1

Имя тема

2

Цель

3

Для корректного распознавания обучающий набор данных должен учитывать разные времена года, темн и св время суток, различныей ракурс съемки и погодные условия. Возможно перекрытие объектов.

4

Набор данные гугл ландмарк отвечает всем этим требованиям. Датасет содержит сложную иерархическую структуру – три уровня вложенности по 15 папок на каждом. Распределение изображений по классам хранится в csv файле. Для возможности итерирования по датасету был разработан класс ГуглЛандмарк.

Так как набор данных содержит 81 тыс классов, даже одна эпоха обучения занимает более 12 часов. К тому же многие изображения не распознать даже экспертному человеку.

5

Поэтому было принято решение создать собтвенный датасет. Набор данных РНДЛандмарк содержит более 2 тыс изображений 46 достопримечательностей Ростова-на-Дону. 70% снимков были выполнены самостоятельно с помощью 6 различных мобильных устройств. 30% изображений получены из поисковых систем и соцсетей.

Также была создана информационная база данных, содержащая..

6

Архитетура резнет101, эффективная для задачи классификации, была обучена в течении 50 эпох с помощью PyTorch. Полученные весы экспортированы в формат tflite через вспомогательные формат onnx и граф pb. Итоговая модель весит 34 гигабайта, что очень много для мобильного приложения. Более легковесные архитектуры фрэймворка PyTorch будут давать менее качественный результат.

7

Хорошие готовые решения предоставляет сервис ML Kit. Но возможности настройки модели из данной библиотеки сильно ограничены.

Фреймворк TensorFlow Lite был разработан специально для работы на мобильных устройствах, сохраняя производительность на компактных моделях. А библиотека TensorFlow Lite Model Maker предоставляет возможность трансферного обучения самых популярных архитектур нейронных сетей для решения задачи классификации изображений.

Это резнет50, основная составляющая которой остаточные блоки, регулирующие глубину нейросети. Мобилнет, построенный на инертированных остаточных блоках. И ЕфишиентНет, оптимально масштабированный по всем трем измерениям.

8

Первичное обучение 7ми моделей происходило в течении 100. Затем по графикам изменения точности и функции потерь было выбрано оптимальное количество эпох для каждой архитектуры. Быстрее всех обучается ЭфишиентНетЛайт0 – 40 эпох, а дольше всех резнет 50 – 80 эпох. Еще были проведены эксперименты с разными размерами батча, но это сущетвенно не изменило точность. Поэтому был оставлен размер батча по усолчания – 256 изображений.

9

Библиотека предоставляет возможность полного и трансферного обучения. Для извлечения признаков из изображения используются веса предобученной модели, а обучается только слой классификации. Сравнив результаты на тестовой выборке, для двух моделей было выбрано трансферное обучение, а для остальных полное.

10

В конечном приложении используется ансамбль из 4х моделей. Резент 50 слишком тяжеловесная – 47 гигабай. Эфишиентлайт3 демонстрирует самую низкую точность, эфишиентнетлайт4 очень близка по точности к эфишиентнетлайт0, но при этом тяжелее в 3 раза.

11

Приложение рндгид состоит из трех экранов. На главном экране приложения можно загрузить изображение из галереи или перейти в камеру и создать снимок.

12

Если приложение запущенно впервые, оно попросит разрешение на доступ к камере.

13

Экран демонстрации отображает результат предсказания. Вверху находится назавание достопримечательности и его изображение. Затем идет краткое описание из информационной базы. По кнопкеОпен ин веб можно открыть полное описание или перейти на сайт объекта. Кнопка шоу он мап переносит на следующий экран.

14

Экран карты демонстрирует расположение объекта на карте города. Для этого в приложение интегрировано гугл мапс сдк.

15

Результаты работы представлены на слайде