

DeepFaceLab — это программа для замены лица в видео с помощью **нейросети**, работающей на графическом ускорителе NVIDIA / AMD / IntelHD Graphics.

Разработчик: iperov

Сайт разработчика: <https://github.com/iperov/DeepFaceLab>

Разрядность: 64bit

Язык интерфейса: Английский

Минимальные системные требования:

Windows 7 и выше

процессор с поддержкой **SSE** инструкций

2Gb ОЗУ с подкачкой

OpenCL-совместимая видеокарта (NVIDIA, AMD, Intel HD Graphics)

Рекомендуемые системные требования:

Windows 7 и выше

процессор с поддержкой **AVX** инструкций

NVIDIA видеокарта с **6GB** видео,

8Gb ОЗУ

Доступные сборки:

DeepFaceLabCUDA9.2SSE для **NVIDIA** видеокарт вплоть до **GTX1080** и любых 64-бит процессоров.

DeepFaceLabCUDA10.1AVX для **NVIDIA** видеокарт вплоть до **RTX** и процессоров с поддержкой **AVX** инструкций.

DeepFaceLabOpenCLSSE для **AMD** видеокарт и любых 64-бит процессоров.

Зеркало со старыми версиями: <https://mega.nz/#F!b9MzCK4B!zEAG9txu7uaRUjXz9PtBgg>

Особенности **DeepFaceLab** :

- **отсутствие** любых предустановок, кроме **драйверов** на вашу видеокарту. Проверено на чистых Windows 7 и 10, а также в VMWare.
- работает на любых **OpenCL**-совместимых видеокартах с памятью от 256Mb и выше, даже довольно старых.
- **автоматически** выбирает «**лучшую**» видеокарту. Так можно работать с офисными приложениями на видеокарте послабее, а тренировка сети будет автоматически на лучшей.
- есть возможность тренировки полностью на **CPU**. 8-е поколение процессоров **Intel** способен выдать неплохой **демонстрационный** результат за 2 дня тренировки.
- все необходимые мета-данные лиц хранятся непосредственно в извлеченных **JPG** файлах
- есть режим **ручного** выделения лиц
- есть **сортировка** извлеченных лиц для быстрого **удаления** ненужных лиц.
- быстрая работа с видео: извлечение, удаление шума, сборка финального видео

Приступая, поймите: эта программа **не гарантирует** идеальной замены лиц во всех случаях! Всё зависит от качества **исходных** данных, совместимости лиц, источников света, и т.д. и т.п. Да и сама технология появилась недавно (февраль 2018), она далека от идеала, т.к. заменяется только лицо, причём без лба и волос. Вы можете **потратить** кучу времени и сил, а в итоге ничего не получить! У вас есть невысокий шанс создания успешного фейка с первой попытки. Только когда вы сделаете **много** попыток создания фейков с разными лицами, тогда вы поймете все **нюансы**.

DeepFaceLab создана на чистом энтузиазме одним человеком. Поэтому, если вы найдете какие-либо ошибки, то отнеситесь к этому с пониманием.

Посмотрите вводную **видео** инструкцию по работе с программой, чтобы получить **общее** представление о рабочем процессе: <https://www.youtube.com/watch?v=K98nTNjXkq8>

Обновите драйвера на вашу видеокарту.

Первый запуск CUDA сборки в чистой папке из торрента будет происходить долго, т.к. компилируются **GPU** программы.

Определим термины **dst** и **src**.

src - это лицо, которое будет использоваться для замены.

dst - это лицо, которое будет заменяться.

DeepFaceLab**workspace** – наша **рабочая** папка для хранения модели целиком (видео, фото, файлы самой программы). Вы можете переименовывать её, чтобы сохранять как резервную копию или работать с другой.

1) clear workspace.bat

Очистит или создаст все папки внутри папки **workspace**.

Помещаете в папку **workspace** ваше видео **data_src.xxx** формата максимум **1080p** - это видео **откуда брать лицо**. Тестовое видео уже содержится в папке workspace.

Помещаете в папку **workspace** ваше видео **data_dst.xxx** формата максимум **1080p** - это видео **где заменить лицо**. Тестовое видео уже содержится в папке workspace.

где **xxx** - любое расширение видео, например mkv, mp4, avi

2) extract images from video data_src.bat

конвертирует видео **откуда брать лицо** в набор PNG в **workspace\data_src**

Enter FPS (?:help skip:fullfps) :

*сколько **кадров** из каждой **секунды** видео извлечь, пропустить = все кадры.*

Output image format? (jpg png ?:help skip:png) :

*Формат выходных изображений **jpg** или **png**. Png извлекает без потерь, но размер изображения в 10 раз больше. Само извлечение также существенно замедляется, особенно если сохраняется на HDD, вместо SSD.*

Здесь мы собираем кадры из которых будем извлекать src лица. Чем их больше тем лучше.

3.1) cut video (drop video on me).bat

опционально. Обрезает видео **где заменить лицо**, на время которое вы укажете.

Нужно перетянуть видео файл на этот .bat файл, при этом оригинальный файл сохраняется, а появляется новый с суффиксом **_cut**

From time (skip: 00:00:00.000) :

*время **начала** обрезки*

To time (skip: 00:00:00.000) :

*время **конца** обрезки*

Specify audio track id. (skip:0) :

*можно указать **аудио** дорожку. Полезно, если например вам нужен другой **язык** из фильма.*

Bitrate of output file in MB/s ? (default:25) :

***битрейт** выходного файла, пропустить = 25.*

Тестовое видео, которое уже лежит в папке, обрезать **не нужно**.

3.2) extract images from video data_dst FULL FPS.bat

конвертирует видео где **заменить лицо** в набор PNG в **workspace\data_dst**

Здесь извлечение только с ПОЛНЫМ FPS, потому что каждый кадр должен быть обработан.

3.other) denoise extracted data_dst.bat

Denoise factor? (1-20 default:5) :

*степень **подавления** шума*

делается перед извлечением **dst** лиц! делает проход по извлеченным видео кадрам, убирая шум, сохраняя **четкими** грани.

Позволяет тем самым сделать финальный фейк более **правдоподобным**, т.к. нейронная сеть не способна сделать детальную текстуру кожи, но грани делает вполне четкими.

Поэтому, если весь кадр будет более "смазан", то и фейк будет казаться более правдоподобным. Особенно актуально для **фильмовых** сцен, которые обычно очень четкие.

4) data_src extract facesbat

производит выборку **конечного набора лиц** из PNG в папку **workspace\data_src\aligned**

Опции bat файлов:

MT, S3FD детекторы.

Для **src** оптимально **MT** детектор.

MT – немного быстрее, производит **больше** ложных лиц.

S3FD – более точный, **меньше** ложных лиц (рекомендуется)

MANUAL - извлечение вручную, полезно для **src** только для **переизвлечения** уже извлеченных лиц, если найдены где-то ошибки с помощью 4.2.other) data_src util add landmarks debug images.bat

Для этого нужно вручную сделать резервную копию **data_src\aligned**, переместить из этой папки лица в **data_src**, сделать ручное извлечение, затем объединить **data_src\aligned** с резервной копией.

GPU

Здесь либо **ALL** (все), либо **Best** (лучший).

Если у вас только **один** GPU, то нет разницы, что выбирать.

Если вы работаете в **офисных** приложениях на **слабом** GPU, а имеется **мощный**, то выбираете **Best**.

Для максимальной скорости на мульти-GPU выбираете **ALL**, но тогда офисная работа на основном GPU может **подтормаживать**.

DEBUG

Записывает в **workspace\data_src\aligned_debug** каждый кадр с **выделенными** лицами и **лицевыми точками**, тем самым можно смотреть работу детекторов.

4.1) data_src check result.bat

просмотр **результатов** выборки лиц с помощью портативной программы XNViewMP.

Здесь ваша цель - **убрать** ненужные лица.

Сначала **пролистываете** скролом и **убираете** те ненужные лица, которые идут подряд **большими** группами. Не нужно удалять мелкие группы. Для этого есть сортировка.

Если нужное лицо перемешано с другими, то запускаете **сортировки** следующих пунктов.

Сортировку по **резкости** в любом случае делаете, потому что мутные лица **src** нужно удалять.

4.2.1) data_src sort by blur.bat

Сортировка **по резкости**. Запускаете и ждете сортировки. Затем смотрите результаты. Самые **мутные** лица будут в конце. Для **src** важно убрать мутные лица.

4.2.2) data_src sort by similar histogram.bat

После этой сортировки лица будут сгруппированы **по содержанию**, так что отсеять ненужные лица теперь намного проще.

Пролистываете скроллом и **удаляете** ненужные лица **группами**.

4.2.4) data_src sort by dissimilar histogram.bat

Эта сортировка оставляет ближе к **концу списка** те изображения, у которых больше всего **похожих**.

Обычно это лица в анфас, которых больше всего, потому что актёр чаще смотрит прямо на камеру либо куда-то в одном направлении в интервью.

Часть с конца списка можете удалить по **усмотрению**.

4.2.5) data_src sort by face pitch.bat

Опциональный пункт. Сортирует лица так, чтобы вначале списка лицо **смотрело вниз**, а к концу списка – **вверх**.

4.2.5) data_src sort by face yaw.bat

Опциональный пункт. Сортирует лица так, чтобы вначале списка лицо **смотрело налево**, а к концу списка – **направо**.

4.2.6) data_src sort by final.bat

Рекомендованный пункт.

Target number of images? (default:2000) :

ввести целевое количество изображений

Автоматически делает **лучшую** финальную выборку лиц из **любого** количества в **целевое** количество.

Применяйте **только** после очистки набора лиц другими пунктами, если у вас осталось более чем 2000 лиц.

Отсеянные лица перемещаются в мусор (aligned_trash)

4.2.other) data_src sort by black.bat

Сортирует по количеству **черных** пикселей в конец списка. Позволяет отсеять лица **вырезанные** экраном.

4.2.other) data_src sort by one face in image.bat

Перемещает в мусор (aligned_trash) все изображения, в которых нашлось **более 1** лица

4.2.other) data_src sort by original filename.bat

Сортирует по оригинальному имени файла

4.2.other) data_src util add landmarks debug images.bat

Опциональный пункт. добавляет к извлеченным **data_src** лицам изображения с нанесенными **лицевыми точками** с суффиксом к файлу **_debug.jpg**

Это позволяет **вручную** отсеять совсем плохие извлечения из вашего набора **src** лиц.

Перед этим сделайте сортировку по **yaw**.

Имеет смысл удалять только те лица, в которых нанесенные **брови** ниже чем реальные, или контур **челюсти** существенно заходит на лицо.

После отсеивания **удалите** вручную все файлы по маске ***_debug.jpg**

4.2.other) data_src util recover original filename.bat

Восстанавливает оригинальное имя файла извлеченного лица.

Итог по извлечению лиц src.

Мутные лица надо удалять.

Лица **закрывающиеся** чем-то (рукой, волосами, и т.д.) - также нужно удалять.

После приведения набора лиц в надлежащий вид делаем **финальную** сортировку sort by final.

Опционально можете провести пункт 4.2.other) data_src util add landmarks debug images.bat

Вы можете собрать **несколько** разных наборов лиц одного актёра и затем использовать их в зависимости от условий лица **dst**, помещая их в папку **data_src\aligned**

5) data_dst extract facesbat

То же, что и п.4, с некоторыми отличиями.

Здесь мы извлекаем лицо, **которое будет заменяться**.

Детектор **MT, S3FD** ?

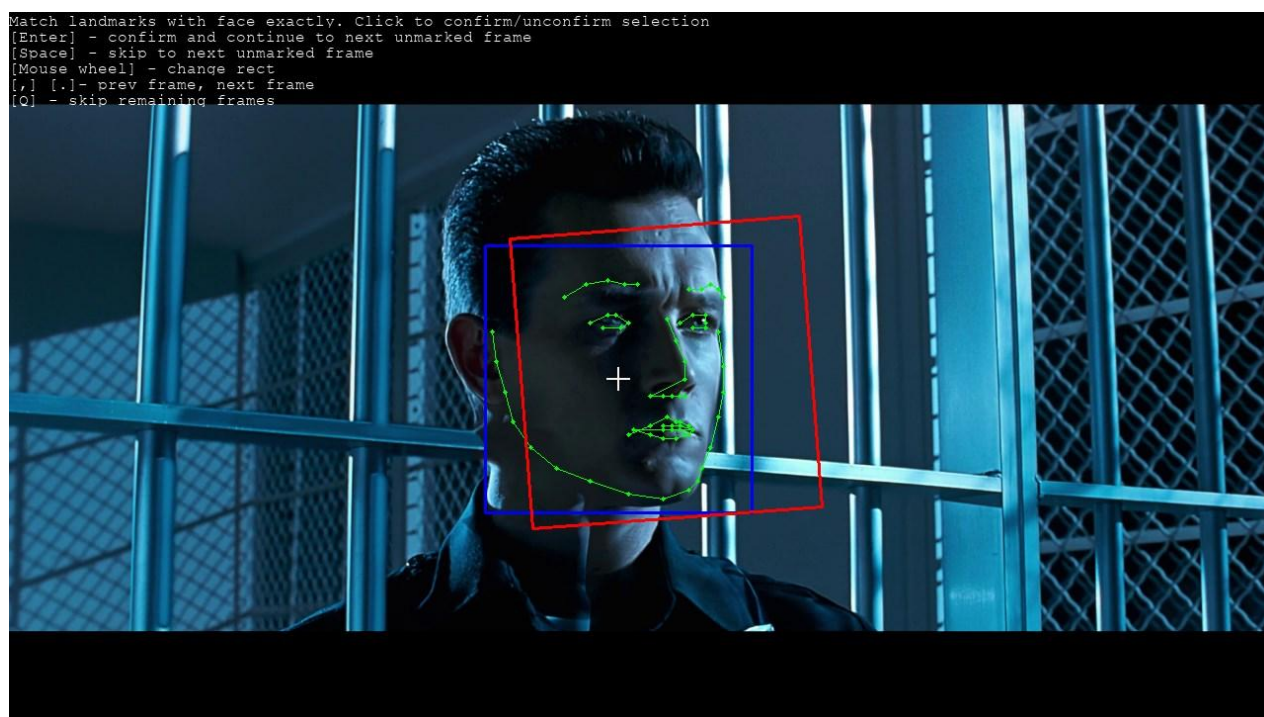
В большинстве случаев - **S3FD**.

Если лицо не определилось в каком-то кадре, то для этого есть опция **+manual fix** - позволяет вручную указать лица на кадрах, где вообще не определилось никаких лиц.

Подвох с **MT+manual fix** в том, что на кадре могут определиться ненужные лица кроме главного, поэтому программа **не предложит** указать лица на этом кадре. В таком случае можете проверить в папке **data_dst\aligned_debug** какие лица обнаружались вообще.

В совсем крайнем случае или для экспериментов есть полностью **ручная** выборка (**manual**), т.е. по каждому исходному кадру **dst** вы **вручную** проходите и указываете лица.

Окно **ручного** исправления лиц.



здесь вам нужно **совместить** зеленые точки с лицом.

Управление:

Enter - подтвердить лицо и следующий кадр.

Пробел - пропустить кадр.

, , - перемещение между кадрами

Колесо мыши - изменять прямоугольник.

Клик – фиксация точек

A – изменение режима точности, влияет на количество FPS

H – скрыть справку

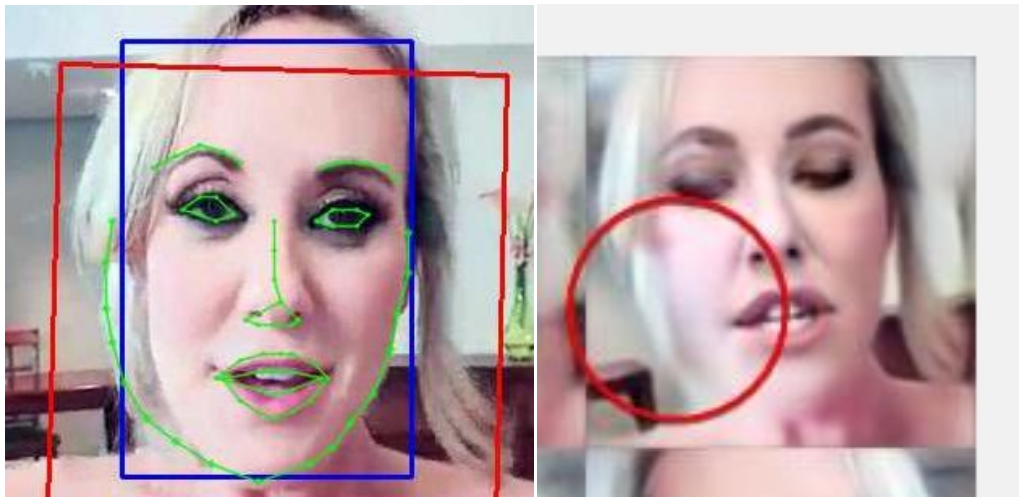
5) data_dst extract faces MANUAL FIX DELETED ALIGNED DEBUG

Очень важный пункт. Напрямую влияет на **качество** итогового фейка.

позволяет переизвлечь те кадры из **dst**, чьи results debug из п.5.1 были **удалены вами**.

Для чего это нужно? Чтобы сделать фейк качественнее, нужно проверять **dst** кадры в папке **aligned_debug**, просмотреть их можно через п.5.1.

Неправильно размеченное **dst** лицо может приводить к значительному **ухудшению** качества:



Если где-то увидите, что **контур** лица существенно **отличается** от реального, например **съехало** на фон, то удаляете эти кадры из папки **aligned_debug** (п.5.1), и запускаете этот пункт.

Произойдет **ручное переизвлечение** удалённых кадров. Видео поясняющее этот процесс: <https://www.youtube.com/watch?v=7z1ykVVCNhM>

5.1) data_dst check results debug.bat

посмотреть все dst кадры с наложенными поверх них предсказанными **контурами** лица

5.1) data_dst check results.bat

Аналогично смотрим результаты выборки лица dst, и удаляем другие **не целевые** лица. А **целевое** лицо, даже **мутные** - оставляем.

5.2) data_dst sort by similar histogram.bat

Если в целевом видео содержатся другие **ненужные** лица, можете произвести эту сортировку, и затем удалить эти лица будет проще.

5.3.other) data_dst sort by original filename.bat

Сортирует по оригинальному имени файла

5.3.other) data_dst util recover original filename.bat

Восстанавливает оригинальное имя файла извлеченного лица.

Итог по извлечению лиц dst.

Ваша цель извлечь **только** целевое лицо (даже мутное) из каждого кадра, удалив все другие лица.

5.4) data_dst mask editor.bat Ручное исправление масок **dst** лиц.

Опциональный и НЕ рекомендуемый пункт, потому что вы можете потратить много времени, но не получить улучшения результата.

ПРИМЕНИМО ТОЛЬКО ДЛЯ SAE МОДЕЛИ с опцией изучения маски либо с опциями изучения стиля.

Внимание: вы можете потратить много времени, но не получить улучшения результата.

Этот режим был создан для улучшения одной из опции конвертатора, но можно также использовать и в своем проекте.

Вы можете вручную подкорректировать маску **dst** лиц, чтобы убрать например препятствия от лица.



Результат конвертации:



Сложные маски труднее тренируются.

Стоит помнить, что в конвертере уже есть режим, который убирает препятствия лица.

Управление в **Mask Editor**.

Левая кнопка мыши – размечать **включающую** маску. Правая кнопка мыши – размечать **исключающую** маску.

Средняя кнопка мыши – завершить текущий **многоугольник**.

Колесо мыши – **отменить** или **повторить** действия с точками. Удерживая **ctrl** – переместится в самое начало или конец истории.

W – **пропустить** и переместить в папку **aligned_skipped** – применимо для сортировки лиц в другую папку. В большинстве случаев вам не нужно это действие.

E – **сохранить** и переместить в папку **aligned_confirmed**. – применимо для сортировки лиц в другую папку. В большинстве случаев вам не нужно это действие.

C – пропустить **в той же** папке.

X – сохранить **в той же** папке.

Q / Z – **предыдущее** изображение.

Зажатый **SHIFT** – ускоряет перемещение по кадрам на 10.

- / + – уменьшение/увеличение размера окна.

ESC – выход.

6) trainbat Тренировка.

Отключаем любые программы, которые могут использовать **видео память**.

Если в процессе **запуска** тренировки появилось много текста, содержащее одно из этих слов:

Memory Error

Allocation

OOM

значит на вашем **GPU** модель не запустилась, и вам нужно **урезать** модель, для этого смотрите описание **опций** моделей.

При первом запуске модели, программа **спросит** о различных **опциях**, которые сохранятся и будут использоваться при **последующих** запусках.

Просто нажимая **Enter** - будут использоваться значения **по-умолчанию**.

Which GPU idx to choose? (skip: system choice) :

*Имея **мульти-GPU** можно тренировать **одну** сцену на **разных** моделях или опциях одной модели **без** клонирования папок. Просто выбираете **индекс GPU** на старте тренировки/конвертации и тогда **файлы** модели в папке **workspace\model** будут содержать префикс этой GPU в имени.*

*Если оставить выбор GPU **по-умолчанию**, то выберется **лучшая GPU** и файлы модели **не** будут содержать **префикс**.*

Enable autobackup? (y/n ?:help skip:%s) :

Включить бэкап файлов моделей. Файлы будут сохраняться в папку model/<>_autobackups/

Write preview history? (y/n skip:n) :

*писать ли **историю** превью на диск*

Choose image for the preview history? (y/n skip:n):

Выбрать изображение для превью истории. [p] – следующее , [enter] - подтверждение

Target iteration (skip:unlimited) :

*целевая итерация, по достижению которой тренировка **остановится**.*

Batch_size (skip:model choice) :

*выбор размера батча - это то **сколько** картинок за раз кормится **нейронной сети** для обучения. По-умолчанию выбирается **низкое** значение, но вы можете **подобрать** это значение самому под свою **видеокарту**. Чем **больше** - тем **лучше**.*

Feed faces to network sorted by yaw? (y/n skip:n) :

*кормит модели **src** лица отсортированные по такому же направлению как и **dst**. Смысл в том, чтобы кормить только нужные лица. Однако пока до конца **не протестировано** хорошо это или плохо.*

Flip faces randomly? (y/n ?:help skip:y) :

*кормит модели все лица **случайно перевернутые** по горизонтали. При выключенной опции финальное лицо будет более **естественным**, но тогда **src** сборка лиц должна **покрывать** все углы поворота.*

Src face scale modifier % (-30...30, ?:help skip:0) :

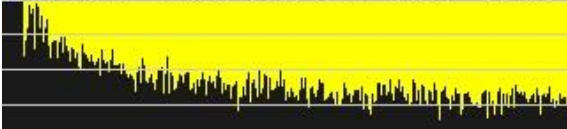

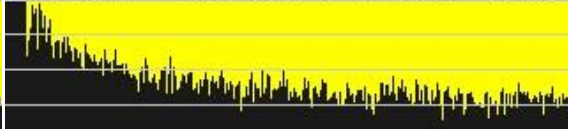

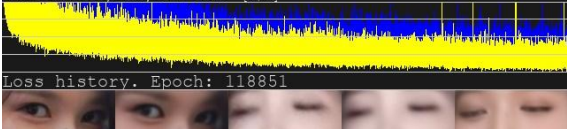

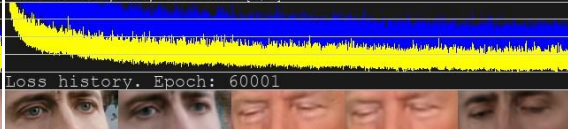

*модификатор масштабирования для **src** лиц. Если **src** лицо более широкое чем **dst** и фейк получился плохим, то имеет смысл немного **уменьшить** это значение.*

Если модель запускается уже второй раз:

Press enter in 2 seconds to override some model settings.

если вы нажмете **Enter** в течение 2х секунд, то появится возможность **заменить** некоторые **опции** модели.



Список моделей для тренировки. Также указано минимальные требования к памяти GPU.

| | |
|-----------------------|---|
| H64 (2GB+) | <p>половина лица с разрешением 64 - это как оригинальная FakeApp или FaceSwap, только лучше за счёт тренировки маски нейросетью + исключаяющей фон вокруг лица + исправленного конвертора. Для видеокарт с видеопамятью 2 и 3Гб данная модель работает в сокращенном режиме, т.е. качество будет хуже чем с 4GB.</p> <div><div><pre>[s]:save [enter]:exit [p]:update [space]:next preview Preview: "src, dst, src->dst" [1/1]</pre><pre>Loss history. Epoch: 102055</pre></div><div><div><pre>[s]:save [enter]:exit [p]:update [space]:next preview Preview: "src, dst, src->dst" [1/1]</pre><pre>Loss history. Epoch: 102093</pre></div></div></div> |
| H128 (3GB+) | <p>как H64 только с разрешением 128. Однако половина лица может плохо обучиться на некоторых условиях света и поворота головы и т.д. Для видеокарт с видеопамятью 3 и 4Gb данная модель работает в сокращенном режиме, т.е. качество будет хуже чем с 5Gb.</p> <div><div><pre>[s]:save [enter]:exit [p]:update [space]:next preview Preview: "src, dst, src->dst" [1/1]</pre><pre>Loss history. Epoch: 118851</pre></div><div><div><pre>[s]:save [enter]:exit [p]:update [space]:next preview Preview: "src, dst, src->dst" [1/1]</pre><pre>Loss history. Epoch: 60001</pre></div></div></div> |

Опции для H64 и H128:

Use lightweight autoencoder? (y/n, ?:help skip:n) :

выбрать урезанную модель. Необходимо для видеокарт с <= 4Gb видео памяти.

| | |
|----------------------------|--|
| <p>DF (5GB+)</p> | <p>модель от dfaker. Полнолицевая модель с разрешением 128, умная функция тренировки лиц, исключая фон вокруг лица.</p>  <p>The screenshot shows the DF model's training interface. At the top, there's a terminal window with commands: [s]:save [enter]:exit, [p]:update [space]:next preview, and Preview: "DF" [1/1]. Below this is a yellow line graph representing the loss history. The text "Loss history. Epoch: 100000" is displayed. The main part of the image is a 4x5 grid of 20 face images, showing a morph from a man's face to Donald Trump's face.</p> |
| <p>LIAEF128 (5GB+)</p> | <p>как DF, только пытается морфировать исходное лицо в целевое лицо, сохраняя черты исходного лица. Морфирование не всегда хорошо, и может сделать вообще не узнаваемое лицо, в таком случае выбирайте DF.</p>  <p>The screenshot shows the LIAEF128 model's training interface. It consists of two side-by-side panels. Each panel has a terminal window at the top with commands: [s]:save [enter]:exit, [p]:update [space]:next preview, and Preview: "src, dst, src->dst" [1/1]. Below the terminal is a yellow line graph for the loss history. The text "Loss history. Epoch: 87423" is on the left and "Loss history. Epoch: 87464" is on the right. The main part of each panel is a 4x5 grid of face images, showing a morph from a man's face to Donald Trump's face. The images in the LIAEF128 grid show more pronounced morphing artifacts compared to the DF grid.</p> |

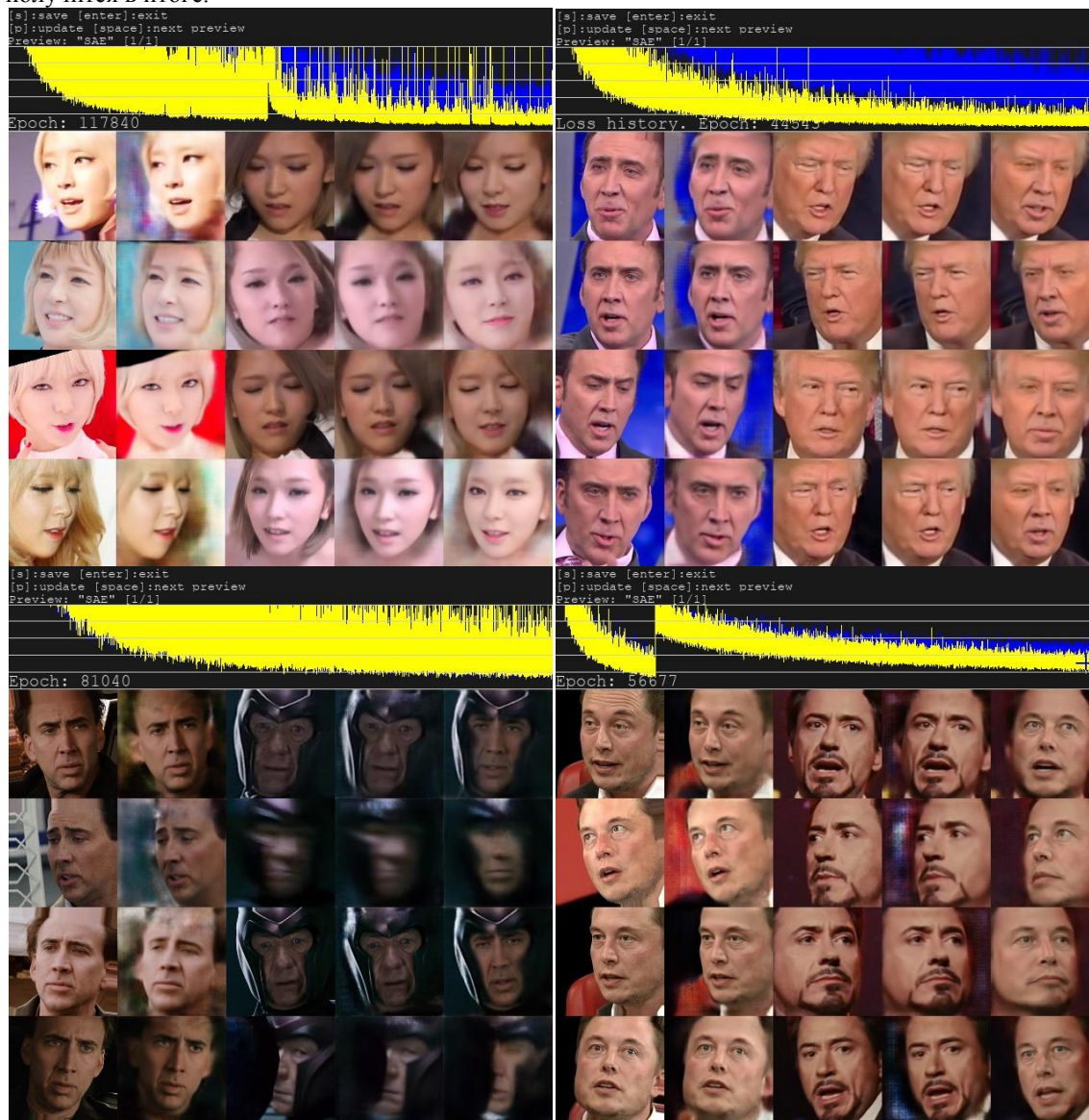
Опции H64, H128, DF, LIAEF128 моделей:

Use pixel loss? (y/n, ?:help skip: n/default) :

позволяет **быстрее** улучшать мелкие **детали** и убрать **дрожание**. Включать только после 20к **итераций**.

SAE
(512MB+)

Самая последняя **лучшая** и **гибкая** модель. Содержит все другие модели. Есть возможность **переносить** стиль лица и освещение, которое будет **изучаться** непосредственно нейронной сетью. При первом запуске можно настроить **различные** параметры. Эти параметры влияют на то, какого размера будет сеть и **запустится** ли на вашей видеокарте, так что можно **подстроить** модель на тренировку как на 512MB так и на 24GB. Конечно же, чем **больше** памяти, тем **лучше** качество получится в итоге.



Пример фейка Cage-Trump: https://www.youtube.com/watch?v=2R_aqHBCIUQ

Пример фейка Elon Musk - Robert Downey jr: <https://www.youtube.com/watch?v=OLWFnPwzgEY>

Опции только для SAE модели:

Resolution (64-256 ?:help skip:128) :

разрешение лица. Больше разрешение - **больше** требуется памяти, **дольше** тренируется модель. Вы можете выбрать любое значение от 64 до 256 **кратное 16**.

Half or Full face? (h/f, ?:help skip:f) :

половинный или *полный* **размер** лица

Learn mask? (y/n, ?:help skip:y) :

учить ли маску. С изучением, маска будет более **сглаженной** и менее дрожащей, иначе будет использоваться грубая. Однако при **стилизованной** тренировке можно обойтись **без** изучения маски.

Optimizer mode? (1,2,3 ?:help skip:%d) :

Эта опция только **NVIDIA** видео-карт. Режим оптимизатора нейронной сети. **1** – по-умолчанию. **2** – позволяет тренировать сеть **x2** большего размера, требует дополнительной ОЗУ памяти. **3** – позволяет тренировать сеть **x3** большего размера, требует еще больше ОЗУ и на 30% медленнее.

AE architecture (df, liae, vg ? :help skip:df) :

тип архитектуры нейронной сети.

AutoEncoder dims (128-1024 ? :help skip:%d) :

*количество **размерностей** сети. Вся информация о лицах будет упакована в эти размерности. Если их не достаточно, то например закрытые глаза не будут распознаны. Больше - лучше, но может не запуститься из-за **нехватки** памяти. Можно уменьшать для достижения работоспособности на вашей **видеокарте**.*

Encoder dims per channel (21-85 ? :help skip:%d)

*количество **размерностей** энкодера сети, помогает модели распознать большие лицевых закономерностей, больше – лучше, но может не запуститься из-за **нехватки** памяти. Можно уменьшать для достижения работоспособности на вашей **видеокарте**.*

Decoder dims per channel (11-85 ? :help skip:%d)

*количество **размерностей** декодера сети, помогает увеличить детализацию, больше – лучше, но может не запуститься из-за **нехватки** памяти. Можно уменьшать для достижения работоспособности на вашей **видеокарте**.*

Remove gray border? (y/n, ? :help skip:n) :

Убрать серую линию по границе предсказанного лица. Требуется больше вычислительных ресурсов.

Use CA weights? (y/n, ? :help skip: %s) :

Использовать ли инициализацию весов типа Convolution Aware. Позволяет достичь больше точности модели, но затрачивает дополнительное время при первой инициализации модели.

Remove gray border? (y/n, ? :help skip:n) :

Убрать серую линию по границе предсказанного лица. Требуется больше вычислительных ресурсов.

Use multiscale decoder? (y/n, ? :help skip:n) :

*использовать ли многомасштабный декодер, позволяет добиться более высокой **четкости**.*

Use pixel loss? (y/n, ? :help skip: n/default)

*позволяет быстрее улучшать мелкие **детали** и убрать **дрожание**. Включать только если качество не улучшается с течением времени. Включение этой опции повышает риск коллапсирования модели.*

Face style power (0.0 .. 100.0 ? :help skip:%.1f) :

*скорость изучения переноса **стиля** лица, таких как **освещение** и **цвет**. Может вызывать **артефакты** при больших значениях. Число с плавающей точкой. 0 - не учить. Включение этой опции повышает риск коллапсирования модели.*

Background style power (0.0 .. 100.0 ? :help skip:%.1f) :

*скорость изучения **фона** вокруг лица. В большинстве случаев **морфит** лицо в **dst**. Число с плавающей точкой. 0 - не учить. Включение этой опции повышает риск коллапсирования модели.*

Apply random color transfer to src faceset? (y/n, ? :help skip:%s) :


*Увеличить вариативность **src** набора лиц с помощью применения переноса цвета на **src** сэмплов из случайных **dst** сэмплов. Это тоже самое как изучение **Face style power**, только перенос цвета более точен, и не требует дополнительных GPU ресурсов. Однако, время тренировки может быть увеличено, из-за того, что нейронной сети придётся учить больше разнообразия **src** набора лиц.*

Enable gradient clipping? (y/n, ?:help skip:%s) :

Включить обрезку градиента, позволяет уменьшить шанс коллапсирования модели, жертвуя скоростью тренировки.

Pretrain the model? (y/n, ?:help skip:n) :

Предтренировать модель с большим количеством различных лиц уже включённых в сборку. Эта техника может помочь тренировать фейк с различающейся формой лица или условий освещения. Лицо будет выглядеть более морфированное. Чем дольше вы тренируете модель, тем больше лицо будет выглядеть морфированным в последствии.

| | |
|--------------------------|--|
| <p>AVATAR (6GB+)</p> | <p>Модель для управления чужим лицом.</p> <pre>[s]:save [enter]:exit [p]:update [space]:next preview [l]:change history range Preview: "AVATAR" [1/1]</pre> <p>Iter: 22604</p>  |
|--------------------------|--|

Как пользоваться AVATAR моделью.

- 1) Помещаете **workspace\data_src.mp4** видео с **квадратным** соотношением сторон, в котором содержится сидящий новостной репортёр, а **фон стабилен** и не перемещается. Долгота видео минут 10-20.
- 2) Прodelьваете **extract images from video data_src.bat** с **FULL fps**
- 3) Помещаете **workspace\data_dst.mp4** видео с **квадратным** соотношением сторон, в котором содержится лицо, которое будет управлять **src** лицом.
- 4) Прodelьваете **extract images from video data_dst FULL FPS.bat**
- 5) Прodelьваете **data_src mark faces S3FD best GPU.bat**
- 6) Прodelьваете **data_dst extract unaligned faces S3FD best GPU.bat**
- 7) Тренируете **train AVATAR.bat** , стадию 1. Настройки **Batch size** на максимальное значение для вашей видеокарты. (пример 32 для 6GB) до приемлемой чёткости лиц.
- 8) Тренируете **train AVATAR.bat** , стадию 2. Настройки **Batch size** на максимальное значение для вашей видеокарты. (пример 4 для 6GB) до приемлемой чёткости лиц.
- 9) Далее как обычно **convert AVATAR.bat**

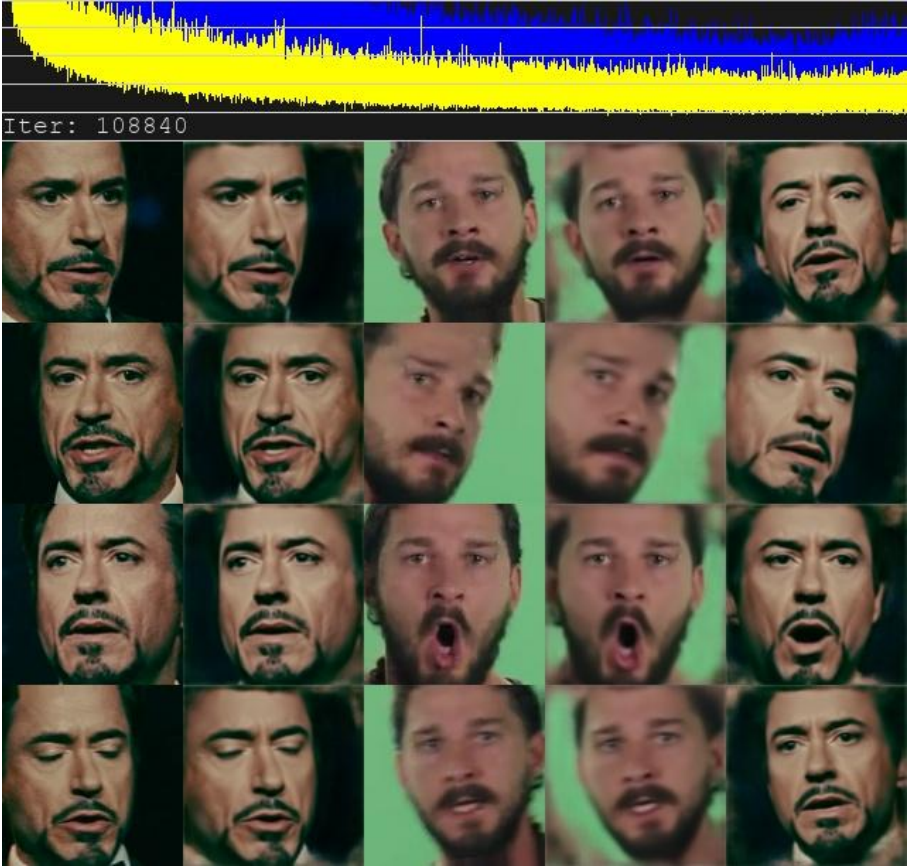
Опции только для AVATAR модели:

Avatar type (0:source, 1:head, 2:full_face ?:help skip:1) :

*Тип аватара. Означает тип конечного изображения. **Source** – значит оно будет такое же как в data_src папке. **Head** – выровненное по носу, охватывающая всю голову. **Full_face** - выровненное по носу, охватывающее всё лицо. Рекомендуется использовать **Head***

Stage (0, 1, 2 ?:help skip:%d) :

*Стадия тренировки. Тренируется сначала 1, потом 2. Можно выбрать стадию 0, в ней будут тренироваться одновременно 2 стадии, но максимальный **batch_size** будет ограничен как для 2-й стадии.*

| | |
|------------------|---|
| SAE HD (6GB+) | <p>Модель для HIGH-END карт с видео памятью от 6ГБ и выше. Модель происходит от SAE, но имеет ряд изменений, а также новую опцию, для достижения максимально возможного качества и правдоподобности лица.</p>  |
|------------------|---|

В процессе тренировки можно выходить через **Enter**, нажав его в окне **Training preview**, и запускать в любое время, модель будет **продолжать** обсчитываться с той же точки.

Тренируем от **24 часов** и больше. Когда результат удовлетворяет - выходим также через **Enter**, нажав его в окне **Training preview**.

Кнопка 'p'(на англ раскладке) в окне **Training preview** обновляет предпросмотр.

В окне **Training preview** также мы видим **кривую ошибки**. Понижаясь, она показывает прогресс тренировки с течением времени.

Итог по тренировке.

Выбираете модель в зависимости от сцены и вашего личного опыта.

7) `convertbat` Наложение лиц.

Выбираете ту **модель** и тот номер **GPU** (если у вас их несколько), с которым тренировали.

Опция **debug** позволяет посмотреть процесс наложения лиц и некоторую техническую информацию по каждому кадру в консоли, нажимаете пробел в окне просмотра.

Основная концепция наложения лиц.



Далее при запуске программа спросит об **опциях**:

Use interactive converter? (y/n skip:y) :

Использовать ли интерактивный режим конвертора. В этом режиме вы можете изменять параметры для каждого кадра и отслеживать изменения в режиме реального времени. Справка по горячим клавишам будет высвечена в начале работы.

Choose mode: (1) overlay, (2) hist match, (3) hist match bw, (4) seamless (default), (5) raw :

Выбор режима наложения лиц.

Режимы

| Overlay | прямое наложение предсказанного лица из нейросети без предварительной обработки | |
|----------------|---|--|
| Hist-match | наложение с уравниванием гистограммы. |  |
| Histh-match-bw | наложение с уравниванием гистограммы из чернобелого канала. | |
| Seamless | наложение методом Пуассона "Poisson image editing." ACM Transactions on Graphics (TOG). Vol. 22. No. 3. ACM, 2003. |  |
| Raw | получить сырые слои для собственной обработки в видео редакторе, например After Effects. |  |

По-умолчанию, если нажать Enter - выберет *seamless*.

Какую выбрать? Зависит от случая. Пробуете все и смотрите результат.

Опции

Seamless hist match? (y/n skip:n) :

Для режима *seamless*. Включить ли уравнивание *гистограммы*.

Masked hist match? (y/n skip:y) :

Для режимов *hist match*, *hist match bw*, *seamless*, указывает, уравнивать ли *гистограмму* по маске лица.

Mask mode: (1) learned, (2) dst, (3) FAN-prd, (4) FAN-dst (5) FAN-prd*dst (6) learned*FAN-prd*FAN-dst (?) help. Default - 1 :

Тип маски.

Learned - **выученная** маска, если вы выбирали опцию **Learn mask**. Контурсы маски смягчены, но иногда могут дрожать.

Dst - **сырая** маска, полученная напрямую из *dst* лиц, поэтому **контурсы** маски дрожащие.

FAN-prd - маска полученная **предтренированной** DEV_FANSEG моделью из **предсказанного** лица. Контурсы лица не дрожат.

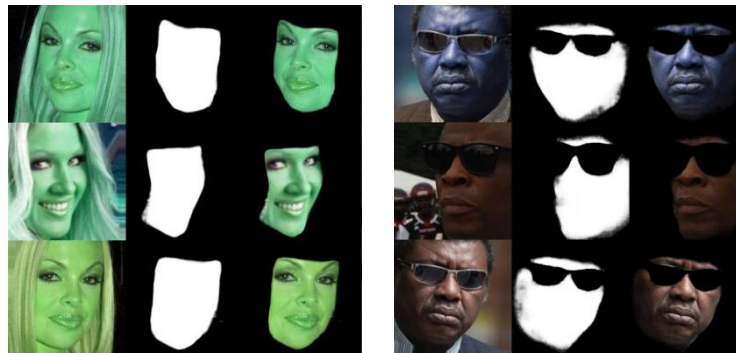
FAN-dst - маска полученная **предтренированной** DEV_FANSEG моделью из *dst* лица. Контурсы лица не дрожат.

FAN-prd*dst - маска полученная перемножением FAN-prd и FAN-dst

`learned*FAN-prd*FAN-dst` – маска полученная перемножением *learned*, *FAN-prd* и *FAN-dst*

Преимущество FAN масок: вы можете получить не дрожащую маску, без изучения её моделью. Также FAN маска убирает препятствия из лица, такие как волосы, пальцы, микрофоны и т.п.

Недостаток FAN масок: могут выдавать артефакты на контурах лиц.

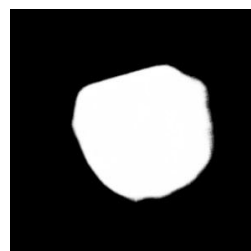
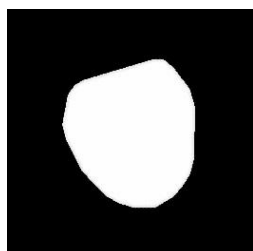
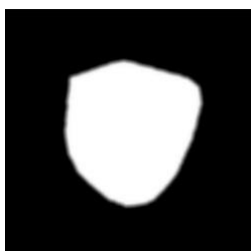


Learned

dst

FAN-prd

FAN-dst



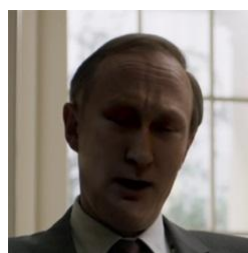
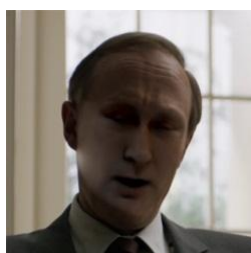
Choose erode mask modifier [-200..200] (default 0) :

Указываете насколько **уменьшить** в размерах *dst* маску. Значение < 0 - расширить маску. Это адаптивное значение, а не абсолютное.

-100 +100



erode mask modifier также позволяет уменьшить мерцания в режиме *seamless*



Choose blur mask modifier [-200..200] (default 0) :

Указываете насколько **сгладить dst** маску. Значение < 0 - уменьшает сглаживание по-умолчанию моделями H64 и H128. Это адаптивное значение, а не абсолютное.

-200 +200



Hist match threshold. [0..255] (default - 255) :

Уменьшение значения **подавляет** артефакты для режимов с уравниванием гистограммы.

(по-умолчанию) 255 231



Choose output face scale modifier [-50..50] (default 0) :

изменить **масштаб** выходного лица в пределах -50+50%. Полезно, когда предсказанное лицо несколько больше оригинала.

-50 +50



Apply color transfer to predicted face? Choose mode (rct/lct skip:None) :

дополнительно **адаптировать** цвет лица, используя **rct** или **lct** алгоритм

Dst лицо Предсказанное лицо



rct lct



Choose sharpen mode:

Выбрать режим увеличения чёткости обычным методом sharpen.

Choose sharpen amount [0..100] :

Какое количество sharpen применить.

Choose super resolution mode:

Выбрать режим «повышения разрешения» для увеличения детальности изображения. Немного улучшает четкость граней, например у глаз.

No Yes



Degrade color power of final image [0..100] (default 0) :

*Степень **деградации** цветности конечной картинки от 0 до 100. Уменьшая общее качество картинки, можно скрыть **недостатки** наложения лица.*

0 100



Export png with alpha channel? [0..1] (default 0) :

*экспортирует только лицо с **альфа каналом** для последующей работы в видео редакторе.*



Итог по наложению лиц.

В начале запускаете с отладкой, пробуя различные параметры и смотрите результат.

Запомнив подходящие значения, запускаете наложение без отладки.

Результат картинок в **workspace\data_dst\merged** - можно использовать самому в **видеоредакторе**, либо **склеить** в видео в п.8

8) converted ... bat Склейка в видео.

Следующие .bat файлы **склеивают** картинки в **видео** с тем же **FPS** и **звуком**, что и **data_dst.mp4** - поэтому **не удаляйте** data_dst.mp4 из **workspace** папки.

to avi.bat

в формат avi

to mp4.bat

в формат mp4

to mp4(lossless+alpha).bat

mp4 без потерь с использованием alpha канала.

to mov(lossless+alpha).bat

mov без потерь с использованием alpha канала. Sony Vegas с использованием QuickTime сможет использовать alpha канал из mov файла.

Всё. Результат в **workspace\result.mp4**

Если результат **не удовлетворил**, можно пробовать разные **опции наложения**, либо продолжать **тренировать** для повышения **четкости**, либо пробовать другую **модель**, либо пробовать другое **исходное** лицо.

Дополнительная информация:

9) util convert aligned PNG to JPG (drop folder on me).bat

позволяет конвертировать папку с извлеченными **PNG** лицами из **старых версий DeepFaceLab** в **JPG**. Для этого перетяните папку с набором лиц прямо на этот **.bat** файл.

Известны случаи, когда **Windows 10** резервирует % **видеопамяти**, даже если видеокарта не используется для вывода картинки. В таком случае придётся чем-то **жертвовать**, чтобы тренировка **запустилась** (batch size, размерности в SAE, итд).

Советы и хитрости.

Узкие лица лучше тренируются на широкие лица. Вот почему фейки с Кейджем так популярны.

Иногда, если нейронная сеть будет получать **dst** лица **только** одного освещения, например, с затенённым носом, при этом **src** лица **не содержат** аналогичного лица с таким же освещением, то нейронная сеть может не понять, куда вообще смотрит лицо на некоторых ракурсах, в результате **предсказанное** лицо будет содержать **неправильный** поворот головы.

В этом случае полезно добавить в тренировочный набор **data_dst\aligned** лица того же персонажа, но извлеченные из других сцен с другим освещением.

Тем самым можно несколько улучшить результат. Но лучше всего, конечно, найти **src** лица с тем же освещением носа.



Продвинутые советы от @GAN-er (только на английском языке):

Tip 1:

You may benefit by starting with a small batch size (within reason) and increasing it later. The reason is that a ****large** batch size will give you a more accurate descent direction but it will also be costlier to calculate**, and when you just start, you care mostly about the general direction; no need to sacrifice speed for precision at that point. There are plenty of sources discussing the batch size, as an example you can check this one:

<https://stats.stackexchange.com/questions/164876/tradeoff-batch-size-vs-number-of-iterations-to-train-a-neural-network>

Tip 2:

Unlike the batch size that the only thing that does is affecting how accurate each step will be as far as the true gradient goes, the dimensions, actually, increase the complexity of your NN. As a rule, ****the more complex a network the better the resulting model****, but since nothing comes for free, ****the more complex the network the more time it will take to converge****.

What you generally want is to ****figure out the max dimensions that you can use**** given your GPU's memory, and your desired max batch size.

You can set the max batch size to something, say K, and then increase the dimensions until you get OOM errors. In the end, you will end up with a triplet, {batch size, ae_dims, ed_dims}

Ideally, you would use 1024 and 85 for your autoencoder and encoder/decoder dimensions, but no card has enough memory for such a configuration even with batch size 1.

Remember that unlike batch size that you can change at will, once you set up the dimensions you can not change them.

Note that ****if you use a complex - high number of dimensions NN, in combination with a small batch size, it will take considerably longer for your model to converge****. So keep that in mind! You will simply have to wait longer, but also you will get a much much better result.

For cards with 11Gb of memory, and for SAE you can try the following settings:

For DF architecture: 12 698 51 For LIAEF architecture: 8 402 47

Tip 3:

If you end up being stuck, i.e. the loss does not go down but for no obvious reason or if you get weird artifacts in some previews before you discard and start from scratch, you may want to flip your DST and SRC for a while. This often is all you need to keep things going again.

Tip 4:

99.995% of your success or failure rate is due to bad SRC or DST sets. This means that 99.995% of your time should be spent in actually ensuring that your sets are well curated. Throwing together a hot podge of material and expecting a decent outcome is guaranteed to result in disappointment. Garbage in, garbage out.

Дополнительное описание SAE.

Эта модель по сути **содержит** все другие модели, если отключить тренировку стиля и `multiscale decoder`.

SAE очень **гибкая**, можно подстроить либо под лучшее **обобщение** лиц сетью, либо под лучшую **четкость** изображения, либо просто чтобы **заработало** на вашей GPU.

Если **src** сборка лиц содержит количество лиц больше чем **dst**, модель может не сойтись. В этом случае используйте опцию **Feed faces to network sorted by yaw**, но только если **dst** лица не содержат волос, закрывающих челюсть, т.к. **yaw** рассчитывается не верно по таким лицам.

Если **src** лицо шире чем **dst**, модель может не сойтись. В этом случае можете попробовать опцию **Src face scale modifier** в -5.

Архитектура **df** делает лицо более похожее на **src**, но если модель не сходится, используйте **liae**.

Если на вашей видеокарте много видеопамати, вы можете выбрать между **большим batch size**, которое улучшает **обобщение** лиц, и **Encoder/Decoder dims** размерностями, которые улучшают **качество** картинки.

Face стиль тренируется, чтобы перенести цвет лица, освещение, макияж. Если он уже хорошо перенесен, то **продолжение** тренировки с **высоким** значением может сделать **артефакты**. Как правило уменьшение значения до 0.1 приводит к стабилизации тренировки.

Background стиль тренируется, чтобы перенести контур лица и окружение. Благодаря контуру лица, подгоняется **src** лицо под контур **dst**.

Как лучше всего тренировать SAE со стилем? Не существует лучшего решения, всё зависит от сцены. Экспериментируйте со **style** значениями, начиная с 10.0 , затем снижая значения до 0.1-2.0 после 15-25к итераций. Включите **write preview history** и отслеживайте **изменения**. Делайте **резервную** копию файлов модели каждые **10к итераций**. Вы можете **откатывать** файлы модели и менять значения, если **что-то пошло не так** в превью-истории.

Работа на AMD/Старых NVIDIA/IntelHD видеокартах.

DeepFaceLab поддерживает старые процессоры без **AVX** инструкций. Для этого используйте **DeepFaceLabCUDA9.2SSE** сборку.

DeepFaceLab поддерживает **AMD**, **IntelHD Graphics** и любые старые **GeForce** видеокарты через использование **OpenCL 1.2** движка.

Для этого используйте **DeepFaceLabOpenCLSSE** сборку.

Работа на **OpenCL** в сравнении с **CUDA** имеет некоторые недостатки:

- извлечение лиц можно запускать только **MT** или **S3FD**, при этом первый проход (**1st pass**) на **MT** всё равно будет работать на **CPU**, но достаточно быстро! **S3FD** полностью работает на **OpenCL** и является предпочтительным методом. Второй проход работает на **OpenCL** без проблем.

- чтобы запустилась **SAE**, её нужно урезать существенно как в **ae_dims**, **ed_ch_dims** так и в **batch size** по сравнению с тем же количеством видеопамяти на **CUDA** версии.

Можете поставить вообще самые **минимальные** значения, а также разрешение **64** и **half face** - в качестве **демонстрации** натренируете фейк худшего качества без проблем даже на **IntelHD Graphics** с 256МБ памяти !

- **каждый** запуск тренировки и после сохранения в процессе тренировки, **OpenCL** будет перезагружать свои подпрограммы, что будет **занимать** некоторое время.

Пожертвование разработке.

Если вам нравится данная программа, вы можете пожертвовать денег для дальнейшего развития.

Yandex.money <https://money.yandex.ru/to/41001142318065>

Paypal https://www.paypal.com/cgi-bin/webscr?cmd=_donations&business=lepersorium@gmail.com&lc=US&no_note=0&item_name=Support+DeepFaceLab&cn=¤cy_code=USD&bn=PP-DonationsBF:btn_donateCC_LG.gif:NonHosted

bitcoin:31mPd6DxPCzbpCMZk4k1koWAbErSyqkAXr

Полезные ссылки

| | |
|--|--|
| Страница разработки | https://github.com/iperov/DeepFaceLab |
| Вводная видео инструкция по работе с программой | https://www.youtube.com/watch?v=K98nTNjXkq8 |
| Мой канал фейков | https://www.youtube.com/channel/UCEtoVzBLuE-ni-54h7kYvsg |
| Популярный канал фейков | https://www.youtube.com/channel/UCUix6Sk2MZkVOr5PWQrtH1g/videos |
| Фейки на реддите | https://www.reddit.com/r/GifFakes |
| Готовые src наборы лиц для DeepFaceLab | https://mega.nz/#F!y1ERHDaL!PPwg01PQZk0FhWLV05_MaQ |
| поиск подходящей знаменитости по загруженному лицу | http://www.pictriev.com |
| поиск подходящей порнозвезды по загруженному лицу | https://findface.sex/ru https://findpornface.com https://pornstarbyface.com http://www.didshedoporn.com |