

Определение дистанции до объекта на видео

Смолин Алексей Владимирович

О себе

- Образование:
 - АлтГТУ, Комплексная защита объектов информатизации, 2003-2008
- Опыт в Сбере:
 - Розничный бизнес/Массовая персонализация, руководитель направления, архитектура/java-разработка
- Город:
 - Барнаул
- Контакты:
 - tg: @alexj_smolin
 - mail: alexj.smolin@gmail.com

Описание проекта

- Применение предобученной нейронной сети (детектора) для определения расстояния до объекта на видео
- Видео может быть произвольным, ограничение только в способности детектора определять заданный класс объектов
- Допускается наличие множества объектов в кадре, применяется фильтрация для отслеживания одного из них
- Обработанное видео, рассчитанные метрики и параметры запуска сохраняются для дальнейшего анализа
- Пример запуска:



- Ссылка на репозиторий: <https://github.com/alexj-smolin/ds-school-project/tree/dev>

Бизнес-логика

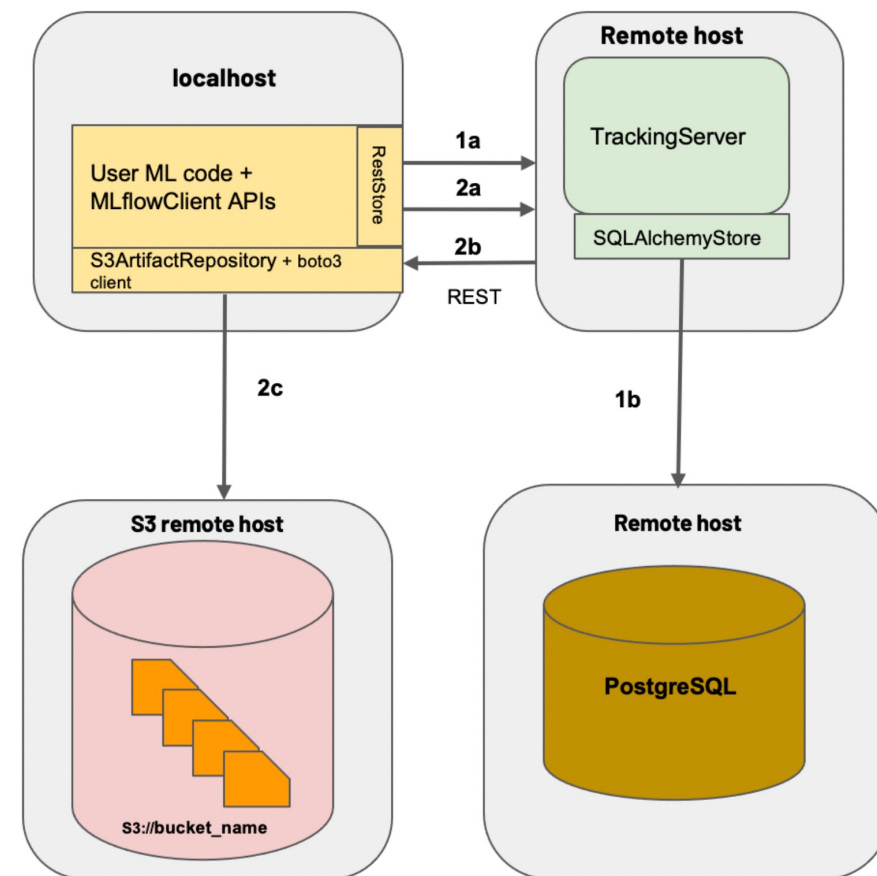
- Покадровая обработка видео с сохранением метрик для каждого кадра
- Из всех объектов в кадре оставляем только наиболее близкий к центру
- Постобработка детектированного объекта:
 - применение сглаживания движения рамки
 - обрезка рамки под размеры исходного объекта (соотношение сторон)
- Различные расстояния рассчитываются на основе реальных размеров объекта и данных с камеры (геометрия, оптика)

Модель данных

- Входные данные:
 - файл с видео
 - характеристики исходного объекта (ширина, высота и класс объекта в терминах используемого детектора)
 - характеристики камеры (фокусное расстояние, ширина и высота матрицы)
 - параметры работы детектора (сложность, минимальная уверенность)
 - параметры постобработки
- Выходные данные:
 - файл с видео + слой с рамкой и основными метриками
 - история изменения различных метрик

Используемые технологии

- Детектор: семейство моделей YOLO (в проекте использовалась модель `yolo-v8-nano`)
- Библиотеки для работы с видео: OpenCV, Torchvision, AV (FFmpeg)
- Трекинг экспериментов: MLFlow + docker (mlflow server, postgres, s3 storage)
- Конфигурация запуска проекта: dotenv



[Сценарий развертки сервера MLFlow](#)

Анализ работы

- Анализ и интерпретация метрик в ноутбуке: <notebooks/EDA.ipynb>
- Направления развития:
 - усовершенствование методов постобработки (переменный фокус, нестабильность рамки)
 - предсказание положения рамки по предыдущим координатам (в случае ошибок детектора)
 - дообучение модели на разные проекции интересующего класса объектов
 - отслеживание множества объектов: идентификация объектов с помощью предсказания положения объекта по истории

Done