современном мире. Улучшение модели, оптимизация вычислений и новые области применения делают YOLO ключевым элементом будущего компьютерного зрения.

Список литературы

1. Евстраткин, К. С. OPENCV: варианты использования компьютерного зрения / К. С. Евстраткин, А. Р. Султанова, А. В. Ерпелев // Цифровые технологии: наука, образование, инновации : Материалы III Международного научного Форума профессорско-преподавательского состава и молодых ученых, Москва, 09 ноября 2020 года / Под редакцией А.В. Олейник, А.А. Зеленского. – Москва: Московский государственный технологический университет "СТАНКИН", 2021. – С. 28-31. – EDN KTLCNJ.

2. Rice Fungal Diseases Recognition Using Modern Computer Vision Techniques / I. V. Arinichev, S. V. Polyanskikh, G. V. Volkova, I. V. Arinicheva // International Journal of Fuzzy Logic and Intelligent Systems. – 2021. – Vol. 21, No. 1. – P. 1-11. – DOI 10.5391/IJFIS.2021.21.1.1. – EDN CPIVQC.
3. Medvedeva, I. A. Automated complex development project based on computer vision technology / I. A. Medvedeva, M. V. Vanslav, M. A. Ragozina // Natural and Technical Sciences. – 2022. – No. 7(170). – P. 177-178. – EDN HONPYU.
4. He K., Zhang X., Ren S., Sun J. Глубокое остаточное обучение для распознавания изображений. CVPR, 2016.
5. Бочкарев С., Лемпитский В. Глубокое обучение в анализе изображений: тенденции и приложения. Распознавание образов, 2021.
6. Минакова, О. В. Повышение эффективности работы в проектах open source на основе архитектурного анализа (на примере проекта Сахана) / О. В. Минакова, И. В. Поцбенева, П. Ю. Гусев // Моделирование систем и процессов. – 2024. – Т. 17, № 1. – С. 84-92. – DOI 10.12737/2219-0767-2024-17-1-84-92. – EDN NZNSKM.
7. Заревич А.И., Макаренко Ф.В., Ягодкин А.С., Зольников К.В. Моделирование поведения мобильных роботов с использованием генетических алгоритмов // Моделирование систем и процессов. – 2022. – Т. 15, № 3. – С. 7-16.
8. Сазонова, С. А. Моделирование процесса диагностики утечек на основе двухальтернативной гипотезы с учетом помех от стохастичности потребления в гидравлической системе / С. А. Сазонова, И. В. Щербакова, Г. И. Сметанкина // Моделирование систем и процессов. – 2024. – Т. 17, № 1. – С. 111-120. – DOI 10.12737/2219-0767-2024-17-1-111-120. – EDN CSKRIZ.
9. Выбор критерия оптимальности при принятии управленческих решений в сложных технических системах / А. В. Скрыпников, И. А. Высоцкая, С. А. Евдокимова [и др.] // Моделирование систем и процессов. – 2024. – Т. 17, № 1. – С. 120-128. – DOI 10.12737/2219-0767-2024-17-1-120-128. – EDN MMIAIH.

References

1. Evstratkin, K. S. OPENCV: Use Cases of Computer Vision / K. S. Evstratkin, A. R. Sultanova, A. V. Erpelev // Digital Technologies: Science, Education, Innovations: Proceedings of the III International Scientific Forum of Faculty and Young Scientists, Moscow, November 09, 2020 / Edited by A. V. Oleinik, A. A. Zelensky. – Moscow: Moscow State Technological University "STANKIN", 2021. – P. 28-31. – EDN KTLCNJ.

2. Rice Fungal Diseases Recognition Using Modern Computer Vision Techniques / I. V. Arinichev, S. V. Polyanskikh, G. V. Volkova, I. V. Arinicheva // International Journal of Fuzzy Logic and Intelligent Systems. – 2021. – Vol. 21, No. 1. – P. 1-11. – DOI 10.5391/IJFIS.2021.21.1.1. – EDN CPIVQC.
3. Medvedeva, I. A. Automated complex development project based on computer vision technology / I. A. Medvedeva, M. V. Vanslav, M. A. Ragozina // Natural and Technical Sciences. – 2022. – No. 7(170). – P. 177-178. – EDN HONPYY.
4. He K., Zhang X., Ren S., Sun J. Deep residual learning for image recognition. CVPR, 2016.
5. Bochkarev S., Lempitsky V. Deep learning in image analysis: trends and applications. Pattern Recognition, 2021.
6. Minakova, O. V. Improving Work Efficiency in Open Source Projects Based on Architectural Analysis (using the Sakhn Project as an Example) / O. V. Minakova, I. V. Potsebneva, P. Yu. Gusev // Modeling of Systems and Processes. – 2024. – Vol. 17, No. 1. – P. 84-92. – DOI 10.12737/2219-0767-2024-17-1-84-92. – EDN NZNSKM.
7. Zarevich A.I., Makarenko Ph.V., Yagodkin A.S., Zolnikov K.V. Modeling the behavior of mobile robots using genetic algorithms // Modeling systems and processes. – 2022. – Vol. 15, No. 3. – pp. 7-16.
8. Sazonova S. A., Shcherbakova I. V., Smetankina G. I. Modeling of the leakage diagnosis process based on a two-alternative hypothesis, taking into account interference from stochastic consumption in a hydraulic system // Modeling of systems and processes. – 2024. – Vol. 17, No. 1. – pp. 111-120. – DOI 10.12737/2219-0767-2024-17-1-111-120. – EDN CSKRIZ.
9. Selection of the optimality criterion in making managerial decisions in complex technical systems / A.V. Skrypnikov, I. A. Vysotskaya, S. A. Evdokimova [et al.] // Modeling of systems and processes. – 2024. – Vol. 17, No. 1. – pp. 120-128. – DOI 10.12737/2219-0767-2024-17-1-120-128. – EDN MMIAIH.