

УДК 004.932

СОВРЕМЕННЫЕ НЕЙРОСЕТЕВЫЕ МОДЕЛИ ДЕТЕКЦИИ ОБЪЕКТОВ

Ульященко А.Ю., Беспалов Т.Д., студенты гр. ПИБ-222, IV курс

Научный руководитель: Тайлакова А.А., доцент

Кузбасский государственный технический университет

имени Т.Ф. Горбачёва

г. Кемерово

Введение

Детекция объектов является одной из ключевых задач компьютерного зрения. Она направлена на обнаружение объектов на изображении или в видеопотоке и определение их пространственного положения с помощью ограничивающих рамок. В отличие от обычной классификации объектов, детекция позволяет одновременно находить сразу несколько объектов разных классов. Данные технологии широко применяются в транспортных системах, медицине, промышленности, системах видеонаблюдения и транспортных системах.

Основные нейросетевые модели детекции объектов

В настоящее время большинство методов детекции основаны на нейросетевых архитектурах. Самыми распространенными из них являются модели YOLO, SSD, RetinaNet и Faster R-CNN. Все перечисленные модели выполняют одинаковую функцию, но имеют отличительные особенности в принципе работы, точности и скорости обработки изображений и видеопотока.

Модель YOLO

YOLO является одной из самых известных одноэтапных архитектур компьютерного зрения. Название модели расшифровывается как You Only Look Once, данный слоган отражает принцип однократного прохода по изображению и видеопотоку. Главное преимущество YOLO заключается в том, что она может использоваться для анализа видеопотока в реальном времени благодаря своей высокой скорости работы. Различные версии данной модели используются в системах дорожного мониторинга, видеонаблюдении и автоматизированном контроле.

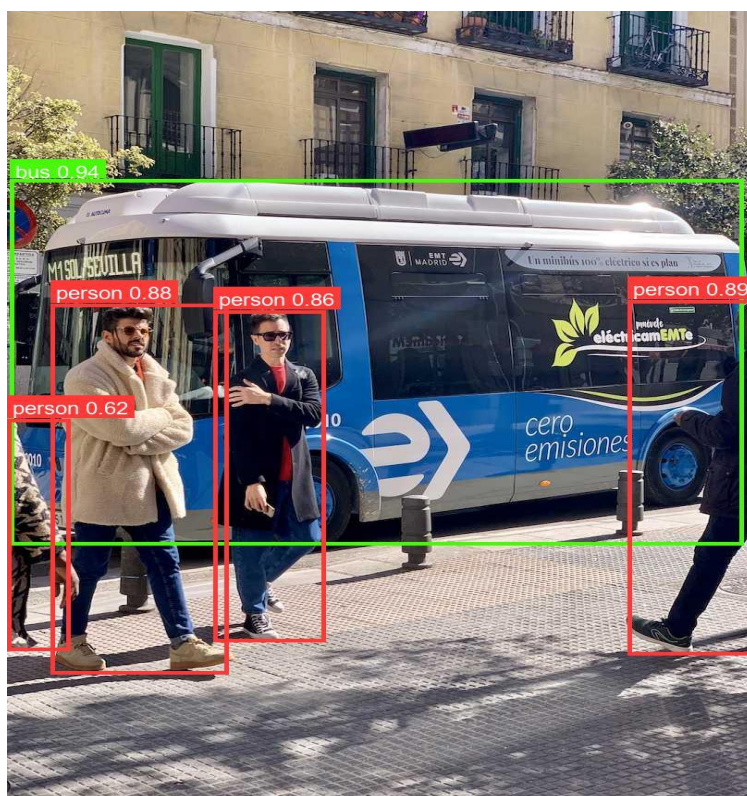


Рисунок 1. Работа модели YOLO11n

Модель SSD

SSD относится к одноэтапным моделям детекции объектов. Она осуществляет классификацию и обнаружение объектов за один проход сети по визуальным данным, что позволяет существенно сократить время обработки кадра. Особенностью SSD является использование нескольких карт признаков различного масштаба, благодаря чему модель способна находить как большие, так и маленькие объекты. Данная модель стала одной из первых удачных попыток объединить неплохую точность и высокую скорость обнаружения. Благодаря этим особенностям SSD часто применяется в системах видеонаблюдения, мобильных и встраиваемых устройствах, а также в обработке видео в реальном времени.



Рисунок 2. Работа модели SSDLite320 MobileNetV3 Large
Модель RetinaNet

RetinaNet также относится к одноэтапным моделям. Её основным отличием стало использование функции сфокусированных потерь, которая уменьшает влияние большого количества простых отрицательных примеров при обучении. Данная функция позволила улучшить качество обнаружения и повысить конкурентоспособность одноэтапных моделей по сравнению с двухэтапными. RetinaNet показала, что быстрые архитектуры могут обеспечивать не только высокую скорость, но и очень хорошую точность распознавания. Данная модель используется в задачах, где важна точность обнаружения, например в анализе аэрокосмических снимков и медицинских изображениях.



Рисунок 3. Работа модели RetinaNet ResNet50 FPN v2

Модель Faster R-CNN

Faster R-CNN относится к двухэтапным моделям детекции. На первом этапе нейронная сеть определяет область изображения, где предположительно находятся объекты, а на втором этапе выполняет их классификацию и уточнение координат для ограничительных рамок. Faster R-CNN является более точной по сравнению с одноэтапными моделями, но уступает им в скорости и требуемой вычислительной мощности. Данная модель чаще всего применяется в задачах, где в приоритете стоит не скорость выполнения, а качество распознавания.



Рисунок 4. Работа модели Faster R-CNN ResNet50 FPN v2

Сравнение моделей детекции объектов

Рассмотренные модели имеют разные особенности. YOLO и SSD направлены на высокую обработку визуальных данных, RetinaNet является компромиссом между скоростью обработки данных и точностью детекции, а Faster R-CNN ориентирована на высокую точность.

Анализ современных моделей показывает, что нет универсального решения. Выбор архитектуры зависит исключительно от поставленной задачи. Если требуется точность, то обычно рассматриваются двухэтапные модели. Если же важна обработка данных в реальном времени, то предпочтение отдается одноэтапным архитектурам.

Заключение

Таким образом, современные нейросетевые модели детекции объектов играют важную роль в развитии компьютерного зрения. Наиболее распространенными решениями остаются YOLO, SSD, RetinaNet и Faster R-CNN. Выбор конкретной модели определяется особенностями решаемой задачи и требований к качеству и скорости обработки данных.

Список литературы:

1. What is Object Detection? // IBM URL: <https://www.ibm.com/think/topics/object-detection> (Дата обращения: 28.03.2026)
2. История YOLO – самой известной архитектуры компьютерного зрения // Хабр URL: <https://habr.com/ru/articles/865834/> (Дата обращения: 27.03.2026)

3. Модель RetinaNet // Машинное и глубокое обучение Онлайн-учебник URL: <https://deepmachinelearning.ru/docs/Neural-networks/Object-detection/RetinaNet> (Дата обращения: 26.03.2026)

4. Модель SSD // Машинное и глубокое обучение Онлайн-учебник URL: <https://deepmachinelearning.ru/docs/Neural-networks/Object-detection/SSD> (Дата обращения: 27.03.2026)

5. Двухстадийные детекторы // Машинное и глубокое обучение Онлайн-учебник URL: <https://deepmachinelearning.ru/docs/Neural-networks/Object-detection/Two-stage-detectors> (Дата обращения: 27.03.2026)