

Презентация

Что у нас было?

Разработать приложение, распознающее виды животных с фотографий. У людей куча фотографий с фото ловушек, и на них могут быть приколисты. Эти фотографии нужно выделить из общей кучи и предсказать животное, которое на нём. Нам надо было найти какую-то подходящую модель, составить датасет и на нём обучить её.

Что мы хотели получить на выходе?

По выходе у нас должен был быть ящик с магией, в который кладешь фотку, и он возвращает разметку — основная идея. Остальное можно развить из неё в дальнейшем.

Что можно было сделать для этого?

Юленька говорила попробовать разные открытые модели, чтобы на датасете обучить их и получить веса. Выбрать из этих моделей подходящую. Собрать датасет с теми животными, которые требуются заказчику, возможно разметить его, чтобы на нём обучать и тестить модель.

Что мы реально сделали для этого?

Мы изучили репозиторий megadetectora, попробовали mmdetection; насобирали открытых датасетов, некоторые фотографии были при этом не размечены; их мы разметили с помощью megadetector'a. Попробовали несколько моделей с сайта с MegaDetector'a, которые могли детектировать и классифицировать или же только детектировать.

К каким выводам мы пришли в процессе?

Форматы датасетов различаются между собой от модели к модели; мы упираемся в вычислительные мощности, в отсутствие GPU; нам не подходит mmdetection (вспомнить почему??? (в частности GPU)); модели ResNet и EfficientNet хорошо работают, но не имеют механизма детекции, их мы запускали в связке с Megadetector'ом; на слишком большом датасете модели обучаются очень долго, поэтому он был усечён; EfficientNet показал лучшие метрики.

Что у нас есть сейчас (то есть конечный продукт)?

Веса для модели; собранный датасет; союз Megadetector'a и EfficientNet; скрипт, который установит всё нужное ПО, запустит детектор и классификатор по очереди и выдаст результат.

Что можно сделать в будущем?

Пользовательский интерфейс или хотя бы его элементы по типу разделить размеченные фотографии по разным папкам, добавить свой датасет или веса; распознавание видеозаписей. Можно в конце удалять каждый раз кропнутые фотки.