



# ДЕЕРФАКЕ В ИССЛЕДОВАНИЯХ ВОСПРИЯТИЯ ЛИЦА

**БАРАБАНИЦКОВ В.А.**

*Московский государственный психолого-педагогический университет (ФГБОУ ВО МГППУ),  
г. Москва, Российская Федерация  
ORCID: <https://orcid.org/0000-0002-5084-0513>, e-mail: [vladimir.barabanschikov@gmail.com](mailto:vladimir.barabanschikov@gmail.com)*

**МАРИНОВА М.М.**

*Московский государственный психолого-педагогический университет (ФГБОУ ВО МГППУ),  
г. Москва, Российская Федерация  
ORCID: <https://orcid.org/0000-0002-8862-4007>, e-mail: [marinovamm@yandex.ru](mailto:marinovamm@yandex.ru)*

Излагается современный метод коллажирования изображений с заменой лиц Deepfake — продукт искусственного интеллекта (AI), с помощью которого можно создавать высококачественные реалистичные видеоролики с поддельным или замененным лицом, без явных следов манипуляций. На основе приложения DeepFaceLab (DFL) поэтапно описывается процесс создания видеоизображений «невозможного лица». Представлены результаты экспериментов по изучению закономерностей восприятия подвижного «невозможного лица» и их отличий в статике и динамике. В качестве стимульного материала выступили две модели виртуальных натурщиков с «невозможным лицом», сконструированные с помощью DFL: видеоизображения химерического лица, правая и левая стороны которого принадлежат разным людям и тэтчерезированное лицо, на котором области глаз и рта повернуты на 180°. Показано, что феномены восприятия «невозможного лица», зарегистрированные ранее в условиях статике (целостность восприятия раздвоенного изображения, эффекты distraction и инверсии), при экспозиции динамических моделей сохраняются и приобретают новое содержание. В отличие от коллажированных изображений оригинальные лица в статике и движении независимо от эгоцентрической ориентации оцениваются позитивно на уровне высоких значений. При всех тестируемых условиях пол виртуального натурщика определяется адекватно, воспринимаемый возраст переоценивается. Оценки эмоций виртуального натурщика по его видеоизображениям дифференцируются на основные (устойчивые) и дополнительные (меняющиеся) состояния, соотношение которых зависит от содержания конкретного эпизода. Технология синтеза изображений Deepfake существенно расширяет возможности психологического исследования межличностного восприятия. Использование цифровых технологий упрощает создание стимульных моделей «невозможного лица», необходимых для углубленного изучения репрезентаций внутреннего мира человека, и формирует потребность в новых экспериментально-психологических процедурах, отвечающих более высокому уровню экологической и социальной валидности.

**Ключевые слова:** IT-Deepfake, DeepFaceLab, межличностное восприятие, видеоизображение лица, невозможное лицо, виртуальный натурщик, динамика и статика выражений лица, химерическое лицо, тэтчерезированное лицо, целостность восприятия, динамический дистрактор, инверсионный эффект, привлекательность лица, базовые эмоциональные экспрессии.

---

**Финансирование.** Работа выполнена в рамках госзадания Министерства просвещения РФ № 730000Ф.99.1.БВ09АА00006.

*Для цитаты:* Барабанищikov В.А., Маринова М.М. Deepfake в исследованиях восприятия лица // Экспериментальная психология. 2021. Том 14. № 1. С. 4—19. DOI: <https://doi.org/10.17759/exppsy.2021000001>



# DEEFAKE IN FACE PERCEPTION RESEARCH

**VLADIMIR A. BARABANSCHIKOV**

*Moscow State University of Psychology & Education, Moscow, Russia*

ORCID: <https://orcid.org/0000-0002-5084-0513>, e-mail: [vladimir.barabanshikov@gmail.com](mailto:vladimir.barabanshikov@gmail.com)

**MARIA M. MARINOVA**

*Moscow State University of Psychology & Education, Moscow, Russia*

ORCID: <https://orcid.org/0000-0002-8862-4007>, e-mail: [marinovamm@yandex.ru](mailto:marinovamm@yandex.ru)

Presents the state-of-the-art Deepfake face replacement image collage method, an artificial intelligence (AI) product that can be used to create high-quality, realistic videos with a fake or replaced face, with no obvious signs of manipulation. Based on the DeepFaceLab (DFL) application, the process of creating video images of an “impossible face” is described step by step. The results of the experiments of studying the perception patterns of the moving “impossible face” and their differences in statics and dynamics are presented. The stimuli were two DFL-generated models of virtual sitters with impossible faces: a video image of a chimerical face, in which the right and left sides belong to different people, and a Tatchered face with the eyes and mouth areas rotated by 180°. It was shown that the phenomena of perception of the “impossible face”, registered earlier under static conditions (integrity of perception of the split image, distraction and inversion effect), are preserved and acquire a new content when dynamic models are exposed. In contrast to the collaged images, the original faces in statics and motion, regardless of egocentric orientation, are evaluated positively at the level of high values. Under all tested conditions the gender of the virtual sitter is determined adequately, the perceived age is overestimated. Estimates of the virtual sitter’s emotions from his video images are differentiated into basic (stable) and additional (changing) states, the ratio of which depends on the content of a particular episode. Deepfake image synthesis technology significantly expands the possibilities of psychological research of interpersonal perception. The use of digital technologies simplifies the creation of “impossible face” stimulus models necessary for in-depth study of representations of the human inner world, and creates a need for new experimental-psychological procedures corresponding to a higher level of ecological and social validity.

**Keywords:** IT-Deepfake, DeepFaceLab, interpersonal perception, video facial imaging, impossible face, virtual sitter, dynamic and static facial expressions, chimeric face, tethered face, perceptual integrity, dynamic distractor, inversion effect, facial attractiveness, basic emotional expressions.

**Funding.** The research was performed within the framework of the state order of the Ministry of Education of the Russian Federation №730000Ф.99.1.БВ09АА00006.

**For citation:** Barabanshikov V.A., Marinova M.M. Deepfake in Face Perception Research. *Экспериментальная психология = Experimental Psychology (Russia)*, 2021. Vol. 14, no. 1, pp. 4–19. DOI: <https://doi.org/10.17759/exppsy.2021000001> (In Russ.).

## Введение

Современные компьютерные программы и технологии цифровой графики позволяют пользователям свободно оперировать с различными классами видеоизображений и создавать их модификации. В области психологии это ведет к развертыванию качественно нового уровня экспериментов, созданию наиболее подходящего стимульного материала, эффективного строя методических процедур, оригинальных способов обработки и анализа полученных данных.



Ярким примером подобных технологий является Deerfake\* — процедура синтеза изображений с использованием искусственного интеллекта (AI). Реализуется случай, когда на материале готовых изображений либо их фрагментов нейросеть собирает по пикселям новый ролик [13]. Программа, созданная в 2014 г. в Стэнфордском университете, используется для решения задач компьютерной графики в кино и игровой индустрии, рекламе, журналистике, корпоративном обучении, организации музейного пространства при проектировании зданий. Наибольшую популярность технология получила благодаря способности подменять внешность людей на фото и видео. В сценах из кинофильмов, ТВ-шоу, выступлений политиков и т. п. лица участников могут быть заменены любыми персонажами; возможны изменения возраста, пола, эмоционального состояния; создаются лица-аватары, которые в действительности не существуют. Конечный продукт применения технологии — дипфейки имеют высокие визуальные и акустические качества, затрудняющие распознавание подделки.

Психологическую науку лицо интересует как важный источник информации о мыслях, переживаниях, особенностях личности, намерениях человека, их проявлениях во вне, восприятии и понимании другими людьми. С этой точки зрения, Deerfake представляется технологией трансформации лица, обеспечивающей новые возможности изучения репрезентаций внутреннего мира человека.

Многие годы в психологии проводят исследования восприятия «невозможного лица» — коллажированных изображений, выполненных на основе элементов или частей, из которых лицо состоит или может состоять, но которые занимают «чужие» места, имеют необычную пространственную ориентацию, рассогласованное содержание или принадлежат разным людям. Разработаны стимульные модели (коллажи) «зеркального» лица, составленного из отдельных (правой либо левой) сторон лица и их реверсий, «композитные» и «химерические» лица, объединяющие в одном изображении верхнюю и нижнюю либо правую и левую половины лиц разных людей, «тэтчерезированное» лицо, содержащее перевернутые рот и глаза и др. [1; 11; 12; 15; 17; 19].

Искусственные объекты восприятия исходно противоречивы и обладают необычными свойствами. Воспринимая «невозможное лицо», наблюдатель попадает в проблемную ситуацию, требующую согласования непривычных пространственных свойств и отношений коллажированного изображения с личным опытом. Возникающие представления о человеке характеризуют взаимосвязь частей и целого в межличностном восприятии, роль эгоцентрической ориентации лица (отдельных частей), закономерности его персонификации, механизмы порождения образа натурщика.

Большинство исследований описываемого типа выполнено на материале статичных изображений человека: фотографий, портретов, рисунков, позволяющих путем коллажирования легко варьировать пространственные отношения лица и менять содержание его частей или элементов. На сегодняшний день этого недостаточно. Современный тренд науки о лице состоит в изучении подвижного, или «живого», лица, его изменений в реальном времени с учетом текущего контекста. Выражения подвижного лица характеризуют активность человека в целом, которая конституирует межличностную ситуацию и регулирует потоки субъект-субъектных трансакций [3]. Именно здесь оказываются полезными информационные технологии Deerfake и их аналоги, позволяющие изучать восприятие видеозображений «невозможного лица» с характеристиками, задаваемыми экспериментатором.

\* От англ. Deep learning и fake — подделка с помощью глубинного обучения.



Цель данной статьи состоит в том, чтобы: 1) описать метод, позволяющий произвольно менять пространственные отношения «живого» лица человека, включенного в процесс коммуникации (мы называем его коллажированием видеоизображений; 2) проиллюстрировать тип исследований «невозможного лица», выполняемого с помощью оригинальной технологии.

## I. Цифровое коллажирование видеоизображений лица

Для конструирования модели виртуального натурщика существует нейросетевое программное обеспечение DeepFaceLab (DFL) – система с открытым исходным кодом для создания дипфейков. Программа рассчитана как на пользователей без знаний о методах глубокого обучения, так и на опытных разработчиков, которые хотят улучшить имеющуюся базу. Последовательность генерации дипфейка состоит из трех основных этапов: извлечение, обучение и конвертация. На этапе извлечения нейросеть поочередно распознает, соотносит, сегментирует и извлекает все изображения лиц из исходного и целевого видеофайлов, генерируя будущую модель. Затем нейросеть обучается и интегрирует извлеченное лицо в целевое видеоизображение. На этапе конвертации система использует обученную модель подмены лица, которую в результате можно настраивать и регулировать самостоятельно, в зависимости от требуемых условий.

Ключевое преимущество DFL среди множества методов и программ по замене лиц состоит в том, что разработчики имеют доступ ко всему коду проекта, что позволяет воспроизводить результаты работы собственной и/или других моделей в своих проектах. Библиотека DeepFaceLab состоит из облегченной версии Keras фреймворка – Leras. В качестве дополнительных преимуществ Leras выступают:

- простой и гибко настраиваемый процесс сборки модели;
- более высокая скорость (в среднем на 20%) обучения моделей;
- возможность контролировать обработку тензоров на более низком уровне, чем это позволяет Keras [9; 14].

Программное обеспечение DeepFaceLab, а также примеры реализации дипфейков находятся по адресу интернет-страницы: <https://github.com/iperov/DeepFaceLab>

Перед установкой DeepFaceLab стоит учесть системные требования, необходимые для корректной работы программного обеспечения.

Минимальные системные требования:

- ОС Windows 7 или выше (64 бит);
- процессор с поддержкой SSE-инструкций;
- оперативная память объемом не менее 2 Гб + файл подкачки;
- OpenCL-совместимая видеокарта (NVIDIA, AMD).

Рекомендуемые системные требования:

- процессор с поддержкой AVX-инструкций;
- оперативная память объемом не менее 8 Гб;
- видеокарта от производителя NVIDIA с объемом видеопамати не менее 6 Гб.

Выполнив переход по вышеуказанной ссылке, следует пролистать вниз до отображения таблицы под названием «Releases», в которой находятся источники загрузки актуальных версий DFL. Для работы под операционной системой Windows, как в нашем случае, потребуется перейти по второй ссылке из таблицы **Windows (Mega.nz)**, внутри которой будут доступны два установочных exe-файла:



1) DeepFaceLab\_NVIDIA\_build\_01\_04\_2021.exe (3.55 Gb) — версия программы для видеокарт NVIDIA высокого уровня производительности;

2) DeepFaceLab\_NVIDIA RTX2080Ti\_and\_earlier\_build\_01\_04\_2021.exe. (2.9 Gb) — версия программы для видеокарт NVIDIA RTX2080Ti и ниже.

Для запуска инструмента с использованием графических процессоров AMD следует зайти в папку «2020», которая находится внутри обозначенной выше ссылки для скачивания ПО, и загрузить оттуда версию DeepFaceLab 1.0.

DFL 2.0 работает с поддержкой CUDA Compute Capability 3.0, данная утилита доступна для скачивания: <https://developer.nvidia.com/cuda-downloads> (2.9 Gb) — после загрузки установочного файла формата .exe необходимо провести установку его компонентов, выбрав настройки по умолчанию.

Наша виртуальная модель создавалась и обучалась на базе видеокарты NVIDIA GeForce Asus TUF-RTX3090 с 24 Gb памяти с использованием DFL 2.0. Также было использовано следующее техническое оборудование: процессор Intel(R) Core(TM) i7-10700KF CPU @ 3.80Ghz, оперативная память 32 Gb, твердотельный накопитель Samsung SSD 970 EVO Plus 500Gb и OS Windows 10. Дальнейшее обсуждение процедуры создания дипфейка будет излагаться на примере цифрового коллажирования видеоизображения химерического лица.

Сконструированная нами модель виртуального натурщика содержала в себе частичное наложение изображений двух женских лиц, изъятых из видеоинтервью. На левую половину лица Натурщицы 1 (целевое видеоизображение) импортировалась аналогичная часть лица Натурщицы 2 (исходное видеоизображение) (рис. 1).

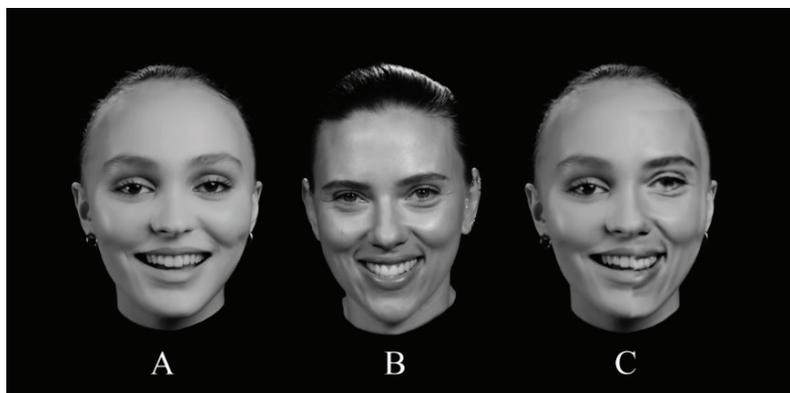


Рис. 1. Стимульные модели в прямом положении: А — оригинальный стоп-кадр с изображением Натурщицы 1 из целевого видео; В — оригинальный стоп-кадр с изображением лица Натурщицы 2 исходного видео; С — синтезированный стоп-кадр с изображением химерического лица

Программа DeepFaceLab 2.0 состоит из нескольких файлов .bat, которые в зависимости от этапа и типа задачи используются для создания дипфейка; они находятся в базовой папке «DeepFaceLab\_NVIDIA» вместе с двумя подпапками:

- «**internal**» — папка содержащая внутренние файлы для работы программы;
- «**workspace**» — папка рабочего пространства, в которой находятся тренировочные модели, видео, наборы данных и окончательные видеовыходы.

В DFL есть возможность использовать два набора данных, т. е. два видеоролика для последующей замены лиц с одного видео на другое.



Видеоматериалы формата .mp4 следует поместить в папку «workspace» и переименовать каждое видео в соответствии с требованиями:

- **data\_dst** – целевое видеоизображение, на которое будет наложено другое лицо;
- **data\_src** – исходное видеоизображение, содержащее лицо, которое будет наложено.

После всех проделанных манипуляций запускается команда: 2) **extract images from video data\_src** – данная функция извлекает кадры из исходного видео data\_src.mp4 и автоматически помещает их в подпапку «data\_src». Команда предполагает следующие настраиваемые параметры:

- **FPS** – частота кадров – по умолчанию (рекомендуется); если значение частоты кадров изменяется, например до 10 кадров в секунду, то это означает, что по итогу будет извлечено меньшее количество изображений, соотносимое с 10 кадрами в секунду);
- **JPG / PNG** – формат извлеченных кадров; формат не влияет на качество дипфейка, однако, вычислительная скорость работы при изъятии jpg выше, чем при изъятии формата png, по этой причине мы рекомендуем прописать команду «jpg».

На этом этапе и далее каждая функция, которая будет выполнена, в конце командной строки всплывшего окошка сообщит о завершении работы текстовой строкой «Done» и попросит для продолжения нажать любую клавишу.

Как только произойдет раскадровка data\_src, следует переход к пункту: 3) **extract images from video data\_dst FULL FPS** – это аналогичная команда для второго набора данных, но в отличие от предыдущей команды она без выбора со стороны пользователя извлекает все кадры из целевого видеофайла data\_dst.mp4 и помещает их в папку «data\_dst». Рекомендуется, как и в прошлом случае, для скорости работы извлекать кадры формата «jpg».

Проделанные шаги ведут к первому этапу подготовки исходного набора данных для дальнейшего машинного обучения. Он заключается в выравнивании граней и создании изображений лиц размером 512 × 512 из извлеченных кадров, находящихся внутри двух папок «data\_src» и «data\_dst». Для того, чтобы это сделать, следует запустить функцию: 4) **data\_src faceset extract** (рис. 2).

Существует несколько вариантов подготовки изображений; в число основных входят:

- 1) **data\_src faceset extract MANUAL** – ручное извлечение и самостоятельная отметка контрольных точек;
- 2) **data\_src faceset extract** – автоматическое извлечение с использованием встроенного алгоритма S3FD (рекомендуется).

Доступные опции:

– выбор зоны охвата извлечения лица в зависимости от типа модели, которую вы хотите обучить:

- a) full face (FF);
- b) whole face (WF);
- c) head (HEAD);

– выбор вычислительного процессора (графический или центральный).

В нашем случае для конструирования стимульной модели с «невозможным лицом» потребовалось выбрать зону охвата WF.

После извлечения запускается команда: 4.1) **data\_src view aligned result** – открывается внешнее приложение, позволяющее просмотреть содержимое папки «data\_src» и проверить ее изображения, удалив лишние, например, с наличием ложных срабатываний, неправильно выравненных исходных лиц и лиц других людей, которые не участвуют в создании модели.

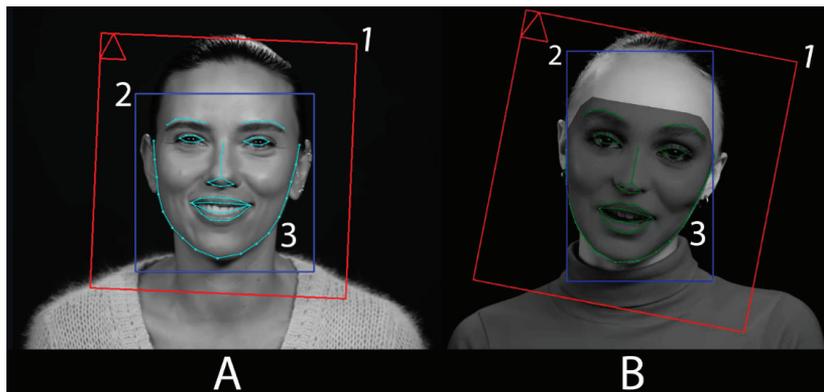


Рис. 2. Снимок экрана, демонстрирующий автоматическое обнаружение граней лица:  
А — обнаруженное лицо из исходного видео при помощи команды `data_src faceset extract MANUAL`;  
В — обнаруженное лицо из целевого видео при помощи команды `data_dst faceset extract`;  
1 и 2 — область поиска лица; 3 — ключевые точки, расположенные вокруг глаз, рта, носа, бровей и по контуру лица

Следующая команда: **4.2) `data_src sort`** — содержит алгоритмы сортировки, упрощающие отбор нежелательных, смазанных лиц или лиц с артефактами в изображении; доступно 15 опций:

- `blur` (размытие);
- `motion_blur` (размытие в движении);
- `face yaw direction` (направление движения лица);
- `face pitch direction` (направление наклона лица);
- `face rect size in source image` (размер прямоугольника лица в исходном изображении);
- `histogram similarity` (подобие гистограммы);
- `histogram dissimilarity` (несходство гистограммы);
- `brightness` (яркость);
- `hue` (оттенок);
- `amount of black pixels` (количество черных пикселей);
- `original filename` (исходное имя файла);
- `one face in image` (одно лицо на изображении);
- `absolute pixel difference` (абсолютная разница пикселей);
- `best faces` — (лучшие лица);
- `best faces faster` (лучшие лица (быстро)).

В процессе создания нашей версии виртуальной модели мы смогли выявить наиболее эффективные алгоритмы сортировки, используемые в следующем порядке: `best faces` → `blur` → `histogram similarity` → `original filename`.

После отбора изображений запускается команда: **5) `data_dst faceset extract + manual fix`** — автоматическое извлечение граней + ручная корректировка либо: **5) `data_dst faceset extract`** — только автоматическое извлечение (рекомендуется).

В этих командах шаги те же, что были проделаны ранее с исходным набором данных. При первом знакомстве с программой мы также советуем использовать автоматическое извлечение, так как данная опция эффективно справляется самостоятельно и не расходует



много времени в отличие от ручной подготовки, для которой также требуются высокие вычислительные мощности компьютера.

После завершения задачи экстракции следует очистить папку «data\_dst» от нежелательных изображений так же, как это было сделано в предыдущих шагах. Для этого запускается соответствующая команда: **5.1) data\_dst view aligned results**, — где будут показаны результаты экстракции, после этого активируется команда: **5.2) data\_dst sort**, — как и в случае с исходным набором лиц, этот инструмент позволяет сортировать все выровненные лица в папке «data\_dst», чтобы легче найти неправильно выровненные лица, ложные срабатывания и т. д. На этом этапе предварительной подготовки заканчивается.

Следующий этап — обучение модели **XSeg и маркировки лиц**. XSeg — это встроенный в DFL инструмент для ручного маскирования целого лица или его частей. Также данный инструмент используется для того, чтобы избежать наличия различных артефактов на итоговом видео, таких как волосы на лице или руки, перекрывающие заменяемое лицо. В нашем случае мы используем этот инструмент для разметки правой половины, на которую будет накладываться соответствующая часть лица другой натурщицы (рис. 3).

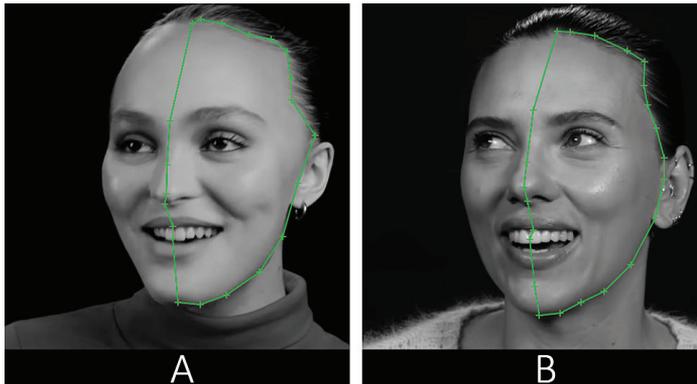


Рис. 3. Снимок экрана с изображением лиц натурщиц из интерфейса XSeg, с поворотом головы и соответствующим выделенным фрагментом половины лица: А — натурщица из целевого видео; В — натурщица из исходного видео

Предварительно обученной встроенной модели XSeg не существует. Необходимо создать свою собственную модель или воспользоваться теми, которые уже были «натренированы» другими пользователями и находятся в свободном доступе.

Спадающие на лицо волосы, руки у лица или другие различные предметы, являются помехами для тренировки накладываемой маски. В зависимости от того, насколько целевое и исходное видео перегружены, зависит количество отметок ключевых точек. Опытным путем мы выяснили, что при первом обучении модели целесообразно сделать минимум 20–30 маркировок вручную, иначе натренированная модель (маска) может содержать в себе различные артефакты и ненужные детали, либо может выглядеть неестественно.

XSeg работает со всеми типами изображений лиц, такими как FF, WF и HEAD, что представляет возможность полного контроля замены частей лица.

Шаги, необходимые для предварительного обучения: **5.XSeg) data\_dst mask — edit** — команда, вызывающая интерфейс XSeg для пометки конечных границ;



**5.XSeg) data\_dst mask – fetch** – копирует грани, содержащие полигоны XSeg, в папку «align\_xseg». Может использоваться для сбора помеченных лиц, и их повторного участия в будущих тренингах по модели XSeg.

**5.XSeg) data\_dst mask – remove** – удаляет помеченные полигоны XSeg из извлеченных кадров, если нужно полностью переобучить модель.

Для data\_src требуется проделать аналогичны функции.

После завершения вышеописанных действий, запускается функция обучения модели – **5.XSeg) train** (рис. 4).

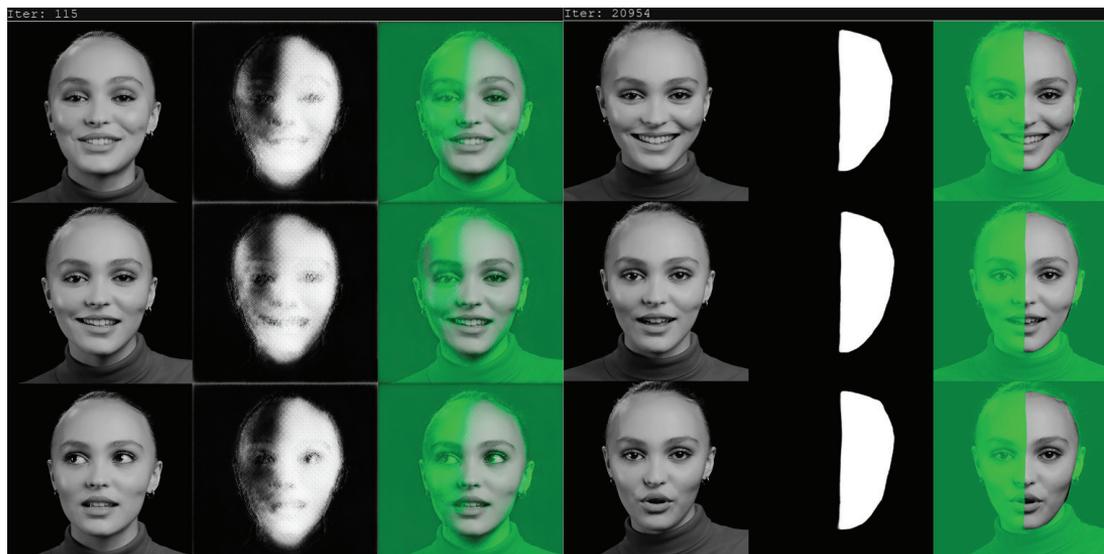


Рис. 4. Снимок экрана, демонстрирующий окно с процессом обучения XSeg и улучшение качества маски в зависимости от количества итераций: изображение слева – 115 итераций, изображение справа – 20954 итерации

Данный этап обучения проходит сравнительно быстро. Рекомендуется каждые 20 тыс. итераций сохранять прогресс и переключаться на функцию **5. XSeg) data\_dst mask – edit** для того, чтобы проконтролировать прогресс обучения. При несоответствии результатов требуется снова вручную отмечать новые полигоны на тех кадрах, которые по каким-либо причинам не устраивают. После редактирования полигонов функция тренировки запускается вновь, количество итераций не обнуляется, а остается с сохраненным прогрессом и продолжает обучение. Рекомендуемое итоговое количество итераций на данном этапе для корректной работы маски – 80–100 тыс.

После завершения тренировки XSeg маски должны быть интегрированы в процесс машинного обучения с помощью запуска команды **5. XSeg) data\_dst trained mask – apply** – это заменит маски, созданные по умолчанию при первом извлечении лиц, на маски, сгенерированные обученной моделью XSeg.

*Примечание:* существуют маски, которые генерируются при первичном извлечении, но они не так эффективны, как маски от XSeg, так как они не распознают артефакты или препятствия на лице модели, в отличие от тренированных масок XSeg. С их помощью можно определить, какая область лица является самым лицом, а что является фоном или дефектом. По этой причине рекомендуется применять XSeg для любых обучаемых моделей.



Применив и сохранив все настройки XSeg следует переход к основному этапу машинного обучения. В настоящее время доступны 2 модели:

1) **SAEHD** (требуется от 6Gb видеопамати): автокодер высокой четкости и качества для высокопроизводительных графических процессоров с не менее 6 Gb видеопамати. Полностью регулируемый;

2) **Quick96** (требуется от 2 до 4 Gb видеопамати): простой режим, предназначенный для графических процессоров низкого уровня с 2–4 Gb видеопамати. Содержит в себе следующие фиксированные параметры:

- разрешение изображений лица  $96 \times 96$  пикселей;
- профиль только анфас.

Quick96 рекомендуется только для видеокарт начального уровня и для новых пользователей, которые хотят протестировать набор данных и получить быстрый результат. В нашем исследовании мы поочередно применяли оба метода машинного обучения. Так, для создания модели химерического лица с использованием упрощенного метода Quick96 понадобилось 25 часов машинного обучения и свыше 1 млн итераций, чтобы получить приемлемый результат. Для того, чтобы получить более качественный результат, при помощи продвинутого способа обучения SAEHD, нам понадобилось свыше 45 часов непрерывного машинного обучения и 200 тыс. итераций. Невзирая на меньшее количество итераций, качество наложенной маски выглядит намного лучше – сравнительный результат представлен на рис. 5, где заметны более плавный переход граней наложенного лица, четкость и детальность морщин, зубов, волосков бровей, передается блеск в глазах из целевого видео.

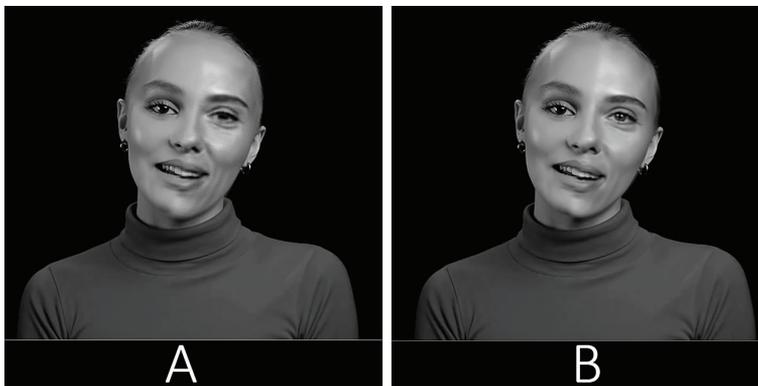


Рис. 5. Виртуальная модель натурщицы, правая сторона лица которой принадлежит другой девушке: А – обученная модель в упрощенном режиме Quick96, свыше 1 млн. итераций, 25 часов обучения; В – обученная модель в продвинутом режиме SAEHD, около 200 тыс. итераций, 45 часов обучения

Для обеих моделей обучения есть общее правило – чем больше количество итераций, тем качественнее итоговый результат.

После проверки всех настроек и выбора подходящей модели обучения, в соответствии с техническими возможностями и личными требованиями, запускается одна из моделей машинного обучения: **6) SAEHD** или **6) Quick96**.

Поскольку Quick96 не настраивается, всплывающее окно команд покажет только один вопрос: что использовать при обучении – ресурсы CPU (центральный процессор) или GPU (графический процессор)? Однако SAEHD предоставляет больше возможностей для гибких настроек.



Следующий этап заключается в редактировании обученной модели на видеовыходе, для этого запускается функция **7) merged Quick96** или **7) merged SAEHD** в зависимости от предыдущего выбора. После этого всплывает окно командной строки с текущими настройками, а также окно предварительного просмотра, в котором показаны все элементы управления, необходимые для работы интерактивного слияния.

Функции, которые могут быть применены в конвертере — смена натренированных масок, цвет накладываемого фрагмента, увеличение или уменьшение резкости, размытие граней наложенного лица, растушевка, размытие в движении, масштабирование изученного лица (больше/меньше), увеличение или уменьшение глубины цвета и т. д. Все настройки являются сугубо индивидуальными, так как каждая модель имеет свои уникальные характеристики — от цвета кожи до технического качества самого изображения.

После того, как произойдет объединение и конвертация всех лиц, будет создана папка под названием «merged», внутри нее находится подпапка «data\_dst», содержащая все кадры, а также подпапка «merged\_masked», которая содержит кадры наложенной маски.

Последний этап — преобразовать весь вышеописанный набор обратно в видео и объединить с исходной звуковой дорожкой из файла data\_dst.mp4.

Для этого потребуется запустить одним из предложенных файлов .bat, которые используют FFmpeg для объединения всех кадров в одно видео следующих форматов: avi, mp4, lossless mp4 или lossless mov:

- 8) merged to avi;
- 8) merged to mov lossless;
- 8) merged to mp4 lossless;
- 8) merged to mp4.

Мы советуем объединять видео в формат mp4, так как это самый оптимальный вариант по качеству видеопотока и скорости конвертации.

После выполнения всех пунктов должны появиться два файла: первый файл с именем result.xxx, (xxx — формат выбранного видео), который, собственно, и считается дипфейком; и второй файл — result\_mask.xxx. Второй файл можно импортировать в программу для видеомонтажа и использовать полученный результат в качестве маски для дальнейшего улучшения замененного лица, не затрагивая остальную часть видео.

Необходимо отметить, что, с точки зрения цели психологического исследования, дипфейк не обязательно должен обладать высокими качественными характеристиками. Стимульной моделью может стать любой промежуточный результат цифрового коллажирования и его последующая модификация.

## II. Оценки видеоизображений «невозможного лица»

В серии экспериментов авторы статьи изучали закономерности восприятия подвижного «невозможного лица» и их отличия от восприятия статичных стимульных моделей. С помощью цифрового коллажирования были сконструированы видео-модели химерического и тэтчерезированного лица, которые в определенном порядке демонстрировались студентам московских вузов [4; 5; 6; 7; 8].

Подвижное химерическое лицо создавалось на основе объединения видеороликов с участием двух актрис, дающих интервью перед камерами. Исходные изображения приводились к единому формату, а на левую половину молодой актрисы (20 лет) накладывалась соответствующая часть поверхности лица актрисы более старшего возраста (35 лет).



Верхняя половина лица объединенного видеоизображения не имела выраженной границы, нижняя — включала небольшое рассогласование поверхностей — излом, играющий роль дистрактора (рис. 6).

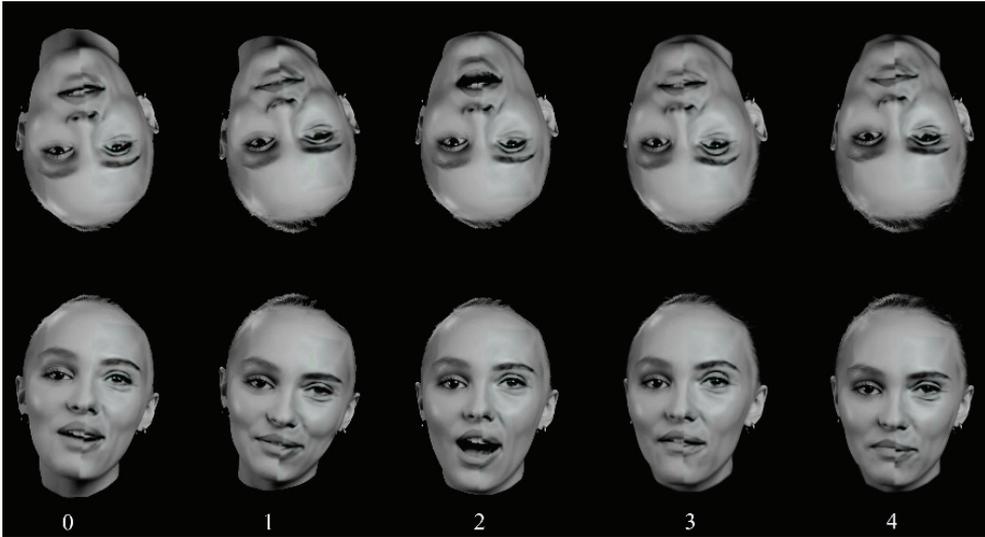


Рис. 6. Раскадровка фрагмента видеоизображения химерического лица; интервал между кадрами — 1 с: внизу — прямая экспозиция,верху — обратная

Подвижное тэтчерезированное лицо конструировалось путем переверота на видеоизображении глаз и рта молодой актрисы (рис. 7). Воспроизводилась иллюзия Маргарет Тэтчер, в соответствии с которой при прямой экспозиции коллажированное статичное лицо воспринимается неестественным, гротескным, переживающим отрицательные эмоции; при инверсии это же лицо выглядит приятным, испытывающим радость [16].

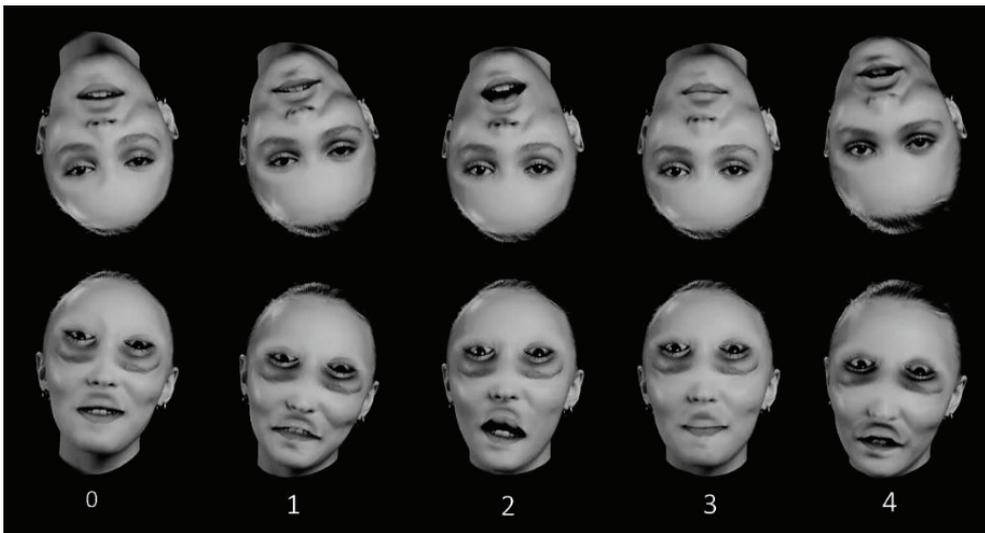


Рис. 7. Раскадровка фрагмента видеоизображения тэтчерезированного лица; интервал между кадрами — 1 с: внизу — прямая экспозиция,верху — обратная



Оценивались воспринимаемые качества стимульных моделей, состояние и характеристика личности виртуальных натурщиков. В ходе экспериментов варьировались статика и динамика, прямая и инвертированная экспозиция стимульных моделей, наличие и отсутствие дистракторов. Длительность каждого предъявления — 15 с. В зависимости от сочетания варьируемых переменных, отдельные экспозиции объединялись в пять относительно самостоятельных эпизодов, которые подвергались перекрестному анализу.

Выполненные исследования позволяют охарактеризовать ряд закономерностей восприятия видеоизображений «невозможного лица» в реальном процессе коммуникации. Согласно полученным данным, феномены восприятия «невозможного лица», зарегистрированные ранее в условиях статики (целостность противоречивых частей, эффекты дистракции и инверсии), при экспозиции динамических моделей сохраняются и приобретают новое содержание. Прямориентированные видеоизображения коллажированного лица по сравнению с фотоизображениями представляются более гармоничными и привлекательными. Инверсионный эффект — снижение чувствительности к деформациям лица — при его перевороте на 180° в динамике выражен сильнее. С введением мультимодальной экспозиции (интонаций звучащей речи) доля положительных оценок возрастает. Динамический дистрактор по-разному включается в процесс формирования образа натурщика, оказывая на него различное влияние. Двойственность восприятия искусственного персонажа снимается иерархизацией отношений сторон либо центра и периферии лицевой поверхности. В первом случае доминантной является правая сторона, во втором — треугольник «глаза—рот». В отличие от коллажированных изображений оригинальные лица в статике и движении независимо от эгоцентрической ориентации оцениваются позитивно на уровне высоких значений. Целостность восприятия раздвоенного лица на видеоизображении подтверждается убежденностью наблюдателей в реальном существовании экспонируемого человека, который адекватно ведет себя в понятной ситуации: привычно двигается, выражает эмоции, что-то рассказывает, учитывает присутствие наблюдателя, бросая на него взгляд и т. п. При всех тестируемых условиях пол виртуального натурщика определяется адекватно, воспринимаемый возраст переоценивается. Оценки эмоций виртуального натурщика по его видеоизображениям дифференцируются на основные (устойчивые) и дополнительные (меняющиеся) состояния, соотношение которых зависит от содержания конкретного эпизода.

### **Заключение**

Резюмируя сказанное, отметим, что Deepfake как технология синтеза изображений существенно расширяет возможности психологического исследования межличностного восприятия. Обнаруживается эффективный метод конструирования высококачественных видеоизображений динамических объектов и систем, включая «живое» лицо и реальные ситуации, в которых оказывается человек. Использование цифровых технологий упрощает создание стимульных моделей «невозможного лица», необходимых для углубленного изучения репрезентаций внутреннего мира человека, и формирует потребность в новых экспериментально-психологических процедурах, отвечающих более высокому уровню экологической и социальной валидности. Описанная методика позволяет учитывать множественность отношений, складывающихся в процессе идентификации личности, и может быть отнесена к системным инструментам изучения и практического использования закономерностей порождения и функционирования образа антропоморфного объекта, лицевая поверхность которого способна принимать разнообразный вид и по-разному включаться в процесс коммуникации.



Благодаря технологическому прогрессу и достижениям в области программирования и машинного обучения, процесс создания дипфейков с использованием полной либо частичной замены лица постоянно совершенствуется, становится проще и доступнее. Высокое качество и скрытость следов манипулирования изображением поднимают проблему его подлинности [10]. Замена лиц участников событий, отображенных на фото и видео, все чаще оказывается для пользователя не только инструментом профессиональной деятельности или развлечением, но и намеренной дискредитацией, а когда-то и шантажом известных политиков, звезд шоу-бизнеса и др. Можно полагать, что верификация содержания цифровых фото и видео как настоящей реальности имеет, наряду с техническими и программными, психологические критерии. Важной становится тема отношения воспринимающего человека к дипфейку, доверия либо недоверия альтернативной реальности.

### **Литература**

1. Барабанчиков В.А. Экспрессии лица и их восприятие. М.: ИП РАН, 2012. 341 с.
2. Барабанчиков В.А. (отв. ред.) Когнитивные механизмы невербальной коммуникации. М.: Когито-Центр, 2017. 359 с.
3. Барабанчиков В.А., Королькова О.А. Восприятие экспрессий «живого» лица // Экспериментальная психология. 2020. Том 13. № 3. С. 55–73. DOI:10.17759/exppsy.2020130305
4. Барабанчиков В.А., Маринова М.М. Воспринимаемые качества и оценка экспрессий подвижного химерического лица. Лицо человека в контекстах природы, технологий и культуры / Отв. ред. К.И. Ананьева, В.А. Барабанчиков, А.А. Демидов. М.: Московский институт психоанализа; Когито-Центр, 2020. 430 с. (В печати).
5. Барабанчиков В.А., Маринова М.М. Восприятие видеоизображений химерического лица // Познание и переживание. 2020. Том 1. № 1. С. 112–134. doi:10.51217/cogexp\_2020\_01\_01\_07
6. Барабанчиков В.А., Маринова М.М., Абрамов А.Д. Виртуальная личность подвижного тэтчеризированного лица // Психологическая наука и образование. 2021. Том 26. № 1. С. 5–18. DOI:10.17759/pse.2021000001
7. Барабанчиков В.А., Маринова М.М., Абрамов А.Д. Виртуальная личность тэтчеризированного лица в статике и динамике // Экспериментальная психология в социальных практиках / Отв. ред. В.А. Барабанчиков, В.В. Селиванов. М.: Универсум, 2020. С. 7–16.
8. Барабанчиков В. А., Маринова М. М., Абрамов А. Д. Иллюзия Маргарет Тэтчер на подвижном лице // Сборник материалов ежегодной конференции по когнитивной науке, посвященной памяти Дж.С. Брунера «Психология познания: низкоуровневые и высокоуровневые процессы» (г. Ярославль, 18–19 декабря 2020 г.). Ярославль: Ярославский государственный университет имени П.Г. Демидова, 2021. (В печати).
9. Anna. DeepFaceLab: инструмент для генерации дипфейков [Электронный ресурс] // Neurohive. URL: <https://neurohive.io/ru/frameworki/deepfacelab-instrument-dlya-generacii-dipejtkov/> (дата обращения: 30.11.2020).
10. Chawla R. Deepfakes: How a pervert shook the world // International Journal of Advance Research and Development. 2019. Vol. 4. № 6, P. 4–8.
11. Demuthova S., Démuth A. Handedness and the Preference of the Visual Field in Face Perception. // European Scientific Journal. 2018. DOI:10.19044/esj.2018.c3p8.
12. Dole M., Méary D., Pascalis O. Modifications of Visual Field Asymmetries for Face Categorization in Early Deaf Adults: A Study With Chimeric Faces // Front. Psychol. 2017. Vol. 8. P. 30. DOI:10.3389/fpsyg.2017.00030
13. Maras M., Alexandrou A. Determining Authenticity of Video Evidence in the Age of Artificial Intelligence and in the Wake of Deepfake Videos. // International Journal of Evidence and Proof. 2018. Vol. 23. P. 255–262. DOI:10.1177/1365712718807226.
14. Perov I., Gao D., Chervoniy N., Liu K., Marangonda S., Ume C., Mr. Dpfks., Facenheim C. F., RPL., Jiang J., Zhang S., Wu P., Bo zhou, Zhang W. DeepFaceLab: A simple, flexible and extensible face swapping framework



[Электронный ресурс]. 2020. // ArXiv. URL: <https://arxiv.org/pdf/2005.05535.pdf> (дата обращения: 13.10.2020).

15. *Stephanie S. A. H. Blom, Henk Aarts & Gün R. Semin*. Lateralization of facial emotion processing and facial mimicry. // *Laterality*. 2020. Vol. 25 № 3. P. 259–274. <http://doi.org/10.1080/1357650X.2019.1657127>

16. *Thompson P.* Margaret Thatcher: A new illusion // *Perception*. 1980. Vol. 9. P. 483–484. <https://doi.org/10.1068/p090483>

17. *Weibert K., Müller V., Sängler J.* Familiarity abolishes right-hemispheric bias in face perception // *Journal of Vision*. 2017. Vol. 17. DOI: 10.1167/17.10.998

18. *Westerlund M.* The Emergence of Deepfake Technology: A Review // *Technology Innovation Management Review*. 2019. Vol. 9, № 11, P. 39–52. <http://doi.org/10.22215/timreview/12>

19. *Williams L.R., Grealy M.A., Kelly S.W., Henderson I., Butler S.H.* Perceptual bias, more than age, impacts on eye movements during face processing // *Acta Psychologica*. 2016. Vol. 164. P. 127–135. DOI:10.1016/j.actpsy.2015.12.012

## References

1. Barabanshchikov V.A. *Ekspressii litsa i ikh vospriyatie* [Facial expressions and their perception]. Moscow: IP RAN Publ., 2012. (In Russ.).

2. Barabanshchikov V.A. (otv. red.) *Kognitivnye mekhanizmy neverbal'noi kommunikatsii*. M.: Kogito-Tsentr, 2017. (In Russ.).

3. Barabanshchikov V.A., Korol'kova O.A. *Vospriyatie ekspressii «zhivogo» litsa* [Perception of “live” facial expressions]. *Ekspperimental'naya psikhologiya = Experimental Psychology*, 2020, Vol. 8, no. 3. pp. 55–73. DOI:10.17759/exppsy.2020130305 (In Russ.).

4. Barabanshchikov V. A., Marinova M. M. *Vosprinimaemye kachestva i otsenka ekspressii podvizhnogo khimericheskogo litsa* [Perceived qualities and evaluation of expressions of the mobile chimeric face]. *Litso cheloveka v kontekstakh prirody, tekhnologii i kul'tury* [The Face of Man in Contexts of Nature, Technology, and Culture] K.I. Anan'eva, V.A. Barabanshchikov, A. A. Demidov (eds.). Moscow: Moskovskii institut psikhoanaliza, Kogito-Tsentr, 2020. (In Press). (In Russ.).

5. Barabanshchikov V.A., Marinova M.M. *Vospriyatie videoizobrazhenii khimericheskogo litsa* [Perception of video images of a chimerical face]. *Poznanie i perezhivanie = Cognition and Experience*, 2020, Vol. 1, no. 1, pp. 112–134. doi:10.51217/cogexp\_2020\_01\_01\_07 (In Russ.).

6. Barabanshchikov V.A., Marinova M.M., Abramov A.D. The Virtual Personality of the Moving Thatcherized Person. *Psikhologicheskaya nauka i obrazovanie = Psychological Science and Education*. Vol. 23, no. 1, pp. 5–18. DOI:10.17759/pse.2021000001 (In Russ.).

7. Barabanshchikov V.A., Marinova M.M., Abramov A.D. *Virtual'naya lichnost' tetcherizirovannogo litsa v statike i dinamike* [Virtual Personality of the Thatcherized Person in Statics and Dynamics] // *Ekspperimental'naya psikhologiya v sotsial'nykh praktikakh* [Experimental Psychology in Social Practices] V.A. Barabanshchikov, V.V. Selivanov (eds). Moscow: Universum, 2020. pp. 7–16. (In Russ.).

8. Barabanshchikov V.A., Marinova M.M., Abramov A.D. *Illyuziya Margaret Tetcher na podvizhnom litse* [The illusion of Margaret Thatcher on a moving face]. *Sbornik materialov ezhegodnoi konferentsii po kognitivnoi nauke, posvyashchennoi pamyati Dzh. S. Brunera*, «Psikhologiya poznaniya: nizkourovnevye i vysokourovnevye protsessy» (g. Yaroslavl', 18–19 dekabrya 2020 g.) [Proceedings of the annual conference on cognitive science in memory of J. S. Bruner “Psychology of cognition: low-level and high-level processes”]. Yaroslavl: Yaroslavl State University. P.G. Demidov, 2021. (In press).

9. Anna. *DeepFaceLab: instrument dlya generatsii dipfeikov* [Electronic resource] [DeepFaceLab: deepfake generation tool]. Neurohive. URL: <https://neurohive.io/ru/frameworki/deepfacelab-instrument-dlya-generatsii-dipfejkov/> (Accessed: 30.11.2020) (In Russ.).

10. Chawla R. Deepfakes: How a pervert shook the world. *International Journal of Advance Research and Development*, 2019. Vol. 4, no. 6, pp. 4–8.

11. Demuthova S., Démuth A. Handedness and the Preference of the Visual Field in Face Perception // *European Scientific Journal*. 2018. <http://doi.org/10.19044/esj.2018.c3p8>

12. Dole M., Méary D., Pascalis O. Modifications of Visual Field Asymmetries for Face Categorization in Early Deaf Adults: A Study With Chimeric Faces. *Front. Psychol*. 2017. Vol. 8, P. 30. <http://doi.org/10.3389/fpsyg.2017.00030>



13. Maras M., Alexandrou A. Determining Authenticity of Video Evidence in the Age of Artificial Intelligence and in the Wake of Deepfake Videos. *International Journal of Evidence and Proof*. 2018, Vol. 23, pp. 255–262. DOI:10.1177/1365712718807226
14. Perov I., Gao D., Chervoniy N., Liu K., Marangonda S., Ume C., Mr. Dpfks., Facenheim C. F., RP L., Jiang J., Zhang S., Wu P., Bo zhou, Zhang W. DeepFaceLab: A simple, flexible and extensible face swapping framework [Electronic resource]. 2020. ArXiv: Computer Vision and Pattern Recognition. URL: <https://arxiv.org/pdf/2005.05535.pdf> (Accessed: 13.10.2020).
15. Stephanie S.A.H. Blom, Henk Aarts & Gün R. Semin. Lateralization of facial emotion processing and facial mimicry. *Laterality*, 2020. Vol. 25, no. 3, pp. 259–274. DOI:10.1080/1357650X.2019.1657127
16. Thompson P. Margaret Thatcher: A new illusion. *Perception*, 1980. V. 9, P. 483–484. DOI:10.1068/p090483
17. Weibert K., Müller V., Sängler J. Familiarity abolishes right-hemispheric bias in face perception. *Journal of Vision*, 2017. vol. 17. DOI: 10.1167/17.10.998.
18. Westerlund M. The Emergence of Deepfake Technology: A Review // *Technology Innovation Management Review*. 2019. Vol. 9, no. 11, pp. 39–52. <http://doi.org/10.22215/timreview/12>
19. Williams L.R., Grealy M.A., Kelly S.W., Henderson I., Butler S.H. Perceptual bias, more than age, impacts on eye movements during face processing. *Acta Psychologica*, 2016. Vol. 164, pp. 127–135. DOI:10.1016/j.actpsy.2015.12.012

#### **Информация об авторах**

*Барабаншиков Владимир Александрович*, доктор психологических наук, профессор, член-корреспондент РАО, директор Института экспериментальной психологии, Московский государственный психолого-педагогический университет (ФГБОУ ВО МГППУ); декан факультета психологии, Московский институт психоанализа (НОЧУ ВО МИП), г. Москва, Российская Федерация, ORCID: <https://orcid.org/0000-0002-5084-0513>, e-mail: [vladimir.barabanshikov@gmail.com](mailto:vladimir.barabanshikov@gmail.com)

*Маринова Мария Михайловна*, помощник директора Института экспериментальной психологии, специалист по УМР, Московский государственный психолого-педагогический университет (ФГБОУ ВО МГППУ); аспирант, Московский институт психоанализа (НОЧУ ВО МИП), г. Москва, Российская Федерация, ORCID: <https://orcid.org/0000-0002-8862-4007>, e-mail: [marinovamm@yandex.ru](mailto:marinovamm@yandex.ru)

#### **Information about the authors**

*Vladimir A. Barabanshikov*, Dr. Sci. in Psychology, Professor, Corresponding Member of Russian Academy of Education, Director, Institute of Experimental Psychology, Moscow State University of Psychology & Education, Moscow, Russia, ORCID: <https://orcid.org/0000-0002-5084-0513>, e-mail: [vladimir.barabanshikov@gmail.com](mailto:vladimir.barabanshikov@gmail.com)

*Maria M. Marinova*, Assistant Director of the Institute of Experimental Psychology, Academic Services Specialist, Moscow State University of Psychology & Education, Moscow, Russia, ORCID: <https://orcid.org/0000-0002-8862-4007>, e-mail: [marinovamm@yandex.ru](mailto:marinovamm@yandex.ru)

Получена 01.11.2020

Received 01.11.2020

Принята в печать 21.12.2020

Accepted 21.12.2020