УДК 004.93
DOI:10.38028/ESI.2022.27.3.005

Построение карт глубины при обнаружении презентационных атак в системах распознавания лиц

Фаворская Маргарита Николаевна, Пахирка Андрей Иванович
Сибирский государственный университет науки и технологий имени акад. М.Ф. Решетнева, Россия, Красноярск, *favorskaya@sibsau.ru*

Аннотация. В настоящее время системы распознавания лиц являются широко распространенным способом биометрической идентификации на практике. Однако такие системы требуют защиты от несанкционированных действий в виде так называемых презентационных атак, когда злоумышленник осуществляет подмену подлинных изображений поддельными изображениями или короткими видеопоследовательностями. В статье предлагается метод обнаружения презентационных атак с использованием глубины сцены без применения специальных датчиков. Задача заключается в том, чтобы усилить тонкую разницу между подлинными и поддельными изображениями. Для этого была обучена и протестирована глубокая сеть, состоящая из блоков свертки центральной разности и многомасштабного модуля внимания. Эксперименты показали, что предварительное преобразование входных изображений лиц в цветовое пространство HSV имеет преимущество по точности обнаружения поддельных изображений. Так, точность обнаружения на собственном наборе данных, наборах данных KITTI и Cityscapes возросла на 3-7% в зависимости от устройств захвата, условий освещения и настроек самого алгоритма.

Ключевые слова: презентационные атаки, распознавание лиц, карты глубины, глубокое обучение

Цитирование: Фаворская М.Н. Построение карт глубины при обнаружении презентационных атак в системах распознавания лиц / М.Н. Фаворская, А.И. Пахирка // Информационные и математические технологии в науке и управлении. – 2022. – № 3(27). – С. 40-48. – DOI :10.38028/ESI.2022.27.3.005.

Введение. Распознавание лиц в системах биометрической идентификации занимает ведущее положение среди других способов идентификации, например, идентификации личности по отпечаткам пальцев и радужной оболочки глаза, которые требуют наличия дорогостоящего оборудования. Распознавание лиц с использованием видеокамер используется для выполнения широкого спектра видеонаблюдения, начиная от неинвазивного городского видеонаблюдения и завершая доступом для ограниченного круга лиц. В последнем случае, как правило, осуществляется двойная или даже тройная биометрическая проверка, не ограничивающаяся анализом визуального представления лица в оптическом диапазоне. Наиболее широкое применение системы распознавания лиц нашли в платежных системах, системах общественной безопасности, контрольно-пропускных системах с невысоким уровнем секретности и т.д. Однако изображения лиц или видеоролики могут быть легко украдены, например, из социальных сетей или файлообменников, что приводит ко всем видам презентационных атак (presentation attacks, PAs) на системы распознавания лиц. Презентационные атаки включают три основных типа: атаки с использованием печати (print attack), атаки воспроизведения (replay attack), под которыми понимается предъявление с помощью электронных средств (смартфонов, планшетов) отдельных изображений или коротких видеороликов, а также атаки с использованием 3D масок (3D mask attack), макияжа или профессионального гримирования. Такие атаки или их совокупность могут нанести значительный ущерб безопасности и имуществу. Таким образом, разработка методов и средств обнаружения поддельных изображений лиц (антиспуфинг, от англ. face anti-spoofing, FAS) является важным направлением исследований для академических кругов и промышленности. В данном исследовании рассматриваются атаки печати и воспроизведения. Атаки с использованием 3D масок, как правило, представляют собой отдельные исследования.

Следует отметить, что с точки зрения конечного пользователя осуществление презентационной атаки может преследовать две цели. Первая цель заключается в том, чтобы обмануть систему и выдать себя за другое лицо, которое имеет доступ. Вторая цель состоит в том,

чтобы выдать себя за другое лицо, хотя изначально пользователь имеет доступ. В обоих случаях основное внимание уделяется предотвращению презентационной атаки, а мотивировка конечного пользователя может быть легко выяснена с помощью базы данных персональной информации или другими способами.

1. Обзор методов антиспуфинга для обнаружения презентационных атак. Развитие антиспуфинговых методов можно разделить на два этапа. Этап, основанный на традиционных методах обработки изображений, предполагал извлечение признаков с участием эксперта и продолжался до этапа появления нейросетевых моделей глубокого обучения. Традиционные подходы, такие, как локальные бинарные шаблоны [1–3], гистограммы направленных градиентов [4–5] и точечные дескрипторы [6], характеризуются плохим обобщением, поскольку качество изображений или видеопоследовательности (в частности, текстура) зависит от конкретной видеокамеры. Методы глубокого обучения, способные автоматически извлекать богатую семантическую информацию, предоставляют надежные решения и в данной области [7–10].

По сравнению с традиционными подходами методы, основанные на глубоком обучении, извлекают информацию об отличительных признаках поддельных образцов поэтапно [7, 9]. В настоящее время существуют около 100 моделей глубоких искусственных нейронных сетей, использующих различные признаки для нахождения поддельных образцов лиц. Однако в данном исследовании мы ориентируемся на методы обнаружения поддельных изображений с использованием глубины. Глубина представляет собой значимый признак, особенно в презентационных атаках, в виде распечатанных фотографий, снимков со смартфонов и ноутбуков, а также в атаках воспроизведения, когда демонстрируется непродолжительный видеоклип. При атаках с использованием недорогих эластичных масок признак глубины фактически теряет значимость и целесообразен только при применении инфракрасных камер.

Кратко рассмотрим известные методы на основе глубины как степени расхождения между реальными и поддельными образцами. Одной их первых значимых работ была статья Атоум и др., опубликованная в 2017 г. [7]. Была построена двухпотоковая сверточная нейронная сеть (СНС), извлекающая локальные признаки спуфинга независимо от их пространственного расположения и строящая карту глубины изображения в целом. Предположение основано на том, что карта глубины отражает изображение реального лица как трехмерного объекта, в то время как изображение поддельного лица является плоскостью при атаках печати или атаках воспроизведения. Однако такое предположение верно для RGB-D систем, когда помимо плоского изображения строится карта глубины с помощью датчика расстояния. Именно таким образом была обучена сеть с использованием больших RGB-D наборов данных внутренних сцен помещений (не изображений лиц) [11]. В работе [8] предложен пространственно-временной подход, основанный на моделировании карт глубины и сигналов дистанционной фотоплетизмографии (remote PhotoPlethysmoGraphy, rPPG) с помощью СНС и рекуррентной нейронной сети соответственно. Для построения карты глубины 2D изображения лица был использован метод плотного выравнивания [12], который оценивает 3D-форму лица. В свою очередь, дистанционная фотоплетизмография – это метод отслеживания жизненно важных сигналов (например, частота сердечных сокращений) без контакта с кожей человека. Для анализа RGB-изображений метод определяет разницу в цвете, используя хроматические составляющие. Отметим, что обычно метод rPPG используется для противодействия атакам маскирования. В работе [13] изучались признаки шума и глубины для обобщенной защиты от спуфинга. Однако использование обычных СНС не обеспечивает создание детализированных шаблонов для оценки глубины.

В работе [9] разработан метод обнаружения презентационных атак по нескольким кадрам на основе оценки величина пространственного градиента между реальными и поддельными изображениями лиц, а также динамики движущихся 3D лиц. Блок остаточного пространственного градиента анализирует контуры, а пространственно-временная информация кодируется в модуле пространственно-временного распространения. Показано, что предложенная функция потерь контраста глубины обеспечивает более точные результаты.

Из краткого обзора видно, что использование данных о глубине оказывает существенное влияние на процесс обнаружения презентационных атак.

2. Модели цветовых искажений при атаках печати и воспроизведения. Артефакты цвета являются одними из основных видов артефактов при атаках печати и воспроизведения. Они обусловлены погрешностями принтеров и видеокамер при воспроизведении цветовой гаммы. Приведем основные артефакты, связанные с искажением цвета:

- цветовая гамма – изображение лица в виде распечатанной фотографии или на цифровом экране моделируется цветовой моделью CMYK- или RGB-преобразованием, что ограничивает цветовую гамму кожи. В результате подлинные изображения лиц имеют более богатую цветовую гамму, чем поддельные изображения;
- распределение цвета – дополнительные цветовые преобразования изменяют цветовое распределение, что приводит к хроматическим различиям по насыщенности цвета, пикам цветовой гистограммы и интервальным распределениям;
- цветовое искажение – поскольку устройства захвата имеют погрешности, поддельные изображения подвергаются цветовому искажению дважды, а подлинные изображения – только один раз, что приводит к более серьезным искажениям цвета поддельных изображений;
- цвет текстуры – атаки печати и воспроизведения искажают изображения кожи, что приводит к несоответствию хроматической текстуры поддельного и подлинного изображений;
- артефакты захвата изображения – поскольку поддельное изображение на фотографии или экране представлено массивом дискретных значений, возникают артефакты в виде муара при захвате видеокамерой поддельного изображения.

В соответствии с законом Ламберта яркость поддельного изображения I^L в точке с координатами (x, y) определяется как

$$I^L(x,y) = K(x,y)L_A, \tag{1}$$

где $K(x, y) \in [0, 1]$ – коэффициент отражения в точке (x, y), $L_A \in [0, \infty)$ – интенсивность окружающего освещения. При фиксированном значении L_A величина $I^L(x, y)$ определяется коэффициентом $K(x, y)$.

Предположим, что интенсивность в заданном пикселе (x, y) подлинного изображения I_R является линейной комбинацией яркости и цветности

$$I_R(x,y) = I_R^L(x,y) + I_R^C(x,y) = K_S(x,y)L_A + C_R(x,y)D_C, \tag{2}$$

где $I_R^L(x,y)$ – яркость подлинного изображения в точке (x, y), $I_R^C(x,y)$ – цветность подлинного изображения в точке (x, y), $K_S(x, y)$ – коэффициент отражения поверхности, $C_R(x, y)$ – цветовая гамма кожи, D_C – параметр искажения цвета, обусловленный устройством захвата.

В случае атаки печати интенсивность в заданном пикселе $I_P(x, y)$ при фиксированном окружающем освещении и искажении от устройства захвата определится следующим образом:

$$I_P(x,y) = K_P(x,y)L_A + I_R(x,y)D_PD_C, \tag{3}$$

где K_P $(x, y) \in [0, 1]$ – коэффициент отражения в точке (x, y) поддельного напечатанного изображения, I_R (x, y) – подлинное изображение лица или подлинное видео, использованное при атаке печати, D_P – параметр искажения цвета при атаке печати.

Подставив выражение (2) в выражение (3) и выделив компоненты яркости и цветности, получим следующие выражения для компонент яркости и цветности при атаке печати:

$$\begin{aligned} I_P^L(x,y) &= \left(K_P(x,y) + K_S(x,y) D_P D_C\right) L_A, \\ I_P^C(x,y) &= C_R(x,y) D_P D_C^2. \end{aligned} \tag{4}$$

Аналогичным образом можно представить выражения для компонент яркости $I_D^L(x,y)$ и цветности $I_D^C(x,y)$ при атаке воспроизведения:

$$\begin{aligned} I_D^L(x,y) &= \left(K_D(x,y) + K_S(x,y) D_T D_C\right) L_A, \\ I_D^C(x,y) &= C_R(x,y) D_T D_C^2, \end{aligned} \tag{4}$$

где $K_D(x, y) \in [0, 1]$ – коэффициент отражения в точке (x, y) экрана дисплея, D_T – параметр искажения цвета при атаке воспроизведения.

Из приведенных выражений (1)–(4) видно, что презентационные атаки вносят искажения в значения яркости и цветности. Далее рассмотрим, каким образом эти искажения проявляются на картах глубины подлинных и поддельных изображений лиц.

3. Построение карты глубины изображения лица. Часто задача обнаружения поддельных изображений лиц рассматривается как задача бинарной классификации. Однако в этом случае не учитывается природа спуфинговых изображений, например, потеря контурной информации, искажение цвета, появление муара и других артефактов. Один из способов преодоления таких проблем основан на разработке методов обнаружения презентационных атак с использованием глубины сцены. Интуитивно понятно, что подлинные изображения лиц имеют более явные признаки глубины, в то время как поддельные изображения лиц при атаках печати и атаках воспроизведения характеризуются наличием только планарной глубины.

Явный способ построения карт глубины основан на применении специальных дополнительных устройств, таких, как RGB-D датчики, стереокамеры, ToF-камеры (Time of Flight) и т.д. Однако более привлекательным является отсутствие дополнительных устройств при наличии соответствующих алгоритмов, усиливающих тонкую разницу между подлинными и поддельными изображениями.

Рассмотрим задачу построения карт глубины изображений лиц без предварительного этапа выравнивания таких изображений, например, по биометрическим точкам. Этап выравнивания необходим для изображений, полученных от систем городского видеонаблюдения. В большинстве наборов данных для обучения глубоких сетей изображения выравнены.

При построении карты глубины за основу взята глубокая сеть из работы [14], которая состоит из блоков свертки центральной разности и многомасштабного модуля внимания. Сеть упрощена за счет удаления многомасштабного модуля внимания. На вход сети подается изображение лица размерностью 256×256×3, а на выходе создается прогнозируемая карта глубины в градациях серого разрешением 32×32 пикселов. При этом среднеквадратичная ошибка потерь (mean square error, MSE) вычисляется по формуле:

$$L_{MSE} = \left\| D_P - D_G \right\|_2^2, \tag{5}$$

где D_P – прогнозируемая карта глубины, D_G – истинная карта глубины.

В качестве классификатора использованы два полносвязных слоя, которые классифицируют построенную карту глубины с помощью логистической функции на два класса. В

качестве функции потерь используется функция энтропии. Отличие предлагаемого подхода состоит в том, что на вход сети подаются не RGB-изображения, а изображения, преобразованные в другие цветовые пространства (в данном случае HSV). Архитектура базовой сети представлена на рисунке 1.

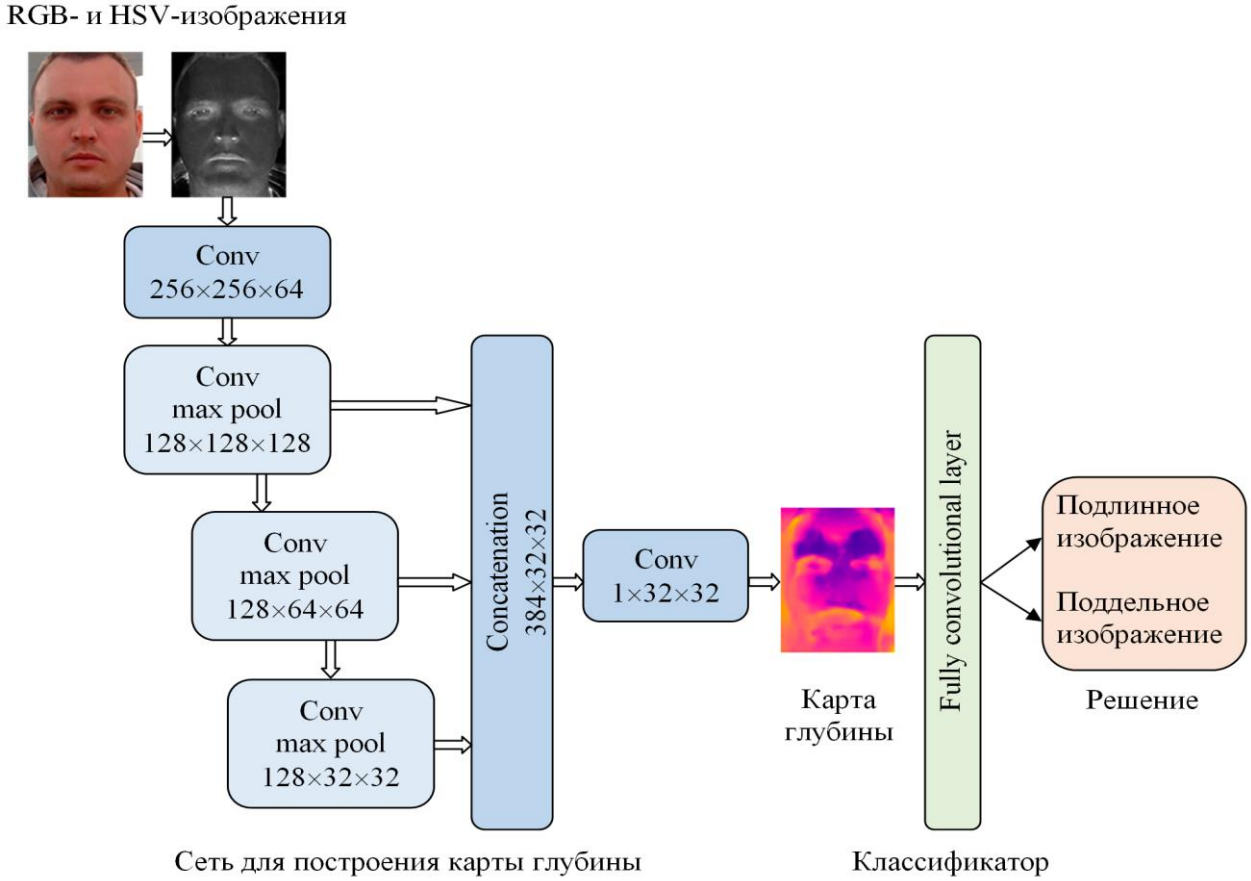

Рис. 1. Архитектура глубокой нейронной сети

4. Экспериментальные исследования. Для генерации карт глубины предложенная модель СНС обучалась на трех различных наборах данных:

- собственный набор данных, включающий 3120 изображений лиц разрешением 256×256 пикселов для 26 различных персон. Примеры карт глубины представлены на рисунке 2а;
- набор данных KITTI [15], включающий 42382 подготовленных стереоизображений разрешением 1242×375 пикселов. Примеры карт глубины приведены на рисунке 2б;
- набор данных Cityscapes [16], содержащий большое количество стерео видеопоследовательностей с аннотациями, записанных в 50 различных городах. Примеры карт глубины представлены на рисунке 2в.

При этом тестирование генерации карты глубины проводилось на наборе данных LCC FASD [17], включающем 1942 оригинальных изображений лиц и 16885 поддельных изображений лиц с применением 83 различных устройств, полученных из разных web-ресурсов.

На рисунке 3 приведены примеры карт глубины для подлинного и поддельных изображений, представленных на рисунке 2, но переведенных в цветовое пространство HSV.

Визуальная оценка свидетельствует, что в большинстве случаев (в зависимости от устройств захвата) карты глубины поддельных изображений, переведенных в цветовое пространство HSV, фиксируют более заметные отличия от карт глубины подлинных изображений, также переведенных в цветовое пространство HSV. При этом точность обнаружения поддельных изображений лиц возросла на 3-7% (таблица 1), что считается хорошим резуль-

татом, поскольку анализ карт глубины не является самостоятельным методом защиты систем биометрического распознавания.

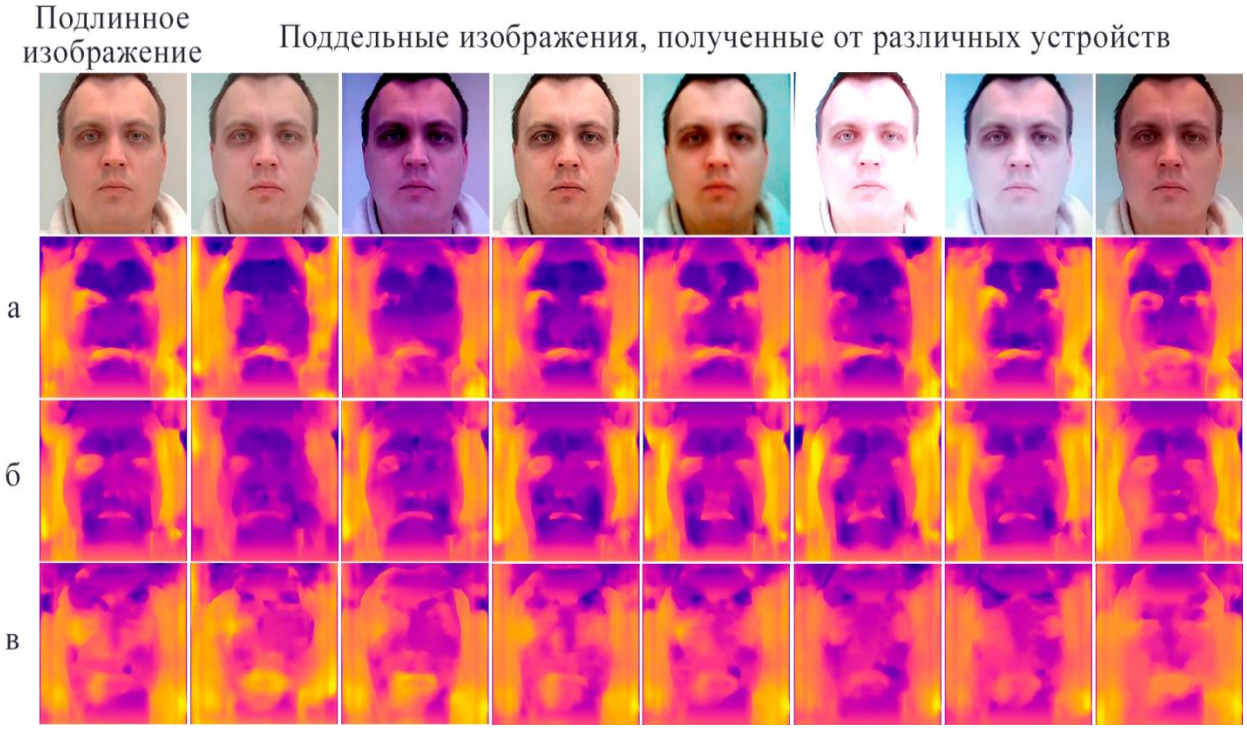

Рис. 2. Примеры карт глубины подлинного и поддельных RGB-изображений: (а) карты глубины изображений из собственного набора, (б) карты глубины изображений из набора KITTI, (в) карты глубины изображений из набора Cityscapes

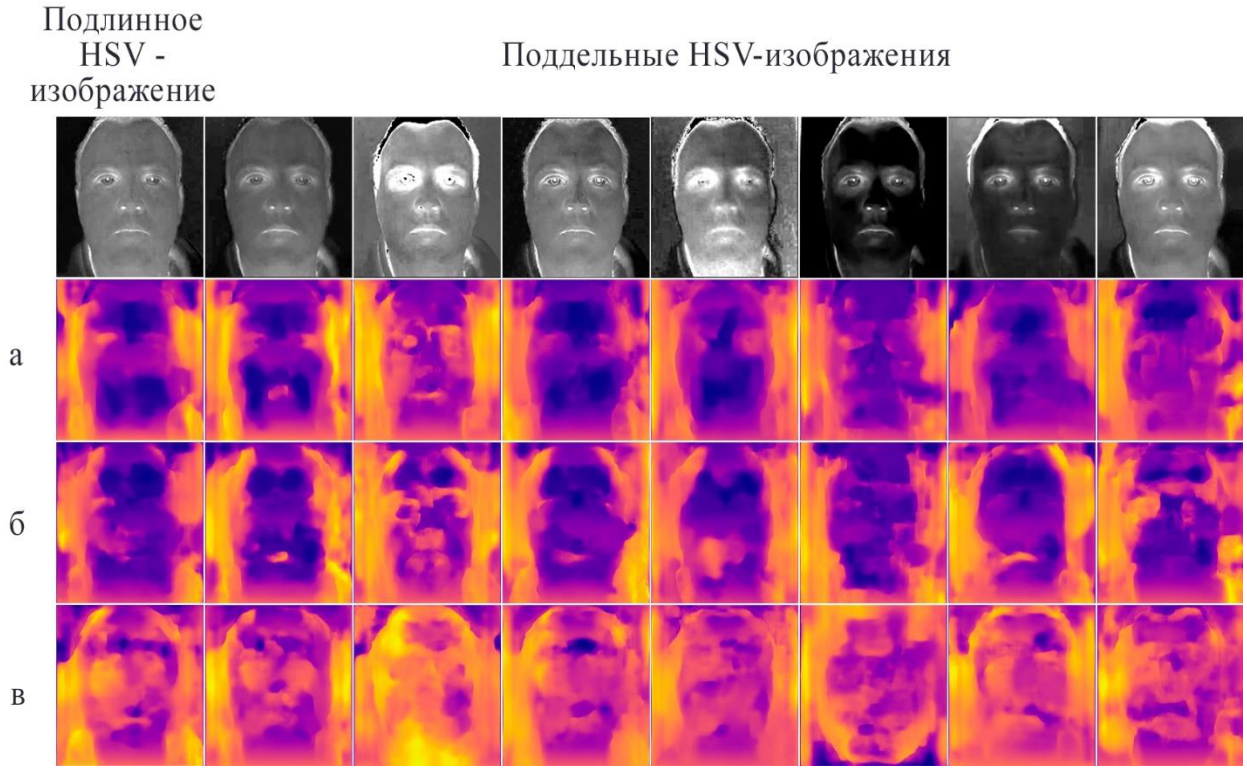

Рис. 3. Примеры карт глубины подлинного и поддельных HSV-изображений: (а) карты глубины изображений из собственного набора, (б) карты глубины изображений из набора KITTI, (в) карты глубины изображений из набора Cityscapes

Таблица 1. Точность обнаружения поддельных изображений при использовании карт глубины для различных наборов данных

Набор данных	Оригинальные изображения, %	В пространстве HSV, %
Собственный набор	84,4	87,4
KITTI	85,3	92,2
Cityscapes	86,3	89,6

Заключение. Проведенное исследование показывает целесообразность применения карт глубины с целью обнаружения презентационных атак в системах распознавания лиц. Карты глубины, построенные с помощью несложной глубокой нейронной сети, повышают точность обнаружения поддельных изображений лиц на 3-7% в зависимости от устройств захвата, условий освещения и настроек самого алгоритма.

Список источников

1. Määttä J., Hadid A., Pietikäinen M. Face spoofing detection from single images using micro-texture analysis // 2011 International Joint Conference on Biometrics (IJCB), IEEE, Washington, DC, USA, 2011, pp. 1-7.
2. De Freitas Pereira T., Anjos A., De Martino J.M., Marcel S. LBP-TOP based countermeasure against face spoofing attacks. Computer Vision - ACCV 2012 Workshops. ACCV 2012. Springer, Berlin, Heidelberg, 2012, vol. 7728, pp. 121-132.
3. De Freitas Pereira T., Anjos A., De Martino J.M., Marcel S. Can face anti-spoofing countermeasures work in a real world scenario? International Conference on Biometrics (ICB), IEEE, Madrid, Spain, 2013, pp. 1-8.
4. Komulainen J., Hadid A., Pietikäinen M. Context based face anti-spoofing. Sixth International Conference on Biometrics: Theory, Applications and Systems (BTAS), IEEE, Arlington, VA, USA, 2013, pp. 1-8.
5. Yang J., Lei Z., Liao S., Li S.Z. Face liveness detection with component dependent descriptor // International Conference on Biometrics (ICB), IEEE, Madrid, Spain, 2013, pp. 1-6.
6. Boulkenafet Z., Komulainen J., Hadid A. Face antispoofing using speeded-up robust features and Fisher vector encoding. Signal Processing Letters, 2016, vol. 24, no. 2, pp. 141-145.
7. Atoum Y., Liu Y., Jourabloo A., Liu X. Face anti-spoofing using patch and depth-based CNNs. International Joint Conference on Biometrics (IJCB), IEEE, Denver, CO, USA, 2017, pp. 319-328.
8. Liu Y., Jourabloo A., Liu X. Learning deep models for face anti-spoofing: binary or auxiliary supervision. In: 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition, IEEE, Salt Lake City, UT, USA, 2018, pp. 389-398.
9. Wang Z., Yu Z., Zhao C., Zhu X., Qin Y., Zhou Q., Zhou F., Lei Z. Deep spatial gradient and temporal depth learning for face anti-spoofing // Conference on Computer Vision and Pattern Recognition (CVPR), IEEE, Seattle, WA, USA, 2020, pp. 5042-5051.
10. Liu S., Yuen P.C., Zhang S., Zhao G. 3D mask face anti-spoofing with remote photoplethysmography // Leibe B., Matas J., Sebe N., Welling M. (eds.) Computer Vision – ECCV, Springer, Cham, LNCS, 2016, vol. 9911, pp. 85-100.
11. Silberman, N., Hoiem, D., Kohli, P., Fergus, R. Indoor segmentation and support inference from RGBD images. In: Fitzgibbon, A., Lazebnik, S., Perona, P., Sato, Y., Schmid, C. (eds) Computer Vision – ECCV, Springer, Berlin, LNCS, 2012, vol. 7576, pp. 746-760.
12. Liu Y., Jourabloo A., Ren W., Liu X. Dense face alignment // International Conference on Computer Vision (ICCV), IEEE, Venice, Italy, 2017, pp. 1619-1628.
13. Jourabloo A., Liu Y., Liu X. Face de-spoofing: Anti-spoofing via noise modeling. In: Ferrari, V., Hebert, M., Sminchisescu, C., Weiss, Y. (eds) Computer Vision – ECCV, Springer, Cham, LNCS, 2018, vol. 11217, pp. 297-315.
14. Yu Z., Zhao C., Wang Z., Qin Y., Su Z., Li X., Zhou F., Zhao G. Searching central difference convolutional networks for face anti-spoofing. Conference on Computer Vision and Pattern Recognition (CVPR), 2020, pp. 5295-5305.
15. The KITTI Vision Benchmark Suite. Available at: http://www.cvlibs.net/datasets/kitti/ (accessed:05.05.2022).
16. The Cityscapes Dataset . Available at: http://www.cityscapes-dataset.com. (accessed:05.05.2022).
17. Timoshenko, D., Simonchik, K., Shutov, V., Zhelezneva, P., Grishkin, V. Large crowdcollected facial anti-spoofing dataset. Computer Science and Information Technologies (CSIT), Yerevan, Armenia, 2019, pp. 208-211.

Фаворская Маргарита Николаевна. *Доктор технических наук, профессор, заведующий кафедрой информатики и вычислительной техники Сибирского государственного университета науки и технологий имени*

М.Ф. Решетнева. Область научных интересов: компьютерное зрение, обработка изображений и видеопоследовательностей, глубокое обучение, распознавание образов. AuthorID: 500950, SPIN-код: 7598-8467, ORCID: 0000-0002-2181-0454, favorskaya@sibsau.ru, Россия, г. Красноярск, пр. им. газ. Красноярский рабочий, 31.

Пахирка Андрей Иванович. *Кандидат технических наук, доцент кафедры информатики и вычислительной техники Сибирского государственного университета науки и технологий имени М.Ф. Решетнева. Область научных интересов: компьютерное зрение, обработка изображений и видеопоследовательностей, глубокое обучение, распознавание образов. AuthorID: 561608, SPIN-код: 1739-9950, pahirka@sibsau.ru, Россия, г. Красноярск, пр. им. газ. Красноярский рабочий, 31.*

UDC 004.93
DOI:10.38028/ESI.2022.27.3.005

Building depth maps for detection of presentation attacks in face recognition systems

Margarita N. Favorskaya, Andrey I. Pakhirka
Reshetnev Siberian State University of Science and Technology, Russia, Krasnoyarsk,
favorskaya@sibsau.ru

Annotation. Currently, face recognition systems are a widespread way of biometric identification in practice. However, such systems require protection against unauthorized actions in the form of so-called presentation attacks, when an attacker replaces genuine images with fake images or short video sequences. The article proposes a method for detecting presentation attacks using the depth of the scene without the use of special sensors. The challenge is to enhance the subtle difference between genuine and fake images. For this, a deep network was trained and tested, consisting of central difference convolution blocks and a multi-scale attention module. Experiments have shown that pre-processing input face images to the HSV color space has an advantage in the accuracy of detecting fake images. Thus, the detection accuracy on our own dataset, KITTI and Cityscapes datasets increased by 3-7% depending on the capture devices, lighting conditions, and settings of the algorithm.

Keywords: presentation attacks, face recognition, depth maps, deep learning

References

1. Määttä J., Hadid A., Pietikäinen M. Face spoofing detection from single images using micro-texture analysis // 2011 International Joint Conference on Biometrics (IJCB), IEEE, Washington, DC, USA, 2011. – C. 1-7.
2. De Freitas Pereira T., Anjos A., De Martino J.M., Marcel S. LBP-TOP based countermeasure against face spoofing attacks. Computer Vision - ACCV 2012 Workshops. ACCV 2012. Springer, Berlin, Heidelberg, 2012, vol. 7728, pp. 121-132.
3. De Freitas Pereira T., Anjos A., De Martino J.M., Marcel S. Can face anti-spoofing countermeasures work in a real world scenario? International Conference on Biometrics (ICB), IEEE, Madrid, Spain, 2013, pp. 1-8.
4. Komulainen J., Hadid A., Pietikäinen M. Context based face anti-spoofing. Sixth International Conference on Biometrics: Theory, Applications and Systems (BTAS), IEEE, Arlington, VA, USA, 2013, pp. 1-8.
5. Yang J., Lei Z., Liao S., Li S.Z. Face liveness detection with component dependent descriptor // International Conference on Biometrics (ICB), IEEE, Madrid, Spain, 2013, pp. 1-6.
6. Boulkenafet Z., Komulainen J., Hadid A. Face antispoofing using speeded-up robust features and Fisher vector encoding. Signal Processing Letters, 2016, vol. 24, no. 2, pp. 141-145.
7. Atoum Y., Liu Y., Jourabloo A., Liu X. Face anti-spoofing using patch and depth-based CNNs. International Joint Conference on Biometrics (IJCB), IEEE, Denver, CO, USA, 2017, pp. 319-328.
8. Liu Y., Jourabloo A., Liu X. Learning deep models for face anti-spoofing: binary or auxiliary supervision. In: 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition, IEEE, Salt Lake City, UT, USA, 2018, pp. 389-398.
9. Wang Z., Yu Z., Zhao C., Zhu X., Qin Y., Zhou Q., Zhou F., Lei Z. Deep spatial gradient and temporal depth learning for face anti-spoofing // Conference on Computer Vision and Pattern Recognition (CVPR), IEEE, Seattle, WA, USA, 2020, pp. 5042-5051.
10. Liu S., Yuen P.C., Zhang S., Zhao G. 3D mask face anti-spoofing with remote photoplethysmography // Leibe B., Matas J., Sebe N., Welling M. (eds.) Computer Vision – ECCV, Springer, Cham, LNCS, 2016, vol. 9911, pp. 85-100.

11. Silberman, N., Hoiem, D., Kohli, P., Fergus, R. Indoor segmentation and support inference from RGBD images. In: Fitzgibbon, A., Lazebnik, S., Perona, P., Sato, Y., Schmid, C. (eds) Computer Vision – ECCV, Springer, Berlin, LNCS, 2012, vol. 7576, pp. 746-760.
12. Liu Y., Jourabloo A., Ren W., Liu X. Dense face alignment // International Conference on Computer Vision (ICCV), IEEE, Venice, Italy, 2017, pp. 1619-1628.
13. Jourabloo A., Liu Y., Liu X. Face de-spoofing: Anti-spoofing via noise modeling. In: Ferrari, V., Hebert, M., Sminchisescu, C., Weiss, Y. (eds) Computer Vision – ECCV, Springer, Cham, LNCS, 2018, vol. 11217, pp. 297-315.
14. Yu Z., Zhao C., Wang Z., Qin Y., Su Z., Li X., Zhou F., Zhao G. Searching central difference convolutional networks for face anti-spoofing. Conference on Computer Vision and Pattern Recognition (CVPR), 2020, pp. 5295-5305.
15. The KITTI Vision Benchmark Suite. Available at: http://www.cvlibs.net/datasets/kitti/ (accessed:05.05.2022).
16. The Cityscapes Dataset. Available at: http://www.cityscapes-dataset.com. (accessed: 05.05.2022).
17. Timoshenko, D., Simonchik, K., Shutov, V., Zhelezneva, P., Grishkin, V. Large crowdcollected facial anti-spoofing dataset. Computer Science and Information Technologies (CSIT), Yerevan, Armenia, 2019, pp. 208-211.

Favorskaya Margarita Nikolaevna. *Doctor of Technical Sciences, Full Professor, Head of the Department of Informatics and Computer Techniques at Reshetnev Siberian State University of Science and Technology. Research interests: computer vision, image and video sequence processing, deep learning, pattern recognition. AuthorID: 500950, SPIN: 7598-8467, ORCID: 0000-0002-2181-0454, favorskaya@sibsau.ru, Russian, Krasnoyarsk, Krasnoyarsky Rabochy ave, 31.*

Pakhirka Andrey Ivanovich. *PhD, Associate Professor of the Department of Informatics and Computer Techniques at Reshetnev Siberian State University of Science and Technology. Research interests: computer vision, image and video sequence processing, deep learning, pattern recognition. AuthorID: 561608, SPIN: 1739-9950, pahirka@sibsau.ru, Russian, Krasnoyarsk, Krasnoyarsky Rabochy ave 31.*

Статья поступила в редакцию 05.08.2022; одобрена после рецензирования 06.09.2022; принята к публикации 16.09.2022.

The article was submitted 08/05/2022; approved after reviewing 09/06/2022; accepted for publication 09/16/2022.