УДК 004.89

ОЦЕНКА МЕТОДОВ БОРЬБЫ С ПОДМЕНОЙ ЖИВЫХ СУЩЕСТВ И ПРЕДЛОЖЕНИЯ РЕШЕНИЙ

[1,2]Нгуен Тхе Кыонг, [1]Сырямкин В.И., [3]Нгуен Чиен Тханг, [1,2]Нгуен Чанг Хоанг Тхуи, [1]Ляшенко Д.

[1]Национальный исследовательский Томский государственный университет, Томск;
[2]Вьетнамский морской университет, Хайфон;
[3]Сайгон-Ханойский коммерческий акционерный банк, Ханой, e-mail: cuongntit@vimaru.edu.vn, svi_tsu@mail.ru, thangnch@gmail.com, trangnht@vimaru.edu.vn, lyashenkodmtriy@gmail.com

Технология распознавания объектов на основе изображений изучается уже более 30 лет. И в последние 10 лет эта технология получила широкое развитие во многих решениях, получивших высокое практическое применение. Однако с развитием цифровых устройств полученные изображения могут предоставляться из самых разных источников: фотографии, сделанные (или снятые видео) непосредственно с живых существ, фотографии, напечатанные на бумаге, или фотографии (видео), отображаемые на экране компьютера или мобильного телефона. Поэтому обнаружение изображения, полученного с камеры системы непосредственно от живого существа или из изображений напечатанных (или отображенных на телефоне, снятых на видео), позволяет определить правильный объект поставленной задачи. Это обнаружение образа играет важную роль в жизни, например в охранной деятельности, в наблюдении за людьми в соответствии с их регистрацией, сделанной ранее, основной целью является обнаружение живых объектов. Статья посвящена проблеме борьбы со спуфингом человеческих лиц с изложением содержания мер, применяемых в банковской системе и в системе мониторинга доступа в соответствии с предыдущей регистрацией там этих лиц. Цель состоит в том, чтобы проверить информацию о клиенте или пользователе, такую как результаты теста подмигивания, ответ на случайные запросы и т.д. В статье дана оценка преимуществ и недостатков этих методов. В заключение в статье предлагается антиспуфинговое решение с использованием 3D-камеры и в сочетании со сверточной нейронной сетью для создания более точной антиспуфинговой системы.

Ключевые слова: обнаружение живого, лицевой антиспуфинг, моргание глаз, генерация лица

EVALUATING METHODS OF ANTI-SPOOFING OF LIVING ENTITIES AND PROPOSE SOLUTIONS

[1,2]Nguyen The Cuong, [1]Syryamkin V.I., [3]Nguyen Chien Thang, [1,2]Nguyen Trang Hoang Thuy, [1]Lyashenko D.

[1]National Research Tomsk State University, Tomsk;
[2]Viet Nam Maritime University, Hai Phong;
[3]SaiGon-HaNoi Commercial Joint Stock Bank, Hanoi, e-mail: cuongntit@vimaru.edu.vn, svi_tsu@mail.ru, thangnch@gmail.com, trangnht@vimaru.edu.vn, lyashenkodmtriy@gmail.com

Object recognition technology based on images has been studied for more than 30 years. And in recent 10 years, this technology has been widely developed with many solutions given with high practical applications. However, with the development of digital devices, the resulting images can be provided from a variety of sources: photos taken (or video captured) directly from living entities, photos printed on paper or photos (video) is displayed on the computer screen or mobile phone. Therefore, the image detection obtained from the camera system directly from the living entity or from images are printed (or displayed on the phone – video) to determine the correct object is the problem posed. This image detection plays an important role in life, for example: in security work, monitoring people in / out in accordance with their registration before, the main purpose is liveness-detection. The article focuses on the issue of anti-spoofing of human faces with the content outlining the measures applied in the banking system, the access monitoring system in accordance with the previous registration of these faces there. The purpose is to validate customer or user information such as: wink test results, response to random challenges, etc. The article evaluates the advantages and disadvantages of these methods. Finally, the article proposes an anti-spoofing solution by using 3D camera and combining with convolutional neural network to create a more accurate anti-spoofing system.

Keywords: liveness detection, face anti spoofing, eye blink, generate face

Обзор

Концепция обнаружения живого: метод, в результате применения которого объект «живое существо» (реальный человек) должен быть обнаружен непосредственно через камеру; изображение, видимое на экране, получаемое через фронтальную камеру, не является фотографией печатного объекта, фотографией, отображаемой на телефоне, или предварительно записанным видео. «Подделка» относится к несанкционированному отображению искусственных копий части биометрических данных, таких как лицо, отпечаток пальца или радужная оболочка, в биометрическую систему для получения незаконного доступа к этой системе (рис. 1) [1].

Рис. 1. Иллюстрация фотографии реального лица (live), изображение лица, напечатанного на бумаге (printed 1, 2), и изображение лица, отображенного на экране (display 1, 2)

«Сопоставлением лиц» называется сравнение лиц, изображения которых представлены в виде фотографий, и лиц, снятых непосредственно для проверки, является ли лицо на двух представленных изображениях лицом одного и того же человека или нет. Эта аутентификационная работа была проведена в рамках различных методов защиты от спуфинга.

Цель этой статьи – изучить решения, созданные для противодействия использованию фальшивых человеческих лиц. Данная статья также включает в себя содержание, посвященное оценке преимуществ и недостатков решений и предложению новых решений для применения 3D-камер. Предложения о методе, подходящем для каждого конкретного условия, также даются пользователям на выбор.

Метод проверки моргания глаз

Пользователь держит телефон и непрерывно около 10 секунд снимает себя на видео, система непрерывно проверяет, моргает он или нет. Если человек на видео моргает, это означает, что он реален, если этот человек не моргает, то будет сообщено об ошибке распознавания. Метод обнаружения моргания измеряет внутреннее движение лица. В конечном итоге фотография не может моргать. Решение проверки моргания глаз состоит в необходимости обнаружения лица с помощью библиотеки «face dlib», на лице будет 68 фокусных точек, с некоторыми точками вокруг глаз, расстояние между точками будет больше, когда глаза открыты, и меньше, когда глаза закрыты (рис. 2) [2, 3]. Для расчета используется следующая формула (eye aspect ratio – EAR):

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

Оценка преимуществ. Система проста, удобна в реализации, может быть установлена на мобильные устройства и не требует установки другого вспомогательного оборудования. С другой стороны, данные можно легко подделать с помощью предварительно записанного видео, если в видео есть мигающий момент, он все равно может быть принят системой, и таким образом тест будет пройден. Эти системы неудобны для пользователей, так как им требуется проведение большого количества попыток для непосредственного обнаружения живых объектов.

Аналогичная техника направлена на расширение зрачков глаза, заставляя экран темнеть, а затем внезапно мигать. Этот метод эффективен при обнаружении фальсификаций.

Метод обнаружения живого, основанный на запросе-ответе

М.Х. Алиа в [4] и Зун и др. в [5] придумали метод запроса пользователя на выполнение некоторой случайной инструкции, а затем проверенного ответа, чтобы подтвердить, были ли инструкции выполнены или нет.

Некоторые из этих проблем могут заключаться в том, чтобы повернуть голову, закрыть глаза, высунуть язык и т.д. Используя эти методы, можно проверить, поворачивает ли пользователь голову в направлении случайного вызова (рис. 3).

Эти вызовы могут повторяться несколько раз в зависимости от уровня безопасности администратора. Для определения направлений вращения лица используется распознавание лиц для проверки точек на лице [6] методом SVM classify [7] или методом CNN classify [8, 9].

Для практической реализации в реальном времени этим системам может потребоваться от пользователей выполнение дополнительных действий, таких как поворот влево при чтении случайных чисел, отображаемых на экране. Система будет иметь дополнительную обработку звука, чтобы подтвердить, соответствует ли он звучанию цифр на экране или нет {один, два, три, четыре,.. девять}.

Оценка. Этот метод легко реализовать на мобильных устройствах, поскольку не требуется никакого дополнительного оборудования. Случайная операция для высокоточной аутентификации. Однако этот метод также имеет проблемы, когда пользователь неправильно выполняет операции, запрашиваемые системой, или система

распознает их с низкой точностью. Иногда пользователям приходится выполнять несколько действий, потому что обученная модель CNN использует слишком высокий порог точности. Эта проблема может быть неприятной для пользователей.

Система использует метод классификации объектов

1. Классификация использует LBP

Область изображения лица подразделяется на несколько областей. Для каждой области вычисляются параметры LBP по следующему методу [10]: Для области изображения 3×3 значение $LBP_{8.1}$ вычисляется следующим образом. Пиксели определяются в центре с восемью окружающими ячейками. Двоичные значения вычисляются данным методом. Для области двоичного изображения, если какое-либо значение ячейки больше значения центральной ячейки, значение ячейки будет равно 1, и в любом другом случае оно будет равно 0. Затем двоичное строковое значение этих восьми ячеек преобразуется в десятичное. На следующем шаге система перепишет это значение с положением соответствующего центрального пикселя в изображении LBP (рис. 4).

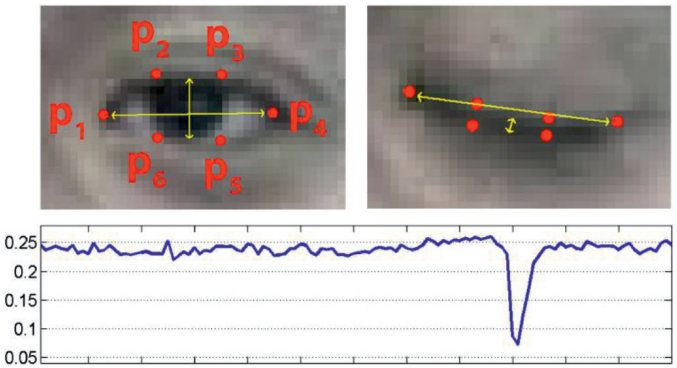

Рис. 2. Вверху слева: Изображение маркеров глаз при открытых глазах. Вверху справа: Изображение глазных ориентиров при закрытых глазах. Внизу: График соотношения сторон глаз с течением времени. Уменьшение соотношения сторон глаза указывает на моргание

Рис. 3. Иллюстрация требований к пользователю

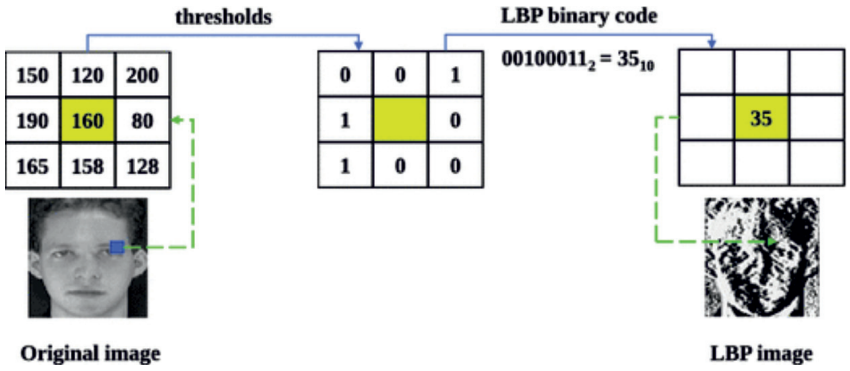

Рис. 4. Базовый пример оператора LBP8,1

После получения двух матриц изображений будет вычислена разница между исходной матрицей изображений и матрицей изображений LBP и получены значения, представляющие глубину изображения. Ожидается, что фотографии, сделанные непосредственно от реальных людей, будут иметь значение LBP, отличное от фотографий, сделанных с плоских фонов, таких как бумажные фотографии или фотографии, отображаемые на экране телефона, компьютера.

Оценка. Этот метод все еще может вводить в заблуждение, поскольку при различных условиях освещения результирующее изображение будет иметь различное значение (например, отсутствие света, избыток света, подсветка) или изображение, напечатанное с высоким качеством. Преимущество этого метода заключается в том, что он не требует дополнительного оборудования, легкий алгоритм может быть запущен даже на мобильных устройствах.

2. Метод использования сети CNN для классификации

Система захватывает несколько изображений лица и проверяет их на изменения и естественное движение. Движущиеся 3D-грани отличаются от 2D-изображений, и сложные алгоритмы в системах с моделями CNN [8, 9] могут обнаружить это различие [1]. Авторы обнаружили повторяющиеся видео и другие дубликаты, используя специальный алгоритм, основанный на текстурах, чтобы отличить записанную версию от реальных людей (рис. 5).

Оценка. Эта модель проста и легка в развертывании. Но склонна к чрезмерной подгонке, если идентификатор не берется из обученной базы данных. Также система изменяется при разных условиях освещения, либо другой угол камеры производит разные изображения; Или как иногда импортированное изображение, даже если оно взято с реальной фотографии, но при низком качестве будет определено как поддельное изображение, в то время как поддельное изображение, импортированное с высоким качеством, все еще может быть распознано как реальное изображение.

3. Метод антиспуфинга с использованием вспышки

Этот метод основан на разнице в отражении света между плоской поверхностью (2D) и выпуклой/вогнутой поверхностью (3D) при использовании вспышки для получения большего освещения. Следовательно, отражение этого света от мишени может быть использовано для идентификации реального лица. Белая область, покрывающая экран, создает соответствующее отражение на лице [11] (рис. 6) [12].

Приведенный выше пример указывает на разницу необработанных пикселей. Таким искусственным способом настоящее лицо отличается от фальшивого из-за разницы в их поверхности. Кадры до и после мигания активности предоставляют шаблоны данных для обучения сети. Вспышка помогает отделить и классифицировать черты лица.

Оценка. При хорошем освещении точность довольно высока – 98,8 %, в шумных условиях все еще достигает 97,3 % [11]. Недостатком этого метода является то, что не все устройства обладают вспышкой. Например, на мобильных устройствах вспышка доступна только с задней камерой, а не с передней. А при ярком солнечном свете на открытом воздухе вспышка малоэффективна.

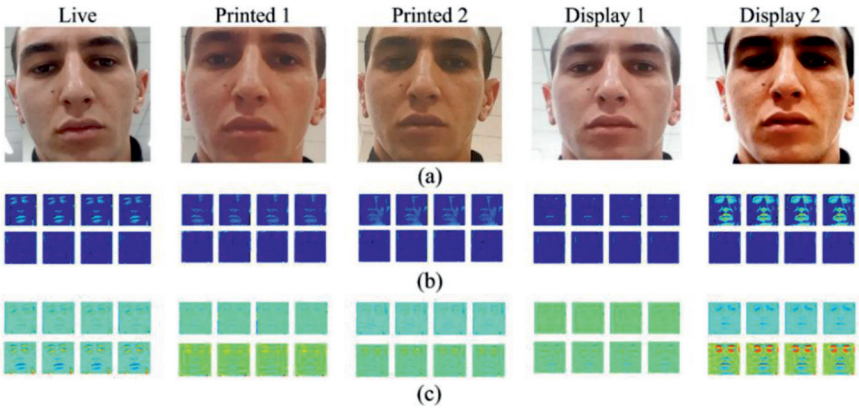

Рис. 5. (a) Образец живого лица и соответствующего ПА в базе данных ОУЛУ. (b) Карты объектов, выводимые первым слоем свертки в CNN с входным сигналом RGB face. (c) Карты признаков, выводимые предлагаемым адаптивным слоем слияния сверточных признаков после первого слоя свертки в CNN с изображениями лиц RGB и DNG в качестве входных данных

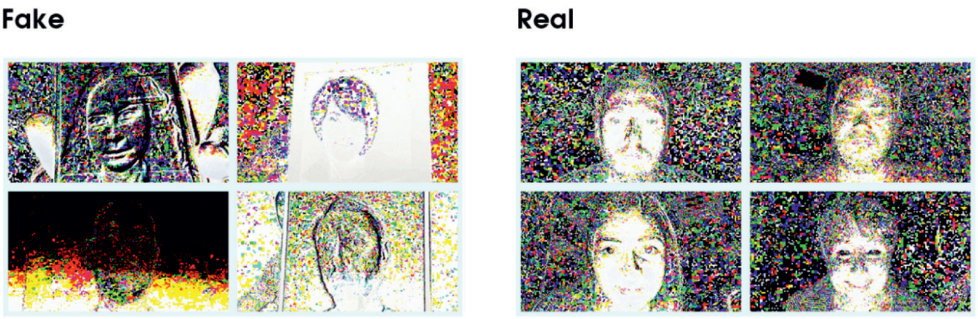

Рис. 6. Разница во вспышке между фальшивым лицом и реальным лицом

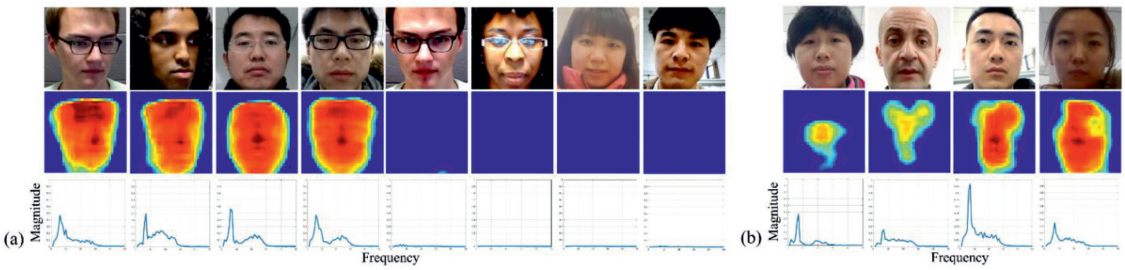

Рис. 7. (a) 8 успешных примеров антиспуфинга и их оценочные карты глубины и сигналы rPPG. (b) 4 примера неудач: первые два являются живыми, а два других – поддельными. Обратите внимание на нашу способность оценивать дискриминационные карты глубины и сигналы rPPG

4. Метод создания 3D карты

Иаозие и др. [13] предложили метод: каждое входное изображение лица делится на два потока, которые являются реальными или поддельными изображениями. Они будут классифицировать, как описано выше, извлекать некоторые области лица, такие как глаза, нос и рот, для обучения модели классификации CNN давать истинную / ложную вероятность.

Реализация обученной модели автоматического эндкодера проходит таким образом, что реальное лицо генерирует матрицу 3D-карты соответствующей формы, а для поддельного лица значение 3D-карты остается другим (рис. 7).

После получения двух значений вероятности двух указанных выше периодов алгоритм суммирует их, чтобы дать общую вероятность того, что входное изображение лица является реальным или ложным изображением.

Оценка. Точное значение довольно высокое, нет необходимости в другом вспомогательном оборудовании. Недостатком является то, что модель пока не подходит для работы на мобильных платформах.

Предложение метода обнаружения фальшивых лиц с помощью специальной камеры

Согласно документу [14], при использовании 3D-камеры в то же самое время не менее двух камер расположено под разными углами. Таким образом, изображения, полученные с реального лица, будут иметь большую разницу (обладают различной глубиной) с поддельными изображениями лица, изображения с 3D-камер будут получены с незначительной разницей.

3D-камера с хорошим качеством обладает обычно четырьмя небольшими камерами. Отсюда полученное изображение будет иметь изображения с областями глубины (с реальными лицами) или нет (в случае с фальшивыми лицами).

Глубина изображения представляет собой различие расстояний между двумя камерами и пикселями на лице. Если фотография сделана с реального человека, то разница в расстоянии будет большой (или иметь выпуклую / вогнутую форму), если фотография поддельная, то эта разница невелика. Экспериментально мы вычислили глубину между пикселем носа и пикселем уха с наибольшей разницей.

Обладая высококачественными камерами, библиотеки SDK обеспечивают хорошую поддержку функций для определения значений разности глубин. Кроме того, эти камеры также имеют потоки для глубинных кадров или потоки для цветных кадров RGB. Поэтому использование этих камер очень удобно для программистов. Если человеческое лицо обнаружено в кадре потока

RGB, то область изображения человеческого лица ссылается на область изображения в глубинном кадре. Если область лица совпадает с областью с соответствующей глубиной, изображение является реальным лицом (рис. 8).

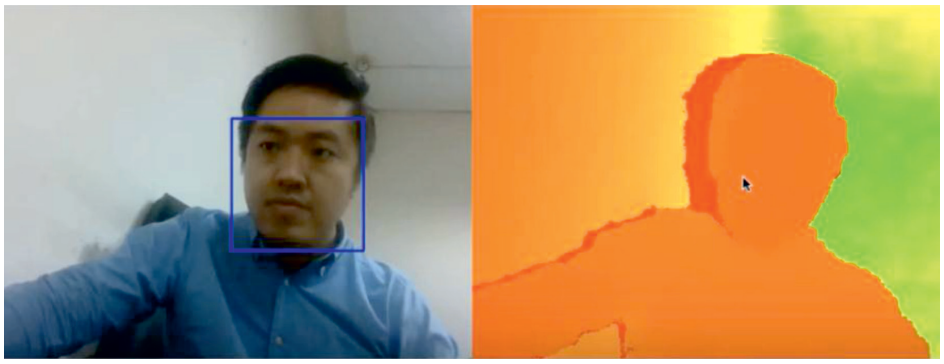

Рис. 8. RGB изображение и изображение глубины 3D камеры

Алгоритм представлен следующим образом:
Вход: Изображение с 3D-камеры
Выход: поддельное лицо / реальное лицо
Алгоритм представлен на рис. 9.

```
1.  Цикл:
2.      imgGray = Считывание изображений с RGB камер – преобразование в оттенки серого;
3.      imgDepth = Считывание изображений с датчика глубины;
4.      Если (обнаруживается человеческое лицо из imgGray):
5.          Построить на лице 68-точечных координат;
6.          Провести соответствие глубине изображения (depth(68));
7.          max = maximum(depth (68));
8.          min = minimum(depth (68);
9.          scale = (max – min);
10.         Если (scale < 1000):
11.             Изображение фальшивого лица;
12.         Иначе:
13.             Изображение реального лица;
14.     Продолжать?:
15.         Да: Переход к циклу;
16.         Нет: Выход из цикла;
17. Конец цикла;
```

Рис. 9. Алгоритм обнаружения реальных / поддельных лиц

Обнаружение человеческих лиц на изображении (линия 3) может использовать Xapp каскадный метод [15] или MTCNN [16]; код, написанный на языке Python, библиотеки в формате OpenCV, dlib, процессор Intel SDK, используя следующие библиотеки или функции: imutils, face_utils, pyrealsense2, numpy, cv2, dlib, pipeline(), config(), stream.depth, stream.color, shape_predictor(), stream.color, align(), align.process(), get_depth_frame(), get_color_frame(), detect_face(), detect_mouth(), shape_to_np().

Результаты после эксперимента: при реальном человеческом лице значение шкалы ∈ составляет [2000, 3500] (рис. 10, а), при нанесении изображения с фальшивым лицом значение шкалы ∈ составляет [0, 300] (рис. 10, б).

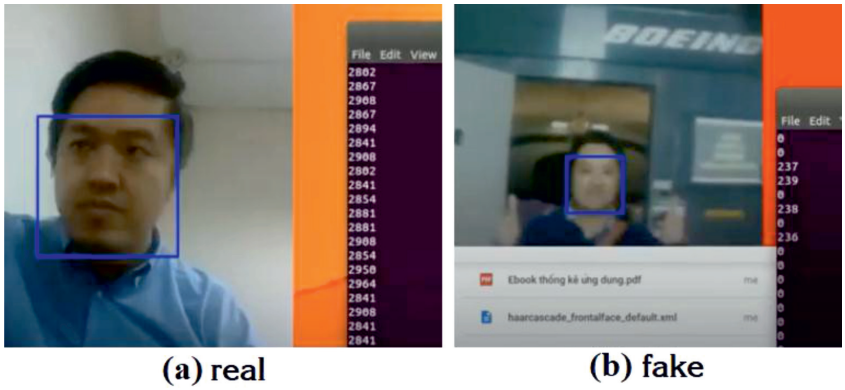

Рис. 10. Экспериментальные результаты по обнаружению реальных человеческих (a – real) и поддельных (b – fake) лиц

Оценка. Преимущества: высокая точность, поддерживаемая датчиком глубины. Недостаток: большая стоимость устройства, а также возможность поддержки текущих функций библиотеки SDK только в операционных системах Windows или Linux и отсутствие их поддержки со стороны Mac OS.

Выводы

Эти методы имеют различные преимущества и недостатки. Благодаря их простоте нет необходимости в дополнительном оборудовании, поэтому метод ответа на запрос по-прежнему является наиболее часто используемым методом, хотя точность его невысока. Метод использования 3D-камер более выгоден из-за большой глубины изображения. Однако существуют ограничения, когда поддельный объект имеет 3D-вид. Предложенный метод применим к тепловизионным камерам [17], так как в тепловизионной камере имеется термодатчик (соответствующий датчику глубины в 3D-камере). Исследователи также могут обратиться к биометрическим методам радужной оболочки глаза [18], биометрии отпечатков пальцев [19], чтобы иметь возможность объединить функции верификации и создать метод, соответствующий их задачам и фактическим условиям.

Список литературы

1. Yasar Abbas Ur Rehman, Lai-Man Po, Mengyang Liu, Zijie Zou, Weifeng Ou, Yuzhi Zhao. Face liveness detection using convolutional-features fusion of real and deep network generated face images. Journal of Visual Communication and Image Representation 59, 2019. P. 574–582. DOI: 10.1016/j.jvcir.2019.02.014.

2. Tereza Soukupova, Jan Cech. Real-Time Eye Blink Detection using Facial Landmarks. 21st Computer Vision Winter Workshop, Rimske Toplice, Slovenia, 2016. [Electronic resource]. URL: https://vision.fe.uni-lj.si/cvww2016/proceedings/papers/05.pdf (date of access: 25.02.2021).

3. Nikitin M.Yu., Konushin V.S., Konushin A.S. Face anti-spoofing with joint spoofing medium detection and eye blinking analysis. Computer Optics. 2019. Vol. 43 (4). P. 618–626. DOI: 10.18287/2412-6179-2019-43-4-618-626.

4. Alia M. Khaled Saad. Anti-Spoofing Using Challenge-Response User Interaction. A thesis Submitted to Department of Computer Science and Engineering, American University In Cairo, 2015. P. 33–43.

5. Ajian Liu, Xuan Li, Jun Wan, Yanyan Liang, Sergio Escalera, Hugo Jair Escalante, Meysam Madadi, Yi Jin, Zhuoyuan Wu, Xiaogang Yu, Zichang Tan, Qi Yuan, Ruikun Yang, Benjia Zhou, Guodong Guo, Stan Z. Li. Cross-ethnicity Face Anti-spoofing Recognition Challenge: A Review. IET Research Journals, 2020 // arXiv:2004.10998v1. P. 1–12.

6. Sagonas C., Tzimiropoulos G., Zafeiriou S., Pantic M. 300 faces in-the-wild challenge: The first facial landmark localization challenge. In Proceedings of the IEEE International Conference on Computer Vision Workshops, 2013. P. 397–403.

7. Alex J. Smola, Bernhard Schölkopf. A Tutorial on Support Vector Regression. Statistics and Computing archive. 2004. Vol. 14. Issue 3. P. 199–222.

8. Zhengzhe Liu, Xiaojuan Qi, Philip Torr. Global Texture Enhancement for Fake Face Detection in the Wild. 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). 2020. DOI: 10.1109/CVPR42600.2020.00808.

9. Нгуен Тхе Кыонг, Сырямкин В.И., Нгуен Чанг Хоанг Тхуи. Модель метода распознавания объектов на изображениях с использованием «сверточной нейронной сети – CNN» // Современные наукоемкие технологии. 2020. № 12 (2). С. 269–280.

10. Maatta J., Hadid A., and Pietikainen M. Face spoofing detection from single images using micro-texture analysis. In Biometrics (IJCB), International Joint Conference, 2011. P. 1–7.

11. Di Tang, Zhe Zhou, Yinqian Zhang, Kehuan Zhang. Face Flashing: a Secure Liveness Detection Protocol based on Light Reflections. Cornell University, Computer Vision and Pattern Recognition, 2018 // arXiv:1801.01949. P. 1–15.

12. MobiDev International Ltd. Anti-spoofing techniques in face recognition [Electronic resource]. URL: https://mobidev.biz/blog/face-anti-spoofing-prevent-fake-biometric-detection (date of access: 25.02.2021).

13. Yaojie Liu, Amin Jourabloo, Xiaoming Liu. Learning Deep Models for Face Anti-Spoofing: Binary or Auxiliary Supervision. Computer Science 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition. DOI: 10.1109/CVPR.2018.00048.

14. Intel Coprporation. 2020 Intel® RealSense™ D400 Series Calibreation Tools user guide [Electronic resource]. URL: https://www.intel.com/design/literature.htm (date of access: 25.02.2021).

15. Viola P., Jones M. Rapid object detection using a boosted cascade of simple features. Proceedings of the 2001 IEEE

Computer Society Conference on Computer Vision and Pattern Recognition. CVPR 2001. [Electronic resource]. URL: https://ieeexplore.ieee.org/document/990517 (date of access: 25.02.2021). DOI: 10.1109/CVPR.2001.990517.

16. Kaipeng Zhang, Zhanpeng Zhang, Zhifeng Li, Yu Qiao. Joint Face Detection and Alignment Using Multitask Cascaded Convolutional Networks. IEEE Signal Processing Letters (Vol. 23), 2016. [Electronic resource]. URL: https://ieeexplore.ieee.org/document/7553523 (date of access: 25.02.2021). DOI: 10.1109/LSP.2016.2603342.

17. Hikvision company // Thermal & Optical Bi-spectrum Network Bullet / Turret Camera – User Manual [Electronic resource]. URL: https://www.hikvision.com/europe/products/Thermal-Products/Security-thermal-cameras/Turret-series/DS-2TD1217-3-V1/ (date of access: 25.02.2021).

18. Soheil Hashemi, Hokchhay Tann, Francesco Buttafuoco, Sherief Reda. Approximate Computing for Biometric Security Systems: A Case Study on Iris Scanning. 2018 Design, Automation & Test in Europe Conference & Exhibition (DATE) [Electronic resource]. URL: https://ieeexplore.ieee.org/document/8342029 (date of access: 25.02.2021). DOI: 10.23919/DATE.2018.8342029.

19. Dileep Kumar, Yeonseung Ryu. A Brief Introduction of Biometrics and Fingerprint Payment Technology // 2008 Second International Conference on Future Generation Communication and Networking Symposia, 2008. [Electronic resource]. URL: https://www.tib.eu/en/search/id/ieee:doi~10.1109%252FFGCNS.2008.11/A-Brief-Introduction-of-Biometrics-and-Fingerprint?cHash=36d527e38bfd5a0996cfdd59766ae9bc (date of access: 25.02.2021). DOI: 10.1109/FGCNS.2008.11.