



ДЕТЕКЦИЯ ЧЕЛОВЕЧЕСКИХ СИЛУЭТОВ НА СНИМКАХ ЛЕСНОГО МАССИВА, ПОЛУЧЕННЫХ С ПОМОЩЬЮ БПЛА

Сергей Земсков

Чемпионат цифрового прорыва 2022 (сезон ИИ)

Описание задачи

- Условие задачи:

На основе данных, полученных с БПЛА, разработайте модель, которая будет находить изображения, на которых присутствуют люди, и будет детектировать их положение на изображении

- Метрика:

$$\text{Result} = 0.6 * \text{Recall} + 0.4 * V_{\text{норм}}$$

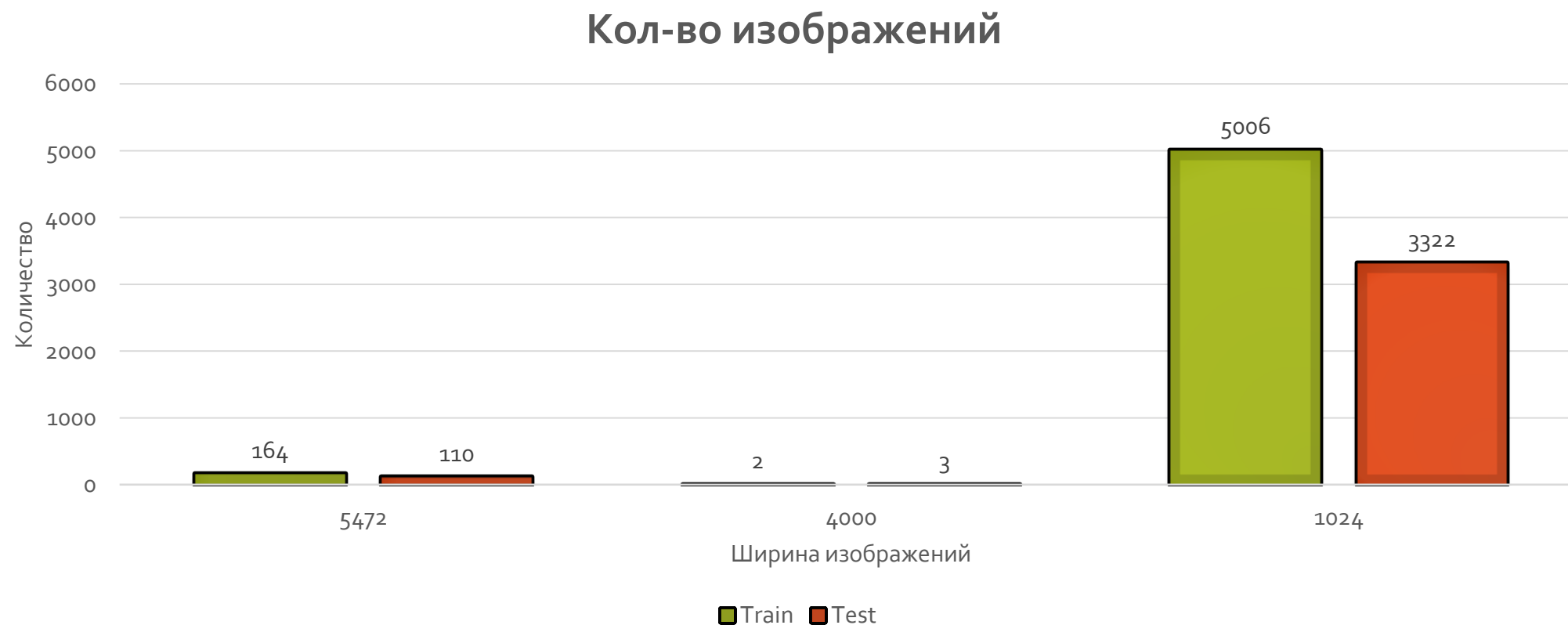
где: $V_{\text{норм}}$ — нормализованное значение разности площади и координат областей, рассчитанное по формуле

$$V_{\text{норм}} = 1 - \sqrt{((x - x_{\text{pred}})^2 + (y - y_{\text{pred}})^2 + (r - r_{\text{pred}})^2) / \text{const}}$$

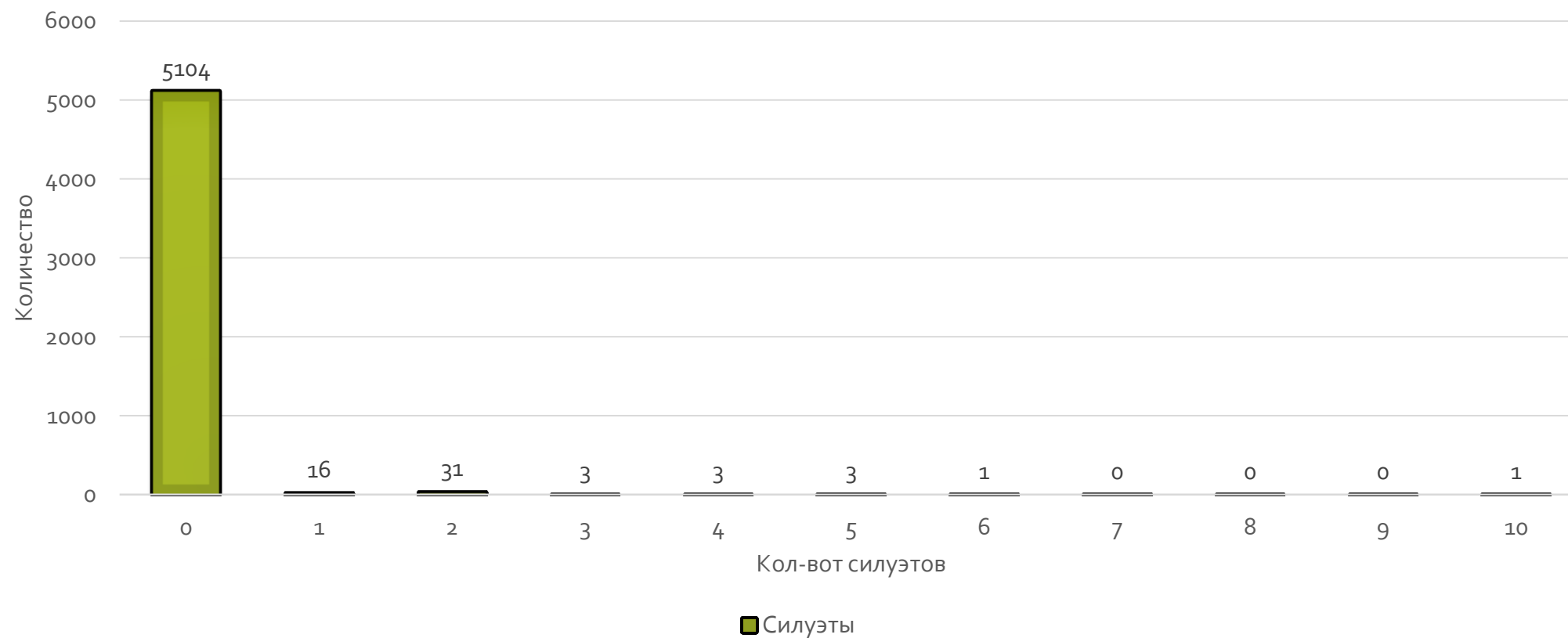
Особенности решения

- Использована модель YOLOv5m на 21M параметров
- Используемые аугментации при обучении: изменение насыщенности цвета, повороты, отражение, переставки кусков изображений и тд. Так как исходные изображения уже повторяются в датасеты с отражением и смещением, поэтому в этой части аугментации были уменьшены.
- В качестве оптимизатора выбран алгоритм AdamW
- Размер изображения уменьшался до 864 пикселей, чтобы батч из 2х изображений уместился в 4Гб памяти GPU.
- Часть изображений без отметок была удалена вручную из тренировочного датасета, во-первых из-за повторяемости, во-вторых для ускорения обучения при ограниченных ресурсах GPU

Кол-во изображений в трейне и тесте



Кол-во силуэтов на каждом изображении в трейне



Этапы разработки

01

Подготовить
разметку
силуэтов на
изображениях

02

Обучить модель

03

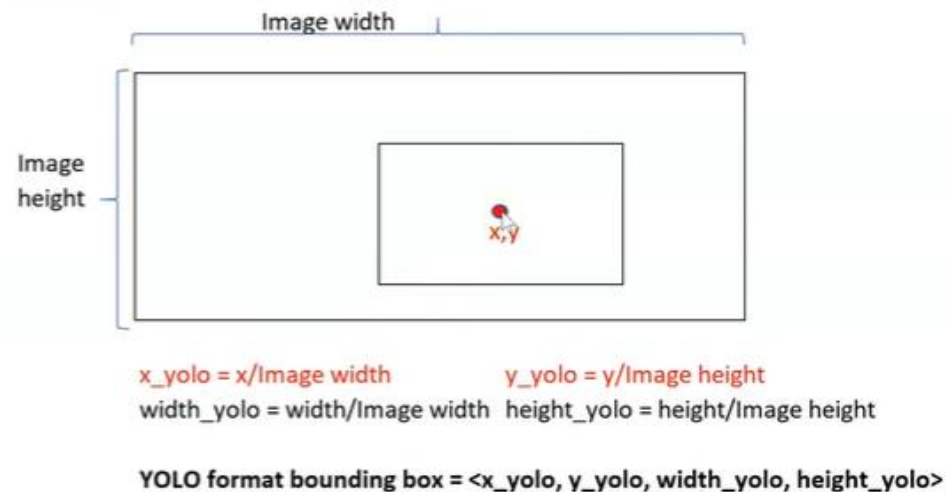
Получить
предсказания
модели

Подготовка разметки

В исходном датасете отметки силуэтов на изображении представлены в формате кол-ва пикселей и радиуса отметки,

```
'[\'{"cx":2549,"cy":2329,"r":80}\', \'{"cx":2437,"cy":2321,"r":74}\']'
```

что потребовало преобразования в формат YOLO (и обратно при предсказании)



Обучение модели

- Запуск пайплайна обучения:

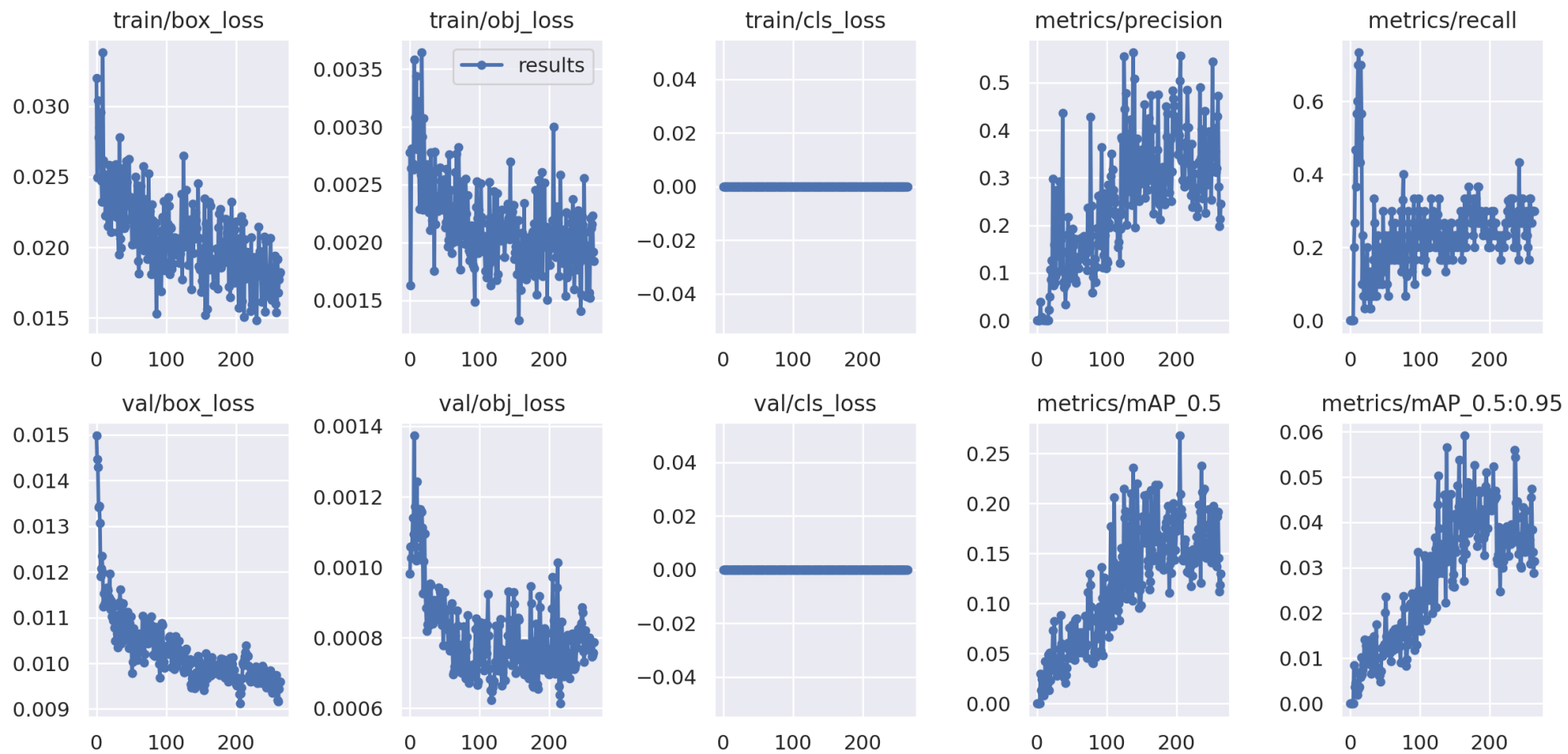
```
!python train.py --img {IMG_SIZE} --batch 2 --epochs 350 --data ../yolo_config.yaml --hyp ../hyp.yaml --weights yolov5m.pt --optimizer 'AdamW'
```

- Гиперпараметры модели (аугментации):

```
hsv_h: 0.015 # image HSV-Hue augmentation (fraction)
hsv_s: 0.7 # image HSV-Saturation augmentation (fraction)
hsv_v: 0.4 # image HSV-Value augmentation (fraction)
degrees: 10.0 # image rotation (+/- deg)
translate: 0.1 # image translation (+/- fraction)
scale: 0.9 # image scale (+/- gain)
shear: 0.15 # image shear (+/- deg)
perspective: 0.0001 # image perspective (+/- fraction), range 0-0.001
flipud: 0.3 # image flip up-down (probability)
fliplr: 0.5 # image flip left-right (probability)
mosaic: 1.0 # image mosaic (probability)
mixup: 0.2 # image mixup (probability)
copy_paste: 0.1 # segment copy-paste (probability)
```

- Для обучения использовалось только 300 изображений от 5162 из трейна
- Для валидации модели изображения были отсечены от 7500.jpg и выбраны 260 изображений (без перемешивания, так как изображения в трейне уже перемешаны)

Динамика обучения и валидации модели



Предсказания модели



Улучшение точности модели

Подходы для улучшения точности предсказаний модели:

- Использование полного набора тренировочного датасета
- Использование дополнительных изображений похожего формата из внешних источников
- Использование более тяжёлой модели (YOLOv5l или YOLOv5m предобученной на изображениях большего разрешения)
- Увеличение размера изображений подаваемых на модель с 864 px до 1024 px
- Улучшение качества разметки исходных изображений

Контакты

- email: sergzemsk@yandex.ru
- telegram: @sergiozemskov
- github: <https://github.com/brut0>