

Отчет по задаче

Реализация алгоритма обнаружения 68 особых точек на лице человека

Для разработки использовался python3.10.

На первом шаге, нужно было создать датасет, для обучения. Было принято решение взять кропы лиц по выходу из детектора dlib без изменения ббокса (без падингов). Первоначально была идея делать кропы с падингами, но из-за неравномерной работы детектора от этой идеи отошли, т.к. получалось слишком маленькое лицо на некоторых экземплярах.

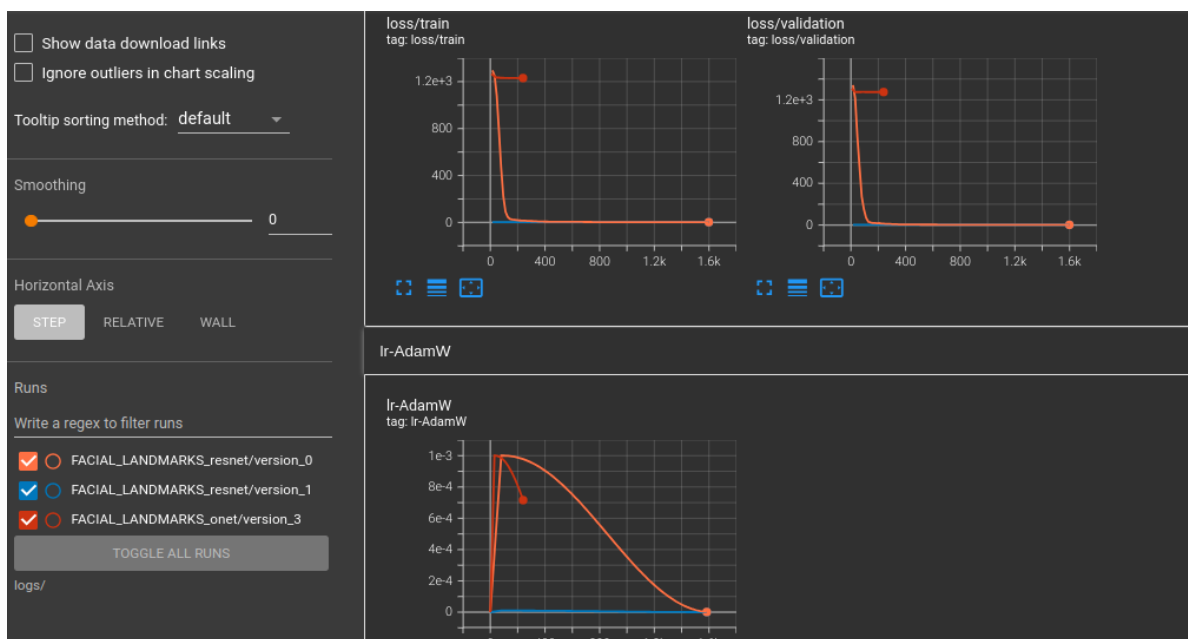
Также было рассмотрено несколько вариантов архитектур: модифицированный **Onet** из статьи и **ResNet18**. В оригинальной статье на вход подавалось изображение размеров **48x48**, но для улучшения качества было решено использовать большее окно: **62x62px**. Также были посчитаны *mean* и *std* метрики по датасету, для улучшения сходимости обучения. В качестве фреймворка обучения был принят *pytorch_lightning*, который позволяет проводить эксперименты детерминировано, т.е. со способностью воспроизвести результат обучения с одинаковыми параметрами.

Функция потерь также была выбрана *Mean squared error (MSE) loss*.

Также для старта обучения была выбрана стратегия *warmup* для незначительных колебаний в начале обучения и лучшей настройки направления изменения градиента на старте.

Весь пайплайн был описан в instructions.md к данному репозиторию https://github.com/lamSVP94/face_alignment/blob/main/instruction.md.

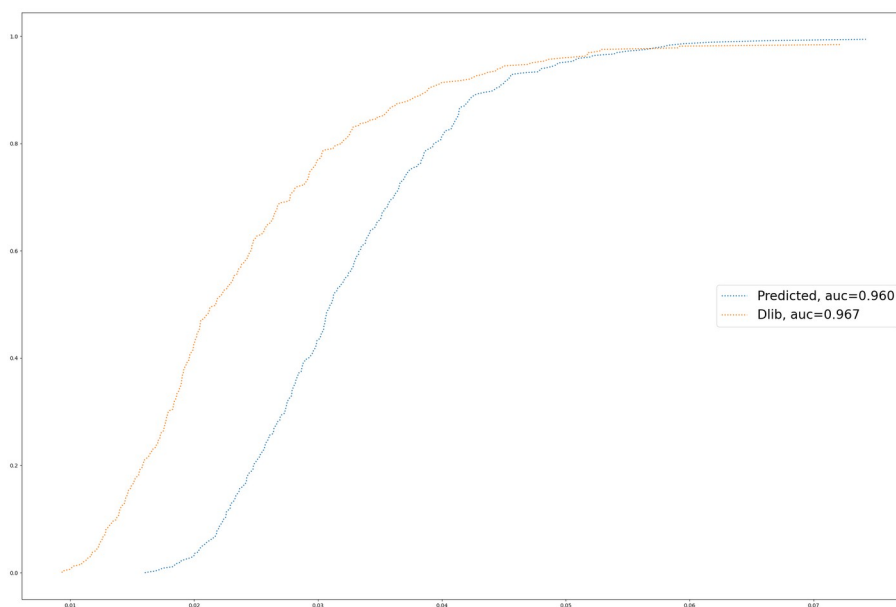
В процессе обучения получены следующие графики изменения loss.



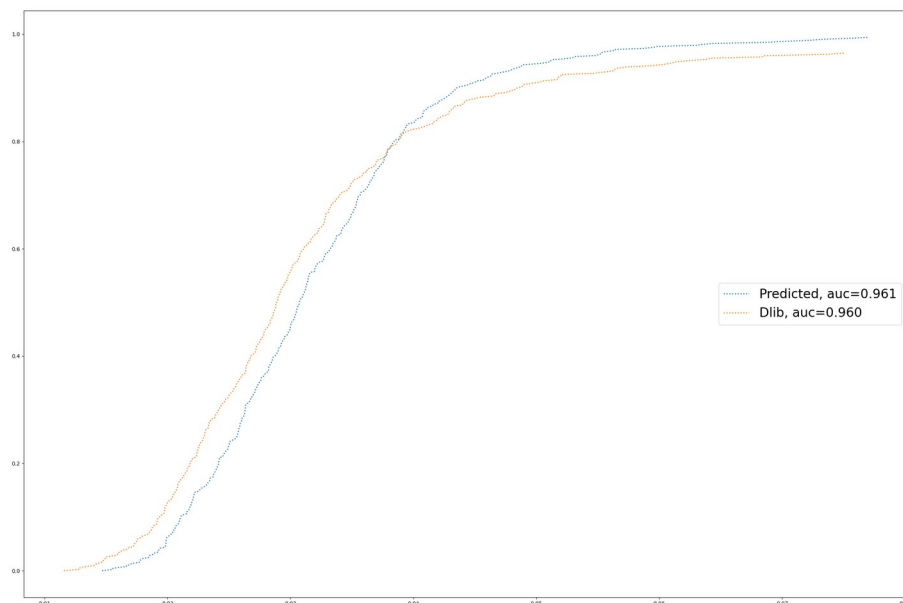
Результаты

Далее в процессе обучения были получены следующие результаты:

На датасете 300W в сравнении с dlib:



На датасете Менро в сравнении с dlib:



В итоге был достигнут порог по точности на датасете Менро. Веса модели доступны по ссылке:

<https://drive.google.com/file/d/1lo-HD8CCAD1nL8uTi3dc4hYTT5LnC2nD/view?usp=sharing>

Идеи по улучшению

- Использование большего входного окна;
- Дополнительные аугментации;
- Более глубокая НС.