

Отчёт по проекту: Анализ объектов в ресторане

Исполнитель: Даниил

Дата: 05.07.2025

1. Задача

Нужно было распознать объекты на видео из ресторана. Я решил сделать не просто детектор, а систему, которая следит за каждым предметом (тарелкой, прибором) и понимает, когда, например, блюдо съели.

2. Подготовка данных

Чтобы модель хорошо училась, я отобрал из видео только самые важные кадры, разметил состояния объектов (например, "полная тарелка" и "пустая") и добавил аугментации, чтобы сделать её устойчивее к разному освещению и движению.

3. Обучение и результат детекции

Обученная модель YOLOv11 показала высокую точность — **общий mAP50-95 достиг 0.817**. Сложнее всего ей дались столовые приборы (cutlery), что ожидаемо из-за их формы и бликов.

Детальные метрики базовой модели:

Класс	Изображений	Объектов	Precision	Recall	mAP50	mAP50-95
all	14	101	0.98	0.974	0.985	0.817
plate_full	10	35	0.991	0.971	0.994	0.889
plate_empty	1	4	0.938	1	0.995	0.877
glass_full	8	8	1	1	0.995	0.844
glass_empty	14	20	1	1	0.995	0.797
cutlery	14	34	0.968	0.899	0.946	0.678

4. Создание системы трекинга

Далее я добавил трекинг, чтобы у каждого объекта был свой уникальный номер (ID). Тут и начались главные инженерные задачи.

- **Проблема:** Трекер работал нестабильно: один и тот же предмет из-за перекрытия рукой мог получить несколько разных ID, а при пересечении вилка и нож могли ошибочно "стать" одним объектом.
- **Решение:** Стандартные подходы не помогали. Чтобы решить проблему в корне, я реализовал полный R&D-цикл:

1. Написал скрипт, который **автоматически нарезал датасет** из тысяч изображений конкретных вилок и ножей из видео.
2. Написал **отдельный пайплайн дообучения**, чтобы натренировать "зрение" трекера безошибочно различать именно эти предметы.
3. **Интегрировал** дообученную модель в итоговый скрипт анализа.

5. Итог

В результате из простой задачи получился полноценный исследовательский проект. Я не только добился высокой точности детекции, но и решил несколько сложных, нетривиальных проблем с трекингом. Финальная система стабильно отслеживает объекты и генерирует лог событий, как и планировалось.