

КЛЮЧЕВЫЕ НАПРАВЛЕНИЯ РАЗВИТИЯ ТЕХНОЛОГИЙ ОБНАРУЖЕНИЯ ОБЪЕКТОВ

Верхов К.А.¹, ассистент, k.verkhov@bsuir.by

Ларькин А.Д.², ассистент, a.larkin@bsuir.by

2025

1. Белорусский государственный университет информатики и радиоэлектроники

2. Белорусский государственный университет информатики и радиоэлектроники

Ключевые слова: компьютерное зрение, обнаружение объектов, машинное обучение, видеоаналитика, датасет.

Аннотация: В статье рассматриваются современные тенденции и перспективы развития технологий обнаружения объектов в компьютерном зрении. Отмечается, что несмотря на достигнутый высокий уровень точности, сопоставимый с человеческим в ряде специализированных задач, остаются проблемы, требующие решения. Каждое из направлений рассматривается как ключевой шаг на пути к созданию более эффективных, гибких и безопасных систем компьютерного зрения.

За последние годы алгоритмы обнаружения объектов проделали колоссальный путь и почти достигли уровня точности человека в некоторых узкоспециализированных задачах. Тем не менее, у области остается целый ряд вызовов, которые будут определять вектор ее развития в ближайшие годы. Далее будут рассмотрены несколько направлений, которые играют особую роль в будущем данной области компьютерного зрения.

Использование AutoML и, в частности, Neural Architecture Search (NAS) является одним из наиболее заметных трендов. Идея в том, чтобы поручить машине самой подбирать оптимальную архитектуру детектора, исходя из задачи и данных. Главная проблема заключается в том, что поиск занимает слишком много ресурсов: требуются огромные вычислительные мощности, что ограничивает практическое применение [1]. Тем не менее, если удастся упростить и ускорить этот процесс, можно получить более эффективные и адаптивные модели без участия человека.

Еще одним развивающимся направлением является создание легковесных детекторов. С ростом популярности «умных» устройств и приложений искусственного интеллекта на мобильных платформах возрастает потребность в легких, энергоэффективных и при этом точных детекторах. Сегодняшние компактные сети уже показывают отличные результаты в задачах классификации, однако в области обнаружения объектов они по-прежнему заметно уступают полноценным моделям, имея разрыв в точности обнаружения более 50% [2]. В ближайшие годы исследователи будут стремиться сократить этот разрыв, чтобы сделать возможным использование

детекторов на смартфонах, дронах и других ограниченных по ресурсам устройствах.

Современные модели требуют миллионов изображений с детальными ограничивающими рамками. Создание такого набора данных требует большого количества ресурсов и времени, поэтому появилась необходимость в других способах обучения. Например, обучение с ограниченным количеством примеров (few-shot learning), а также обучение на слабой разметке, где у изображения есть лишь общий ярлык класса, без точного указания положения объекта [3]. Эти подходы могут существенно снизить затраты на создание датасетов и открыть путь к более широкому применению технологий.

Еще одно важное направление связано с переносом знаний. Это поможет использовать модели, обученные на аннотированных изображениях конкретной исходной задачи, для решения отдельной, но связанной целевой задачи [1]. Отсутствие необходимости создания нового объемного датасета для обучения новой модели, позволяет экономить ресурсы и делает технологии более гибкими.

Следующим направлением является трехмерное обнаружение объектов. Оно особенно актуально для автономного транспорта. Машина должна не просто распознавать объект на картинке, а точно понимать его форму и положение в пространстве. Хотя современные алгоритмы показывают высокую точность, до уровня надежности, сравнимого с человеком, еще далеко [4]. А в условиях, где цена ошибки может стоить человеческой жизни, этот разрыв становится критически важным.

Обнаружение объектов на видео также является важным направлением. Большинство существующих методов работают с отдельными изображениями, не учитывая взаимосвязь между кадрами [1]. Но в реальном мире объекты движутся, и информация о динамике может существенно повысить качество распознавания. Использование информации о пространственно-временных связях между кадрами позволит приблизить компьютерное зрение к полноценному пониманию происходящего.

Список использованных источников

1. Zaidi, S.S.A. A survey of modern deep learning based object detection models / Zaidi, S.S.A., Ansari, M.S., Aslam, A., Kanwal, N., Asghar, M. and Lee, B. // Digital Signal Processing – Vol. 126, N. 1 – 2022 – P. 103514.
2. Mittal, P. A comprehensive survey of deep learning-based lightweight object detection models for edge devices / Mittal, P. // Artificial Intelligence Review – Vol. 57, N. 242 – 2024.
3. Dong, N. Incremental-detr: Incremental few-shot object detection via self-supervised learning / Dong, N., Zhang, Y., Ding, M. and Lee, G.H. // Proceedings of the AAAI Conference on Artificial Intelligence – Vol. 37, N. 1 – 2023 – Pp. 543-551.

4. Mao, J. 3D object detection for autonomous driving: A comprehensive survey / Mao, J., Shi, S., Wang, X. and Li, H. // International journal of computer vision – Vol.131, N. 8 – 2023 – Pp. 1909-1963.