



31 мая 2023 г.

Анализ видео-фрагмента интервью с целью получения лучшего
изображения лица
Interview Video Analysis to Obtain the Best Portrait
Программный проект

Выполнила:

Мэн Сыфэй группы БПМИ 208

Принял руководитель проекта:

Передерин Д.А.,

Приглашенный преподаватель, доцент

Школы востоковедения ФМЭиМП НИУ ВШЭ



Описание предметной области и актуальность работы

- Автоматическое извлечение лучшего изображения лица из видеозаписи является важной и сложной задачей.
- Отсутствует веб-приложение, способное генерировать превью видео на основе качества лица.



Цель работы

Цель работы заключается в создании алгоритма¹ и веб-приложения², способного обрабатывать видеоролик, предоставленный пользователем, и возвращающего наилучшее изображение лица из видео.

$$O_{\text{final}} = \alpha O_{\text{beauty}} + \beta O_{\text{ear}} + \gamma O_{\text{gaze}} + \theta O_{\text{blur}} - \delta O_{\text{mar}}$$

$$O_{\text{final}} \rightarrow \max,$$

где O_{beauty} - оценка привлекательности лица, O_{ear} - уровень открытости глаз, O_{gaze} - оценка направления взгляда, O_{blur} - степень размытия кадра, O_{mar} - уровень открытости рта, и $\alpha, \beta, \gamma, \theta, \delta$ - соответствующие нормировочные коэффициенты.

¹Код: <https://github.com/mengsifei/3CourseProject>

²Веб-приложение: <http://1462839-cj10416.tw1.ru/>



Задачи работы

Задачи:

- Изучение научной литературы по проблеме
- Выбор подходящих методов
- Реализация алгоритма и разработка веб-приложения

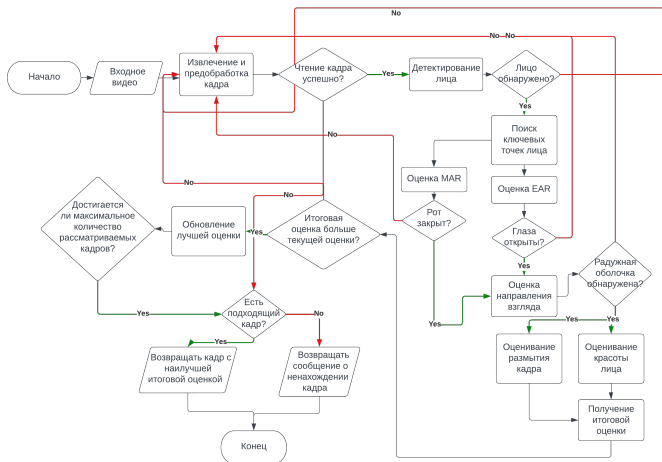


Функциональные требования

- Возможность выбрать самый красивый кадр, в котором глаза человека открыты, рот закрыт или слегка приоткрыт.
- Возможность загрузить видеозаписи в разных форматах.
- Возможность настроить размер лучшего кадра.
- Возможность скачать результат в формате jpg.
- Возможность остановить процесс выполнения алгоритма.



Архитектура алгоритма

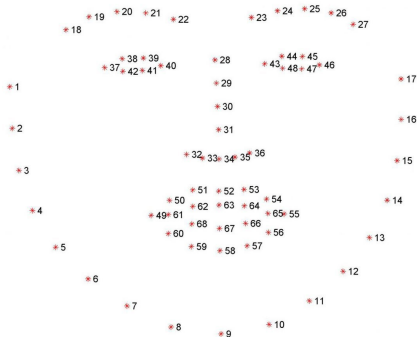


Алгоритм получения лучшего изображения лица из видео



Описание предлагаемого метода

Детектирование лиц и поиск ключевых точек лица



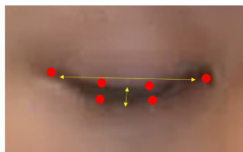
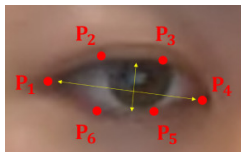
68 ключевых точек лица

- Детектирование лиц:
`get_frontal_face_detector()`
- Поиск ключевых точек лица:
`shape_predictor()`



Описание предлагаемого метода

Проверка открытости глаз



EAR открытого и закрытого глаза

$$EAR = \frac{\|p_2 - p_6\| + \|p_3 - p_5\|}{2\|p_1 - p_4\|}$$

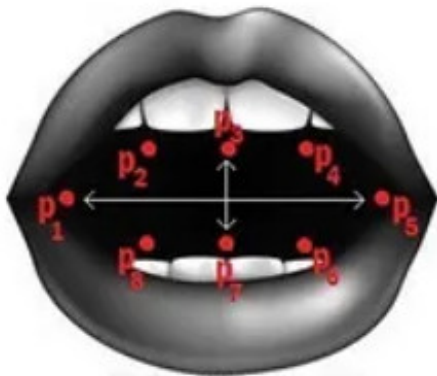
- Пороговое значение^a = 0.25

^aSathasivam S., Mahamad A. K., Saon S., Sidek A., Som M. M. и Ameen H. A. "Drowsiness Detection System using Eye Aspect Ratio Technique". B: IEEE Student Conference on Research and Development (SCORED) (2020)



Описание предлагаемого метода

Проверка открытости рта



MAR открытого рта

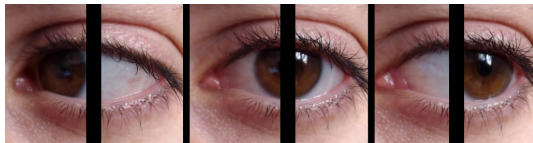
$$MAR = \frac{\|p_2 - p_8\| + \|p_3 - p_7\| \|p_4 - p_6\|}{2\|p_1 - p_5\|}$$

- Пороговое значение = 0.09



Описание предлагаемого метода

Направление взгляда



Площадь склеры левой и правой части глаза
при разных сторонах взгляда

- Проверка направления взгляда основана на отношении площадей склеры левой и правой части глаза.
- Если отношение приближается к единице, то взгляд направлен прямо вперед.



Описание предлагаемого метода

Оценка степени размытия кадра

- Метод основан на вычислении дисперсии лапласиана кадра.
- Преимущества
 - Высокая скорость
 - Удобство реализации с помощью библиотеки OpenCV



Описание предлагаемого метода

Датасет SCUTFBP-5500



(a) Оценка=4.3



(b) Оценка=1.3



(c) Оценка=4.5



(d) Оценка=1.5

- Набор состоит из 5500 фронтальных изображений лиц в возрасте от 15 до 60 лет
- Каждое изображение помечено оценкой красоты в диапазоне от 1 до 5
- Каждое изображение было оценено 60 добровольцами в возрасте от 18 до 27 лет
- Среднее значение всех оценок считалось истинным значением

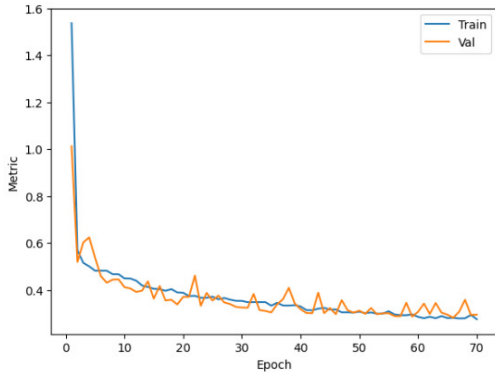
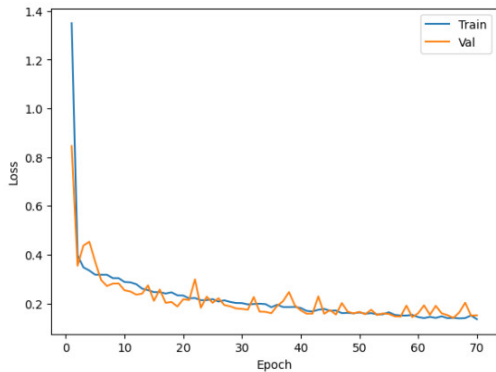


Обучение модели

- Аугментации изображения:
 - RandomRotation
 - GaussianBlur
 - RandomHorizontalFlip
- Функция потерь: SmoothL1Loss с параметром $\text{beta} = 0.4$
- Метрики качества модели:
 - Коэффициент Пирсона
 - MAE
 - RMSE
- Используется адам оптимизатор с длиной шага 10^{-4} и коэффициентом L2 регуляризации, равным 10^{-3} .
- Модель обучается в течение 70 эпох с размером батча 32



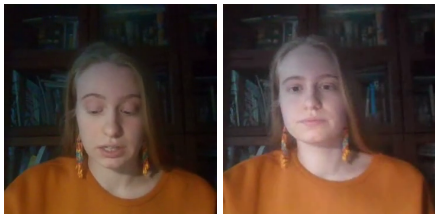
Обучение модели



SmoothL1Loss и MAE в процессе обучения модели EfficientNet V2



Результаты модели



(a) Оценка=3.16

(b) Оценка=3.85

Оценка привлекательности лица с разными выражениями

	EfficientNetV2	AlexNet ^a
PC	0.8469	0.8298
MAE	0.2927	0.2938
RMSE	0.3782	0.3819

Таблица: Сравнение качества моделей на тестовой выборке

^aLingyu Liang, LuoJun Lin, Lianwen Jin, Duorui Xie и Mengru Li.
"SCUT-FBP5500: A Diverse Benchmark Dataset for Multi-Paradigm Facial
Beauty Prediction". B: arXiv:1801.06345 (2018).



Описание системы с точки зрения пользователя

Веб-приложение

- Бэкенд веб-приложения реализован с помощью Flask
- Фронтенд веб-приложения реализован с помощью Bootstrap
- Веб-приложение доступно по ссылке
<http://1462839-cj10416.tw1.ru/>
- Документация доступна по ссылке
<http://1462839-cj10416.tw1.ru/about>



Веб-приложение



Описание системы с точки зрения пользователя

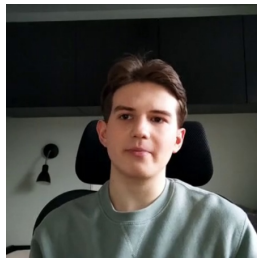
Для запуска веб-приложения на локальном компьютере

- Код доступен по ссылке
<https://github.com/mengsifei/CourseProject/tree/main>
- Алгоритм
 1. Проверить версии пакетов, указанные в файле requirements.txt
 2. Выполнить команды
 - 2.1 `./bootstrap.sh`
 - 2.2 `./start.sh`



Тестирование и анализ результатов

В рамках оценивания качества работы алгоритма были проверены 100 видеозаписей от студентов Школы Востоковедения НИУ ВШЭ.





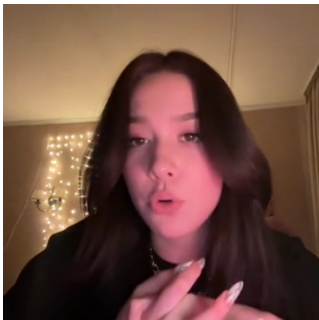
Сравнение времени выполнения алгоритма

	Длительность видео	Время выполнения
1	4 мин. 12 сек.	20.8 сек.
2	3 мин. 11 сек.	28.14 сек.
3	3 мин. 27 сек.	20.94 сек.
4	3 мин. 19 сек.	19.71 сек.
5	3 мин. 14 сек.	23.46 сек.
6	2 мин. 19 сек.	14.66 сек.
7	7 мин. 23 сек.	30.84 сек.
Среднее	3 мин. 87 сек	22.65 сек

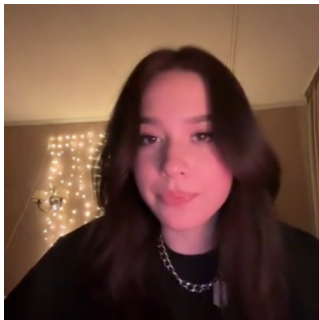
Таблица: Время выполнения алгоритма



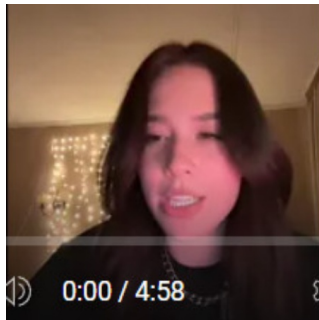
Сравнение с известными аналогами



(a) Olfu79 (Генерация случайных кадров)



(b) Project (Проверка по разным критериям)



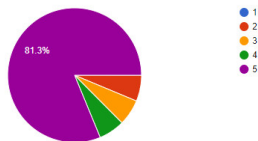
(c) Youtube (С помощью DNN модели)

	Youtube	Project	Olfu79
Время обработки (сек.)	90.68	27.14	< 1

Таблица: Сравнение времени выполнения разных алгоритмов



Направления дальнейшей работы



Распределение оценки качества
веб-приложения

Обнаруженные проблемы

- Алгоритм протестирован только на видео с одним человеком
- Проблема с наличием макияжа

Перспективы:

- Обучить отдельные модели для:
 - Распознавания ключевых точек лица,
 - Повышения разрешения изображения выбранного кадра.
- Добавить русскоязычную версию приложения.
- Добавить функции:
 - Загрузить сразу несколько видео для отработки.
 - Получить видео по ссылке.



Список использованных источников

- Lingyu Liang, LuoJun Lin, Lianwen Jin, Duorui Xie и Mengru Li. "SCUT-FBP5500: A Diverse Benchmark Dataset for Multi-Paradigm Facial Beauty Prediction". B: arXiv:1801.06345 (2018).
- Sathasivam S., Mahamad A. K., Saon S., Sidek A, Som M. M. и Ameen H. A. "Drowsiness Detection System using Eye Aspect Ratio Technique". B: IEEE Student Conference on Research and Development (SCORED) (2020).
- Suwarno и Kevin. "Analysis of Face Recognition Algorithm: Dlib and OpenCV". B: JITE (Journal Of Informatics And Telecommunication Engineering) 4.1 (2020), с. 173—184.
- Mingxing Tan и Quoc V. Le. "EfficientNetV2: Smaller Models and Faster Training". B: International Conference on Machine Learning (2021).