

doi: 10.17586/2226-1494-2022-22-3-547-558

УДК 004.932.2

Метод генерации масок на изображениях лиц и системы их распознавания

Георгий Александрович Кухарев¹, Елена Витальевна Рюмина^{2✉},
 Никита Александрович Шульгин³

^{1,3} Санкт-Петербургский государственный электротехнический университет «ЛЭТИ» им. В.И. Ульянова (Ленина), Санкт-Петербург, 197376, Российская Федерация

² Санкт-Петербургский Федеральный исследовательский центр Российской академии наук, Санкт-Петербург, 199178, Российская Федерация

¹ gakukharev@etu.ru, <https://orcid.org/0000-0003-2188-2172>

² ryumina_ev@mail.ru✉, <https://orcid.org/0000-0002-4135-6949>

³ nikitashulgin@gmail.com, <https://orcid.org/0000-0002-1884-5832>

Аннотация

Предмет исследования. Исследована проблема распознавания лиц в масках. Показано, что разнообразные по форме, текстуре и цвету реальные маски создают проблемы для современных систем распознавания изображений лиц. Одна из таких причин — отсутствие необходимых реальных наборов обучающих данных. Создание новых данных на базе простых методов формирования масок на изображениях лиц может решить эту проблему. **Метод.** Предложен оригинальный метод, включающий генерацию различных типов, форм и цветов масок непосредственно на оригинальной текстуре изображений лиц. При этом учитывалось формирование масок на лицах отдельных людей, лицах в групповых фото и в сценах с потоками людей. **Основные результаты.** На основе 100 оригинальных изображений лиц из базы «CUHK Face Sketch Database» создана тестовая база, которая включает в себя более 20 000 изображений лиц с масками, доступных для использования. Выполнены эксперименты по распознаванию лиц тестовой базы в рамках реализованных четырех систем: трех современных на базе «глубокого обучения» и одной детерминированной — на базе косинус-преобразования. Оценена результативность этих систем, интерпретированы полученные результаты распознавания лиц в масках и отмечены маски, которые были проблемой для выбранных систем. **Практическая значимость.** Предложенный метод генерации масок может быть использован для создания баз данных и тестовых баз изображений с масками. Полученные результаты будут полезны исследователям и специалистам в области обработки и анализа изображений.

Ключевые слова

генерация масок, распознавание изображений лиц, антропометрия лица, нейросети, экстрактор признаков, метрика косинусного подобия, критерий минимального расстояния

Благодарности

Работа выполнена при поддержке проекта фонда РФФИ № 20-04-60529-вирусы, а также частично в рамках бюджетной темы № FFZF-2022-0005.

Ссылка для цитирования: Кухарев Г.А., Рюмина Е.В., Шульгин Н.А. Метод генерации масок на изображениях лиц и системы их распознавания // Научно-технический вестник информационных технологий, механики и оптики. 2022. Т. 22, № 3. С. 547–558. doi: 10.17586/2226-1494-2022-22-3-547-558

Method for generating masks on face images and systems for their recognition

Georgy A. Kukharev¹, Elena V. Ryumina^{2✉}, Nikita A. Shulgin³

^{1,3} Saint Petersburg State Electrotechnical University “LETI”, Saint Petersburg, 197376, Russian Federation

² Saint Petersburg Federal Research Center of the Russian Academy of Sciences (SPC RAS), Saint Petersburg, 199178, Russian Federation

¹ gakukharev@etu.ru; <https://orcid.org/0000-0003-2188-2172>

² ryumina_ev@mail.ru✉, <https://orcid.org/0000-0002-4135-6949>

³ nikitashulgin@gmail.com, <https://orcid.org/0000-0002-1884-5832>

© Кухарев Г.А., Рюмина Е.В., Шульгин Н.А., 2022

Abstract

The problem of masked face recognition is investigated. It is shown that real masks of various shapes, textures and colors have become a problem for state-of-the-art face recognition systems. A reason for this is the lack of the necessary real training datasets. Creation of new data based on simple methods of forming masks on face images could solve this problem. An original method is proposed including the generation of various types, shapes, and colors of masks directly on the original texture of face images. The formation of the masks on the faces of individuals, on faces in group photos, and in scenes with streams of people was taken into account. Based on 100 original face images from the CUHK Face Sketch Database, a test database was created that includes more than 20,000 masked faces images which available for use. Experiments were carried out to recognize faces from the test database within the implemented four systems, among which three are state-of-the-art systems based on “deep learning” and one is deterministic system based on the cosine-transform. The performance of these systems was evaluated, the obtained results of masked face recognition were interpreted, and the masks that were a problem for selected four systems were noted. The proposed mask generation method can be used to create corpora and test databases of images with masks. The obtained results will be useful to researchers and specialists in the field of image processing and analysis.

Keywords

masks generation, face images recognition, facial anthropometry, neural networks, feature extractor, cosine similarity metric, minimum distance criterion

Acknowledgements

The work was supported by the RFBR project No. 20-04-60529-viruses, and also partially under the budget topic No. FFZF-2022-0005.

For citation: Kukharev G.A., Ryumina E.V., Shulgin N.A. Method for generating masks on face images and systems for their recognition. *Scientific and Technical Journal of Information Technologies, Mechanics and Optics*, 2022, vol. 22, no. 3, pp. 547–558 (in Russian). doi: 10.17586/2226-1494-2022-22-3-547-558

Введение

Лица людей в медицинских (гигиенических) масках с 2019 года стали реальностью нашей повседневной жизни, вошедшей в эпоху «COVID» пандемии. При этом контроль ношения масок остается необходимой процедурой в общественных местах.

Задача контроля наличия медицинской (гигиенической) маски на лицах стала первым серьезным вызовом и проблемой для ее решения. Еще задолго до «COVID» пандемии на эту проблему обратили внимание в Китайской Академии Наук [1], поскольку маски в Китае использовали для защиты людей от дыма и гари атмосферного воздуха. И, если для медицинских (гигиенических – белых и светлых) масок на лицах анфас эта задача решалась относительно легко, то для замаскированных лиц в реальных ситуациях и условиях («Masked Faces in the Wild» [1]) поиск решений достигался очень трудно. Эти решения опирались на сбор и создание специальных тестовых баз изображений лиц в масках и проведение множества исследований, учитывающих специфику и характеристики таких изображений [2]. Кроме задачи контроля наличия защитной маски на лицах, существует задача распознавания изображений лиц (РИЛ) в масках. Решению задач РИЛ в последнее время также уделяется большое внимание [3].

Отметим, что стандартные гигиенические светлого-голубые медицинские маски — не единственное средство индивидуальной защиты. Для целей защиты люди активно используют и тканевые маски. Тканевые маски могут быть разной текстуры, формы и цветовой гаммы (однотоновой или многотоновой), с QR-кодами и надписями, а также с изображениями морд различных животных и голов птиц. Использование таких масок стало проблемой для современных систем наблюдения и систем контроля доступа (например, в метро, аэропортах и других объектах массового нахождения

людей), построенных на базе методов машинного обучения и, в том числе методов глубокого обучения.

Для создания надежных систем контроля наличия защитных масок и РИЛ в них, были собраны новые наборы изображений лиц в масках и базы с искусственными масками [4]. Например, на платформах MaskTheFace¹ и AIZOOTech/FaceMaskDetection² «маски выбирались» из специального репозитория масок и буквально «пристраивались» к нижней области лица — от перiorбитальной области до конца подбородка.

В отличие от изображений лиц в медицинских гигиенических масках, лица с реальными цветными текстурными масками практически не распознавались в ранних системах глубокого обучения. При этом хуже всего распознавались лица с масками черного, синего-черного-серого, темно-серого и красного цветов, маски с изображениями животных и птиц, закрывающие всю нижнюю часть лица. Об этом свидетельствует отчет NIST [5]. Кроме этих типов масок, также не всегда распознавались лица в прозрачных защитных масках и щитах на перiorбитальной области, что возникало из-за их отражающей засветки приемной фотоаппаратуры.

Цель и содержание работы

Разнообразие вариантов используемых на практике масок, привели к тому, что в вопросах РИЛ в масках все еще нет единого мнения, решений и рекомендаций. Исходя из этого, в настоящей работе будут рассмотрены маски только на лицах анфас и в сценариях, используемых в системах контроля доступа (доступ к финан-

¹ MaskTheFace [Электронный ресурс]. Режим доступа: <https://sites.google.com/view/masktheface/home/> (дата обращения: 20.03.2022).

² AIZOOTech/FaceMaskDetection [Электронный ресурс]. Режим доступа: <https://github.com/AIZOOTech/FaceMaskDetection/actions/> (дата обращения: 20.03.2022).

совым ресурсам, закрытым лабораториям, контроль пассажиропотока в аэропортах и вокзалах и, наконец, контроль потока туристов и беженцев на пограничных пунктах [6]). Люди, участвующие в этих сценариях, хотят, чтобы «система их узнала» и, поэтому положение лиц анфас является для этих людей не только мотивированным, но и необходимым для формально правильного интерактива с системой контроля доступа.

Предложен метод генерации различных форм, типов и цветов масок и метод их размещения/формирования на лицах, который реализуется непосредственно на оригинальной текстуре изображений лиц. Такой способ реализации создает высокую реалистичность положения и органичность нахождения масок на лицах. При этом маски легко формируются как на отдельных портретах людей и выделенных кадров видео, так и на групповых фото и лицах в сценах с потоками людей.

Корректное размещение/формирование масок на лицах основано на ключевых точках лица, которые вычисляются в онлайн-режиме, и используемых мнемонических моделей масок. Предложенный метод формирования масок не зависит от размеров изображений-оригиналов, фенотипа лиц и фона изображения, а также от точного положения лиц анфас. Метод может применяться с видимыми отклонениями от положения лиц анфас. Эти отклонения могут выражаться в плоском и пространственном поворотах лиц, которые обычно присутствуют на групповых снимках и сценах с потоками людей. Применение метода возможно со многими известными базами изображений лиц.

Отметим, что метод не требует использования специальных репозиторий масок и применяет изображения 6 форм масок, 10 оригинальных цветных тканевых текстур и 15 текстур типовых масок (15 штампов масок). Текстура генерируемых масок образуется путем замены или перестановки компонент (R, G и B) цветной оригинальной тканевой текстуры, другими компонентами R, G и B — из штампов масок, из другой цветной тканевой текстуры.

В работе представлены варианты сгенерированных таким образом масок на лицах, которые практически не отличаются от реальных масок (REAL MASKS) и значительно превосходят поддельные маски (например, «печатные маски» [5] или SIMULATED MASKS [7]). В качестве исходных изображений-оригиналов использованы, как пример, 100 изображений лиц из базы «CUHK Face Sketch Database» (CUFS)¹. На их основе сгенерировано более 20 000 изображений лиц в масках различных типов, форм, текстуры и цвета². Выполнены эксперименты по распознаванию лиц из базы CUFS в рамках четырех систем, среди которых три — современные на базе глубокого обучения и одна — детерминированная на базе косинус-преобразования. Оценена результативность выбранных систем, интерпретирова-

ны полученные результаты распознавания лиц в масках и отмечены маски, которые были проблемой для всех систем.

Последние достижения по распознаванию изображений лиц в масках на базе глубокого обучения

Сравнительный анализ новых решений проблемы РИЛ в масках был рассмотрен на конференции по компьютерному зрению (The International Conference on Computer Vision — 2021 [8, 9]). Организаторы конференции получили сотни заявок на участие, что подтвердило актуальность и необходимость поиска решений проблемы РИЛ в масках. Представим краткие результаты рассмотренного анализа и наилучших решений проблемы РИЛ.

Отметим, что при решении проблемы РИЛ без масок сложилась четкая тенденция использования методов глубокого обучения [10], которые также применяются для РИЛ в масках [11]. Решение проблемы РИЛ в/без масок состоит из двух основных компонентов. Первым компонентом является детектор лиц, их сравнение представлено в работе [2]. Исследование показало, что RetinaFace [12] лучше остальных работает на изображениях лиц в масках в реальных ситуациях и условиях. Вторым компонентом — экстрактор признаков и соответствующая для него предварительная обработка (например, выравнивание лиц, нормализация изображений и т. д.). Из [11] прослеживаются два наиболее популярных экстрактора: на основе семейства ResNet [13] и VGG-16 [14]. Эти и подобные экстракторы признаков предварительно обучаются на крупных наборах данных изображений лиц и предоставляются в открытый доступ, например, в библиотеках DeepFace [15], Keras-VGGFace³ и др. Учитывая вышесказанное в настоящей работе в качестве нейросетевых систем использованы экстракторы признаков: ArgFace, ResNet и VGG-16 [12–14].

Метод генерации и размещения/формирования масок на изображениях лиц людей

На первом (предварительном) этапе из исходной сцены детектируется и выделяется лицо человека, а далее на нем определяются координаты антропометрических точек лица. Определение координат основано на алгоритмах, рассмотренных в работах [16, 17]. Программная реализация алгоритмов представлена, например, в библиотеках OpenCV⁴ и Dlib⁵. Пример полученных 68-и антропометрических точек показан на рис. 1, а. Здесь и далее используются изображения лиц из базы CUFS. Из исходного набора в 68 координат

¹ CUHK Face Sketch Database (CUFS) [Электронный ресурс]. <http://mmlab.ie.cuhk.edu.hk/archive/facesketch.html> (дата обращения: 20.03.2022).

² База изображений фронтальных лиц в масках [Электронный ресурс]. <https://hci.nw.ru/ru/pages/masked-frontal-faces> (дата обращения: 13.04.2022).

³ Keras-vggface [Электронный ресурс]. Режим доступа: <https://github.com/rcmalli/keras-vggface> (дата обращения: 20.03.2022).

⁴ OpenCV [Электронный ресурс]. Режим доступа: <https://opencv.org/> (дата обращения: 20.03.2022).

⁵ Dlib [Электронный ресурс]. Режим доступа: http://dlib.net/face_landmark_detection.py.html (дата обращения: 20.03.2022).

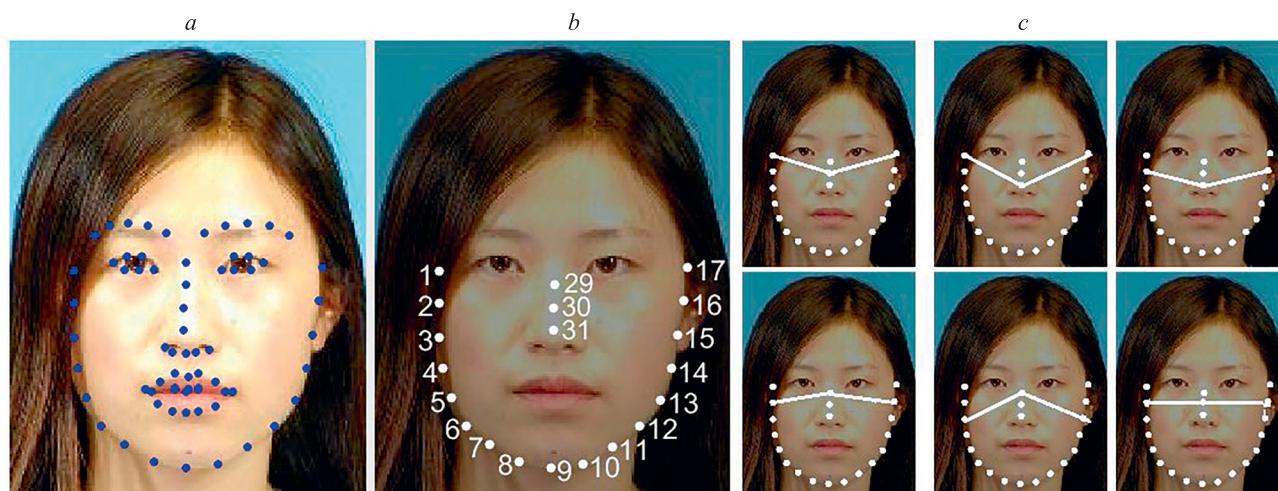


Рис. 1. Идеи метода генерации масок на области лица: исходное изображение лиц с ключевыми точками (a); 20 ключевых точек, определяющих границы контура масок (b); шесть контуров масок, полученных по выбранным ключевым точкам (c)

Fig. 1. Ideas for the mechanism of the masks appearance in the face area: initial face images with keypoints (a); 20 keypoints defining the border of the mask contour (b); six contours of masks coming from selected keypoints (c)

антропометрических точек использовано только 20 координат: три из них на линии симметрии носа (с номерами 29–31) и 17 — с номерами 1–17, охватывающими овал лица. 20 ключевых точек показаны на рис. 1, b.

Верхняя часть контура каждой формируемой маски определяется выбранной координатой на линии симметрии носа и двумя координатами с левой и правой стороны овала лица. Оставшиеся координаты овала лица определяют нижнюю часть контура формируемой маски.

Шесть вариантов контуров масок показаны на рис. 1, c. Каждый контур масок можно mnemonicски описать следующими цепочками точек: {29, 3–15, 29}; {29, 2–16, 29}; {30, 2–16, 30}; {30, 1–17, 30}; {31, 1–17, 31}; {30, 2–16, 30}. При этом запись типа «3–15» включает все промежуточные точки овала лица — от начальной (3) до конечной (15). Используя такую mnemonicскую модель, можно легко отобразить контур формируемой маски в прямоугольной «матричной системе координатных осей», в которой: начало координат находится в левом верхнем углу; вертикальная ось

нумеруется сверху вниз; а горизонтальная ось — слева направо. Такая система координатных осей соответствует столбцам и строкам пикселей на лицах анфас, что позволяет просто и точно сформировать на нем не только контур маски, но и саму маску.

Процесс формирования маски на лице представлен на рис. 2 и включает следующие этапы:

- определение границ поля маски по четырем ключевым точкам: первая — на горбинке носа по линии симметрии лица, вторая — на нижней точке подбородка и две точки на охвате овала по линии козелка уха (рис. 2, a);
- замещение поля маски автоматически сгенерированной текстурой маски с отображением на ней ребра и тени (возникающей при боковом освещении лица), а также размещение ключевых точек исходной mnemonicской модели для проверки возможности размещения формируемой маски (рис. 2, b);
- прорисовка контура маски, для отображения областей лица, выходящих за внешние границы маски, и укрытых текстурой (рис. 2, c);

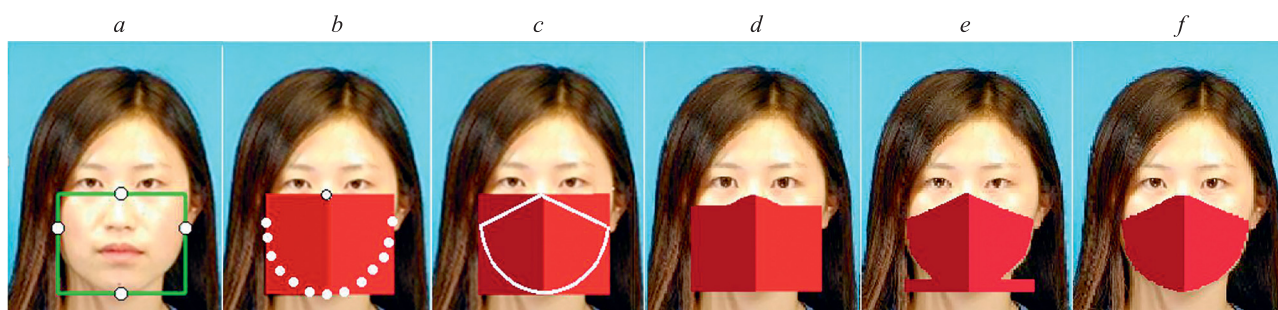


Рис. 2. Процесс формирования маски: определение границ поля для маски (a); замещение области поля текстурой маски, отображение ребра и теней маски, а также размещение точек контура маски (b); прорисовка контура формируемой маски (c); восстановление пикселей изображений лиц, укрытых текстурой (d–f)

Fig. 2. The process of mask formation: determination of the field boundaries for the mask (a); replacing the field area with the mask texture, displaying the edge and shadows of the mask as well as placing the points of the mask outline (b); drawing the contour of the mask being formed (c); restoration of face images pixels covered with a texture (d–f)

— восстановление тех пикселей изображений лиц, которые укрыты текстурой, выходящей за внешние границы контура маски. При этом само восстановление пикселей реализуется как замена значений пикселей текстуры на значения пикселей исходного изображений лиц (рис. 2, *d-f*).

В предлагаемом решении расчет координат пикселей выполняется по методу определения положения точки относительно многоугольника. В основе метода лежит алгоритм 112 «Position of point relative to polygon» [18]. После его проверки и уточнения он был представлен в отечественных сборниках алгоритмов в языках АЛГОЛ [19] и Бейсик [20]. Соответствующая процедура на языке пакета MATLAB, реализующая указанный алгоритм, представлена в работе [21].

Рассмотрим случай, когда в поле границ маски (рис. 2, *a*) расположен многоугольник — контур маски (рис. 2, *c*). В этом случае используются пиксели, находящиеся в поле границ маски, но не входящие в ее контур. Исходные данные для расчета координат интересующих пикселей — координаты контура маски, представленные в форме мнемонической модели. Весь набор просматриваемых координат пикселей определяется габаритными размерами поля маски. Просматривая пиксели в прямоугольной сетке (от левого верхнего угла до нижнего правого угла поля маски), найдем все координаты пикселей, удовлетворяющие условию решаемой задачи. Например, поле маски в представленном случае на рис. 1, *a* имеет размер 89×114 пикселей, при размере исходного изображения 250×200 . А число восстанавливаемых пикселей составляет 3307 из просмотренных 10 146.

Перейдем к этапу восстановления пикселей. Число замен пикселей при этом будет равно 3307. Процесс восстановления (по факту замены) пикселей лица представлен в последовательности скриншотов на рис. 2, *d-f*.

В рамках предлагаемого метода формирования масок можно изменять текстуру и цвета масок, фон портрета и кожи лица, а также форму масок (в соответствии с мнемонической моделью), тип масок (непрозрачные, прозрачные, с закрытой областью глаз прозрачным щитком) и положение масок в нижней области лица. Изображения лиц в маске могут быть представлены не только в исходном формате (например, RGB), но и в любых других пяти форматах (RBG, GRB, GBR, BRG, BGR), основанных на перестановках компонент R, G и B. Это позволяет в 6 раз расширить цветовую гамму отдельно изображений лиц и масок, а также изображения лиц в масках при одном и том же наборе исходных изображений без масок — т. е. эталонов.

На рис. 3 представлены некоторые варианты типа масок, их формы, текстур и цвета, а также варианты различной цветовой гаммы кожи лиц и фона. Более полный набор масок представлен по адресу¹.

Отметим также, что при небольших *плоских* и *пространственных поворотах* лица в рамках матричной

системы координат, контуры масок и сами маски формируются без искажений (рис. 4).

На рис. 5 представлены изображения лиц при плоских и пространственных поворотах в масках (SIMULATED MASKS) из баз данных [7, 22].

При сравнении изображений лиц в масках (рис. 4–5), видно, что предложенный в настоящей работе метод генерации масок корректно и строго по овалу лица размещает маски в отличие от других state-of-the-art методов.

В предлагаемую базу включены изображения лиц в масках с размытой или сглаженной текстурой лица, а также с наложенным на них шумом. Полученные изображения имитируют реальные условия получения исходных данных в условиях удаленности от камеры, дождя и/или снега в рамках систем уличного наблюдения. Фактически, получены изображения лиц в масках при полностью искаженной области лица. Отметим, что в работе представляет интерес только область лица, не укрытая маской, но она также не соответствует аналогичной области на лице эталона из-за изменения оригинальных значений пикселей при сглаживании и зашумлении открытых областей лица. Заметим, что в существующих базах данных изображений лиц в масках вряд ли можно встретить такие наборы обучающих и тестовых данных.

Детерминированная система, использованная в эксперименте

Выше было отмечено, что современные системы РИЛ реализуются на методах глубокого обучения. Однако для РИЛ применимы и простые системы на основе детерминированных методов. Класс простых систем (Simple Face Recognition Systems, Simple FaReS) представлен в монографии [23]. В настоящей работе модель детерминированной системы имеет следующий вид:

$$\text{GUFS}(100/1/\text{NUM})\{\text{2DDCT}: M \times N \rightarrow P \times P/\text{zigzag}\} \times \times [\text{KMP}/L1/\text{rank}=1], \quad (1)$$

где $\text{GUFS}(100/1/\text{NUM})$ — база изображений лиц GUFS , включающая 100 эталонов (по 1 эталону на класс) и NUM тестовых изображений лиц в масках (при этом NUM кратно 100); $\text{2DDCT}: M \times N \rightarrow P \times P/\text{zigzag}$ — блок обработки изображений лиц: изображение размером $M \times N$ подвергается двумерному косинус-преобразованию, с выбором признаков методом «zigzag» по P диагоналям спектральной матрицы размером $P \times P$ и общим числом признаков, равным $(P(P+1)/2)-1$; $[\text{KMP}/L1/\text{rank}=1]$ — классификатор: реализует процедуру классификации по критерию минимального расстояния, с метрикой $L1$ и рангом 1.

Эта простая для реализации система показала самую высокую результативность РИЛ в масках по сравнению с другими Simple FaReS [22]. Результат контрольной проверки распознавания 14×100 изображений лиц в масках при 100 эталонах представлен на рис. 6. В соответствии с формулой (1), NUM состоит из 1400 изображений, $P \times P \geq 15 \times 15$, а число признаков для $P = 15$ составило 119.

¹ База изображений фронтальных лиц в масках [Электронный ресурс]. <https://hci.nw.ru/ru/pages/masked-frontal-faces> (дата обращения: 13.04.2022).



Рис. 3. Примеры типа масок, их формы, текстур и цвета, а также варианты изменения цветовой гаммы кожи лиц и фона изображения

Fig. 3. Examples of textures, colors, shapes, and types of masks as well as options for changing the color scheme of the skin of faces and the background of the image



Екатерина Березовская – ведущая новостей первого канала: https://www.instagram.com/katya_berezovskaya/

Рис. 4. Результат формирования масок при плоских и пространственных поворотах лиц, в том числе и на лицах из видеофайлов: линии ребра на масках отображают степень отклонения лиц от позиции анфас

Fig. 4. The result of the formation of masks for flat and spatial rotations of faces including faces from video files: edge lines on the masks show the degree of deviation of faces from the full-face position



Рис. 5. Изображения лиц при плоских и пространственных поворотах в масках из баз данных [7, 22]

Fig. 5. Masked faces images for flat and spatial rotations of faces from databases [7, 22]

Эталон	Разная текстура масок						Разные формы и моноцвет масок			Шум 40 %	Два метода зашумления	
	1	2	3	4	5	6	7a	7b	8	9	13	14
100 ИЛ												
2DDCT	100 % Все 1400 ИЛ в масках были распознаны правильно и соотнесены к своим классам											

Рис. 6. Результат контрольной проверки распознавания 100×14 изображений лиц в масках

Fig. 6. The result of the reference check of the recognition of 1400 images of masked faces

Нейросетевые системы, использованные в эксперименте

В нейросетевых системах РИЛ в масках использован детектор лиц RetinaFace. Эксперименты проведены с помощью трех экстракторов признаков ArgFace (на базе ResNet-34, представленном в библиотеке DeepFace), ResNet-50 и VGG-16 из библиотеки Keras-

VGGFace). На вход экстракторы принимают изображения лиц, приведенные к квадратному разрежению. Для сохранения пропорций лица недостающие пиксели заполняются нулями. Кроме того, для каждого экстрактора признаков требуется своя предварительная обработка. Так, для ArgFace выполняется линейная нормализация пикселей изображений лиц и приведение его к разрешению 128×128 . Для ResNet-50 и

VGG-16 проводится канальная нормализация пикселей (центрирование) изображений лиц и получение разрешения 224×224 . В качестве классификатора использован критерий минимального расстояния с метрикой косинусного подобия.

Экспериментальные результаты

В таблице приведены экспериментальные результаты, полученные с помощью предложенных систем РИЛ в масках, где названия нейросетевых систем соответствуют применяемым в них экстракторам признаков. В качестве показателя результативности РИЛ в масках выполнена оценка верно идентифицированных относительно общего числа использованных тестовых изображений лиц. Для эксперимента из более 20 000 сгенерированных изображений лиц в масках было отобрано 16 000.

Выбор данных масок обусловлен фактами, отмеченными исследователями при решении задач РИЛ для некоторых типов, форм и цвета масок [5], включая проблему зашумленности изображений лиц, удаленности лица от камеры, и использования популяций лиц в масках.

Сгенерированные изображения лиц можно разделить на группы следующим образом:

- 1) монохромные маски белого, красного, серого, черного, а также черно-синего цветов, медицинские и цветные текстурные;
- 2) лица в масках (п. 1) с наложением шума 40 % от максимальной яркости изображений [23];

- 3) лица в масках (п. 1) с имитацией удаленности от камеры [23];
- 4) лица в масках (п. 1) с комбинацией изменений по пп. 2 и 3;
- 5) различные преобразования изображений лиц в масках (де-идентификация, сжатие, сдвиг и растяжение с различной степенью [23]);
- 6) лица в масках различных цветов, текстур, в том числе с мордами животных;
- 7) лица в масках (п. 5) с перестановкой компонент (R, G, B) — всего 6 групп.

Результаты экспериментов показали, что детерминированная система Simple FaReS значительно превосходит по результатам нейросетевые системы. Незначительный спад точности наблюдается при комбинировании двух трансформаций (шум и имитация удаленности от камеры) изображений лиц в масках. Более существенный спад наблюдается при применении де-идентификации к изображениям лиц в масках: точность РИЛ составила 84 %, что на 10 % меньше, чем точность нейросетевой системы ArgFace. Эффективность детерминированной системы РИЛ в масках особенно явно видна на искаженных изображениях лиц (трансформации номер 8–10), что обусловлено использованием пространства признаков на основе косинус-преобразования, которое не так значительно (в отличие от яркостных признаков) реагирует на искажения и шумовые изменения текстуры изображений лиц. Кроме того, в пространстве признаков косинус-преобразования мера подобия изображений лиц между признаками оригинала и лица с маской, выше,

Таблица. Результаты распознавания изображений лиц, полученные в рамках использованных систем

Table. Results of the masked face recognition obtained using the proposed systems

Номер	Тип маски/трансформации	Число изображений лиц	Точность, %			
			Simple FaReS	ArgFace	ResNet-50	VGG-16
1	Белая маска	100	100	90	95	89
2	Красная маска	100	100	93	89	86
3	Серая маска	100	100	90	90	90
4	Черная маска	100	100	94	95	78
5	Черно-синяя маска	100	100	89	95	82
6	Текстурная маска	100	100	91	95	98
7	Медицинская маска	100	100	89	98	95
8	Шум (тип маски номер 1–7)	700	100	4	6	9
9	Имитация удаленности от камеры (тип маски номер 1–7)	700	100	6	28	10
10	Комбинация изменений по трансформациям номер 8 и 9 (тип маски номер 1–7)	700	99	1	1	2
11	Де-идентификация	2000	84	94	85	85
12	Различные текстурные маски	1600	100	95	98	98
13	Перестановки компонент	9600	100	91	97	96
14	Средняя точность по всем видам преобразований с изображениями лиц в масках (трансформации номер 8–11)	—	96	26	30	27
15	Средняя точность по всем группам (номера 1–13)	—	99	71	75	71

чем в пространстве исходных данных (яркость пикселей). Данный результат положительно влияет на работу классификатора в детерминированной системе. Это подтверждает рис. 7, на котором представлены результаты оценки мер подобия «своего изображения лица» (рис. 7, *a*) и «чужого изображения лица» (рис. 7, *b*) с тремя видами изображений в белых масках с искаженной текстурой. Мера подобия вычислена как взаимная фазовая корреляция в пространстве исходных (яркостных) признаков (Cross Phase Correlation of Image, CPCI) и признаков косинус-преобразования (Cross Phase Correlation of Feature, CPCF), а также расстояния (Distance) в пространстве признаков косинус-преобразования. Аналогичные результаты с примерами и объяснениями такого эффекта получены в работе [24].

На основании анализа нейросетевых систем заметим, что лучшей системой по средней результативности РИЛ по всем группам тестовых данных является ResNet-50, остальные системы уступают ей лишь на 4 %. При этом, опираясь на результаты системы ResNet-50, видно, что ни одна из представленных типов масок не распознается на 100 %. Хуже всего распознаются маски красного и серого цветов (точность

составила не более 90 %). Отметим, что нейросетевые системы совершенно не работоспособны в случаях, когда изображения лиц в масках имеют размытую или сглаженную текстуру лица, а также наложенный на них шум. РИЛ с искаженной текстурой лица в рамках сверточных нейросетей приводит к очень низкому результату из-за использования в них сверточных фильтров. Отклик сверточных фильтров в конечных результатах свертки приводит к совершенно другим минимаксным решениям и не только по значениям, но и по их положению – вплоть до хаотического их представления на выходе нейросети. Поскольку «хаос» несравним со значениями, полученными по эталонам, задача РИЛ с искаженной текстурой лица не может быть полностью решена в рамках сверточных нейросетей. Эта проблема нейросетевых систем может быть частично решена при применении предложенного метода генерации масок на изображении лиц людей. Использование данного метода позволит создать необходимые объемы обучающих данных на уже известных наборах данных (например, VGG-Face [25]) для дообучения современных экстракторов признаков на основе семейства ResNet и VGG-16.

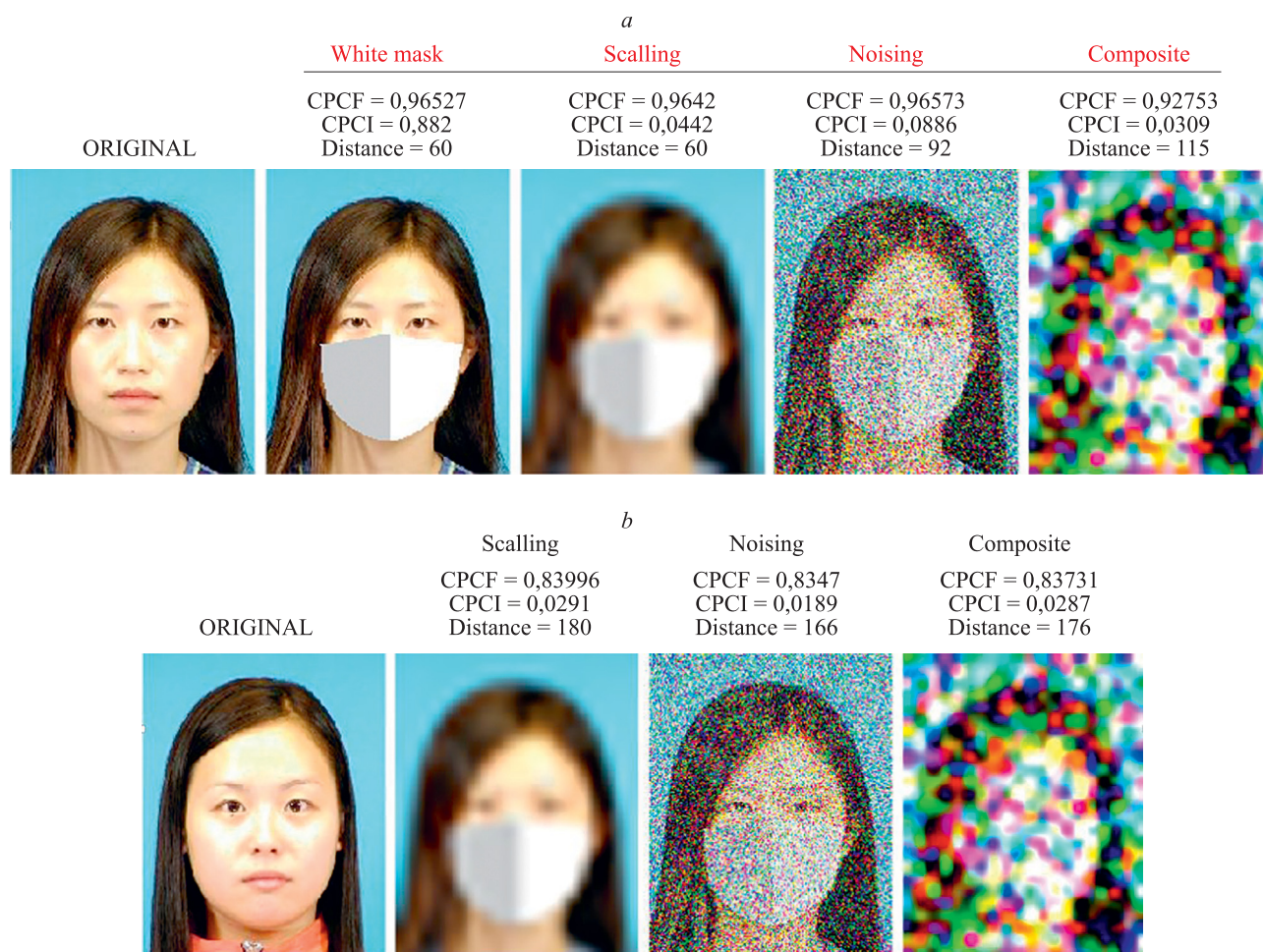


Рис. 7. Оценки подобия между своим (*a*) и чужим (*b*) изображениями лица и тремя видами изображений лиц в искаженных масках

Fig. 7. Estimates of similarity between one's own (*a*) and another's (*b*) face images and three types of images of faces in distorted masks

Заключение

Исследована проблема распознавания изображений лиц в масках при условии, что эти маски наложены на оригинальные изображения лиц и, таким образом, становятся поддельными (SIMULATED MASKS) по отношению к реальным маскам (REAL MASKS) на лицах людей. Показано, что необходимость выполнения этого исследования связана, во-первых, с разнообразием типов, форм, текстуры и цвета реальных масок, используемых для гигиенической защиты лиц, и, во-вторых, с отсутствием соответствующих этим маскам реальных наборов обучающих данных. Отмечено, что оба эти факта создали проблемы для распознавания изображений лиц в подобных масках в рамках современных нейросетевых систем.

Как один из вариантов решения этой проблемы, предложен метод генерации масок на оригинальные изображения лиц людей. Метод включает генерацию и размещение на лицах различных типов масок (медицинских, цветных текстурных и прозрачных) — как на отдельных портретах, так и на групповых фото и лицах в сценах с потоками людей. Показано, что корректное размещение масок на изображении лица основано на ключевых точках лица, вычисляемых в онлайн-режиме, использовании мнемонических моделей масок и реализации процесса формирования масок непосредственно на текстуре изображений лиц. При этом предложенный метод не зависит от размеров изображений-оригиналов и их фенотипа, а также точного положения лиц

АНФАС, что позволяет использовать предложенный метод со многими известными базами изображений лиц. Представлены варианты сгенерированных масок, которые практически не отличаются от реальных масок (REAL MASKS) и значительно превосходят поддельные маски (например, «печатные маски» или SIMULATED MASKS). В качестве исходных изображений-оригиналов использованы, как пример, 100 изображений лиц из базы CUFS. На их основе сгенерировано более 20 000 изображений лиц в масках. Выполнены эксперименты по распознаванию лиц из этой базы в рамках четырех систем: детерминированной системы с экстрактором признаков на базе косинус-преобразования и классификатором по минимуму расстояния с метрикой L1; трех нейросетевых систем, отличающиеся экстракторами признаков: ArgFace, ResNet-50 и VGG-16 — в качестве классификатора используется критерий минимального расстояния с метрикой косинусного подобия.

Предложенный метод генерации масок может быть использован для создания наборов данных и тестовых баз изображений лиц с масками. При этом используемые экстракторы признаков в нейросетевых системах распознавания изображений лиц могут быть дообучены на новых сгенерированных корпусах, что позволит частично решить существующие проблемы распознавания на основе нейросетей. Исследованные системы в дальнейшем могут быть использованы в локальных и потоковых системах контроля доступа. Работа будет полезна исследователям и специалистам в области обработки и анализа изображений лиц с масками.

Литература

1. Ge S., Li J., Ye Q., Luo Z. Detecting masked faces in the wild with LLE-CNNs // *Proc. of the 30th IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. 2017. P. 426–434. <https://doi.org/10.1109/CVPR.2017.53>
2. Ryumina E., Ryumin D., Ivanko D., Karpov A. A novel method for protective face mask detection using convolutional neural networks and image histograms // *International Archives of the Photogrammetry Remote Sensing and Spatial Information Sciences*. 2021. V. XLIV-2/W1-2021. P. 177–182. <https://doi.org/10.5194/isprs-archives-XLIV-2-W1-2021-177-2021>
3. Косулин К.Э., Карпов А.А. Методы аудиовизуального распознавания людей в масках (обзорная статья) // *Научно-технический вестник информационных технологий, механики и оптики*. 2022. Т. 22. № 3. (в печати).
4. Hsu G.-S.J., Wu H.-Y., Tsai C.-H., Yanushkevich S., Gavrilova M. Masked face recognition from synthesis to reality // *IEEE Access*. 2022. V. 10. P. 37938–37952. <https://doi.org/10.1109/ACCESS.2022.3160828>
5. Ngan M.L., Grother P. J., Hanaoka K.K. Ongoing Face Recognition Vendor Test (FRVT) Part 6A: Face recognition accuracy with masks using pre-COVID-19 algorithms: NISTIR 8311 / National Institute of Standards and Technology U.S. Department of Commerce. 2020. 58 p. <https://doi.org/10.6028/NIST.IR.8311>
6. Gorodnichy D., Yanushkevich S., Shmerko V. Automated border control: Problem formalization // *Proc. of the IEEE Symposium on Computational Intelligence in Biometrics and Identity*. 2014. P. 118–125. <https://doi.org/10.1109/CIBIM.2014.7015452>
7. Huang B., Wang Z., Wang G., Jiang K., He Z., Zou H., Zou Q. Masked face recognition datasets and validation // *Proc. of the 18th IEEE/CVF International Conference on Computer Vision*. 2021. P. 1487–1491. <https://doi.org/10.1109/ICCVW54120.2021.00172>
8. Deng J., Guo J., An X., Zhu Z., Zafeiriou S. Masked face recognition challenge: The insightface track report // *Proc. of the 18th IEEE/CVF*

References

1. Ge S., Li J., Ye Q., Luo Z. Detecting masked faces in the wild with LLE-CNNs. *Proc. of the 30th IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2017, pp. 426–434. <https://doi.org/10.1109/CVPR.2017.53>
2. Ryumina E., Ryumin D., Ivanko D., Karpov A. A novel method for protective face mask detection using convolutional neural networks and image histograms. *International Archives of the Photogrammetry Remote Sensing and Spatial Information Sciences*, 2021, vol. XLIV-2/W1-2021, pp. 177–182. <https://doi.org/10.5194/isprs-archives-XLIV-2-W1-2021-177-2021>
3. Kosulina K.E., Karpov A.A., Methods for audiovisual recognition of people in masks. *Scientific and Technical Journal of Information Technologies, Mechanics and Optics*, 2022, vol. 22, no. 3, in press. (in Russian)
4. Hsu G.-S.J., Wu H.-Y., Tsai C.-H., Yanushkevich S., Gavrilova M. Masked face recognition from synthesis to reality. *IEEE Access*, 2022, vol. 10, pp. 37938–37952. <https://doi.org/10.1109/ACCESS.2022.3160828>
5. Ngan M.L., Grother P. J., Hanaoka K. K. *Ongoing Face Recognition Vendor Test (FRVT) Part 6A: Face recognition accuracy with masks using pre-COVID-19 algorithms. NISTIR 8311*. National Institute of Standards and Technology U.S. Department of Commerce, 2020, 58 p. <https://doi.org/10.6028/NIST.IR.8311>
6. Gorodnichy D., Yanushkevich S., Shmerko V. Automated border control: Problem formalization. *Proc. of the IEEE Symposium on Computational Intelligence in Biometrics and Identity*, 2014, pp. 118–125. <https://doi.org/10.1109/CIBIM.2014.7015452>
7. Huang B., Wang Z., Wang G., Jiang K., He Z., Zou H., Zou Q. Masked face recognition datasets and validation. *Proc. of the 18th IEEE/CVF International Conference on Computer Vision*, 2021, pp. 1487–1491. <https://doi.org/10.1109/ICCVW54120.2021.00172>
8. Deng J., Guo J., An X., Zhu Z., Zafeiriou S. Masked face recognition challenge: The insightface track report. *Proc. of the 18th IEEE/CVF*

- International Conference on Computer Vision. 2021. P. 1437–1444. <https://doi.org/10.1109/ICCVW54120.2021.00165>
9. Zhu Z., Huang G., Deng J., Ye Y., Huang J., Chen X., Zhu J., Yang T., Guo J., Lu J., Du D., Zhou J. Masked face recognition challenge: The webface260m track report // *arXiv*. 2021. arXiv.2108.07189. <https://doi.org/10.48550/arXiv.2108.07189>
 10. Adjabi I., Ouahabi A., Benzaoui A., Taleb-Ahmed A. Past, present, and future of face recognition: A review // *Electronics*. 2020. V. 9. N 8. P. 1188. <https://doi.org/10.3390/electronics9081188>
 11. Alzu'bi A., Albalas F., Al-Hadhrani T., Younis L.B., Bashayreh A. Masked face recognition using deep learning: A review // *Electronics*. 2021. V. 10. N 21. P. 2666. <https://doi.org/10.3390/electronics10212666>
 12. Deng J., Guo J., Ververas E., Kotsia I., Zafeiriou S. RetinaFace: Single-shot multi-level face localisation in the wild // *Proc. of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. 2020. P. 5202–5211. <https://doi.org/10.1109/CVPR42600.2020.00525>
 13. He K., Zhang X., Ren S., Sun J. Deep residual learning for image recognition // *Proc. of the 29th IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. 2016. P. 770–778. <https://doi.org/10.1109/CVPR.2016.90>
 14. Simonyan K., Zisserman A. Very deep convolutional networks for large-scale image recognition // *Proc. of the 3rd International Conference on Learning Representations (ICLR)*. 2015. P. 1–14.
 15. Serengil S.I., Ozpinar A. LightFace: A hybrid deep face recognition framework // *Proc. of the IEEE Innovations in Intelligent Systems and Applications Conference (ASYU)*. 2020. P. 9259802. <https://doi.org/10.1109/ASYU50717.2020.9259802>
 16. Viola P., Jones M.J. Robust real-time face detection // *International Journal of Computer Vision*. 2004. V. 57. N 2. P. 137–154. <https://doi.org/10.1023/B:VISI.0000013087.49260.fb>
 17. Kazemi V., Sullivan J. One millisecond face alignment with an ensemble of regression trees // *Proc. of the 27th IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. 2014. P. 1867–1874. <https://doi.org/10.1109/CVPR.2014.241>
 18. Shimrat M. Algorithm 112: position of point relative to polygon // *Communications of the ACM*. 1962. V. 5. N 8. P. 434. <https://doi.org/10.1145/368637.368653>
 19. Агеев М.И., Алик В.П., Марков Ю.И. Библиотека алгоритмов 1016–1506. М.: Радио и связь, 1978. 128 с. (Библиотека технической кибернетики).
 20. Гричишин Я.Т., Ефимов В.И., Ломакович А.Н. Алгоритмы и программы на Бейсике: учебное пособие для студентов педагогических институтов по физико-математическим специальностям. М.: Просвещение, 1988. 160 с.
 21. Щеголева Н.Л., Кухарев Г.А. Простой алгоритм классификации линейно неразделимых данных // *Естественные и технические науки*. 2012. № 1(57). С. 358–364.
 22. Loey M., Manogaran G., Taha M.H.N., Khalifa N.E.M. A hybrid deep transfer learning model with machine learning methods for face mask detection in the era of the COVID-19 pandemic // *Measurement*. 2021. V. 167. P. 108288. <https://doi.org/10.1016/j.measurement.2020.108288>
 23. Кухарев Г.А., Каменская Е.И., Матвеев Ю.Н., Щеголева Н.Л. Методы обработки и распознавания изображений лиц в задачах биометрии / под ред. М.В. Хитрова. СПб.: Политехника, 2013. 388 с.
 24. Cao Q., Shen L., Xie W., Parkhi O.M., Zisserman A. VGGFace2: A dataset for recognising faces across pose and age // *Proc. of the 13th IEEE International Conference on Automatic Face & Gesture Recognition (FG)*. 2018. P. 67–74. <https://doi.org/10.1109/FG.2018.00020>
 25. Кухарев Г.А., Мауленов К.С., Щеголева Н.Л. Защита изображений лиц от распознавания в социальных сетях: способы решения и их перспективы // *Научно-технический вестник информационных технологий, механики и оптики*. 2021. Т. 21. № 5. С. 755–766. <https://doi.org/10.17586/2226-1494-2021-21-5-755-766>
- International Conference on Computer Vision, 2021, pp. 1437–1444. <https://doi.org/10.1109/ICCVW54120.2021.00165>
 9. Zhu Z., Huang G., Deng J., Ye Y., Huang J., Chen X., Zhu J., Yang T., Guo J., Lu J., Du D., Zhou J. Masked face recognition challenge: The webface260m track report. *arXiv*, 2021, arXiv.2108.07189. <https://doi.org/10.48550/arXiv.2108.07189>
 10. Adjabi I., Ouahabi A., Benzaoui A., Taleb-Ahmed A. Past, present, and future of face recognition: A review. *Electronics*, 2020, vol. 9, no. 8, pp. 1188. <https://doi.org/10.3390/electronics9081188>
 11. Alzu'bi A., Albalas F., Al-Hadhrani T., Younis L.B., Bashayreh A. Masked face recognition using deep learning: A review. *Electronics*, 2021, vol. 10, no. 21, pp. 2666. <https://doi.org/10.3390/electronics10212666>
 12. Deng J., Guo J., Ververas E., Kotsia I., Zafeiriou S. RetinaFace: Single-shot multi-level face localisation in the wild. *Proc. of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2020, pp. 5202–5211. <https://doi.org/10.1109/CVPR42600.2020.00525>
 13. He K., Zhang X., Ren S., Sun J. Deep residual learning for image recognition. *Proc. of the 29th IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2016, pp. 770–778. <https://doi.org/10.1109/CVPR.2016.90>
 14. Simonyan K., Zisserman A. Very deep convolutional networks for large-scale image recognition. *Proc. of the 3rd International Conference on Learning Representations (ICLR)*, 2015, pp. 1–14.
 15. Serengil S. I., Ozpinar A. LightFace: A hybrid deep face recognition framework. *Proc. of the IEEE Innovations in Intelligent Systems and Applications Conference (ASYU)*, 2020, pp. 9259802. <https://doi.org/10.1109/ASYU50717.2020.9259802>
 16. Viola P., Jones M.J. Robust real-time face detection. *International Journal of Computer Vision*, 2004, vol. 57, no. 2, pp. 137–154. <https://doi.org/10.1023/B:VISI.0000013087.49260.fb>
 17. Kazemi V., Sullivan J. One millisecond face alignment with an ensemble of regression trees. *Proc. of the 27th IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2014, pp. 1867–1874. <https://doi.org/10.1109/CVPR.2014.241>
 18. Shimrat M. Algorithm 112: position of point relative to polygon. *Communications of the ACM*, 1962, vol. 5, no. 8, pp. 434. <https://doi.org/10.1145/368637.368653>
 19. Ageev M.I., Alik V.P., Markov Iu.I. *Algorithms 1016-1506 Library*. Handbook. Moscow, Radio i svjaz' Publ., 1978, 128 p. (in Russian)
 20. Grichishin Ia.T., Efimov V.I., Lomakovich A.N. *Algorithms and Programs in BASIC*. Moscow, Prosveshnenie Publ., 1988, 160 p. (in Russian)
 21. Shchegoleva N.L., Kukharev G.A. A simple classification algorithm for linearly inseparable data. *Natural and Technical Sciences*, 2012, no. 1(57), pp. 358–364. (in Russian)
 22. Loey M., Manogaran G., Taha M.H.N., Khalifa N.E.M. A hybrid deep transfer learning model with machine learning methods for face mask detection in the era of the COVID-19 pandemic. *Measurement*, 2021, vol. 167, pp. 108288. <https://doi.org/10.1016/j.measurement.2020.108288>
 23. Kukharev G.A., Kamenskaya A.I., Matveev Y.N. *Methods of facial images processing and recognition in biometrics*. St.Petersburg, Politehnika Publ., 2013, 388 p. (in Russian)
 24. Cao Q., Shen L., Xie W., Parkhi O.M., Zisserman A. VGGFace2: A dataset for recognising faces across pose and age. *Proc. of the 13th IEEE International Conference on Automatic Face & Gesture Recognition (FG)*, 2018, pp. 67–74. <https://doi.org/10.1109/FG.2018.00020>
 25. Kukharev G.A., Maulenov K.S., Shchegoleva N.L. Protecting facial images from recognition on social media: solution methods and their perspective. *Scientific and Technical Journal of Information Technologies, Mechanics and Optics*, 2021, vol. 21, no. 5, pp. 755–766. (in Russian). <https://doi.org/10.17586/2226-1494-2021-21-5-755-766>

Авторы

Кухарев Георгий Александрович — доктор технических наук, профессор, профессор, Санкт-Петербургский государственный электротехнический университет «ЛЭТИ» им. В.И. Ульянова, Санкт-Петербург, 197376, Российская Федерация,  18037842200, <https://orcid.org/0000-0003-2188-2172>, gakukharev@etu.ru

Рюмина Елена Витальевна — младший научный сотрудник, Санкт-Петербургский Федеральный исследовательский центр Российской академии наук, Санкт-Петербург, 199178, Российская Федерация,  57220572427, <https://orcid.org/0000-0002-4135-6949>, ryumina_ev@mail.ru

Шульгин Никита Александрович — студент, Санкт-Петербургский государственный электротехнический университет «ЛЭТИ» им. В.И. Ульянова, Санкт-Петербург, 197376, Российская Федерация, <https://orcid.org/0000-0002-1884-5832>, nikitashulgin@gmail.com

Authors

Georgy A. Kukharev — D. Sc., Full Professor, Saint Petersburg State Electrotechnical University “LETI”, Saint Petersburg, 197376, Russian Federation,  18037842200, <https://orcid.org/0000-0003-2188-2172>, gakukharev@etu.ru

Elena V. Ryumina — Junior Researcher, Saint Petersburg Federal Research Center of the Russian Academy of Sciences (SPC RAS), Saint Petersburg, 199178, Russian Federation,  57220572427, <https://orcid.org/0000-0002-4135-6949>, ryumina_ev@mail.ru

Nikita A. Shulgin — Student, Saint Petersburg State Electrotechnical University “LETI”, Saint Petersburg, 197376, Russian Federation, <https://orcid.org/0000-0002-1884-5832>, nikitashulgin@gmail.com

Статья поступила в редакцию 14.04.2022

Одобрена после рецензирования 26.04.2022

Принята к печати 30.05.2022

Received 14.04.2022

Approved after reviewing 26.04.2022

Accepted 30.05.2022



Работа доступна по лицензии
Creative Commons
«Attribution-NonCommercial»