Федеральный исследовательский центр «Информатика и Управление»
Российской академии наук

На правах рукописи

Ефимов Юрий Сергеевич

Методы детектирования подделок в биометрических системах на мобильном устройстве

05.13.17 – Теоретические основы информатики

ДИССЕРТАЦИЯ
на соискание учёной степени
кандидата технических наук

Научный руководитель
д. т. н.
Матвеев Иван Алексеевич

Москва – 2022

Оглавление

Введение

Биометрия представляет собой совокупность поведенческих, анатомических и физиологических характеристик, используемых для решения задачи идентификации человека. Этим термином также иногда называют и сам алгоритм решения задачи, и систему, в составе которой он находится. В основе технологий лежит свойство уникальности биометрической характеристики человека (индивидуума). В результате большого количества исследований сформировался список разнообразных типов таких характеристик, которые также называются *биометрическими модальностями*: изображения лица, радужной оболочки глаза (РОГ), рисунки папиллярных линий пальца, вен сетчатки глаза, особенности походки, голоса и др [66]. Упомянутый список постоянно пополняется, поскольку исследования в данной области в настоящий момент вызывают большой интерес как в индустриальных отраслях, так и в государственных структурах.

Задача подтверждения личности человека или задача аутентификации — бесспорно является актуальной для многих исследовательских групп по всему миру, как академических, так и индустриальных. Современные разработки в области автоматизации и цифровизации такой процедуры находят применение в различных системах контроля доступа, к которым можно отнести паспортные столы, контрольно-пропускные пункты и т.п. В подобных сценариях также возникает потребность в решении задачи идентификации, то есть определения личности человека путём поиска среди соответствий в объёмной базе данных. За последние два десятилетия произошёл ряд существенных прорывов в областях цифровой обработки и анализа изображений вкупе со значительным увеличением производительности вычислительных средств и систем компьютерного зрения [49, 71, 79, 132]. Столь заметный прогресс сыграл важную роль в процессе развития методов автоматического выделения и распознавания сложной биометрической информации о живых существах, которая является крайне из-

менчивой, слабо формализуемой и плохо моделируемой; позволил ставить и эффективно решать соответствующие задачи. Суммарный вклад исследователей в данной области формирует одну из молодых областей прикладной математики, биометрическую идентификацию [6].

Изображение РОГ, видеообраз лица и папиллярный узор пальца относятся к списку биометрических характеристик, стандартизованных ICAO для применения в международных электронных паспортах [65]. Помимо этого, первые две упомянутые модальностей применяются для осуществления автоматического паспортного контроля в ряде крупных аэропортов западных стран, таких как США, Канада, Великобритания и Нидерланды. Радужная оболочка глаза получила широкое внедрение в системах безопасности ОАЭ [61]: сообщается о 62 триллионах сравнений биометрических шаблонов, полученных в течение последних 10 лет. Более 1 млрд жителей Индии предоставили государственным органам миграционного контроля изображения своих лиц, РОГ и отпечатков пальцев рук в рамках программы UIDAI [133]. Аналитического агентство Business Wire сообщает, что объём рынка систем распознавания лиц составил \$5.1 млрд. в 2021 году, а к 2028 году увеличится более чем в 2 раза — до \$12.6 млрд. [39]. Рост интереса к биометрическим способам идентификации также наблюдается в связи с недавней пандемией вируса COVID-19 и потребности в уменьшении числа личных контактов.

С распространением технологий распознавания личности стала проявляться необходимость устойчивой защиты систем от попыток подлога и взлома чужими биометрическими шаблонами, которые в современной литературе называются *спуфингом (spoofing)*. Подавляющее большинство применяемых систем позволяет решать задачу идентификации с высокой точностью, но является уязвимым к попыткам обмана системы подделками. Для распространенных модальностей изображений лица и РОГ попытки подлога осуществляются в основном при помощи физических артефактов с изображением жертвы, таких как бумажные распечатки фотографий, записи с экранов или объемных объектов,

повторяющих трехмерное строение лицевой или периокулярной областей жертвы. Для повышения надежности биометрических систем необходима разработка дополнительных модулей, осуществляющих проверку подлинности пользователя на входном изображении. Подходы к решению данной задачи называются методами определения *живости* или *анти-спуфингом*.

Удобство биометрических методов аутентификации по сравнению с ПИН-кодами, паролями, смарт-картами и иными способами защиты информации создает потребность в их применении в составе современных мобильных устройств, которые предоставляют универсальные возможности по совершению банковских операций, личной переписки и обмена личными данными. Значительная часть появившихся на рынке в последнее десятилетие смартфонов оборудована компактными сенсорами распознавания. С каждым годом повышаются требования к безопасности пользовательских данных. Растет доля устройств, применяющих биометрическую аутентификацию. Эти факторы создают потребность в исследовании способов обмана систем распознавания методом подлога и разработке устойчивых и масштабируемых методов определения живости для мобильных систем.

Актуальность темы исследования. Способы проверки подлинности разного рода информации в обществе прошли долгий путь эволюционного развития от механических замков, ключей, систем печатей и кодовых фраз до подходов автоматизированной и автоматической аутентификации. Жизнь современного человека на регулярной основе включает разного рода верификации тех или иных персональных данных: осуществление финансовых транзакций, приобретение товаров и услуг, доступ к устройствам и сервисам, процедуры идентификации личности при пересечении границ и др.

Наблюдающийся в последнее десятилетие прирост мощности вычислительных устройств, совершенствование систем регистрации и обработки цифровых изображений, параллельное накопление значительных объёмов данных и развитие систем компьютерного зрения, машинного и, в особенности, глубокого

обучения позволили совершить значительный рывок технологий *биометрической идентификации*. Ключом доступа в данном случае выступает уникальная *биометрическая характеристика человека (БХЧ)* или *биометрическая модальность*. К популярным модальностям часто относят: папиллярный рисунок пальцев и ладони, изображение венозного русла кисти и ладони, особенности голоса, почерка, походки, изображения радужной оболочки и сетчатки глаза, изображения и форму лица.

Практически каждую из упомянутых БХЧ можно искусственно воспроизвести и предъявить биометрической системе с целью получения доступа к личной информации путем обмана. Различные модальности обладают различной сложностью подделывания, зависящей как от возможностей получения копии БХЧ, так и от сложности её воссоздания в условии ограниченных ресурсов. Процедура подделывания биометрических систем называется *спуфингом* (spoofing), а задача детектирования подлога — задачей определения *живости* или *антиспуфингом*.

Рост точности и производительности биометрических систем приводит к расширению области применения технологий автоматического распознавания человека. Современные мобильные устройства предоставляют пользователю широкий спектр возможностей по хранению значительных массивов данных, ведению личной и деловой переписки, осуществлению финансовых операций, доступу к защищённым цифровым ресурсам и др. В последнее десятилетие производители начали внедрять в смартфоны методы биометрической аутентификации как альтернативу паролям или цифровым кодам для ограничения доступа к персональной информации и повышения удобства использования [13, 51],

До событий 2020 года, связанных с пандемией COVID-19 и вызванным экономическим кризисом, наблюдались высокие темпы роста безналичных платежных транзакций согласно World Payments Records 2021 [135]. Суммарное их количество составило порядка 785 млрд. в 2020 году, и эксперты уверены в скором возвращении прежней скорости их прироста в связи с адаптацией общества

к новым реалиям и постепенному возврату к нормальной жизни. Реалии пандемии COVID-19 подстегнули интерес к безналичным расчетам, в том числе совершаемым при помощи повсеместно распространённых мобильных устройств. По данным WPR в настоящий момент суммарная доля мобильных транзакций в мире составляет 45%, а в период 2020-2025 гг. ожидается прирост их доли в 21.5%. Платёжные системы в современных смартфонах требуют идентификации пользователя (при помощи ПИН-кода и др.), причём набирают популярность биометрические методы, предоставляемые со стороны коммерческих банков [1].

Помимо рынка мобильных устройств отмечается рост спроса к цифровизации и персонификации бытовых услуг и сервисов. К таковым относятся концепции «умного дома» (Smart Home), виртуальных помощников (Smart Assistant и др.), модель «интернета вещей» (Internet of Things) и многое другое. Практическое применение таких элементов быта не требует непосредственного участия пользователя, подразумевая т.н. некооперативное распознавание. Подавляющее большинство перечисленных выше приложений требуют присутствия системы автоматической идентификации/аутентификации личности.

Как следствие, расширение области применения технологий биометрической идентификации и аутентификации порождает множество актуальных задач, требующих решения в связи с растущими потребностями современного общества в высоком уровне безопасности личных данных и удобстве применяемых в быту услуг и сервисов. Среди таких задач выделяется проблема определения живости участника процедуры распознавания и обнаружения попыток взлома системы при помощи искусственно созданных БХЧ (т.н. подделок), поскольку именно этот компонент системы в первую очередь определяет уровень защиты, достигаемый при её использовании.

Особенно актуальной эта задача является мобильных биометрических приложений по ряду причин. Системы некооперативного распознавания в мобильных устройствах и приложениях требуют удобства и быстродействия для поль-

зователя, а также устойчивости к изменчивости окружения и самой БХЧ. В результате происходит ужесточение ограничений на средства регистрации изображений, применяемые алгоритмы распознавания и противодействия подложным попыткам входа. От мобильной биометрической системы требуется возможность работы в режиме реального времени при низком количества ошибок ложного недопуска *(False Rejection Rate — FRR)*, даже для входных данных низкого качества. При этом сохраняется потребность в высоком уровне предоставляемой защиты, в том числе и от взлома при помощи поддельных БХЧ, что соответствует низкому количеству ошибок ложного допуска *(False Accept Rate — FAR)*. Наконец, реализация системы распознавания зачастую происходит на устройствах с сильно ограниченными вычислительными ресурсами.

Среди наиболее известных исследовательских групп: Cambridge University, Великобритания (J. Daugman); Michigan State University, США (A. Ross, X. Liu); , Warsaw University of Technology, Польша (A. Czajka), University of OULU (J. Komulainen); Institute of Automation of the Chinese Academy of Sciences, КНР (T. Tan), Idiap Research Institute, Швейцария (S. Matcela), в том числе и несколько российских: Федеральный Исследовательский центр «Информатика и управление» РАН (д.т.н. И.А. Матвеев), МГУ им. Ломоносова (д.ф-м.н. А.С. Крылов), Пензенский государственный университет (д.т.н. А.И. Иванов); НИИЦ БТ МГТУ им. Н. Э. Баумана; Институт систем обработки изображений РАН и др. Область биометрического распознавания в настоящее время также привлекает значительное количество коммерческих компаний, занимающихся как непосредственно созданием соответствующих технологий, так и вынужденных разрабатывать собственные решения для растущих потребительских нужд в сфере безопасности данных. В качестве примеров успешного внедрения получаемых решения для первой группы можно привести российские компании NTechLabs и VisionLabs, ко второй — крупные банки Сбербанк, ВТБ, Tinkoff.

Наиболее уязвимыми с точки зрения возможности спуфинга модальностями для мобильных систем являются изображения радужной оболочки глаза и

видеообраза лица ввиду сравнительно небольшой сложности процедуры подделывания. Поэтому актуальными направлениями развития области определения живости в настоящее время являются: разработка высокопроизводительных методов анти-спуфинга для видеообраза лица в условиях некооперативного распознавания при помощи смартфона; разработка высокопроизводительных методов обнаружения подделок при помощи вспомогательных сенсоров, таких как мобильная стереокамера; разработка новых методов противодействия новым способам взлома мобильных систем распознавания по РОГ; разработка методов выделения границ радужки на изображениях как высокого, так и низкого качества.

Цели и задачи диссертационной работы:

В работе были поставлены следующие **цели**:

- Создать методы и алгоритмы для автоматического определения живости пользователя и обнаружения попыток взлома при помощи подделок в системах распознавания по видеообразу лица, оборудованных единственной камерой для съемки в видимом спектре излучения, способные обрабатывать каждое изображение с частотой поступления кадров на мобильном устройстве, удовлетворяющие критериям ошибок: уровень ложных недопусков не более 3% при уровне ложного допуска не более 1%;

- Создать методы и алгоритмы выявления подделок лица при использовании пары камер с малым стереобазисом, способные обеспечивать защиту от распространённых видов атак и имеющие достаточное для мобильных приложений быстродействие;

- Разработать методы и алгоритмы поиска границ радужной оболочки глаза для входных данных как высокого, так и низкого качества, характерного для мобильных биометрических систем;

- Создать методы и алгоритмы определения живости для мобильной си-

стемы распознавания по радужке, способные обеспечивать защиту, в том числе, от ранее не исследованных видов взлома при помощи подделок.

Для достижения поставленных целей были решены следующие **задачи**:

- Исследование и разработка методов определения живости человека по видеообразу лица, удовлетворяющих критериям, необходимым для обеспечения возможности их применения в мобильном устройстве;
- Исследование и разработка методов обнаружения подделок лица человека с применением стереоинформации, извлекаемой при помощи камеры мобильного устройства с малым стереобазисом;
- Исследование и разработка методов поиска области радужки на изображении низкого качества путем аппроксимации ее границ окружностями с возможностью применения для растров, извлекаемых в условиях применения мобильной биометрической системы;
- Исследование и разработка методов противодействия подделкам изображений радужки для мобильных систем распознавания;
- Сбор и разметка баз данных, в которых представлены изображения и последовательности изображений, реализующие приведенных выше задачи;
- Создание среды, программных средств и проведение вычислительных экспериментов по определению работоспособности перечисленных методов с опорой на собранные базы данных;
- Создание программных средств (библиотеки и демо-приложений) для апробации реализованных методов на мобильном устройстве.

Научная новизна.

- Предложен новый эффективный метод защиты от подделывания в системах распознавания по видеообразу лица, обладающий многостадийной структурой и способный работать на мобильном устройстве с ограниченными вычислительными возможностями в режиме реального времени в сценариях изменяющихся условий окружения;

- Предложен новый надёжный метод защиты от подделывания изображения лица для мобильных систем, оборудованных стереокамерой с малым стереобазисом, обеспечивающий защиту от распространённых видов атак;

- Предложен новый высокопроизводительный метод аппроксимации границ радужки с применением методологии глубокого обучения, допускающий применение для изображений как высокого, так и низкого качества;

- Разработан новый надёжный метод определения живости радужки на изображении глаза, способный обнаруживать новые ранее не использовавшиеся виды взлома системы распознавания при помощи подделокю

Теоретическая и практическая значимость. Результаты, изложенные в диссертации, используются в мобильных устройствах, выпускаемых компанией Samsung Electronics Co. Ltd. Среди устройств — флагманские модели, выпускаемые компанией в период с 2018 по 2021 гг.: смартфоны Samsung Galaxy S9/S9+, смартфон Samsung Galaxy Note9, планшет Samsung Galaxy Tab S4, смартфоны Samsung Galaxy S10e/S10/S10+, смартфоны Samsung Galaxy Note10/Note10 Ultra, смартфоны Samsung Galaxy S20/S20+/S20 Ultra, смартфоны Samsung Galaxy Note20/Note20 Ultra, смартфоны Samsung Galaxy Fold/Z Fold2/Z Fold3, смартфоны Samsung Galaxy S21/S21+/S21 Ultra.

Положения, выносимые на защиту:

- Выделены специфические качества методов определения живости видеообразов лица и радужки в системах биометрического распознавания, используемых в мобильных устройствах, описаны основные ограничения и

требования, предъявляемые к алгоритмам определения живости и защиты от подделок;

- Разработан и внедрён многостадийный метод определения живости по видеообразу лица для пользователей смартфонов, оборудованных единственной фронтальной камерой; предложена методология сбора репрезентативной базы данных и с ее помощью получена база изображений лиц в условиях, имитирующих применение системы распознавания человеком в повседневной жизни, осуществлена программная реализация метода;

- Описаны и исследованы виды подделок лица, которые могут быть обнаружены при использовании мобильных стереокамер с малым базисом, предложена методология сбора и с её помощью получена собрана база стереоизображений подлинных лиц и подделок, предложен метод защиты от взлома с высокой обобщающей способностью, произведено тестирование на открытой базе стереоизображений лиц;

- Выделена группа методов поиска границ радужки на изображении для мобильных биометрических приложений, разработан и программно реализован нейросетевой метод решения задачи, произведена его оценка и сравнение с описанными в литературе решениями;

- Описаны и исследованы новые способы изготовления подделок радужки, собрана база данных изображений подлинных и искусственных образцов, разработаны метод распознавания живости глаза, устойчивый к новым видам подделок глаз, произведено его сравнение с аналогами из литературы по качеству решения задачи и производительности.

Степень достоверности и апробация результатов. Достоверность результатов обеспечивается обширным анализом работ в области исследования, описанием проведённых экспериментов, их воспроизводимостью, апробацией результатов на практике. Основные результаты диссертации докладывались на

следующих конференциях: 20-я Всероссийская конференция с международным участием «Математические методы распознавания образов» (ММРО-2021), Москва, 2021; 64-я Всероссийская научная конференция МФТИ, Москва, 2021; International Conference on Pattern Recognition and Artificial Intelligence, Montreal, Canada, 2018; 19-я Всероссийская конференция с международным участием «Математические методы распознавания образов» (ММРО-2019), Москва, 2019; Intelligent Data Processing Conference, Gaeta, Italy 2018; Intelligent Data Processing Conference, Barcelona, Spain, 2016; XXI Международная научно-техническая конференция студентов, аспирантов и молодых учёных «Научная сессия ТУСУР», Томск, 2016.

Публикации. Материалы диссертации опубликованы в 17 печатных работах, из них 6 в журналах из списка ВАК и индексируемых в WoS, Scopus.

Личный вклад автора. Содержание диссертации и основные положения, выносимые на защиту, отражают персональный вклад автора в опубликованные работы. Подготовка к публикации полученных результатов проводилась совместно с соавторами, причём вклад диссертанта был определяющим. Все представленные в диссертации результаты получены лично автором.

Структура и объем диссертации. Диссертация состоит из введения, 5 глав, заключения и библиографии. Общий объём диссертации 136 страниц, из них 115 страниц текста, включая 27 рисунков. Библиография включает 156 наименований на 18 страницах.

Глава 1

Методы биометрии

1.1. Обзор подходов биометрического распознавания

Биометрия (или биометрика) — область знаний, изучающая методы и средства измерения и формализации поведенческих, анатомических и физиологических характеристик человека с целью их использования для решения задач верификации или идентификации его личности. Этим термином называют как сам алгоритм решения таких задач, так и систему, в составе которой он находится. Результаты измерения упомянутых характеристик называются *Биометрической характеристикой человека (БХЧ)* или *биометрической модальностью*. Подразумевается, что процедура распознавания личности реализуется путем сравнения полученной БХЧ с набором ранее зарегистрированных, совокупность которых называют биометрическим эталоном или шаблоном.

Все БХЧ могут быть поделены на две группы: физиологические (статические) и поведенческие (динамические) [9]. Биометрические модальности могут быть отнесены к одной из двух групп: физиологические или поведенческие. Первая категория как правило содержит статические или малоизменяющиеся во времени характеристики фенотипа человека, в то время как ко второй относят динамические измерения, присущие индивидуумам. Принято считать наиболее распространенными следующие БХЧ:

1. Физиологические БХЧ:
    а. Видеообраз лица: овал, форма, размер отдельных деталей, геометрические параметры (расстояние между его определенными точками), узор подкожных кровеносных сосудов и др.;
    б. Структура радужной оболочки глаза;

в. Структура кровеносных сосудов на сетчатке глаза;

г. Особенности папиллярного узора одного или нескольких пальцев, ладони: параметры минуций (координаты, ориентация), параметры пространственно-частотного спектра и др.;

д. Особенности строения ладони: геометрия (ширина, длина, высота пальцев, расстояние между определенными точками), неровности складок кожи, рисунок вен, папиллярный рисунок ладони и др.;

е. Особенности уха: форма (контур, наклон, козелок, противокозелок, форма и прикрепление мочки), геометрические параметры уха (расстояние между определенными точками) и др.;

2. Поведенческие БХЧ:

а. Голосовые характеристики: тембр, частотный спектр и др.;

б. Динамика походки;

в. Рукописный почерк и др.

В процессе распознавания участвует автоматизированный алгоритм принятия решений, который также называют биометрической системой. Компоненты модуля биометрического распознавания должны реализовывать следующие операции [9]:

- регистрации выборки БХЧ от конкретного пользователя;
- формирование вектора биометрических данных из выборки БХЧ;
- формирование биометрического вектора признаков;
- сравнение биометрических векторов признаков с эталонами (шаблонами);
- принятие решения о соответствии сравниваемых БХЧ;
- формирование результата о достижении идентификации (верификации);

- принятие решения о повторении, окончании или видоизменении процесса идентификации (верификации).

Первоочередной задачей при проектировании биометрической системы считается определение источника БХЧ. Идеальная БХЧ должны быть всеобщной, уникальной, постоянной и измеримой. Всеобщность требует присутствия биометрической характеристики у каждого человека. Уникальность подразумевает невозможность существования двух индивидуумов с одинаковыми параметрами БХЧ. Постоянство означает отсутствие зависимости черт биометрической модальности от времени. Измеримость характеризует возможность извлечения БХЧ для каждого человека при помощи некоторого устройства или сенсора.

Реальные БХЧ не идеальны и это ограничивает их применение. Тем не менее, к наиболее оптимальным для практического использования с учетом упомянутых требований можно отнести видеообраз лица, изображение радужки, отпечаток пальца [7].

1.2. Биометрические приложения

Повсеместно известными и привычными методам аутентификации являются: кодовые фразы и пароли в банковских системах, ПИН-коды пластиковых карт и мобильных устройств, электронные ключи доступа и др. Альтернативой таких подходам является применение биометрических технологий, которые позволяют повысить безопасность доступа к конфиденциальным данным без создания неудобств для пользователя.

В настоящее время биометрические методы применяются наиболее часто в следующих целях:

- Автоматизированной проверки личности при пересечении границ государств, при входе на охраняемый объект, предоставление доступа к личным устройствам, хранилищам данных, электронным ресурсам, банковским ячейкам, вкладам и т.п.

- Подтверждения и контроля финансовых операций: снятие наличных в банкомате, онлайн-платежи и т.п.

неинвазивность биометрических технологий совместно с возможностью повышения уровня безопасности при аутентификации личности создает спрос на их внедрение в самые разные области современного быта. Наиболее активный интерес возникает со стороны сферы государственного контроля границ. Характерными примерами являются обязательство сдачи биометрических данных при получении заграничного паспорта гражданами российской федерации (изображения лица и текстура отпечатков пальцев) и внедрение универсальных электронных карт за рубежом (ID). Не менее значительный спрос на биометрические решения создает финансовая область. Так, крупные банки внедряют дополнительные проверки личности, основанные на технологиях распознавания лиц и голоса, а крупные производителей мобильных устройств и операционных систем (Apple, Google, Samsung) создают и внедряют бесконтактные системы оплаты при помощи смартфонов (Apple Pay, Android Pay, Samsung Pay), выступая в роли посредников. В настоящее время в развитых странах большинство бытовых транзакций осуществляется при помощи мобильных устройств. Во многом этому способствует внедрение бесконтактных систем оплаты, цифровизация транспорта и рост популярности служб онлайн-заказов и доставки товаров потребления в крупных городах.

1.3. Общее строение биометрических систем

Классические методы идентификации по биометрическим чертам имеют в большинстве случаев общее строение, Рис. 1.1.

Первичный этап (блок 1) как правило соответствует регистрации цифрового образа БХЧ, к примеру, изображения радужки или видеообраза лица. Далее (блок 2) осуществляется оценка качества полученного цифрового слепка с точки зрения его пригодности для выделения областей интереса и отсечения шумовой

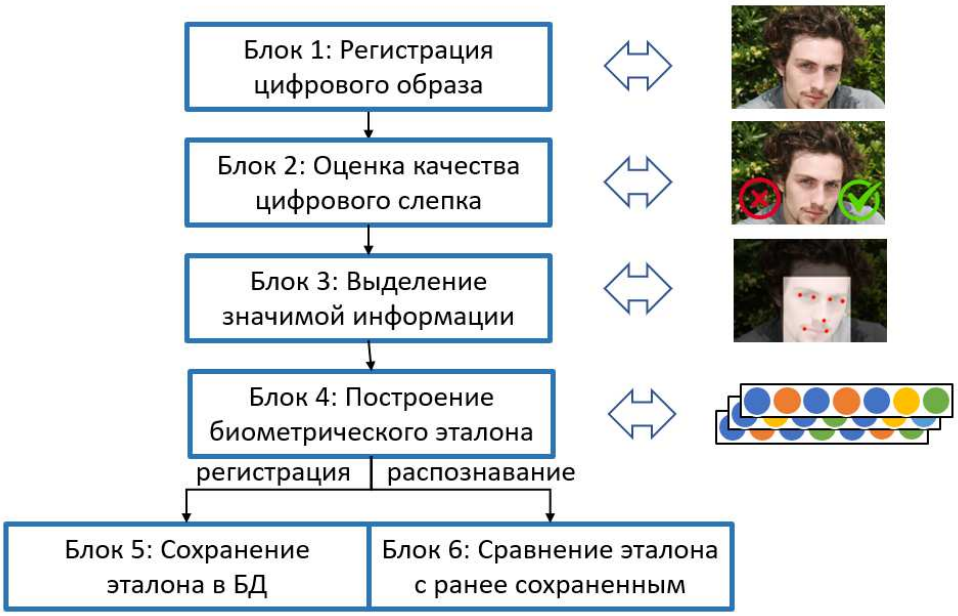


Рис. 1.1. Обобщенное строение методов биометрической идентификации.

информации, а также для формирования биометрического эталона. Подсистема проверка живости БХЧ как правило считается компонентом данного модуля. Эта часть системы может как состоять из нескольких смысловых этапов, так и быть распределенной между иными компонентами. Последующий этап (блок 3) отвечает за исключение фоновой и шумовой информации из цифрового образа БХЧ. К примеру, в случае с изображением радужки данные блок выделяет область ее текстуры между зрачком и склерой, отсекая при этом побочную информацию в виде век, ресниц, бликов и т.п. Для видеообраза лица аналогичная процедура выделяет искомую область и выполняет фронтализацию растра лицевой области: приведение наблюдаемой позы головы к фронтальной. Затем, после подавления побочной информации, осуществляется построение биометрического эталона (блок 4). Наконец, в случае сценария регистрации, он заносится в БД (блок 5), а в случае сценария верификации — сравнивается ранее попавшими в БД эталонами (блок 6).

1.4. Подделывание биометрии

Применимость биометрических систем идентификации на практике во многом определяется их возможностями по обнаружению предъявленных им подделок. Саму попытку такой фальсификации в литературе часто называют *спуфингом* или *спуфинг-атакой (spoofing attack, presentation attack)*. Практически любая БХЧ может быть имитирована искусственно с наперед заданным уровнем качества при наличии достаточных финансовых и временных ресурсов [91].

Основной сложностью при создании подделок является доступ к биометрическим данным жертвы. которые требуется имитировать. В зависимости от вида БХЧ, при извлечении информации о ней в системе может применяться как доступные в быту сенсоры (камеры видимого диапазона света), так и узкоспециализированное оборудование (дактилоскопические сенсоры, инфракрасные камеры с активной подсветкой). К примеру, в случае с видеообразом лица для создания подделки достаточно получить снимок жертвы в высоком разрешении и при естественном равномерном освещении, что не составляет большого труда в современных реалиях при большой распространенности социальных сетей и медиа. При этом регистрация термограммы лицевой области требует не только кооперации со стороны жертвы, но и специальной и как правило дорогостоящей аппаратуры, характеристики сенсора которой должны быть близки к таковым в атакуемой биометрической системе.

Значительные усилия исследователей сосредоточены на противодействии так называемым «спуфинг-атакам» (spoofing attack, presentation attack). Задачу детектирования подделок часто называют «анти-спуфингом» или определением живости. При решении подобной задачи как правило строится логичное предположение о том, что наблюдаемые при помощи сенсора биометрической системы характеристики подделки будут отличаться от таковых у истинной попытки распознавания вследствие различий их физических свойств. Такие характеристики могут быть статическими, т.е. проявляющимися в каждый момент времени, и

динамическими, т.е возникающими в процессе распознавания пользователя.

Стоит упомянуть, что существует два вида спуфинг-атак по их предназначению. К первой группе относятся попытки сокрытия личности участника процесса распознавания за счет искусственного искажения наблюдаемых БХЧ. Вторую группу формируют способы имперсонификации жертвы злоумышленником. В данной работе рассматриваются лишь методы второй группы, как представляющие наибольшую опасность для бытового применения биометрических систем, в то время как первая группа спуфинг-атак вызывает большой интерес исследователей в области судебной экспертизы и криминалистики [91].

Попытки подделывания как правило осуществляются при помощи статичных объектов, в отличии от живого и как правило подвижного участника процедуры распознавания. Решение задачи в таком случае естественным образом упрощается, если допускается прямое взаимодействие с субъектом: при попытке идентификации или верификации требуется выполнить набор движений, определяемых самой биометрической системой. Методы детектирования подделок, использующие упомянутую выше интуицию, называются кооперативными. Подобные подходы заметно повышают устойчивость системы к спуфинг-атакам [3], но снижают применимость системы в целом, вследствие повышения времени распознавания и возникновения дополнительных сложностей для пользователей из-за необходимости взаимодействия.

Некоторые различия между подлинными и искусственными БХЧ могут быть слабо определяемыми при формировании цифрового образа сенсорами биометрической системы. В таком случае допускается добавление в нее дополнительного источника информации, а именно вспомогательного сенсора, позволяющего наблюдать характерные для подделок артефакты, отличающие их от живых субъектов. С аналогичной целью иногда осуществляется комбинирование модальностей внутри одной биометрической системы. К примеру, системы распознавания по видеообразу лица могут быть дополнены камерой определения глубины сцены, которая позволяет использовать информацию о геометрии

для достоверного недопуска плоских подделок вида распечаток или видеоповторов на цифровых дисплеях. В то же время, сама геометрия лица может быть рассмотрена как отдельная БХЧ. Подобные методы определения подделок при помощи вспомогательной информации от дополнительных сенсоров формируют группу, которая в литературе называется «hardware-based». Существенным недостатком таких подходов, однако, считается повышение стоимости биометрической системы в целом, что может быть критично для ее массового применения в быту. Рост цены вызван не только расходами на дополнительную аппаратуру, но и возрастающими издержками при интеграции компонентов системы в одно целое и поддержке такой комбинированной структуры.

Сложность решения задачи определения спуфинг-атак связана в первую очередь с тем, что невозможно заранее определить набор возможных способов подделывания системы, учесть все возможные слабые места используемых подходов и аппаратных средств. При разработке метода детектирования живости исследователь ограничен доступными ресурсами и выборкой данных, в то время как возможности злоумышленника при наличии длительного доступа к системе можно считать неограниченными, поскольку требуется найти единственную уязвимость. Кроме того, имеющиеся в открытом доступе наборы данных для тестирования методов детектирования спуфинга, как правило, предоставляют скудный набор примеров подделок, а самостоятельный сбор объемной базы примеров истинных и искусственных БХЧ является крайне трудоемким и затратным процессом. Поставленную можно считать решаемой лишь при ограничении набора рассматриваемых способов подделывания. К примеру, допускается рассмотрение лишь сравнительно недорогих спуфинг-атак, стоимость создания которых сопоставима с ценностью информации, доступ к которой ограничивается при помощи биометрической системы. Так, с целью взлома возможно создание высококачественной подвижной маски-копии лица жертвы из материалов, близких к живым тканям по характеристикам. Для системы распознавания по видеообразу лица попытка распознавания подобной подделкой будет почти на-

верняка успешной, но стоимость создания подобной искусственной БХЧ может превышать несколько тысяч долларов США. Технологии создания столь качественных дубликатов лица не являются массовыми и практически недоступны обывателю. Вероятность создания и применения столь затратной атаки является практически нулевой для бытовых биометрических систем.

Таким образом, при создании методов определения живости для той или иной БХЧ требуется учитывать сценарии использования системы и присущие таковым ограничения удобства применимости, а также диапазон способов подделывания, противодействие которым будет целесообразно с точки зрения стоимости их создания и вероятности их возникновения.

1.5. Противодействие подделкам в мобильной биометрии

В данной работе предлагается рассматривать методы детектирования подделок для БХЧ изображения радужки в ИК диапазоне и видеообраза лица человека. Эти модальности можно считать одними из наиболее распространенных в мобильных биометрических приложениях и в то же время связанными общей идеей наличия наблюдаемой при помощи некоторой камеры характеристики, подверженной изменчивости в зависимости от условий окружения.

Среди представленных на рынке смартфонов, оборудованных биометрическими системами, можно выделить несколько групп по частоте применения той или иной БХЧ. К наиболее популярным можно отнести методы идентификации личности по видеообразу лица и отпечатку пальца. Первая из упомянутых БХЧ естественным образом может быть получена при помощи фронтальной камеры современного смартфона, широко распространенного элемента конструкции ввиду развития социальных сетей. Ряд устройств с операционной системой Google Android впервые получили возможность идентификации пользователя по снимку лица в 2011 году [12]. Различные виды сенсоров для получения рисунка папиллярных линий пальца человека также получили широкое распростра-

нение начиная с 2014 года с релизом Apple iPhone 5s [14], несмотря на первые попытки внедрения технологии в 2011 году [97].

К менее популярным биометрическим модальностям, применяемым в мобильных устройствах, стоит отнести изображение радужки глаза в коротковолновом инфракрасном (ИК) диапазоне [41, 95, 116] и геометрию объема лица [13, 51]. Успешные попытки внедрения данных технологий произошли в 2015 и 2017 годах соответственно и были усовершенствованы в следующих поколениях выпущенных устройств. Системы распознавания по геометрии лица и их вариации в зависимости от типа сенсора, применяемого для извлечения информации о глубине сцены, имеют популярность у производителей крупных брендов, выпускаются и поддерживаются по настоящее время.

К группе редких биометрических признаков, применяемых в современных смартфонах, можно отнести системы распознавания по рисунку вен руки человека. Единственная попытка коммерциализации подобного решения случилась в 2019 году [81] и не получила широкого распространения.

Особенности взаимодействие и широта возможных сценариев применения мобильного устройства формируют дополнительные требования к применяемой в нем биометрической системе с подсистемой детектирования подделок (Рис. 1.2). Постоянно изменяющиеся условия среды не должны существенно влиять на точность решения задачи анти-спуфинга. Мобильное устройство может применяться в солнечную погоду, в помещении, в полумраке с низким уровнем освещения. Пользователь может быть предъявлен системе с носимыми очками, головными уборами, медицинскими масками и иными перекрытиями областей лица и глаз. Тряска рук, возможные изменения направления взгляда, моргания и иные особенности поведения пользователя устройства в повседневной жизни не должны влиять на удобство использования биометрической системы [7]. Высокая точность решения задачи определения живости должна достигаться в режиме реального времени на мобильном устройстве с существенными ограничениями вычислительных ресурсов и доступной для использования

памяти.

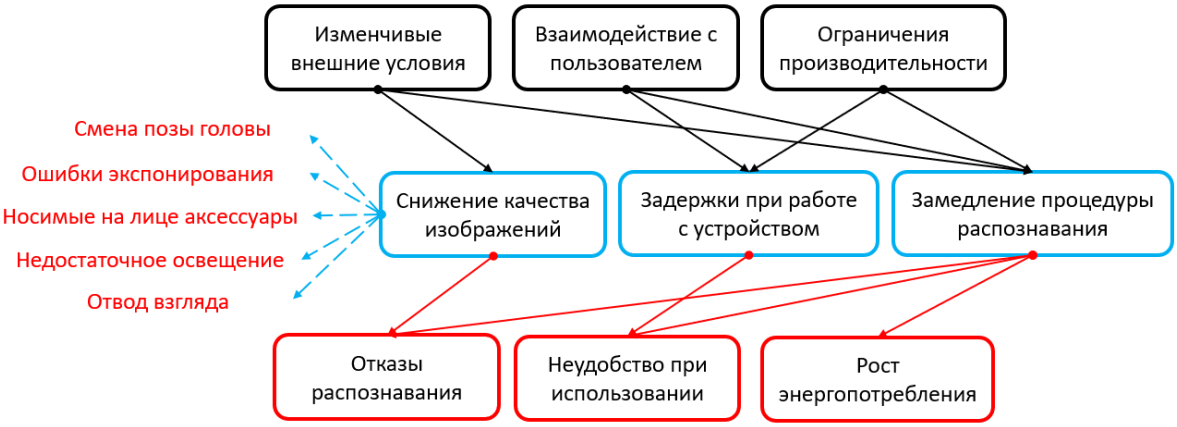


Рис. 1.2. Основные проблемы при определении живости БХЧ с мобильного устройства

Описанные ограничения приводят к ухудшению качества изображений, используемых в подсистемах детектирования подделок. В результате набор наблюдаемых различий между искусственными и подлинными БХЧ сокращается, размывается граница между классами «живой» и «подделка». Требование высокой производительности в условиях ограниченных вычислительных ресурсов сужает диапазон доступных для применения в мобильных биометрических системах методов.

Как правило не допускается применение кооперативных подходов, основанных на взаимодействии с пользователем, поскольку суммарное время отклика системы при попытке распознавания превышает комфортное с точки зрения удобства использования устройства (1-2 секунды).

Аналогично, встраивание вспомогательных сенсоров для определения живости БХЧ не столь широко распространено в мобильных биометрических системах ввиду возрастающей стоимости смартфона, которая во многом определяет его конкурентоспособность и привлекательность для покупателей. Более того, подобные вспомогательные технические элементы могут вызвать проблемы компоновки корпуса устройства, что негативно влияет на дизайн, чрезвычайно важную с точки зрения характеристику устройства.

Таким образом, подсистема обнаружения попыток подделывания в мобильных биометрических приложениях требует разработки некооперативных методов, устойчивых к изменчивости условий окружения при съемке участника процедуры распознавания. В большинстве случаев такие приложения не допускают применения вспомогательных аппаратных решений для определения живости человека, вынуждая исследователей полагаться на уже присущий мобильным устройствам набор сенсоров для извлечения биометрической информации. Используемые алгоритмы должны обладать низкой вычислительной сложностью для возможности внедрения в мобильные биометрические системы.

1.6. Выводы к первой главе

Произведен обзор наиболее распространенных биометрических характеристик и методов распознавания человека. Рассмотрены основные области применения и перспективы развития биометрических решений. Приведена общая схема биометрической идентификации и место подсистемы обнаружения подделок в ее составе. Описаны общие группы подходов к построению подсистем детектирования подделок. Рассмотрены наиболее широко используемые БХЧ для мобильных систем идентификации человека. Описана специфика решения задачи определения живости с учетом особенностей и сценариев применения современных мобильных устройств. Выделены основные требования, предъявляемые к методам обнаружения подделок в мобильных биометрических приложениях:

- некооперативность применяемых решений;
- переиспользование доступных аппаратных средств без внедрения дополнительных сенсоров;
- устойчивость к высокоизменчивым входным данным;
- низкая вычислительная сложность используемых алгоритмов.

Глава 2

Определение живости лица в мобильных системах

Основанные на биометрии системы идентификации личности показали высокую надежность в решении соответствующих задач обеспечения безопасного доступа в самых разных областях человеческой деятельности. Значительную часть жизни современного человека занимает взаимодействие с личным мобильным устройством, позволяющим совершать самые разнообразные действия, от ведения личной переписки до осуществления рабочей и финансовой деятельности. В настоящее время подавляющее большинство доступных на рынке смартфонов оснащены компактными биометрическими системами, которые призваны упростить доступ к функционалу личного устройства для владельца и в то же время защитить хранящиеся на нем данные от чужих глаз. Стоит упомянуть, что далеко не все устройства могут обеспечить полную защиту личных данных на смартфоне от несанкционированного копирования в случае взлома опытными хакерами в обход программной оболочки. Подразумевается, что биометрическая система создает дополнительный уровень противодействия при взломе неспециалистами, не имеющими знаний о скрытых уязвимостях мобильных операционных систем.

2.1. Особенности мобильных биометрических систем

К наиболее распространенным БХЧ, применяемым в современных персональных устройствах (смартфонах, плантешах и т.д.), относят рисунок папиллярных линий пальца и видеообраз лица. Обе из упомянутых модальностей позволяют решать задачу аутентификации с высочайшей точностью, но первая из них требует внедрения в мобильное устройство дополнительного сенсора для

считывания отпечатков пальцев, что вызывает повышение суммарной стоимости и проблем встраивания нового технического элемента в дизайн. При этом вторая из упомянутых БХЧ может быть получена при помощи фронтальной камеры видимого света практически на каждом современном смартфоне. Данный функциональный элемент обрел широкую популярность в последнее десятилетие как следствие распространения социальных медиа и личных видеозвонков и возросшей пропускной способности сотовых сетей.

Ввиду популярности лица и отпечатка пальца как биометрических модальностей особое внимание уделяется безопасности таких систем с точки зрения противодействия спуфингу или взлому при помощи подделок. Системы распознавания по лицу демонстрируют высокую степень уязвимости по отношению к подобным попыткам несанкционированного доступа ввиду простоты создания искусственных копий БХЧ. Распространенность социальных сетей позволяет найти набор качественных фотографий лица, а для создания достаточно эффективной подделки требуется лишь распечатать его в цвете на плотной фотобумаге или продемонстрировать на дисплее высокого разрешения. В СМИ упоминаются случаи успешного обмана встроенных технологий распознавания по лицу для ряда современных смартфонов [109, 117]. Более того, обширные исследования специалистов в области информационной безопасности [42, 139] демонстрируют результаты, показывающие наличие уязвимости к спуфингу практически во всех современных смартфонах с установленной операционной системой Android и единственной фронтальной камерой для получения видеообраза лица.

В случае со сканером отпечатка пальца создание искусственной копии злоумышленников осложняется по ряду причин. Во-первых, необходимо некоторым образом получить качественный рисунок папиллярных линий одного из пальцев жертвы, что требует близкого контакта и определенных навыков. Во-вторых, воспроизвести этот рисунок при помощи пластилина, гипса или иных вспомогательных материалов. Наконец, стоит учесть различия набора уязвимо-

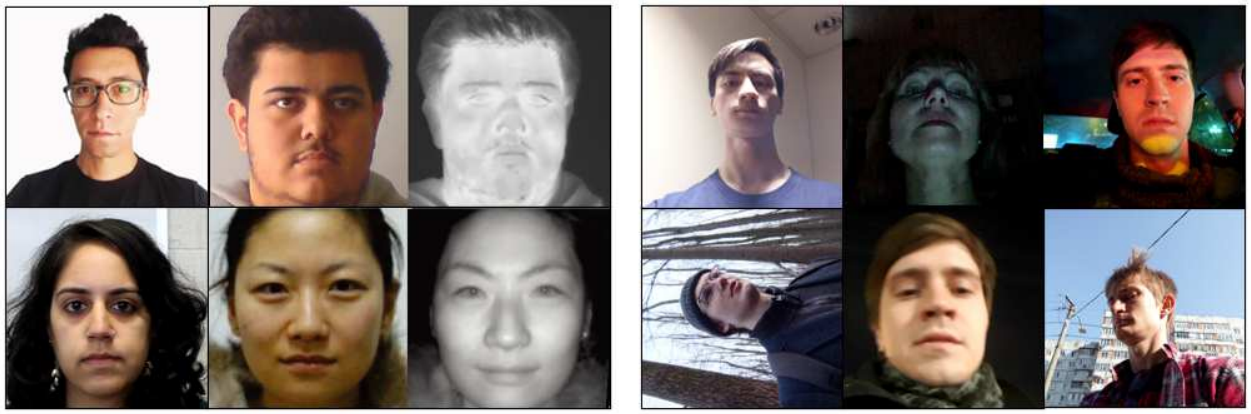


Рис. 2.1. Примеры изображений полученных в сценарии стационарной системы распознавания, в т.ч. с применением вспомогательных сенсоров (слева), и растров, демонстрирующих значительную изменчивость условий съемки, характерную для мобильных биометрических систем (справа)

стей тех или иных типов дактилоскопических сенсоров. Иными словами, взлом системы распознавания по отпечатку пальца требует от взломщика наличия богатого практического опыта работы с подобными системами, в то время как обман системы распознавания по лицу кажется весьма простым и интуитивно понятным процессом [91]. Поэтому наиболее актуальной модальностью для разработки методов анти-спуфинга является видеообраз лица.

Применение мобильной биометрической системы предполагает учёт специфики взаимодействия с пользователем: требуется учитывать поведенческие особенности человека и возможные значительные изменения окружения. Попытки распознавания могут происходить при ходьбе или при наличии тремора рук, что приводит к дрожанию устройства и размытию кадра; пользователь может носить очки, головные уборы, маски; смартфон допускается располагать как вблизи лица, так и на расстоянии вытянутой руки, как в портретной, так и в ландшафтной ориентации; наблюдаемая поза головы способна существенно видоизменяться. Уровень освещенности может сильно различаться в различных локациях применения устройства: от $10^{-}4$ люкс в полутьме без источников света до более 10^5 люкс под прямыми солнечными лучами. Эти факторы негативно влияют на качество биометрических данных (Рис. 2.1) и, как следствие,

на точность идентификации и способность детектирования спуфинг-атак [119].

Вдобавок для мобильных устройств существуют требования удобства применимости в повседневной жизни. Встраиваемые биометрические системы также должны предоставлять простой интерфейс взаимодействия с пользователем и обеспечивать высокую скорость распознавания, которая во многом определяется вычислительной сложностью композиции применяемых алгоритмов (Рис. 1.2. Также процесс аутентификации должен потреблять минимальное количество энергии устройства и допускать обработку входных данных с частотой поступления кадров. Это создает компромисс между сложностью встраиваемых алгоритмов и энергопотреблением, существенно ограничивает диапазон применимых методов детектирования подделок. К примеру, кооперативные подходы к определению живости лица не могут быть встроены в мобильные биометрические системы.

Стоит также учитывать особенности мобильных систем защиты личных данных, подобных паролям и пин-кодам. С недавних пор к этому списку стали добавляться и сами алгоритмы идентификации вместе с обрабатываемой ими биометрической информацией. Главное требование к таким системам — отсутствие прямого доступа к ним извне из операционной системы устройства и ее периферии. В настоящее время существуют технологии, реализующие необходимые ограничения на практике, как правило в виде системы на чипе (SOC, System on Chip). Это понятие подразумевает подсистему аппаратной начинки устройства в виде части центрального процесса, оборудованную специальной независимой от основной операционной системой. Среди наиболее широко распространенных — TrustZone от ARM [15]. Применения методов биометрического распознавания внутри SoC осложняется из-за особенностей таких систем: уменьшенная пиковая тактовая частота процессора, ограниченный объем доступной оперативной памяти и в некоторых случаях невозможность использования многопоточности.

2.2. Классификация способов взлома

Процесс несанкционированного доступа в систему биометрического распознавания при помощи искусственно созданного материального артефакта, содержащего образ человека, чей биометрический шаблон сохранен в базе зарегистрированных пользователей, называется спуфингом (spoofing) или атакой PA (presentation attack). Попытки взлома различаются по типам артефактов, степени сложности получения изображения пользователя и уровню подготовки злоумышленника.

Возможны следующие виды физических артефактов:

- Бумажные: распечатки изображения лица жертвы на обычной или фотобумаге повышенной плотности. Возможны модификации: создание прорезей для глаз или рта, вырезание по контуру силуэта лицевой области.
- Электронные: выведенное на некоторый дисплей изображение или видеозапись лица пользователя.
- Объемные: созданные из некоторого материала маски, повторяющие геометрию и текстуру лица жертвы. Тип используемого материала определяет диапазон подвижности подобного вида атак: пластиковые или гипсовые артефакты позволяют точно передать текстуру лица, силиконовые маски могут давать подвижность лицевой области. Отличаются высокой временной и материальной стоимостью создания.

Подделки лица могут отличаться источниками биометрической информации:

- Уровень I: статическое изображение лица пользователя обычного качества. Пример: изображение из социальных сетей или полученное при помощи скрытой фотосъемки.

- Уровень II: видеозапись или изображение лица высокого качества. Пример: изображения личной фотосессии в контролируемых условиях освещенности.

- Уровень III: видеозапись с набором лицевых движений: повороты головы, моргание или изображение лица в другой модальности. Примеры: тепловой снимок, карта глубины, ближний инфракрасный спектр и т.д. Такие данные практически невозможно получить без длительной слежки за жертвой.

Лица, осуществляющие взлом, могут иметь разный уровень осведомленности о работе биометрических систем и подсистем определения живости:

- Уровень I: Обыватель, не обладающий опытом работы с биометрическими системами.

- Уровень II: Пользователь, имеющий общее представление о методах детектирования подделок, типах используемых сенсоров и скрытых проверках.

- Уровень III: Эксперт или группа экспертов в области биометрической идентификации и информационной безопасности.

В совокупности с учетом вышеописанного можно составить следующую классификацию спуфинг-атак:

- Базовая категория (Б): распечатанная или статичная и показанная на экране фотография, как с видимыми так и с невидимыми краями подделки. При создании не требует высокого уровня экспертизы и специализированного оборудования. Источник лицевой информации: уровень I. Минимальная экспертиза взломщика: уровень I.

- Продвинутая категория (П): демонстрируемая на экране высокого разрешения видеозапись пользователя или бумажная маска на основе фото высокого разрешения. Для создании бумажной маски возможно вырезание

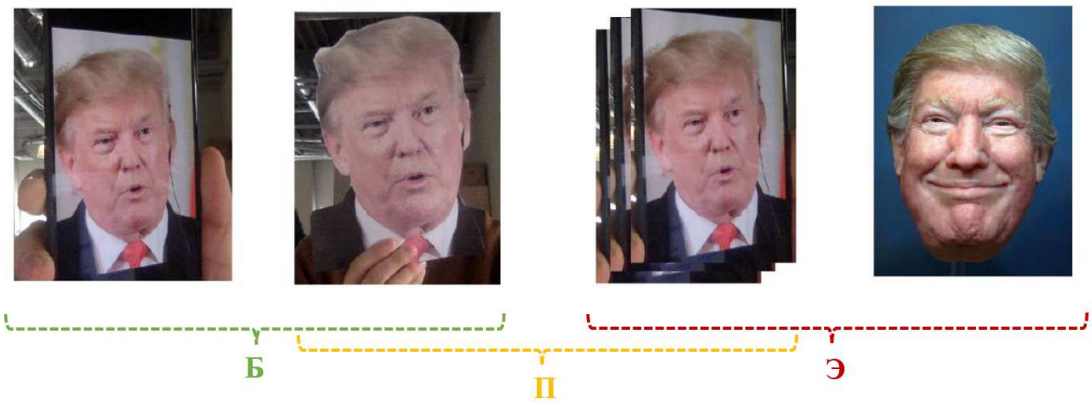


Рис. 2.2. Образцы поддельных лиц разных категорий.

лицевой области по контуру силуэта или воспроизведение грубой трехмерной геометрии лица при помощи бумаги. Подобный тип атак неосуществим без достаточного уровня экспертизы взломщика и высокого качества исходной фотографии/видеозаписи жертвы. Источник лицевой информации: уровень II. Минимальная экспертиза взломщика: уровень II.

- Экспертная категория (Э): данные из дополнительных модальностей и объемные маски высокого качества изготовления. По сравнению с предыдущим уровнем требуются значительные временные и финансовые затраты на подготовку и создание реалистичного артефакта. Источник лицевой информации: уровень II. Минимальная экспертиза взломщика: уровень III.

Подавляющее большинство видов атак, упоминаемых в литературе и в материалах открытых баз изображений, относится к категориям Б и П (табл 2.1,Рис. 2.2 Причина такой неравномерности в том, что оставшаяся группа Э является маловыгодной для практического взлома: необходима высокая степень экспертизы злоумышленника вкупе с применением дорогостоящего и слабо представленного на рынке оборудования.

По этой же причине на практике в требования заказчиков при составлении технического задания для алгоритмов определения живости в составе мобильных систем распознавания по лицу входит устойчивое противодействия поддел-

Метка	Описание	Категория
PR	Печатная фотография без изменений. Возможна как глянцевая, так и матовая печать.	Б
PC	Силуэт лицевой области, вырезанный по контуру из печатной фотографии. Возможные изменения: вырезание области глаз или рта для наложения поверх настоящего лицо и совершения движений в этих регионах.	П
PP	Вырезанный регион лицевой области, содержащий, к примеру, области глаз и щек, для наложения поверх настоящего лица. Необходима проверка успешного прохождения таким артефактом этапа сравнения с сохраненным в системе шаблоном.	П
SM	Неподвижное изображение лица, выведенное на экран смартфона.	Б
SI	Неподвижное изображение лица, выведенное на экран высокого разрешения: планшета, ноутбука, телевизора.	Б
SV	Видеозапись движений лица, выведенная на экран высокого разрешения или смартфона.	П
MP	Объемная бумажная маска низкого качества лица жертвы.	П
MA	Объемная пластиковая маска лица жертвы, повторяющая геометрию и текстуру области интереса.	Э
MS	Объемная силиконовая маска лица жертвы, повторяющая геометрию и текстуру области интереса.	Э

Таблица 2.1. Распространенные виды спуфинг-атак

кам лишь первых двух категорий. В качестве примера можно привести список требований к системам распознавания по лицу мобильной операционной систе-

мы Google Android [93]. В списке рассматриваемых видов атак содержатся лишь примеры подделок базовой и продвинутой категорий: фотографии лица, демонстрация лица на экране высокого разрешения и вырезанные по контуру силуэта лицевой области распечатки.

Стоит отметить, что реализация биометрических систем в составе Google Android согласно упомянутому документу подразумевает введение ПИН-кода устройства раз в 72 часа даже для самых устойчивых к спуфингу решений (Biometric Class 3, formerly strong). Это говорит о некотором недоверии по отношению к безопасности существующих на рынке мобильных систем распознавании по лицу при помощи фронтальной камеры, и в то же время свидетельствует о наличии спроса на разработку более совершенных алгоритмов определения живости в индустрии.

2.3. Обзор методов детектирования подделок

Одна из первый работ [120] посвященных проблеме анти-спуфинга была опубликована в 2002 году и постулировала основную мотивацию последующих исследований: определение живости путем распознавании физиологической информации как признака подлинности биометрического шаблона.

Ранние разработки в области детектирования поддельных БХЧ появились задолго до эпохи повсеместного использования методов глубокого обучения и были основаны либо на текстурном анализе определенных областей интереса региона лица, либо на обнаружении действий и движений объекта и кооперативных подходах. Так в [82] предлагается использовать спектральный анализ для построения классификатора, способного различать изображения подлинных лиц и двумерных распечатанных масок. Отличия в освещенности объемных объектов по сравнению с плоскими проявляются в области низких частот, в то время как область высоких частот будет отражать несходство детализации искусственно созданных подделок по отношению к текстуре живого лица.

в качестве классификатора применялся метод опорных векторов (SVM). Большую популярность в тот период развития анализа изображений имели текстурные дескрипторы, например, LBP (local binary pattern) [100]. На основе этого способа извлечения высокочастотных признаков была предложена работа [90], также использующие метод опорных векторов для построения решающего правила. Существуют также решения [73], комбинирующие подходы к построению признакового описания при помощи LBP и частотный анализ изображения лица. Упомянутые методы имеют значительное преимущество в виде низкой вычислительной сложности, что позволяет применять их с учетом ограничений мобильной биометрической системы, но в настоящий момент не позволяют достичь высокой обобщающей способности на более объемных выборках данных с большим разнообразием спуфинг-атак.

Популярным подходом к определению живости лица является использование динамических признаков, извлекаемых из последовательности изображений. В одной из ранних работ [68] по данной теме предлагается анализировать изменчивость периокулярной области лица для серии последовательных кадров. Органы зрения отличаются большой подвижностью (движения глаз и век) даже на сравнительно небольших временных интервалах, что позволяет использовать эту особенность для противодействия статическим подделкам разного вида. Ряд работ [101, 102, 126] посвящен непосредственно выявлению морганий путем моделирования паттернов естественных изменений степени открытости век у человека при помощи Conditional Random Fields (CRF) и сравнения с таковыми у искусственно созданных примеров БХЧ. Анализ движений области глаз часто применяется в контексте кооперативной биометрической идентификации. В нескольких работах [19, 77, 130] было предложено использование различных методов расчета оптического потока для последовательности кадров. Его компоненты позволяют извлекать сложные динамические признаки лицевой области и оценивать грубую карту глубины, что является достаточным условием для построения устойчивой защиты от двумерных статичных подде-

лок. Ряд работ [80, 137] предлагает извлекать признаки формы лица но набору кадров при помощи трехмерной реконструкции лицевой области и без опоры на расчет оптического потока. К недостаткам подобных методов стоит отнести их высокую вычислительную сложность. Более экзотический подход [74] к извлечению динамических признаков живости предполагает использование программно-регулируемого изменения фокуса камеры биометрической системы при съемке подряд идущих кадров. Предполагается, что положение пользователя в кадре не успеет существенно измениться, а отличия в степени размытия переднего и заднего планов для объемных лиц и плоских подделок может быть использовано для детектирования спуфинг-атак базового уровня сложности. В качестве ограничений динамических подходов в контексте мобильных приложений стоит упомянуть снижение удобства применения биометрической системы при росте количества последовательных кадров, необходимых для принятия решения в связи с возрастающим временем отклика.

Зачастую при построении алгоритмов оценки живости видеообраза лица исследователи полагаются на анализ фона сцены [55, 104, 146] с целью детектировать неестественное искажение геометрии кадра и присутствие краев и артефактов, характерные для изображений подделок. Подобные решения часто используются как комплементарные к методам определения подлинности по лицевой области и динамическим признакам.

В последние годы сверточные нейронные сети показали высокую производительность применительно к задачам распознавания изображений по сравнению с иными уже существующими подходами. Описанные выше подходы к определения живости лица были во многом идейно воспроизведены с использованием методологии глубокого обучения.

В литературе описано множество [17, 26, 83, 148] способов анализа текстуры лицевой области при помощи нейросетевых классификаторов, построенных, как правило, с помощью логистической функции потерь. Полученные результаты показали высокое качество решения задачи при валидации на данных

внутри одной и той же базы изображений, содержащей как правило общий набор типов подделок и видов камер в обучающей и тестовой выборках. При этом решения демонстрировали [149] крайне низкие результаты при тестировании на непересекающихся базах с примерам подделок или типов сенсоров, не встреченных сетью при обучении. Проблема обобщаемости при кросс-тестировании получаемых решений по мнению исследователей вызвана рядом причин, важнейшая из которых — ограниченные по объему и вариабельности данные открытых баз изображений подделок. Создание таких баз является крайне трудоемким процессом и требует от исполнителей высокого уровня экспертизы в области. Ресурсов академических исследовательских групп по сравнению с индустрией зачастую недостаточно для выполнения поставленной задачи, в то время как коммерческие базы данных отсутствуют в открытом доступе. Второй по важности причиной является склонность нейросетевых решений к переобучению на текстурных особенностях изображений, что особенно проявляется на небольших выборках. Сенсор каждой камеры обладает уникальным видом высокочастотных шумов, связанных с работой подсистем экспозиции и фокусировки [23]. Данные текстурные особенности проявляются по-разному на изображениях подлинных лиц и подделок в зависимости от условий окружений (уровень и направление освещения), что связано с различиями в альбедо поверхностей, представленных перед камерой. Наконец, стоит отметить, что зачастую поставленная задача рассматривается как задача бинарной классификации на классы «живое лицо»/«подлог», которая в процессе обучения нейронной сети решается минимизацией логистической функции потерь методом градиентного спуска. При этом проблема определения живости в биометрических системах является более многогранной: для предсказания требуется учет характеристик типа подделки, уровня освещенности кадра и изменчивости видов лица человека [149]. В связи с этим многие современные подходы к построению нейросетевых методов анти-спуфинга предлагают широкий спектр [47, 69, 75, 87, 127] способов регуляризации путем введения дополнительных функций потерь.

Методы глубокого обучения применялись также для детектирования подделок по динамическим признакам набора кадров лицевой области [46, 85, 138, 142, 150]. Получаемые результаты демонстрировали прирост качества решений по сравнению с однокадровыми подходами при валидации в рамках одной базы данных, но сохраняли проблему снижения производительности при кросс-тестировании. Особенностью данной группы методов является высокая вычислительная сложность, что затрудняет их адаптацию для ограничений мобильных биометрических систем.

2.4. Метод определения живости лица для мобильного устройства

В данной главе представлено решение для обнаружения подделок видеообраза лица в мобильных биометрических приложениях, отвечающее соответствующим требованиям точности и быстродействия. Уникальными характеристиками метода являются: учет атрибутов лица и контекста кадра, позволяющий осуществлять предварительную оценку живости и отбраковывать заведомо некорректные примеры с учетом особенностей взаимодействия пользователя с мобильным устройством; многостадийная структура алгоритма; а также комбинированное решение для детектирования подделок по набору регионов интереса входного изображения. Упомянутые особенности позволяют применять метод в режиме реального времени в условиях значительной изменчивости окружения.

2.4.1. Структура метода определения живости

Общая схема предлагаемого решения дана на Рис. 2.3. Основная идея заключается в построении системы обнаружения явных и характерных попыток подделывания алгоритма распознавания с целью раннего отказа от дальнейшей верификации. При этом в случае набора истории N подряд идущих кадров с меткой «подделка», возможно прерывание текущей попытки верификации цели-

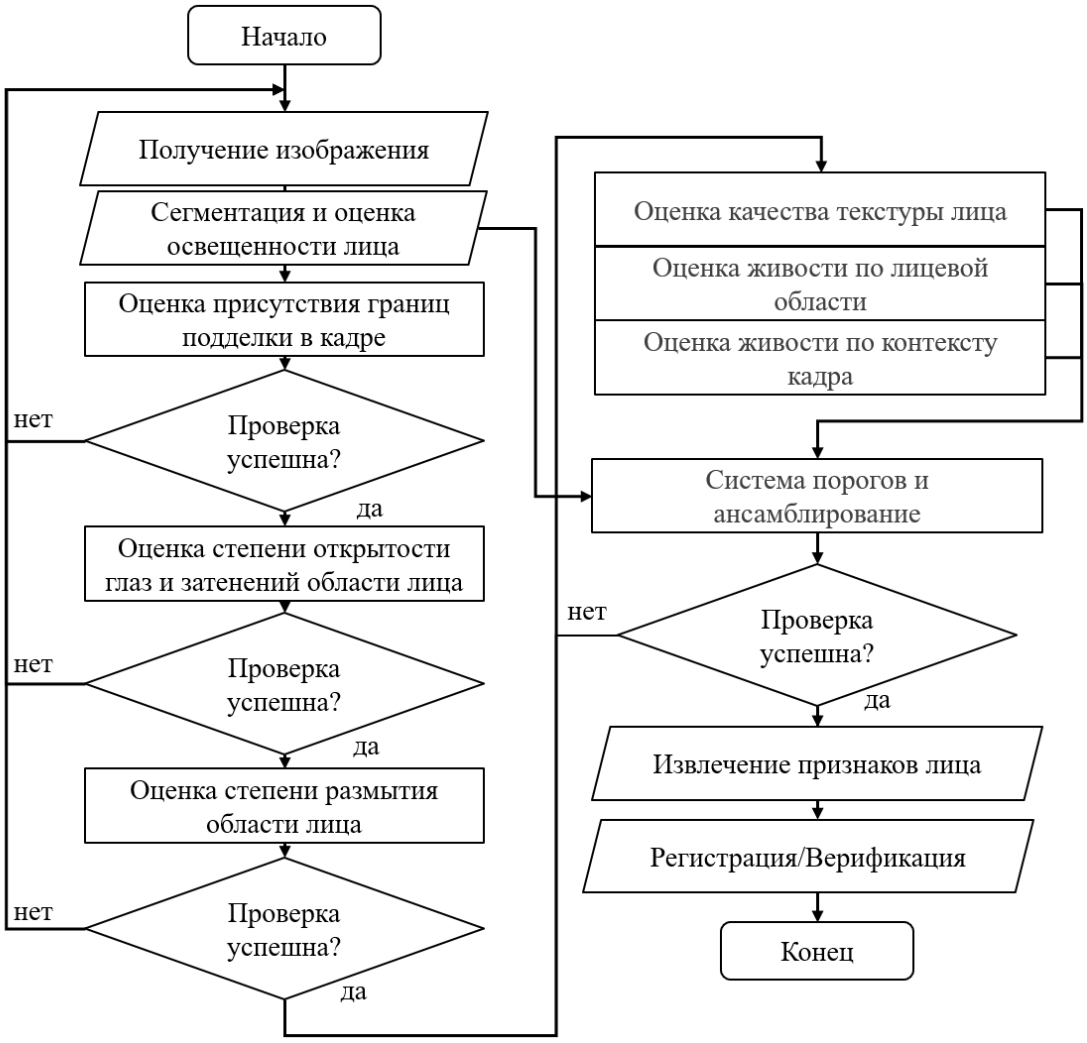


Рис. 2.3. Схема многостадийного алгоритма определения живости лица.

ком. Вычислительно сложные алгоритмы определения живости в таком случае предполагается применять только после успешного прохождения всех упомянутых базовых проверок. Используемые методы ранних отказов от распознавания должны при данном сценарии применения обладать крайне низкими значениями доли ложных отказов (FRR, False Rejection Rate), чтобы применение их последовательной комбинации не оказывало значительное влияние на общую статистику отказов от распознавания.

Во множестве описанных в литературе работ задача детектирования подделок очень часто рассматривается как задача бинарной классификации на два класса «подлинный пример» и «подделка». Зачастую такая задача решается путем обучения нейросетевой модели или ансамбля таких моделей на выборке

значительного объема, содержащей большое количество примеров подделок разных видов и настоящих лиц с большой вариабельностью освещения. При такой процедуре практически невозможно контролировать, что получаемые признаковые описания соотносятся с реальными артефактами, содержащимися в изображениях подделок (неестественные размытие, блики, границы); модель склонна попадать в локальные оптимумы и демонстрировать эффект переобучегтя, что порой приводит к появлению ошибок классификации (ложных пропусков) даже для простейших примеров подлога [149].

Мотивация построения многостадийной структуры алгоритма определения живости с ранними отказами заключается в разбиении задачи на набор блоков, каждый из которых отвечает за обнаружении одного из явных артефактов, характерных для подделок. Во-первых, это позволяет применять менее глубокие и вычислительно сложные архитектуры нейронных сетей. Во-вторых, снижает вероятность ложного пропуска подделки. Наконец, полученную систему проще поддерживать в случае необходимости подстройки под новый тип входных данных. При обнаружении некорректной работы возможно отслеживание ее причин до конкретного модуля, который возможно без существенных затрат обновить за счет переобучения с выбором наилучшей модели по качеству решения на отложенной валидационной выборке. В сценарии применения единственной глубокой нейросетевой модели для определения живости, процесс перенастройки будет существенно более затратным по времени и ресурсам, поскольку потребуется как соответствие прежним критериям производительности на валидации, так и удовлетворение новым требованиям.

2.4.2. Учёт условий окружения

Применение алгоритма определения живости для мобильной биометрической системы требует учета большой изменчивости условий окружения при построении комбинированного решения. В случае обнаружения сложного сценария распознавания, например, попытки верификации в темном помещении,

допускается изменение решающего правила с целью снизить уровень ложных отказов системы. Подобный механизм адаптации возможен, поскольку значительная часть поддельных лиц будет заметно отличаться от подлинных в полутьме с единственным источником освещения в виде экрана смартфона. Тем не менее данное утверждение не верно для примеров подлога путем показа дисплеев с изображениями лиц.

Яркостные характеристики входных изображений сами по себе могут быть использованы для оценки характеристик условия окружения. Однако такая оценка будет некорректной в случае попыток подделывания при помощи снятого в полумраке изображения лица на небольшом ярком экране или распечатанной на бумаге сцены с человеком, снятой в темном помещении, но предъявленной при естественном свете. Вдобавок, данный подход не может быть эффективно применен на практике вследствие присутствия в алгоритмах обработки сигнала с сенсора фронтальной камеры процедуры автоматического подбора экспозиции.

Тем не менее, подобранные параметры экспонирования (время открытия затвора t_{exp} и коэффициент усиления c_{gain}) могут быть получены в режиме реального времени для текущего кадра. Уровень освещенности b (brightness) снимаемой сцены прямо пропорционален упомянутым параметрам:

$$b \sim t_{\text{exp}} c_{\text{gain}}. \tag{2.1}$$

Некоторые современные версии операционной системы Android позволяют получить значение параметра b во внутренней шкале измерений напрямую с камеры без промежуточных расчетов. Вдобавок, большинство современных смартфонов оборудовано датчиком освещенности, которые позволяет грубо оценивать эту величину для передней панели устройства.

Совместное использование упомянутых характеристик позволяет построить систему адаптивной подстройки решающих правил для сложных сценариев применения биометрической системы.

2.4.3. Живость по границам подделки

Процесс спуфинга мобильной системы распознавания злоумышленником уровня экспертизы I или II допускает появление краев подделки в области видимости сенсора камеры, будь то границы распечатанной фотография (PR) или рамка дисплея иного устройства (SM,SI,SV), Табл. 2.1, в то время как попытки доступа реальным человеком как правило не будут содержать явных артефактов подобного рода. Данный признак может быть использован для внедрения дополнительного уровня защиты от взлома путем построения нейросетевого бинарного классификатора, оценивающего вероятность присутствия неестественных пограничных регионов в контексте кадра.

Наличие краев подделки внутри сцены является низкочастотной информацией, для обнаружения которой не требуется высокое разрешение входного растра. В результате для обучения нейронной сети будет достаточно информации, извлекаемой из результата масштабирования входного кадра к низкому разрешению, что позволит существенно уменьшить количество операций с плавающей точкой (FLOPs, Floating-Point Operations) прямого прохода сети для любой применяемой архитектуры. Вдобавок допустимо исключить цветовую информацию с целью снижения вероятности переобучения под цветовые характеристики конкретной камеры мобильного устройства и перевести данные из цветового пространства RGB в черно-белое представление. Экспериментальным путем было определено, что для входного кадра с соотношением сторон 1x1 для решения поставленной задачи будет достаточно монохромного растра с размерами 80×80. В случае неквадратных размеров входного кадра допускается масштабирование наименьшей из сторон к значению 80 пикселей с сохранением пропорций.

Стоит учесть, что выборку обучающих примеров должны составлять изображения подлинных лиц и лишь та часть примеров подделок (PR,SM,SI), Табл. 2.1, в сценах которых наблюдается хотя бы один край подделки. Формирование

подвыборки подобных примеров возможно путем экспертной разметки соответствующей части базы спуфинг-атак. При этом возникает риск переобучения на задних планах как для живых лиц, так и для подделок, поскольку на практике невозможно обеспечить полный диапазон всех возможных естественных фонов, в том числе содержащих прямые контрастные линии.

Предлагается искусственно нарастить выборку поддельных лиц путем синтеза дополнительных примеров в исходном разрешении данных, получаемых с камеры устройства. Подобные подходы описаны в литературе [55, 147], но применяются как правило для повышения обобщающей способности методов определения живости в целом, без опоры на конкретные характеристики некоторых спуфинг-атак. Синтез примеров для распечатанных подделок типа PR не составляет труда путем наложения региона исходного растра с лицом на некоторое вспомогательное случайное изображение естественного фона из общедоступных баз данных [141] или подвыборки класса «настоящее лицо» непосредственно. Имитация изменчивости расстояния и положения подделки возможна путем случайной перспективной трансформации региона лица с последующим размытием пропорционально размеру региона для имитации потери резкости, а также случайно яркостной коррекцией. Малозаметные артефакты и неточности подобной трансформации будут пренебрежимо малы при последующем масштабировании к низкому разрешению для обучения сети.

Для синтеза поддельных примеров SM и SI содержащих рамки смартфонов, планшетов или стационарных дисплеев, предлагается создать вспомогательную выборку изображений-шаблонов для ряда популярных устройств такого рода в разных цветовых расцветках для большей вариабельности. Задний план и область экрана таких изображений предполагается выделять разными цветами для возможности последующего маскирования и наложения в соответствующие регионы растров случайных сцен фона и примеров лиц, содержащихся в обучающей выборке для задачи, подобно технологии хромакей (chroma key), Рис. 2.4.

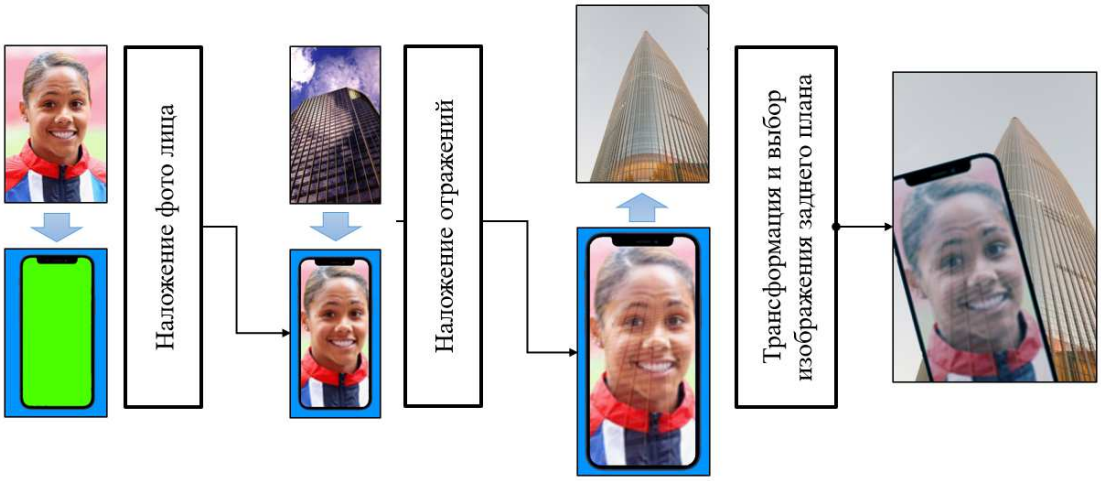


Рис. 2.4. Схема построения синтетических примеров для оценки присутствия границ поддельки.

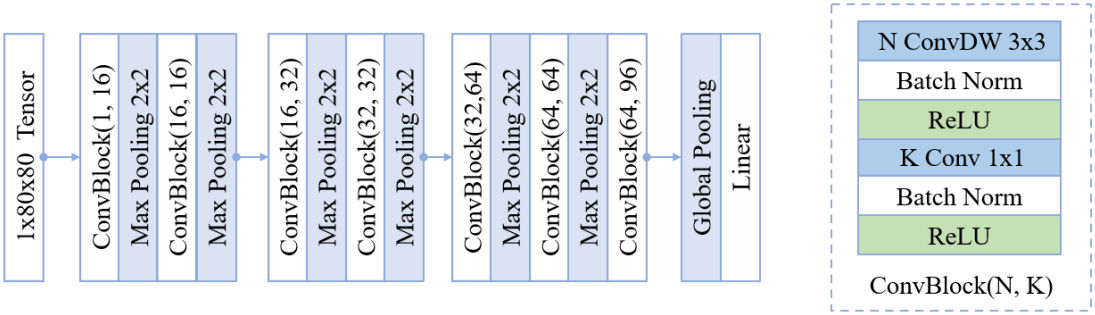


Рис. 2.5. Архитектура нейронной сети для обнаружения подделок с краями.

Синтезированные примеры всех упомянутых типов поддельных данных предполагается добавить в обучающую выборку в пропорции 1:2 и сравнить качество полученных решений для исходного и расширенного набора данных на отложенном тесте. Для отсечения наименее правдоподобных результатов синтеза применялся детектор лиц [153]: в случае отказа искусственно полученный растр не был добавлен в обучающую выборку.

в качестве архитектуры модели была выбрана неглубокая архитектура, подобная [63] по строению блоков, Рис. 2.5. Разрешение входного изображения, количество слоев и их параметров подобраны таким образом, чтобы обеспечить медианное время обработки одного изображения до пары миллисекунд на одном ядре современного мобильного устройства и лишь незначительно уменьшить

Условия освещенности	Тип данных			
	REAL	SM	PR	SI
IN	54172	14200	13661	14512
OUT	31340	4220	12367	4644
LI	22930	5351	1359	5690
Всего:	108442	23771	27387	24864
Синтетика	-	+12871	+13693	+12210

Таблица 2.2. Описание базы изображений для детектирования границ подделки. Рассмотрены условия освещенности: IN — естественное освещение в помещении; OUT — яркий солнечный свет вне помещения; LI — низкий уровень освещения в помещении.

скорость отклика системы целиком.

Описание используемой обучающей и тестовой выборок дано в табл. 2.2 При их формировании учитывалось требование балансировки примеров основных классов и подклассов по типам подделок и условий освещенности. Выборки не пересекаются по личностям присутствующих на изображениях людей, чтобы избежать переобучения на биометрические характеристики.

Обучение модели производилось в течении 300 эпох, с целью регуляризации применялись аугментации случайной коррекции яркости, насыщенности и контраста, а также наложение случайного пуассоновского шума. Аугментации поворота изображения не рассматривались, поскольку примеры с различной ориентацией самих подделок содержались в искусственно синтезированной подвыборке, и применение подобной трансформации может приводить к неестественным искажениям сцен. Результаты обучения модели (Табл. 2.3) и демонстрируют прирост производительности модели при обучении на расширенной обучающей выборке. Подкласс примеров подлинных лиц для сценария распознавания в естественном помещении показывает снижение точности классификации ввиду присутствия в базе данных сцен, снятых в офисных помещениях, содержащих на фоне значительное количество контрастных линий и перепадов

Обучающая выборка	Точность классификации, %					
	IN	OUT	LI	PR	SM	SI
Базовая	96.5	97.2	98.7	95.7	98.2	96.4
Расширенная	**98.4**	**99.3**	**99.1**	**97.9**	**99.2**	**98.0**

Таблица 2.3. Результаты тестирования моделей на отложенной выборке для подклассов подлинных и поддельных примеров.

яркости, похожих на края подделок. Подкласс примеров взлома при помощи распечаток лиц также демонстрирует сниженную производительность по сравнению с другими типами подделок ввиду присутствия сцен, в которых фон фотографии сливается с задним планом кадра.

2.4.4. Живость по степени размытости региона лица

Мобильные биометрические системы распознавания для данной БХЧ используют фронтальную камеру для съемки. Попытки взлома такой системы при помощи подделок, содержащих образ лица на сравнительно небольших дисплеях, наподобие таковых у смартфонов или планшетов, часто сопряжены с необходимостью зафиксировать лицевой регион в области видимости камеры непосредственно вблизи от нее.

Применение фронтальной камеры в системах идентификации в современных смартфонах часто осуществляется без автоматического определения фокусного расстояния оптической системы с целью повышения скорости отклика при попытке распознавания. Мобильные сенсоры захвата изображения в большинстве случаев оборудованы системой фокусировки, подбирающее оптимальное положение линзы полным перебором таким образом, чтобы контрастность некоторого региона входного растра была максимальной. Поэтому процесс подстройки фокусного расстояния может занимать несколько секунд, что является критичным с точки зрения удобства применения устройства пользователем. Стоит отметить, что во фронтальных камерах современных смартфонах начинают по-

являться более оптимальные с точки зрения производительности альтернативы упомянутому подходу, так называемые системы автофокуса путем определения фазы (PDAF, phase detection auto-focus), позволяющие осуществлять подстройку параметров оптики за сотни миллисекунд, но их распространение ограничено более высокой стоимостью таких решений. В связи с перечисленными особенностями, мобильные биометрические системы часто фиксируют фокус камеры при попытке распознавания соответствующим некоторому среднему ожидаемому расстоянию от устройства до лица пользователя. Это приводит к тому, что в некоторых сценариях использования лицевая область на изображении может оказаться размытой, в частности, при попытках взлома при помощи артефактов, расположенных вблизи устройства. Подобные попытки представляют значительную опасность, поскольку как правило не содержат явно неестественных областей, соответствующих краям подделок, расположенных вне области видимости камеры.

Заметный смаз региона лица на входном изображении может быть обнаружен простейшим нейросетевым классификатором живости при условии, что некоторая часть обучающей выборки примеров класса «подделка» содержат подобный дефект, и при этом ничтожно малая часть примеров класса «подлинное лицо» не содержит его. В противном случае возможны ложные пропуски моделью поддельных растров, похожим по степени размытия на подлинные вследствие переобучения. В случае использования обширной базы видеопоследовательностей, собранной при имитации попытки распознавания человеком в естественных условиях в разнообразных условиях освещенности, практически невозможно гарантировать отсутствие смазов в лицевой области. Слабо контролируемые небольшие случайные движения рук или головы участников и идентификация при даже медленной ходьбе неизбежно приведут к возникновению размытия. Полученные в результате крупномасштабной съемки изображения как правило подвергаются полуавтоматической обработке с целью выделения региона лица и ключевых точек для последующего построения базы данных

разметки и анализом части извлекаемых регионов лица человеком-экспертом для отбраковки некорректных кадров. Но при такой процедуре бывает затруднительно выявить все примеры смаза, который может быть слабо заметен при просмотре изображений. Более того, восприятие степени размытия может различаться для разных людей-экспертов, что затрудняет фильтрацию некорректных примеров.

В таком случае для выявления подобных искажений изображения лица и для повышения безопасности системы против спуфинг-атак, расположенных вблизи смартфона, целесообразно разработать алгоритм оценки качества, позволяющий определять степень смаза в области интереса. Подобный метод может быть реализован в виде нейросетевого бинарного классификатора, оценивающего вероятность присутствия неестественного размытия. Однако, процесс обучения подобного решения сопряжен с рядом сложностей, затрудняющих его практическую реализацию. Во-первых, как упоминалось выше, процесс создания бинарных меток для процедуры обучения затруднителен по причине трудности интерпретации искажений такого рода. Во-вторых, ошибки в бинарной разметки обучающей выборки могут привести к переобучению сверточной нейронной сети на малой части некорректных примеров и привести к ошибкам ее применения на практике и появлению уязвимостей. Наконец, степень размытости в области лица напрямую зависит от расстояния до объекта съемки, которое вследствие этого требуется учитывать при принятии решения.

Чтобы справиться с перечисленными трудностями допустимо использовать подходы самоконтролируемого обучения (self-supervised learning). Предлагается свести задачу бинарной классификации к задаче регрессии некоторой степени размытости лицевого региона в виде вещественного числа, чтобы впоследствии на валидационной выборке подобрать систему порогов в зависимости от расстояния до объекта. При таком подходе в целом не требуется присутствие поддельных примеров в обучающей выборке, допускается воспользоваться примерами класса «настоящее лицо». К каждому изображению предлагает-

Рис. 2.6. Строение нейронной сети для предсказания уровня размытия.

ся применять фильтр Гаусса со случайно выбранным вещественным значением дисперсии σ из интервала от 0 до $\sigma_0 = 3$. Полученный в результате растр может быть обработан неглубокой сверточной сетью, обучаемой на предсказание выбранного значения σ путем минимизации среднеквадратичной ошибки между истинным и предсказанным значением. Для такого сценария обучения не требуется обширная и разнообразная выборка с корректной разметкой, выбор целевой значения целевой переменной из равномерного распределения $U(0, \sigma_0)$ обеспечивает широкий диапазон изменчивости обучающих данных.

Для обучающей выборки были отобраны порядка 100000 изображений подлинных лиц в трех условиях окружения: естественное освещение, яркий солнечный день при съемке вне помещения и полутемное помещение с уровнем освещенности не превышающем 1-2 люкс, в котором освещение региона лица возникает в основном за счет яркости подсветки экрана устройства. В качестве области интереса для нейросетевого классификатора был выбран квадратный регион лица, определяемый на этапе сегментации. Диапазон применимости мобильной биометрической системы подразумевает возможность распознавания устройства с расстояния вытянутой руки (50-60см). В таком случае используемый алгоритм сегментации с учетом разрешения изображений, получаемых от камеры, извлекает квадратные области со стороной, не превышающей 128 пикселей. С целью исключения артефактов, возникающих при повышении разрешения растра, данное значение было выбрано в качестве пространственной размерности входных данных. Каждый регион лица большего размера, возни-

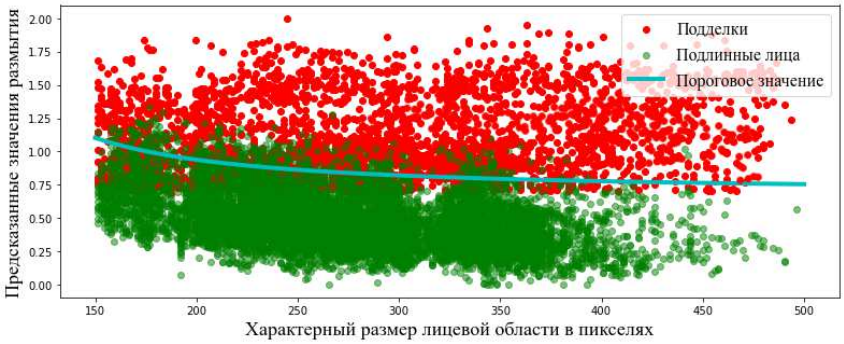


Рис. 2.7. Совместное распределение предсказанных степеней размытия и характерных размеров лица на изображении.

кающий при попытках распознавания с более близкого расстояния, будет билинейно масштабироваться к разрешению 128 пикселей для обработки сетью. Для снижения вероятности переобучения под высокочастотные особенности камер, используемых при съемке базы изображений, цветности входных растров была преобразована из формата RGB в монохромный.

Строение используемой архитектуры дано на Рис 2.6. При обучении применялись аугментации случайного поворота на 90 градусов и случайной яркостной коррекции для регуляризации и повышения обобщающей способности. Для валидационной выборки помимо примеров изображений настоящих лиц были отобраны подделки вида распечаток (PR), снятых без видимых границ вблизи мобильного устройства, экранов смартфонов (SM) и планшетов (SI). Распределение значений предсказанных степеней размытия на отложенной выборке, содержащей примеры размытия для поддельных лиц, в зависимости от размеров квадратной области лица и его типа дано на Рис. 2.7. В ходе исследований было выявлено, что для обнаружения размытия требуется адаптивный порог, зависящий от характерного размера лица в пикселях на изображении. Подобранная форма зависимости значений порога представлена на Рис. 2.7 голубой линией.

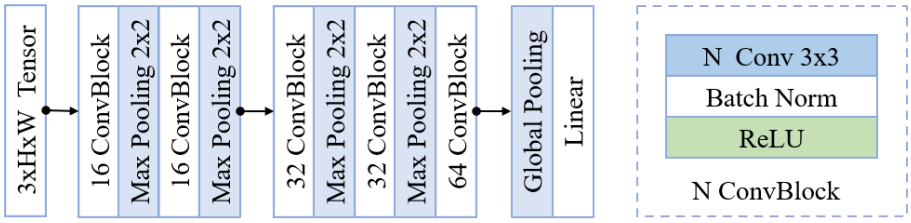


Рис. 2.8. Общая архитектура применяемых для оценки качества лица моделей.

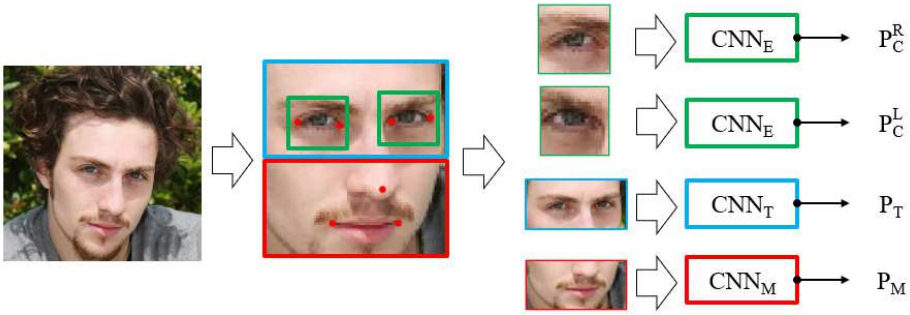


Рис. 2.9. Схема определения атрибутов лица.

2.4.5. Оценка состояния лицевой области

Сценарии повседневного применения биометрической системы с мобильного устройства подразумевают высокую степень изменчивости лицевой области. В частности, допустимы возникновения перекрытий головными уборами, масками или темными непрозрачными очками, что затрудняет верификацию пользователя и снижает количество доступной для детектирования подделок информации, содержащейся в ней. Некоторые попытки подделывания при помощи показа лиц на экранах мобильных устройств или дисплеев также содержат блики, перекрывающие информативные области. В дополнение, требуется учитывать присутствие закрытых глаз, которое может свидетельствовать о попытке верификации вне ведома владельца устройства в случае, если таковой находится в состоянии сна, что создает угрозу безопасности личных данных [53].

С целью обнаружения упомянутых искажений предлагается применить неглубокие нейросетевые классификаторы (Рис. 2.8, позволяющие оценивать вероятности перекрытия верхней области лица P_{T}, области рта P_{M}, а также вероятность присутствия закрытых левого и правого глаз $P_{\mathrm{C}}^{\mathrm{R}}$ и $P_{\mathrm{C}}^{\mathrm{L}}$. Суммарно

требуется обучение четырех моделей, каждая из которых будет применяться для соответствующего региона интереса, извлекаемого из фронтализованного растра области лица и масштабируемого к сравнительно малому разрешению с целью сокращения вычислительной сложности. Фронтализация позволит повысить устойчивость к углам поворота головы в кадре, а небольшой размер получаемых регионов достаточен для сравнительно несложных задач обнаружения артефактов или заметных и измеримых характеристик лица. Пример извлечения регионов интереса приведен на Рис. 2.9. Каждая из моделей обучается отдельно на соответствующей подвыборке при помощи логистической функции потерь. Полученные вероятности рассматриваются как меры качества изображения лица и используются для отказа от идентификации для кадра, в котором присутствуют оба закрытых глаза или перекрытия периокулярной или ротовой областей.

2.4.6. Комбинированный метод детектирования подделок

Описанных ранее методы обнаружения артефактов, характерных для некоторых искусственных примеров рассматриваемой БХЧ, недостаточно для создания системы защиты от более искусно выполненных попыток взлома (Табл. 2.1, таких как вырезанный регион лицевой области (PC или PP), а также от образцов с невидимыми в поле зрения камеры краями, содержащихся в области фокуса камеры.

Для детектирования таких случаев необходим нейросетевой модуль определения живости, который должен учитывать следующую информацию, представленную на входном растре:

- особенности текстуры лица человека: артефакты печати, малозаметные блики, муаровый узор и т.д.;
- неестественные паттерны в лицевой области: видимые обрезанные края двумерной маски лица или отсутствие глубины резкости головы и фона и

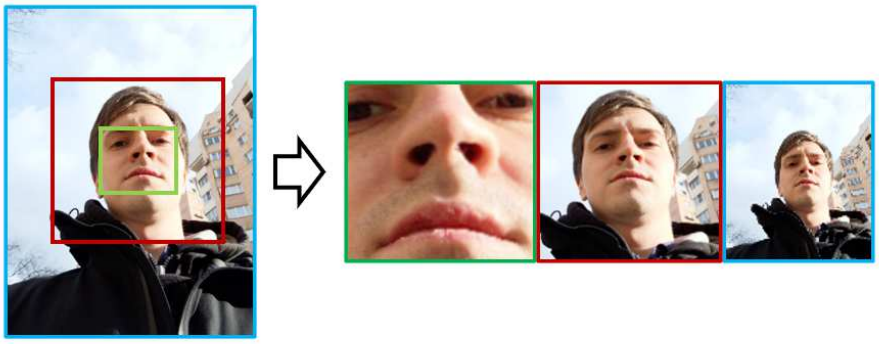


Рис. 2.10. Анализируемые регионы входного изображения.

т.д.;

- нестандартный контекст снимаемой сцены: присутствие кистей человека в кадре, неестественный силуэт и т.д.

Таким образом, для детектирования подделок требуется анализ соответствующих регионов изображения человека при попытке распознавания, Рис. 2.10:

- образец текстуры лица с центром в области носа;
- регион вокруг головы, частично захватывающий задний план, область шеи и плеч;
- весь кадр целиком.

Первый из перечисленных регионов с целью сохранения информации об исходной текстуре лица должен извлекаться без масштабирования, в то время как пространственное разрешение оставшихся регионов может быть снижено с целью уменьшения вычислительной сложности применений нейронных сетей при сохранении производительности. Таким образом, для решения задачи определения живости по одному изображению лица человека требуется обучение как минимум трех сверточных нейронных сетей или блоков, по одному на каждый из описанных регионов интереса. Для построения решающего правила допустимо усреднение получаемых от классификаторов вероятностей, взвешенная сумма

или комбинация упомянутых моделей в одну мультимодальную с несколькими входными слоями. Однако, для практического применения вариант построения ансамбля путем взвешенной суммы предсказаний с подбором значений весов на отложенной валидации обладает рядом преимуществ. Во-первых, в таком случае менее затратно поддерживать и обновлять каждую из сетей по отдельности в случае обнаружения критических случаев некорректной работы. Во-вторых, снижается вероятность переобучения под конкретные комбинации трех регионов, присутствующие в обучающей выборке.

В настоящее время не существует достаточно объемных баз данных изображений или видеозаписей лиц и соответствующих подделок, полученных в сценариях применения мобильной биометрической системы и содержащих значительную изменчивость параметров окружения и видимого образа пользователя. По этой причине для проведения исследований и разработки была собрана собственная база данных. Ее образцы представляют собой видеообразы лица или соответствующих подделок категорий Б и П длительностью от 3 до 10 секунд, Табл. 2.1, полученных при помощи нескольких мобильных устройств с различающимися по характеристикам фронтальными камерами. Процесс съемки базы включал в себя имитацию повседневного применения смартфона в различных условиях освещенности: при ярком солнечном свете, засветках сбоку и сверху, в естественном освещении в помещении, а также в полумраке (1-2 люкс) с включенной подсветкой экрана устройства. Полученные видеообразы лиц включали изменения поз головы человека, направления взгляда, присутствие носимых аксессуаров (очки, маски, головные уборы), а также различные расстояния расположения устройства относительно пользователя. Поддельные образцы также были получены с вариацией освещения, расстояния и пространственной ориентации. Детальная информация о собранной БД представлена в Таб. 2.4. Данная база данных и ее подвыборке применялись для обучения моделей определения живости, в том числе и в экспериментах, описанных ранее.

Для описанных регионов интереса была выбрана общая архитектура моде-

Параметры	REAL	PR	PC	PP	SI	SV
Количество видеороликов, тыс. шт.	830	189	101	40	192	188
Количество сценариев освещенности	5	3				
Количество кадров в видеоролике	30 – 90	20 – 40				
Расстояние съемки	25 – 60 (см)					
Количество пользователей	1220					
Расовая принадлежность	Азиаты & Африканцы & Европеоиды					
Разрешение матрицы камеры	1280 × 960					

Таблица 2.4. Параметры базы данных видеопоследовательностей, использованной в экспериментах.

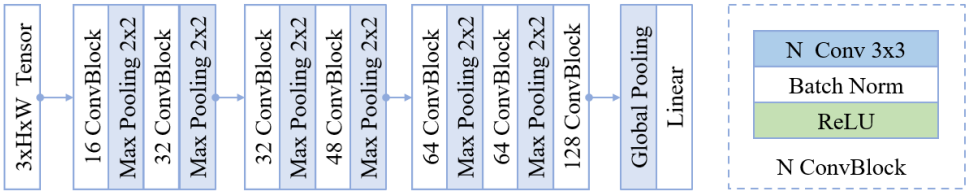


Рис. 2.11. Строение нейронной сети BaseNet.

ли *BaseNet*, представленная на Рис. 2.11. В данной модели применялись обычные сверточные блоки с батч нормализацией (Batch Normalization), демонстрирующие прирост обобщающей способности для решения задач детектирования подделок [149] по сравнению с подходом MobileNet [63]. Для повышения производительности прямого прохода нейронной сети применялась 8-битная квантизация операций свертки [21]. Размер входных данных для региона лица, получаемого без масштабирования, был подобран с учетом минимального возможного размера прямоугольной области, определяемой на этапе сегментации и составил 128 пикселей. Для регионов контекста лицевой области и кадра целиком соответствующие размеры были подобраны с учетом требований производительности сети и составили 140×140 и 128×170 соответственно.

Полное медианное время обработки входных изображения для соответствующих нейронных сетей на одном ядре мобильного процессора Qualcomm

Меры качества,%		Основной алгоритм	+ Поиск границ	+ Лицевые атрибуты	+ Оценка размытия
Кадры	FAR	1.54	1.03	0.98	0.51
	FRR	4.33	4.91	5.10	5.32
	EER	2.89	-	-	-
Видео	FAR,	1.42	0.89	0.83	0.92
	FRR	2.04	2.55	2.64	2.84
	FRR	1.90	2.55	2.64	2.84

Таблица 2.5. Качество обнаружения подделок многостадийным методом.

Snapdragon 888 составило 8, 11 и 13 мс.

2.4.7. Результаты применения многостадийного алгоритма

В соответствии с принятыми в области определениями и понятиями, описанными, например, в [149], а также стандартах ISO/IEC 30107-1:2016 и ГОСТ Р 58624.1-2019, для оценки характеристик системы обнаружения подделок были выбраны следующие:

- FAR (False Accept Rate) — доля изображений подделок, ошибочно классифицированных как живые, также называемый в литературе APCER (Attack Presentation Classification Error Rate);
- FRR (False Reject Rate) — доля изображений живых образцов, ошибочно классифицированных как подделки, иногда также называемый в литературе BPCER (Bona-fide Presentation Classification Error Rate);
- EER (equal error rate) - равный уровень ошибок, при котором FAR=FRR;

Полученные результаты по точности классификации набора данных подлинных и поддельных примеров предложенным решением даны в таблице 2.5. Описанные меры качества оценены при помощи системы автоматического тестирования на отложенной валидационной подвыборке описанной ранее базы данных (Табл. 2.4). Подвыборка состоит из 10% полной выборки и из уникаль-

ных субъектов общей базы, не содержащихся в обучающих наборах используемых методов.

Процедура тестирования применялась без учета результатов систем извлечения и сравнения признаков и опиралась лишь на выходные данные модуля образмеривания (сегментации) лиц. Было рассмотрено два сценария тестирования:

1. Покадровое формирование предсказаний живости без учета принадлежности изображений видеопоследовательностям;
2. Оценка живости для видеопоследовательностей с экспоненциальным сглаживанием предсказаний и ранним отказом от дальнейшей обработки в случае обнаружения последовательности $N = 10$ кадров, помеченных как «подделка» модулями предварительной оценки качества 2.4.1;
3. Вычисление показателей точности классификации (Таб. 2.5).

В режиме обработки видеопоследовательностей таковые считались определенными как «подлинные», если и только если хотя бы один кадр был определен как таковой и при этом не произошло раннего отказа от продолжения процедуры. Это объясняет достаточно высокие значения характеристики FRR для покадрового режима, которые частично компенсируются учетом всех кадров видеообраза лица.

Анализ быстродействия Применимость предлагаемого подхода для сценариев распознавания с помощью мобильного устройства была исследована при помощи смартфона, оснащенного процессором Qualcomm Snapdragon 888 (2.84 GHz, Quad-core) в однопоточном режиме. Медианное время выполнения составило 10 и 32 (мсек) для операций первой и второй стадии (2.4.1, Рис. 2.3) соответственно.

2.5. Выводы ко второй главе

Исследованы подходы к построению модулей противодействия взлому при помощи подделок в составе систем биометрического распознавания по видеообразу лица, предназначенные для использования в мобильных устройствах и учитывающие характерные для такого сценария применения ограничения и особенности. Предложены, протестированы и внедрены:

1. новая многостадийная структура метода детектирования подделок, основанная на применении блоков обнаружения набора характерных для искусственных БХЧ артефактов, позволяющая выполнять определение живости человека при помощи устройства с существенно ограниченными вычислительными ресурсами в режиме реального времени (≈ 25 кадров/сек.), удовлетворяющая критериям ошибок: $FRR \leq 3\%$ при $FAR \leq 1\%$.

2. алгоритм раннего обнаружения поддельных примеров, позволяющий:

   - осуществлять разностороннюю оценку атрибутов лица человека на изображении с целью обнаружения наиболее распространенных способов подделывания;
   - производить анализ условий окружения при помощи вспомогательных данных с установленных в мобильное устройство сенсоров вспомогательных для их учета при принятии решения о живости представленного системе лица;

Глава 3

Определение живости лица при помощи стереокамеры

В отличие от систем, использующих трудно воспроизводимые биометрические признаки, такие, как рисунок отпечатка пальца или текстура радужки, изображение лица человека несложно получить с целью создания подделок. Множество систем распознавания лиц использует изображения в видимом диапазоне, что позволяет осуществлять спуфинг-атаку на такую систему при помощи обычной качественной фотографии, показываемой на цифровом экране или распечатанной на принтере высокого разрешения (Глава 2).

Большинство представленных на рынке мобильных устройств с системой распознавания по лицу, дающей доступ к личной информации пользователя и осуществлению платежных операций, оборудовано дополнительными датчиками [13, 51] для обеспечения высокого уровня безопасности против спуфинг-атак, как правило сенсорами определения глубины. При этом небольшая группа устройств [52, 118] снабжена парой фронтальных камер, позволяющих оценивать глубину снимаемых сцен алгоритмами стереозрения. По сравнению с большинством более совершенных датчиков, используемых для оценки глубины сцены, дополнительная фронтальная камера вносит небольшую добавочную стоимость в систему по ряду причин. Во-первых, упомянутые сенсоры применяют технологию активной подсветки для получения карты глубины, что требует установки источника этой подсветки и приемника — дополнительной камеры, зачастую восприимчивой к ИК-излучению, что делает ее совмещение с основной фронтальной камерой невозможным. Во-вторых, два добавочных датчика требуют для установки дополнительное пространство на передней панели мобильного устройства, большую часть которой занимает сенсорный дисплей. Поэтому разумно исследовать возможности применения фронтальных стерео-

камер мобильных устройств для решения задачи *анти-спуфинга* в системах распознавания по лицу.

3.1. Обзор методов определения живости лица с помощью аппаратных средств

Подделки лица человека можно разделить на три группы по способу их создания: распечатки изображения лица, цифровые изображения или видео и лицевые маски [149]. Сложность детектирования подделок во многом определяется качеством используемых материалов и устройств. Важнейшим принципом обнаружения подделки лица является определение трехмерных характеристик видимой сцены. Первые два способа здесь отсекаются, третий становится значительно более трудоемким.

Описанные в литературе подходы методы анти-спуфинга можно разделить на две группы: использующие дополнительное оборудование (сенсоры глубины [129], камеры в ближнем инфракрасном диапазоне (ИК-камеры) [125], тепловые камеры [121]) и основанные исключительно на программной обработке входного изображения. Методы первой группы позволяют решать задачу детектирования подделок с высокой точностью, однако их применение на практике существенно увеличивает стоимость системы. Вторую группу можно разделить на две подгруппы: кооперативные и некооперативные методы. Кооперативные методы требуют выполнения определенных движений лицом и/или его частями в соответствии с запросом системы, что повышает уровень безопасности, но раздражает пользователя и увеличивает время отклика системы.

Системы биометрической идентификации в мобильных устройствах, таких, как смартфоны и ноутбуки, должны иметь малое время отклика, возможность работать на ограниченных вычислительных ресурсах, допускать применение в разнообразных и неконтролируемых условиях съемки. В случае с изображением лица человека в видимом спектре изменчивость условий съемки делает

возможным появление практически идентичных низкокачественных изображений настоящих лиц и подделок [149]. Эти факторы существенно ограничивают набор подходов к решению задачи.

Задача построения карты глубины сцены по изображениям стереокамеры является классической [5]. Ее решение, как правило, состоит из нескольких этапов: калибровка стереопары, построение карты смещений (диспаратности) и последующее построение карты глубин с учетом параметров калибровки.

Более современный подход – применение сверточных нейронных сетей [27, 72, 152]. При наличии достаточно большого по объему и разнообразию набора входных изображений можно решать как задачу оценки смещений [27], так и извлечения глубины изображения в разнообразных условиях освещенности [72]. Недостатками этих подходов применительно к их использованию в мобильных устройствах являются высокие вычислительные затраты и необходимость большого стереобазиса, такого, как, например, в беспилотных автомобилях. Описаны менее ресурсоемкие подходы [155]. Для предсказания глубины сцены авторы предлагают задействовать как информацию от мобильной стереопары с малым расстоянием между центрами камер, так и от подсистемы фазового фокуса одной из камер (PDAF, phase detection auto focus).

Большинство описанных в литературе методов, использующих стереоизображения для антиспуфинга, так или иначе пытаются извлечь информацию о глубине изображения или видео лица человека для определения его живости [128]. Эти подходы имеют преимущество перед однокадровыми, особенно для изображений, полученных в разнообразных условиях съемки. Как правило, такие алгоритмы основаны на классических или нейросетевых классификаторах, обрабатывающих признаки карты смещений или карты глубин. В [113] авторы предлагают способ построения приблизительной карты диспаратности при помощи небольшой сверточной нейронной сети, предобученной на целевом домене. После этого этапа обучается вторая нейронная сеть, использующая признаковые описания, которые получены от первой.

В [84] для предсказания метки класса предлагается нейронная сеть, построенная по принципу автокодировщика и состоящая из двух частей. Кодирующая часть нейронной сети извлекает промежуточные признаки, которые затем используются декодирующей частью для регрессии значений диспаратности. Результат обработки входных изображений декодером затем подается в небольшую сверточную нейронную сеть для классификации. Обе части нейронной сети обучаются совместно как на регрессию истинных значений диспаратности, так и на предсказание правильной метки класса, что повышает обобщающую способность полученной нейронной сети. Аналогичный подход был применен в ряде работ по детектированию подделок и привел к повышению точности итоговых решений [87]. Недостатком этого подхода является вычислительная сложность, поскольку для предсказания метки класса требуется пропустить входную пару изображений через обе части нейронной сети.

Все упомянутые работы используют обучающие выборки, полученные при помощи стереокамер с большим стереобазисом (более 4 см), что повышает устойчивость и точность восстановления трехмерных признаков. Однако типичные стереокамеры мобильных устройств имеют расстояния между центрами сенсоров не более 2 см.

3.2. Обнаружение подделок при помощи стереозрения

Выбор функции потерь

Детектирование подделок – это задача бинарной классификации, при решении которой при помощи методов машинного обучения часто применяют логистическую функцию потерь или перекрестную кросс-энтропию:

$$L_0 = \frac{1}{K}\sum_{i=1}^{N}\left(-y^i \log a^i - (1-y^i)\log(1-a^i)\right) \to \min\,, \tag{3.1}$$

где $y^i \in \{0;1\}$ – метка класса, a^i – ответ алгоритма на i-м примере обучающей выборки, имеющей размер K.

Пусть классификатором является нейросеть $\mathbb{N}(I_0, I_1; w)$ с набором весов w. Ее предсказания соответствуют вероятности принадлежности лица на стереоизображении к положительному классу (в данном случае – к классу «настоящее лицо»):

$$a^i = \mathbb{N}(I_0^{(i)}, I_1^{(i)}; w) = P(y^i = 1; w). \tag{3.2}$$

Представим упомянутую сеть в виде композиции подсетей, осуществляющих извлечение признаков из пары изображений $\mathbb{N}_{\text{f}}(I_0, I_1; w_{\text{f}})$ и предсказание метки класса $\mathbb{N}_{\text{o}}(I; w_{\text{o}})$:

$$\mathbb{N}(I_0, I_1; w) = \mathbb{N}_{\text{o}}(\mathbb{N}_{\text{f}}(I_0^{(i)}, I_1^{(i)}; w_{\text{f}}); w_{\text{o}}). \tag{3.3}$$

В работе подразумевается, что при генерации метки класса итоговая модель должна опираться на отличия глубины сцен, содержащих настоящие и поддельные лица. При этом процедура минимизации функции потерь не гарантирует того, что модель научится извлекать релевантные и устойчивые признаки для корректных предсказаний на отложенных данных. Более того, как показывают эксперименты, оптимизация лишь перекрестной кросс-энтропии не обеспечивает хорошей обобщающей способности результата обучения в случаях ограниченных по разнообразию и размеру обучающих выборок.

В литературе описаны способы повышения обобщающей способности нейронных сетей за счет многоцелевого обучения [115]. Оптимизация осуществляется для суммы нескольких функций потерь, соответствующих разным, но связанным между собой подзадачам. В результате при прохождении входного сигнала через обученную нейронную сеть в ней возникают промежуточные представления, порождающие более устойчивое признаковое описание для решения каждой из подзадач, в том числе и основной. Обобщающая способность повышается за счет того, что процедура обучения не позволяет весам нейронной сети оказаться в тривиальном для данного набора данных локальном оптимуме.

Применение данного подхода описано для задачи детектирования подделок среди цветных изображениям лиц [48, 87]. Полученные модели показывали

лучшее качество по сравнению с их аналогами с единственной классификационной функцией потерь.

Предлагается использовать вспомогательную функцию потерь, которая позволит нейронной сети извлекать информацию о глубине представленной на стереоизображении сцены. Для этого требуется добавить в модель подсеть $\mathbb{N}_{\text{a}}(I; w_{\text{a}})$, которая предсказывает карту принадлежности пикселей A к переднему плану с помощью признакового описания, полученного подсетью $N_{\text{f}}(.,.; w_{\text{f}})$. Каждый элемент A может принимать только значения 0 или 1. Примеры таких карт для истинного и поддельного лица показаны на Рис. 3.1.

а б

Рис. 3.1. Карты принадлежности пикселей переднему плану: (а) настоящего лица; (б) подделки.

Для каждого пикселя первого из входных изображений I_0 с координатами $m \in [1; M], n \in [1; N]$ нейронная сеть $\mathbb{N}_{\text{a}}(.; w)$ должна также предсказать вероятность:

$$b_{mn} = \mathbb{N}_{\text{a}}(\mathbb{N}_{\text{f}}(I_0, I_1; w_{\text{f}}); w_{\text{a}})_{mn} = P(A_{mn} = 1; w_{\text{a}}, w_{\text{f}}). \tag{3.4}$$

В таком случае внутренние представления обученной нейронной сети будут содержать в себе информацию, связанную с особенностями глубины изображений и величинами смещений между левым и правым изображениями стереопары. Для моделирования принадлежности пикселей предлагается исполь-

зовать сигмоидную функцию активации и для обучения применять логистическую функцию потерь 3.1 для каждого из пикселей по отдельности:

$$L_1^i = \frac{1}{MN} \sum_{m=1}^{M} \sum_{n=1}^{N} \left(-A_{mn}^i \log b_{mn}^i - (1 - A_{mn}^i) \log(1 - b_{mn}^i) \right) \to \min \,, \tag{3.5}$$

$$L_1 = \frac{1}{K} L_1^i \,. \tag{3.6}$$

Итоговая функция потерь определяется как сумма: $L = L_0 + L_1$.

Предварительная обработка входных данных Большое значение для качества работы классификатора имеет предварительная подготовка данных, минимизирующая различия в условиях регистрации. Здесь рассматриваются ректификация, определение ориентации лица и выбор цветового пространства.

Калибровка и ректификация. Первым этапом предобработки изображений стереопары является *ректификация* – приведение изображений к некоторому стандартному виду путем компенсации искажений, вносимых индивидуальными особенностями камер. Ректификация производится на основании калибровочных данных, (матрица внутренних параметров камеры и коэффициенты дисторсии) могут быть извлечены из памяти устройства [50]. Каждая камера современных мобильных устройств калибруется на производственных линиях на специальной установке, как правило, еще до окончательной сборки корпуса. К сожалению, при сборке, транспортировке и эксплуатации положения сенсоров и/или линз камер может измениться, что приводит к несоответствию калибровки действительным параметрам камер [38]. Дефекты калибровки незаметны в большинстве приложений, однако являются существенными для фотограмметрии. Повторная калибровка на стороне пользователя нежелательна даже в автоматизированном режиме, поскольку это снижает удобство и вносит значительный риск некорректного выполнения.

По этой причине в данной работе используются изображения, полученные без ректификации.

Определение ориентации лица. Особенностью использования мобильных устройств является то, что их ориентация при распознавании может быть различной: портретной и ландшафтной. Детектирование подделок происходит после этапа образмеривания.

Среди точек, полученных при образмеривании, содержатся положения центров глаз $\mathbf{p}_R = (x_R, y_R)$ и $\mathbf{p}_L = (x_L, y_L)$, эти координаты можно использовать для определения ориентации входного растра. Угол наклона прямой, соединяющей зрачки к оси OX, равен

$$\alpha_{LR} = \operatorname{arctg}\left(\frac{y_R - y_L}{x_R - x_L}\right)$$

и определяет ориентацию R входного кадра (рис. 3.2):

$$R = \begin{cases} 0, |\alpha_{LR}| < \frac{\pi}{4}\ , \\ 90, \frac{\pi}{4} < \alpha_{LR} < \frac{3\pi}{4}\ , \\ 180, |\alpha_{LR}| > \frac{3\pi}{4}\ , \\ 270, -\frac{3\pi}{4} < \alpha_{LR} < -\frac{\pi}{4}\ . \end{cases} \tag{3.7}$$

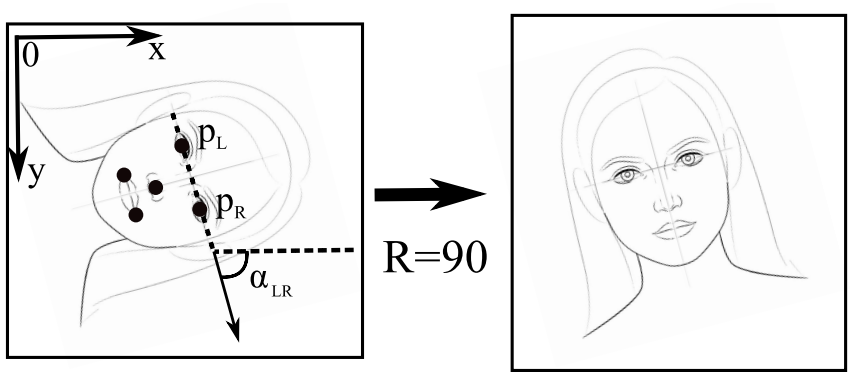


Рис. 3.2. Схема компенсации ориентации лица на входном растре

Значение R задает угол, на который требуется повернуть исходный растр против часовой стрелки, чтобы ориентация лица на нем стала естественной, строго «подбородком вниз».

Выбор цветового пространства. Цветовые каналы RGB-представления изображений сильно скоррелированы, поэтому для повышения обобщающей

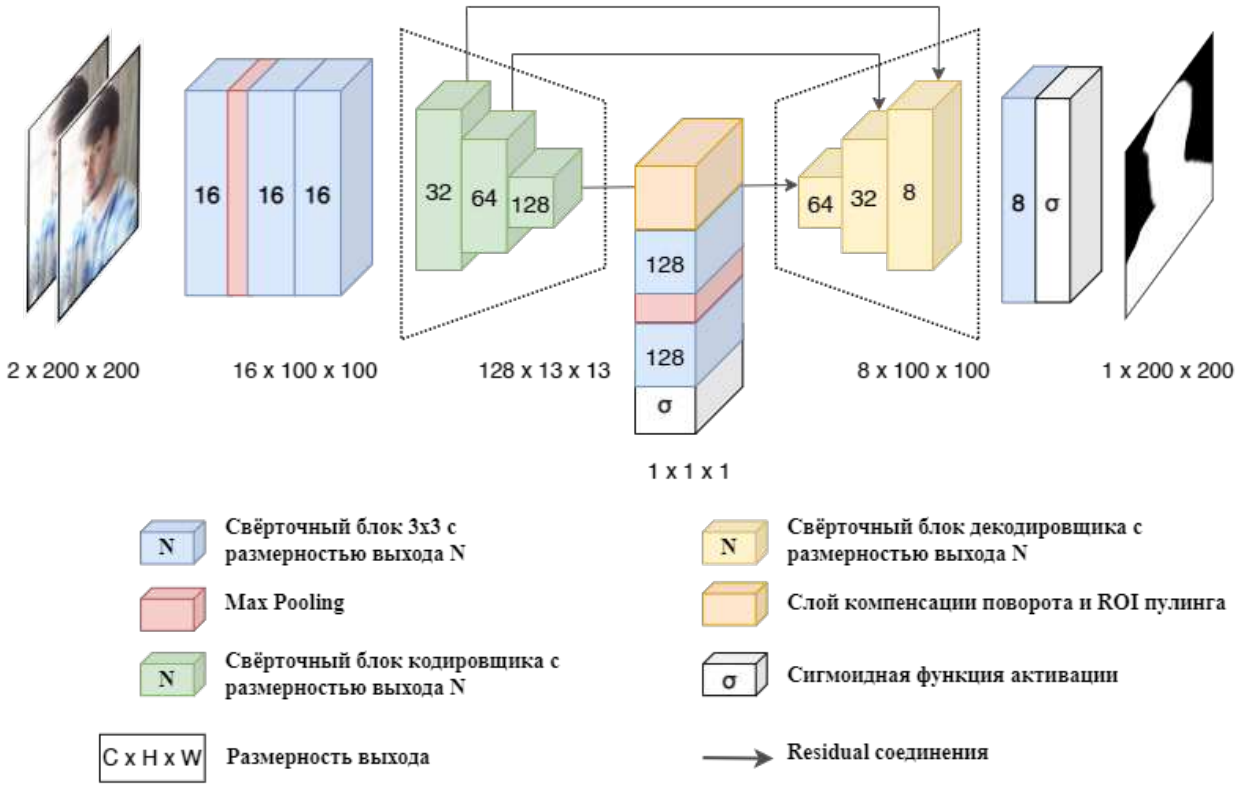


Рис. 3.3. Архитектура нейросетевого решения

способности нейронных сетей предлагается применять либо иные цветовые пространства, либо избавляться от цветности вовсе [17]. В работе изображения перед подачей в нейронную сеть преобразовываются к одноканальному (монохромному) представлению. К полученным парам растров применяется алгоритм блочного приведения гистограмм [8], чтобы избавиться от искажений, вносимых расхождением автоэкспозиций стереокамер.

Архитектура решения

В качестве основы для построения нейронной сети была выбрана сравнительно легковесная архитектура семейства UNet [114] с добавлением остаточных блоков [57] для улучшения сходимости. Модель устроена по принципу автокодировщика, составленного из комбинаций сверточных блоков и операций сокращения пространственной размерности (пулинга). Общее строение представлено на рис. 3.3, архитектуры блоков кодирующей и декодирующей частей сети – на рис. 3.4. Каждый элемент блок-схемы содержит название операции, ее параметры и размер результирующего тензора. Символом s' обозначена вели-

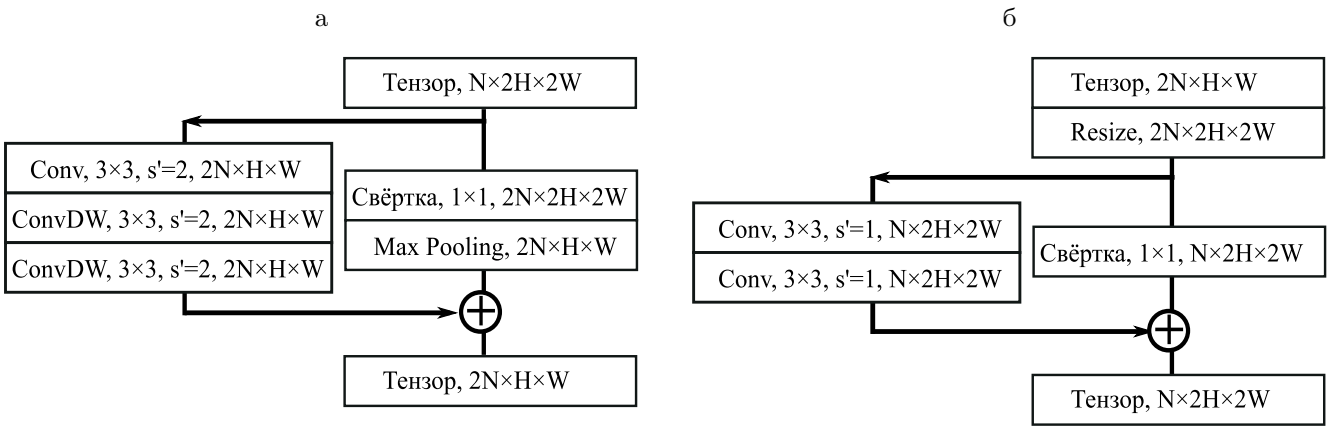


Рис. 3.4. Строение используемых сверточных блоков: кодирующей (а) и декодирующей (б) компонент.

чина смещения фильтра операции свертки. Большая часть сверточных блоков кодировочной части построена по принципу, описанному в [63] для повышения производительности архитектуры. Строение используемых сверточных блоков дано в Таблице 4.1.

Пары изображений подаются в модель в исходной их ориентации (как они получены с сенсоров камер), чтобы сохранить горизонтальным направление смещений соответственных пикселей и упростить задачу сегментации переднего плана для декодирующей части. При этом промежуточное признаковое описание лица после обработки входного сигнала кодировщиком может быть представлено в некорректной ориентации 3.7. Чтобы упростить классификацию, имеет смысл повернуть это признаковое описание пространственно на угол, кратный $\pi/2$, таким образом, чтобы линия уровня глаз соответствовала ориентации $R = 0$. На рис. 3.3 эта операция обозначена как слой компенсации поворота.

Преобразование карты признаков.

В сетях, построенных по принципу автокодировщиков, нейроны внутреннего представления обычно имеют большие рецептивные поля, охватывающие значительные связные области пикселей входного изображения. Такие представления с малой пространственной размерностью и большим количеством кана-

лов содержат богатое агрегированное признаковое описание. В случае предсказания карты глубины или аналогичных задач с попиксельным предсказанием некоторых значений важно при построении ответа с помощью декодирующей части сети использовать внутреннее представление целиком. Однако задача детектирования подделок является локальной: содержательная часть признаков «живости» пространственно локализована в области лица на исходном растре с поправкой на операции пулинга. Отличия подделок от настоящих лиц содержатся именно в особенностях карты глубины вокруг лицевой области: у настоящего лица в этой области присутствует резкий перепад по отношению к заднему плану и плавные перепады на переднем плане, а у поддельного перепад к фону и на переднем плане отсутствует. Более того, учет информации от заднего плана может привести к переобучению в связи с ограниченным размером обучающей выборки.

При этом подача растра лицевой области напрямую в нейронную сеть нецелесообразна, так как часть информации о соотношении глубины фона и переднего плана может потеряться. Более того, лицо на изображении может занимать различную площадь ввиду разнообразия расстояний съемки, поэтому будет необходимо приведение входных данных к общему размеру, что может исказить исходную карту смещений между правым и левым растром в стереопаре. В работе предлагается отобразить область лица в промежуточном представлении входного сигнала нейронной сети в результирующий тензор T_{o} фиксированного размера $S \times S$ для его последующей обработки блоком предсказания метки класса. Для отображения используется билинейная интерполяция.

Область лица задается в данном случае как прямоугольник $C = (x, y, w, h)$, где x, y — положение центра прямоугольной области, w, h — ее ширина и высота соответственно. Значения параметров вычисляются из положения ключевых точек на лице:

$$x = \frac{1}{4}(x_L + x_R + x_{ML} + x_{MR}), \tag{3.8}$$

$$y = \frac{1}{4}(y_L + y_R + y_{ML} + y_{MR}), \tag{3.9}$$

$$w = h = 2\sqrt{(x_R - x_L)^2 + (y_R - y_L)^2} \ . \tag{3.10}$$

В данной работе разрешение результирующего представления области лица $S = 10$ выбрано с учетом средних параметров области интереса, определенных на обучающей выборке.

3.3. Экспериментальные результаты

Предложенный метод протестирован на различных наборах стереоизображений как общедоступных, так и собранных вручную. Большое внимание уделено получению изображений с разнообразными условиями регистрации.

Формирование базы изображений.

В литературе описано несколько баз стереоизображений для задачи детектирования подделок, полученных при помощи полноразмерных стереопар с большим стереобазисом [84, 113]. Набора данных для мобильных приложений и поставленной задачи в открытом доступе авторы не нашли. Тем не менее, известна обширная база изображений Holopix [64], полученная при помощи мобильной стереокамеры и содержащая большое разнообразие типов сцен, среди которых присутствует класс «селфи», соответствующий классу «настоящих лиц» в контексте данной работы. При съемке были использованы два вида сенсоров: со стереобазисом в 12 и 5 мм. Выборка содержит изображения различных разрешений, снятые в разнообразных условиях. Среди 50 тыс. растров этой выборки лишь 1052 можно отнести к классу «селфи» (selfie, self-photo), который подразумевает наличие единственного лица в кадре и его расположение на расстоянии от 20 до 60 см до камеры (рис. 3.5).

Для полуавтоматического отбора изображений применялся алгоритм детектирования лиц [153]. Упомянутый набор данных допустимо использовать в

(а) (б) (в)

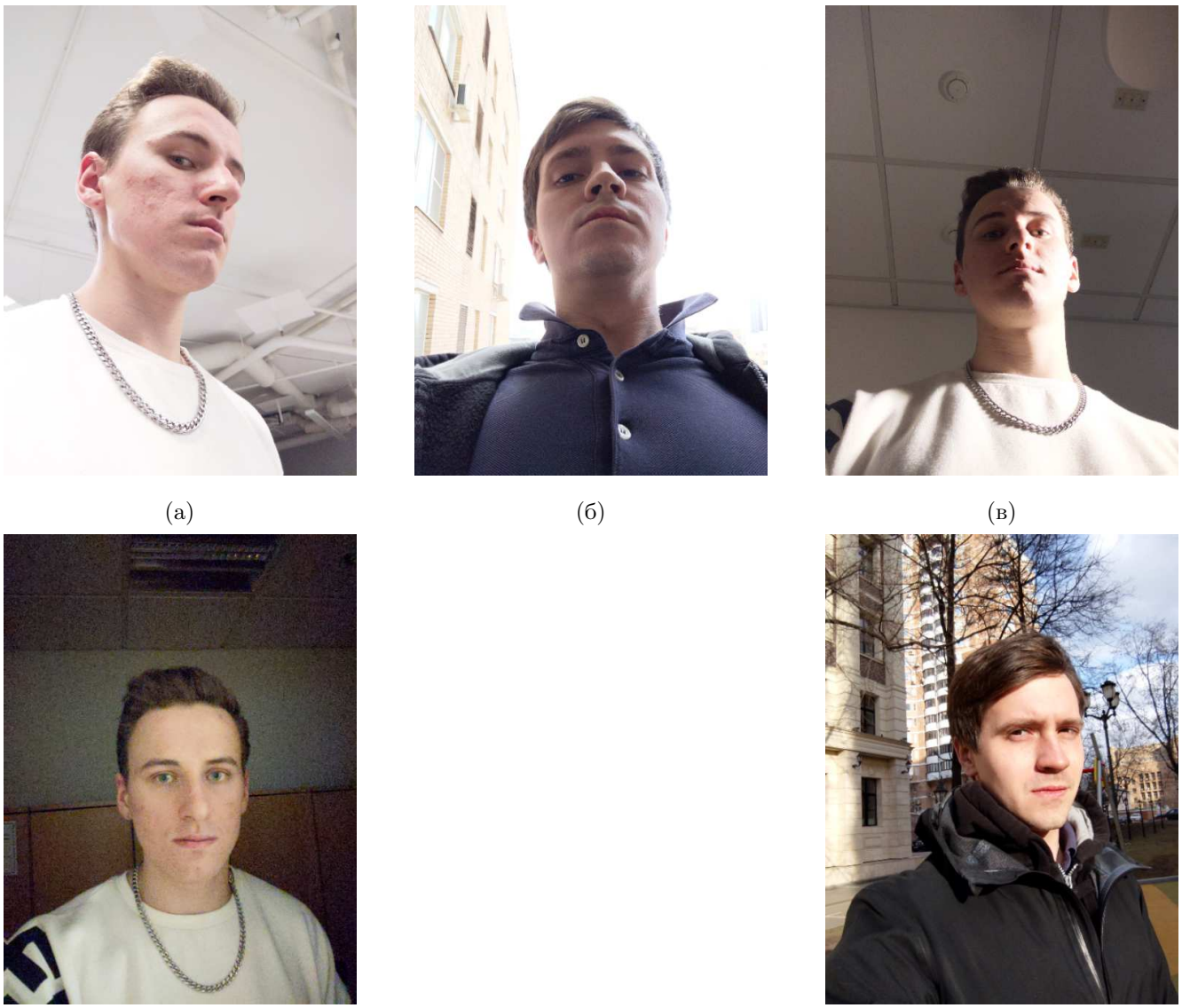
Рис. 3.5. Примеры изображений выборки Holopix.

качестве отложенной тестовой выборки для проверки обобщающей способности алгоритма.

По причине нехватки открытых баз изображений был осуществлен дополнительный сбор данных при помощи Google Pixel 3 – одного из смартфонов с двойной фронтальной камерой, позволяющей получать изображения с обоих сенсоров одновременно. Выборка изображений лиц получена от 90 участников обоих полов. Каждому участнику предлагалось принять участие в съемке трех сценариев освещенности: естественное освещение в помещении (E1), яркая засветка с одной из сторон или всей сцены целиком (E2, E3 или E5) и съемка в полутемном помещении (E4). Примеры изображений даны на Рис. 3.6

Полученные фотографии частично были использованы для создания изображений подделок следующих типов: распечатка лица (PR, printed), лицо на экране высокого разрешения (SI, screen image) и лицо на небольшом дисплее мобильного устройства (SM, smartphone). Примеры изображений подделок даны на рис. 3.7. Сбор изображений подделок происходил как минимум в двух условиях освещенности: при достатке (E1 или E5) и недостатке света (E4).

Съемки каждого участника и подделок его лица осуществлялась в двух ориентациях мобильного устройства: портретной и ландшафтной. Применялись два расстояния до сенсора камеры: порядка $25 - 30$ см, что соответствует ком-



(а) (б) (в)

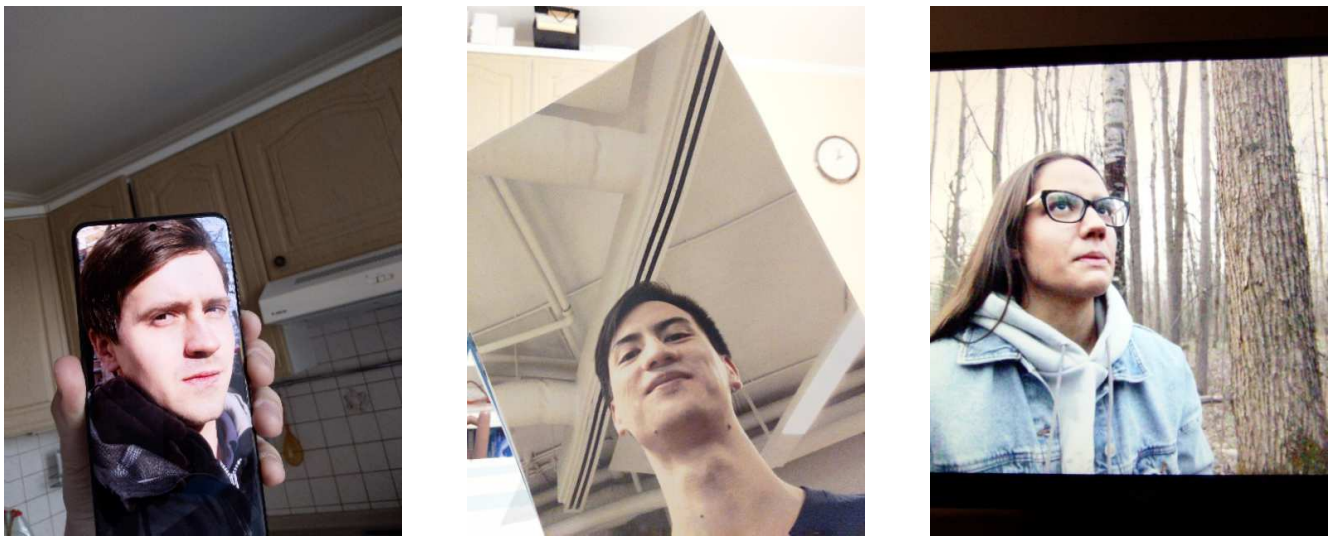
(г) (д)

Рис. 3.6. Рассмотренные в собранной базе изображений условия окружения.

фортному положение смартфона относительно глаз, и порядка 45 – 60 см – положение смартфона в вытянутой руке. Подробное описание полученного набора изображений приведено в табл. 3.1.

Разбиение на обучающую и валидационную выборки осуществлялось в пропорции 7 к 3. При этом изображения одного и того же участника помещались лишь в одну из подвыборок.

Положения лиц и ключевых точек на каждом из полученных изображений определены при помощи метода [153]. Для формирования бинарных карт принадлежности пикселей переднему плану, описанных ранее, использован один



(а) (б) (в)

Рис. 3.7. Исследуемые типы подделок.

из методов вычисления оптического потока с его последующей бинаризацией [88]. Грубые ошибки определения оптического потока, возникающие вследствие некорректной работы камеры мобильного устройства и/или съемки в сложных условиях освещенности, исключены из обучающей выборки.

Границы применимости метода.

Как известно, разрешающая способность стереопары по глубине dZ, согласно эпиполярной геометрии [11], зависит от нескольких параметров: величины стереобазиса B, фокусного расстояния используемых сенсоров f, погрешности измерения смещений d и самой глубины данной точки Z:

$$dZ = Z^2 \frac{d}{fB}. \tag{3.11}$$

Мобильное устройство Google Pixel 3 имеет следующие характеристиках камер: $B = 10\text{мм} = 0.01\text{м}$, разрешение сенсоров $W \times H = 2448 \times 3264$ точек, $f = 4.5\text{мм} = 0.0045\text{м}$, апертура 1/1.8. Оценочный размер сенсора по разрешению и значению апертуры составляет порядка $\omega = 5\text{мм} = 0.005\text{м}$. Таким образом, один пиксель на выходном изображении имеет физический размер $\omega/W = 2 \cdot 10^{-6}\text{м}$. При этом в целях повышения скорости обработки входных изображений нейронной сетью пространственное разрешение требуется сокра-

Таблица 3.1. Описание использованной базы изображений

Тип освещенности	Количество изображений			
	Настоящие лица	Подделки		
		SM	PR	SI
E1	11326	5667	5844	7117
E2	7840	1839	8923	3006
E3	10980			
E4	3240			
E5	10335	4555	4617	4822
Всего	43721	12221	19384	15936

тить до некоторого

$$W_{\text{CNN}} = \frac{W}{s}\,, \tag{3.12}$$

где $s > 1$ – коэффициент масштабирования, который можно определить в рамках поставленной задачи.

В данной работе предполагается, что обученная нейронная сеть должна уметь отличать настоящее лицо от плоского поддельного на расстоянии от 20 до 60 см при помощи информации о глубине сцены, которая содержится в стереоизображениях. Характерный размер головы человека можно определить как 20 см, поэтому величина dZ не должна превосходить это значение, чтобы потенциально извлекаемая карта глубины могла различать видимое лицо на фоне близкого плоского объекта позади, в худшем случае при $Z = 0.6$м.

Точность определения смещений d задается равной 1 пикселю на растре пониженного разрешения W_{CNN}, т.е.:

$$d = \frac{s\omega}{W} = 2 \cdot 10^{-6} s. \tag{3.13}$$

Требуется выполнение следующего неравенства:

$$Z^2 \frac{s\omega}{fBW} < dZ \leftrightarrow s < \frac{fBW \cdot dZ}{\omega Z^2}. \tag{3.14}$$

Подставляя ранее определенные значения, получаем

$$s \lessapprox 12.24. \tag{3.15}$$

В таком случае наименьший допустимый размер используемого изображения в пикселях равен

$$\frac{W}{s} = \frac{2448}{12.24} = 200\ .$$

Это значение было выбрано в качестве разрешения входных изображений для нейронной сети.

Ограничение сверху на разрешающую способность также определяет невозможность применения предлагаемого метода для детектирования подделок в виде плоских масок. Для детектирования этого типа спуфинга требуется разрешение $dZ \approx 5\text{см} = 0.05\text{м}$, дабы извлекать информацию о геометрии лица и его частей. Тогда требуется использовать коэффициент масштабирования $s_{\text{mask}} = \frac{1}{4}s \lessapprox 3.06$, что определяет минимальный размер изображения в 816 пикселей. Применение нейронных сетей в условиях видеопотока с таким пространственным разрешением в реальном времени на маломощных мобильных вычислительных устройствах в условиях ограничений внутренней защищённой ОС [15] затруднительно.

Процедура обучения.

Все эксперименты с обучением нейронных сетей проводились на обучающей части выборки с контролем на валидационной части. Модели обучались методом стохастического градиентного спуска с адаптивным моментом. Первоначальное значение темпа обучения составляло 0.001 и уменьшалось экспоненциально в 0.9 раз каждые 10 эпох. Обучение каждой модели проводилось на протяжении 128 эпох. Для предсказания метки класса при валидации использовались лишь кодирующий и классификационный блоки нейронной сети. Декодирующий блок не участвовал, поскольку его предназначение – лишь регуляризация обучения.

Во время обучения для повышения устойчивости модели к вариациям входных данных применялись аугментации случайной яркостной коррекции, наложения случайного пуассоновского шума, случайного аффинного поворота на угол до 20° относительно оптического центра изображения, случайного отображения по горизонтали и извлечения случайного региона фиксированного размера вокруг области лица, меньшего, чем размер исходного изображения.

Помимо этого, в отдельном эксперименте была применена операция случайного обнуления смещений между пикселями пары: для некоторых примеров вне зависимости от метки класса одно из них приравнивалось к другому, результату присваивалась метка «поддельное лицо» и вспомогательная маска 3.4 заполнялась нулями. Интуиция подобного подхода состоит в регуляризации процедуры обучения нейронной сети. В результате описанного преобразования построенные в модели признаковые описания должны опираться на особенности карт смещений, а не на текстурные характеристики растров. Нейронная сеть, обученная таким образом, помечена как “RZ — RandomZero”.

Сравнение модификаций предлагаемого подхода.

Поскольку осуществить сравнение качества решения предлагаемого метода с аналогами затруднительно ввиду различий области применения и источников входных данных, решено осуществить сравнение с базовыми алгоритмами, не использующими стереоинформацию.

В работе рассматривается несколько способов построения нейросетевого классификатора для решения задачи антиспуфинга. Самый простой способ – использование лишь одного изображений из стереопары в черно-белом режиме для предсказания метки класса. Такая модель применена в качестве базового алгоритма с именем “Base”. Этот подход можно усложнить, добавив интерполяцию карты признаков вокруг региона лица. Модели с такой модификацией обозначены как “ROI”.

Основной способ предсказания метки класса – использование пары изображений в черно-белом режиме без добавления декодирующего блока (разд. 3.2)

и без применения интерполяции признаков региона лица (разд. 3.2). Этот класс моделей имеет наименование “Stereo”. Далее этот подход можно развить, добавив соответствующие модификации, первую из которых предлагается обозначить через “Aux”.

Результаты вычислительных экспериментов с разными модификациями предлагаемого подхода даны в табл. 3.2.

Таблица 3.2. Качество решения на отложенной выборке

Модель	Значения мер, %		
	APCER	BPCER	EER
Base	**0.12**	41.31	12.54
Base+ROI	0.58	15.06	4.82
Stereo	2.35	13.02	4.95
Stereo+Aux	0.89	**2.9**	1.89
Stereo+Aux+ROI	0.57	3.01	1.45
Stereo+Aux+ROI+RZ	0.23	5.24	**1.24**

Основной мерой качества в задаче детектирования подделок считается значение равной ошибки классификации (EER). Значения APCER и BPCER отражают склонность моделей к присваиванию метки класса «1» или «0» на пороге принятия решения 0.5.

По итогам вычислительного эксперимента наилучшая точность решения на валидационной выборке достигается для модели типа “Stereo” со всеми упомянутыми выше модификации. Базовые модели этого типа чаще присваивают входным изображениям метку класса «подделка». При этом добавление модификаций, призванных получить более информативные признаки для решения поставленной задачи, действительно повышает обобщающую способность моделей, что отражается на итоговых мерах качества.

Модели типа “Stereo” позволяют получить лучшее решение задачи по сравнению с моделями типа “Base”. При этом добавление модификации “ROI” да-

ет возможность повысить производительность до уровня базовой модели типа “Stereo” без модификаций. При этом модели “Stereo” используют больше информации при работе. Скорее всего, это связано с тем, что модификация “ROI” снижает склонность сети к переобучению на признаках заднего плана, а текстурной информации лицевой области растров достаточно, чтобы достичь сравнительно высокой точности решения на валидационной выборке.

В качестве отложенной тестовой выборки была использована подвыборка набора изображений [64]. Для оценки выбрана модель “Stereo+Aux+ROI+RandomZe порог принятия решения был принят равным порогу меры EER: 0.38. В результате 970 из 1052 пар растров было помечено алгоритмом как «настоящее лицо». Это соответствует точности классификации в 92.2%. Среди ошибок классификации большую часть (51 пример) составляют пары, содержащие изображение лица на большом расстоянии и снятые на сенсоры со стереобазисом в 5мм, что можно понять из визуализации карты смещений для этих примеров. Оставшиеся ошибочные предсказания содержат, напротив, лица, снятые с очень близкого расстояния (менее 20 см), и лица, снятые со значительными бликами от солнца в кадре.

Оценка производительности.

Предлагаемая модель для классификации в режиме тестирования использует лишь кодирующий и предсказательный блок описанной нейронной сети. Суммарное количество операций умножения и сложения в данных блоках модели составляет порядка 63.2 MFlop. Медианное время выполнения на одном ядре процессора Qualcomm Snapdragon 888 составляет 65 мс. Применение 8-битной квантизации весов и активаций обученной нейронной сети методом [21] позволяет сократить время выполнения до 23 мс за счет применения целочисленной арифметики и оптимизации операций свертки.

3.4. Выводы к третьей главе

Предложен метод определения спуфинг-атак в мобильных системах распознавания по лицу с применением пары камер с малым стереобазисом. Он заключается в использовании сверточной нейронной сети небольшого размера, обученной со специальной функцией потерь. Предлагаемый подход достигает высоких показателей точности детектирования подделок, сравнимых с описанными в современной литературе аналогичными подходами, в том числе на данных открытой базы стереоизображений. От известных аналогов предлагаемый метод отличается малым временем выполнения на современных мобильных процессорах, поэтому может быть применен для детектирования подделок в биометрических системах с малыми вычислительными ресурсами.

Глава 4

Поиск границ радужки на изображении глаза

Выделение (сегментация) региона радужной оболочки глаза на изображении — неотъемлемая составляющая любого алгоритма распознавания по данному биометрическому признаку. Ошибки на данном этапе приводят к повышению уровня ошибок принятия решений системы в целом, снижая в итоге ее надежность и удобство применения [108]. Значительная часть описанных в литературе методов предполагают использование системы в условиях мало изменяющихся условий регистрации изображения. Для подобного сценария применения системы допускается применение основанных на эвристиках классических методов, которые позволяют достичь необходимой точности локализации. Расширение диапазона применения рассматриваемой биометрической технологии распознавания требует разработки более гибких и устойчивых к изменчивым условиям съемки подходов к сегментации радужки.

Вариативные сценарии применения пользователем мобильного устройства в быту расширяют диапазон рассматриваемых качественных характеристик регистрируемых системой изображений, что, в свою очередь, существенно влияет на свойства извлекаемой биометрической информации.

Стоит упомянуть границы изменчивости условий съемки, связанных с освещенностью, уровень которой варьируется в диапазоне от 10^{-4} в темном помещении или в ночное время суток до 10^5 (лк) при воздействии прямых солнечных лучей в середине дня. Сценарий мобильного распознавания по РОГ допускает наличие различных источников света с самым разнообразным расположением относительно лица субъекта съемки. Тени, блики и отражения способны внести заметные искажения в наблюдаемую текстуру биометрического признака. Внешние условия в процессе распознавания способны спровоцировать моргание, сильное прищуривание или резкие движения глаз, что может привести к

Рис. 4.1. Примеры растров области глаза, извлекаемых в мобильной системе распознавания: причина (снизу) и результат (сверху)

деградации качества изображения в целом. Реакция на свет может вызывать изменение размеров зрачка пользователя устройств и повлечь деформацию структуры радужки. Более детальное описание влияющих на распознавание условий внешней среды дано в [99, 111, 131, 154]. Примеры изображений радужки, зарегистрированных при воздействии упомянутых внешних факторов, даны на Рис. 4.1.

4.1. Обзор существующих методов

В последние два десятилетия было предложено значительное количество решений задачи сегментации или образмеривания области пикселей радужки на изображении глаза. Значительная их часть нашла свое применение в коммерческих решениях и позволила достичь высокой точности не мобильных биометрических систем. Некоторые из алгоритмов выделения радужки можно заслуженно называть классическими.

Упомянутые подходы можно разбить на несколько групп, связанных общей методологией решения задачи:

- Использование интегро-дифференциального оператора, впервые описанного в [35]. В предположении, что зрачок и радужку можно аппроксимировать двумя концентрическими окружностями, имеет смысл применение оператора, позволяющего выделять радиально-симметричные структуры. Основанные на данной идее методы позволяют достичь высокой точности

локализации границ радужки, но обладают чрезмерно высокой для большинства приложений вычислительной сложностью [6].

- Методы, основанные на анализе гистограммы изображений с последующей бинаризацией для поиска областей зрачка и радужки и их образмеривания. Подобные решения обладают низкой вычислительной сложностью и хорошо зарекомендовали себя в применении к изображениям высокого качества, полученных в контролируемых условиях [31, 134], но демонстрируют снижение точности в более естественных сценариях [105, 106].

- Моделирование границ РОГ параметрически заданными кривыми и определение оптимальных параметров этих кривых при помощи методологии Хафа (Hough) путем анализа т.н. массивов-аккумуляторов. Среди множества описанных в рамках упомянутой методологии стоит выделить [20, 25, 107, 140, 4]. К достоинствам данной группы методов стоит отнести низкую вычислительную сложность, которая зачастую достигается путем применения ряда эвристик [2]. В то же время, подход Хафа менее устойчив к шуму во входных данных по сравнению, например, с методами, применяющими интегро-дифференциальный оператор и его модификации.

Перечисленные методы и их вариации составляют значительную часть работ, посвященных задаче сегментации изображения РОГ. Многие авторы зачастую предлагают способы повышения точности решения за счет применения специальной предварительной обработки входного растра [103, 151], вспомогательных решающих правил, позволяющих во многих случаях компенсировать недостатки применяемых подходов [24, 35, 89, 96, 156, 6, 10]. В общем случае решение задачи сегментации можно представить в виде алгоритма, составленного из нескольких блоков, представленных на Рис. 4.2. Работа [6] предлагает подробный анализ и классификацию классических подходов, упомянутых выше.

Рис. 4.2. Обобщенная последовательность шагов классических методов сегментации изображения радужки.

Распознавание изображений методами машинного обучения быстро развивается в последние годы вследствие увеличения объема доступных данных и роста производительности вычислительных систем. Среди упомянутой группы подходов выделяется область глубокого обучения, демонстрирующая наилучшее качество возможных решений при наличии обширных и разнообразных обучающих выборок. Глубокие сверточные нейронные сети как частный пример методов упомянутой области начиная с 2012 года применяются для множества задач компьютерного зрения и позволяют превзойти по точности не только ранее разработанные классические подходы, но в некоторых случаях даже человека [49, 71, 79, 132].

Подходы глубокого обучения постепенно нашли употребление и в алгоритмах биометрических систем, в частности, для задачи сегментации изображения РОГ. Так в работе [86] впервые было показана возможность сегментации радужки для изображений, снятых в разнообразных условиях, при помощи достаточно объемной сверточной нейронной сети. Задача была сформулирована в виде попиксельной бинарной классификации: входному растру требовалось поставить в соответствие соразмерную бинарную маску, значение каждого пикселя в которой определяет класс объекта. К примеру, допускается помечать положительным классом область радужки, а отрицательным — фоновую информацию. Авторы приводят сравнение двух подходов к решению такой задачи: «patch based» и «end-to-end». Первый предполагает обучение сети с использованием небольших фрагментов исходного изображения, каждому из которых в зависи-

мости от его расположения поставлена в соответствие метка принадлежности к положительному классу. Второй предлагает напрямую обучать нейронную сеть предсказывать соразмерную бинарную маску с применением логистической функции потерь.

Работа [16] предлагает иной способ построения решения при помощи «patch-based» подхода. При этом в работе [124] того же года были показаны преимущества «end-to-end» решения на фоне упомянутой альтернативы. Авторы экспериментально показали, что применение «patch-based» метода ухудшает качество сегментации. Успешный пример применения актуальной для задач сегментации архитектуры SegNet [18] с регуляризацией при помощи техники дропаут (dropout) [136] дан в [67]. Внедрение более современных методик построения нейронных сетей для попиксельной сегментации [78] позволяет существенно повысить качество получаемых решений для вариативных входных данных. Также предлагается [45] расширение «end-to-end» подхода: единственная нейронная сеть призвана решать задачи локализации, выделения и сравнения признаков одновременно.

Упомянутые подходы позволяют достигать весьма высокого качества решения задачи локализации области радужки, однако их вычислительная сложность затрудняет их применение в системах малой производительности, таких, как мобильные устройства. Использование «end-to-end» подхода связано с необходимостью построения нейронной сети из двух составляющих: кодирующей исходное изображение в некоторое промежуточное представление и декодирующей, выполняющей в некотором смысле обратную операцию. Вторая компонента требует большого количества операций сложения и умножения чисел за счет операций повышения размерности при помощи сверточных слоев, что вносит существенный вклад в общую вычислительную сложность.

Альтернативой «end-to-end» подходам являются методы, позволяющие оценить параметры аппроксимирующих границы радужки кривых. Задача подбора таких параметров не требует предсказания карты результатов, соразмер-

ной входному изображению, а лишь предсказания набора вещественных чисел. Такой подход требует меньшего количества слоев и обучаемых параметров в применяемой модели. Как следствие, вычислительная сложность подхода сокращается как минимум в два раза, поскольку для CNN таковая растет практически линейно в зависимости от количества использованных слоев. Так в [40] предлагается предсказывать координаты центра зрачка при помощи небольшой сверточной нейронной сети. При этом полученное решение может быть затруднительно применять на практике в мобильных устройств за счет неоптимальной архитектуры модели относительно вычислительной сложности. К тому же, задача локализации решена лишь частично, без детектирования внешняя граница радужки.

При аппроксимации региона радужной оболочки окружностями точность решения будет заведомо ниже по сравнению с «end-to-end» подходами, хотя бы потому, что ее истинная форма отличается от идеально округлой [6]. Тем не менее, легковесная нейросетевая модель, способная давать грубую оценку положения пикселей РОГ на изображении может упростить последующее применение более вычислительно сложных подходов, которым в таком случае потребуется лишь уточнить искомое решение задачи.

4.2. Аппроксимация границ радужки методами глубокого обучения

В данной работе предлагается метод аппроксимации границ радужки окружностями с допустимой ошибкой не более 5% от диаметра ее области. В основе метода лежит применение сверточной нейронной сети с небольшим количеством параметров, обучаемой при помощи функции потерь, которая применяется для задач классификации. Решение задачи детектирования радужной оболочки осуществляется как последовательность шагов: сперва применяются две сверточные нейронные сети для определения параметров ограничивающих

окружностей, затем при помощи эвристического подхода оценивается качество полученных предсказаний.

Постановка задачи.

Входными данными для предлагаемого метода является изображение, на котором заведомо содержится глаз. Без ограничения общности будем предполагать, что входное изображение I есть квадратный растр размером $W \times W$ пикселей, внешняя и внутренняя границы радужной оболочки на котором задаются двумя окружностями: (X_i, Y_i, D_i) и (X_p, Y_p, D_p), где X, Y — координаты центра окружности диаметром D в левосторонней системе координат с центром в верхнем левом углу изображения. Индексы i и p соответствуют границам «склера-радужка» и «радужка-зрачок». Значения координат центров и диаметров обеих окружностей при этом удовлетворяют неравенствам:

$$\begin{cases} \min\{X, W-X, Y, W-Y\} < \dfrac{D}{2}\ , \\ D \leqslant W \leqslant 4D\ . \end{cases} \tag{4.1}$$

Результатом работы описанного метода являются параметры аппроксимирующих границы радужной оболочки окружностей:

1. x_i, y_i, d_i — внешняя граница, «радужка-склера»;
2. x_p, y_p, d_p — внутренняя граница, «радужка-зрачок».

Предлагается считать, что параметры окружностей были определены корректно, если абсолютная ошибка детектирования не превышает $\alpha = 5\%$ истинного диаметра радужки на изображении:

$$\begin{aligned} &|x_i - X_i| < \alpha D_i, \qquad |x_p - X_p| < \alpha D_i, \\ &|y_i - Y_i| < \alpha D_i, \qquad |y_p - Y_p| < \alpha D_i, \\ &|d_i - D_i| < \alpha D_i, \qquad |d_p - D_p| < \alpha D_i. \end{aligned} \tag{4.2}$$

Для оценки точности работы метода применяется набор или выборка из N изображений, для каждого из которых возможно определить, удовлетворя-

ют ли предсказанные на нем значения $x_i^k, y_i^k, d_i^k, x_p^k, y_p^k, d_p^k$ неравенствам (4.2). Для обозначения истинных и предсказанных параметров окружностей для k-го по порядку изображения из выборки используются верхние индексы. Совокупность упомянутых параметров применяется для расчета значений следующих мер качества.

1. Средняя абсолютная ошибка оценки диаметра:

$$MAE(D, d) = \frac{1}{N} \sum_{k=1}^{N} |D^k - d^k|. \tag{4.3}$$

2. Распределение относительных расстояний между предсказанным и истинным центром окружности:

$$H(\alpha) = \frac{1}{N} \left| \left\{ k : \frac{\rho^k}{D^k} < \alpha, k \in \overline{1, N} \right\} \right|, \tag{4.4}$$

где $\rho^k = \sqrt{(x^k - X^k)^2 + (y^k - Y^k)^2}$, $|\{\cdot\}|$ — мощность множества.

3. Распределение относительных ошибок детектирования параметров окружности:

$$Q(\alpha) = \frac{1}{N} \left| \left\{ k : \frac{l^k}{D^k} < \alpha, k \in \overline{1, N} \right\} \right|, \tag{4.5}$$

где $l^k = |x^k - X^k| + |y^k - Y^k| + |d^k - D^k|$.

Переход к задаче классификации.

В соответствии с постановкой задачи параметры границ радужки могут быть рассчитаны с относительной ошибкой до 5%. Поэтому отсутствует необходимость обрабатывать входные изображения в полном разрешении, и решение задачи может быть получено без потерь точности при помощи масштабированного к меньшему разрешению изображения при условии корректности условий (4.1). Размер такого изображения может быть выбран как минимальный, позволяющий притом допустить относительную ошибку не более 5%. В предположении, что диаметр радужки принимает свое наименьшее возможное значение

и с учетом неравенств (4.1), после масштабирования изображения глаза к размеру $W \times W$ диаметр будет равен $D = W/4$. Тогда максимальная допустимая ошибка аппроксимации будет составлять

$$\Delta_{\max} = \alpha D_i = 0.05\frac{W}{4} = \frac{W}{80}. \tag{4.6}$$

Таким образом, чтобы максимально возможная ошибка составляла не менее одного пикселя, требуется выполнение неравенства $W \geq 80$. Поэтому значение W выбрано равным 80 пикселям, что соответствует наименьшему допустимому значению. Поскольку неравенство (4.1) всегда выполняется, параметры (x, y, d) могут оцениваться как целочисленные. Это означает, что каждый параметр может принимать конечное множество целочисленных значений от 0 до $W-1$. Каждое из возможных изображения глаз может быть отнесено к одному из $W = 80$ классов: $\left\{C_k^X\right\}_{k=0}^{W-1}$, где C_k^X — класс изображений, для которых верно $X \in [k; k+1)$. Аналогично, можно отнести изображения к одному из классов $\left\{C_k^Y\right\}_{k=0}^{W-1}$ значений Y и классов $\left\{C_k^D\right\}_{k=0}^{W-1}$ значений D.

Таким образом, задача определения параметров x, y и d может быть сведена к решению трех задач классификации:

1. на классы $\left\{C_k^X\right\}_{k=0}^{W-1}$;
2. на классы $\left\{C_k^Y\right\}_{k=0}^{W-1}$;
3. на классы $\left\{C_k^D\right\}_{k=0}^{W-1}$.

Если для входного изображения I верны следующие утверждения: $I \in C_i^X$, $I \in C_j^Y$, $I \in C_k^D$, тогда $x = i$, $y = j$, $d = k$ будут корректными параметрами аппроксимации радужки в смысле условий 4.2.

Решение задачи классификации.

Для решения задачи классификации в данной работе предлагается использовать сверточную нейронную сеть с небольшим количеством параметров в соответствие с идеями, предложенными в [63]. Особое строение сверточных слоев, описанное в этой статье, упрощает применение подобных нейронных сетей

в мобильных устройствах. Основным компонентом, из которых состоит предлагаемая сверточная сеть, является блок операций, описанный в табл. 4.1.

Слой	Размер ядра	Шаг	Размер входного тензора
Depth-wise свертка	3×3	s	$N \times K \times K$
Batch normalization	-	-	$N \times \frac{K-3}{s} + 1 \times \frac{K-3}{s} + 1$
Активация ReLu	-	-	$N \times \frac{K-3}{s} + 1 \times \frac{K-3}{s} + 1$
Свертка	1×1	1	$N \times \frac{K-3}{s} + 1 \times \frac{K-3}{s} + 1$
Batch normalization	-	-	$M \times \frac{K-3}{s} + 1 \times \frac{K-3}{s} + 1$
Активация ReLu	-	-	$M \times \frac{K-3}{s} + 1 \times \frac{K-3}{s} + 1$

Таблица 4.1. Строение блока $MobileConvBlock(M, s)$

В дальнейшем предлагается называть такую последовательность операций MobileConvBlock или MCB, где M и s — параметры MCB, которые определяют соответственно количество каналов тензора после применения блока и шаг применения фильтра в первой операции свертки «depth-wise».

Строение приведенной в данной работе сверточной нейронной сети дано в табл. 4.2.

Гиперпараметры ее архитектуры были выбраны в серии экспериментов как те, которые позволили достичь наилучшей точности решения задачи детектирования на валидационной выборке.

Описанное в табл. 4.2 строение нейронной сети применяется как для детектирования внешней границы радужной оболочки, так и для внутренней. Будем в дальнейшем называть первую модель IrisModel, а вторую — PupilModel.

Применение метода к детектированию.

В данной работе предлагается определять параметры границ радужной оболочки в два этапа, Рис. 4.3.

Этап 1. Определяются параметры внешней границы радужной оболочки, что позволяет сократить область поиска окружности зрачка на втором этапе.

Слой	Размер входного тензора
Свертка 3×3	$1 \times 80 \times 80$
MCB(16, 2)	$16 \times 78 \times 78$
MCB(32, 1)	$16 \times 38 \times 38$
MCB(32, 2)	$32 \times 36 \times 36$
MCB(64, 1)	$32 \times 18 \times 18$
MCB(64, 2)	$64 \times 16 \times 16$
MCB(64, 1)	$64 \times 7 \times 7$
MCB(64, 1)	$64 \times 5 \times 5$
Global Average Pooling	$64 \times 3 \times 3$
Полносвязные слои для x, y, d	$64 \times 1 \times 1$
SoftMax для x, y, d	80

Таблица 4.2. Строение нейронной сети для аппроксимации границ радужки

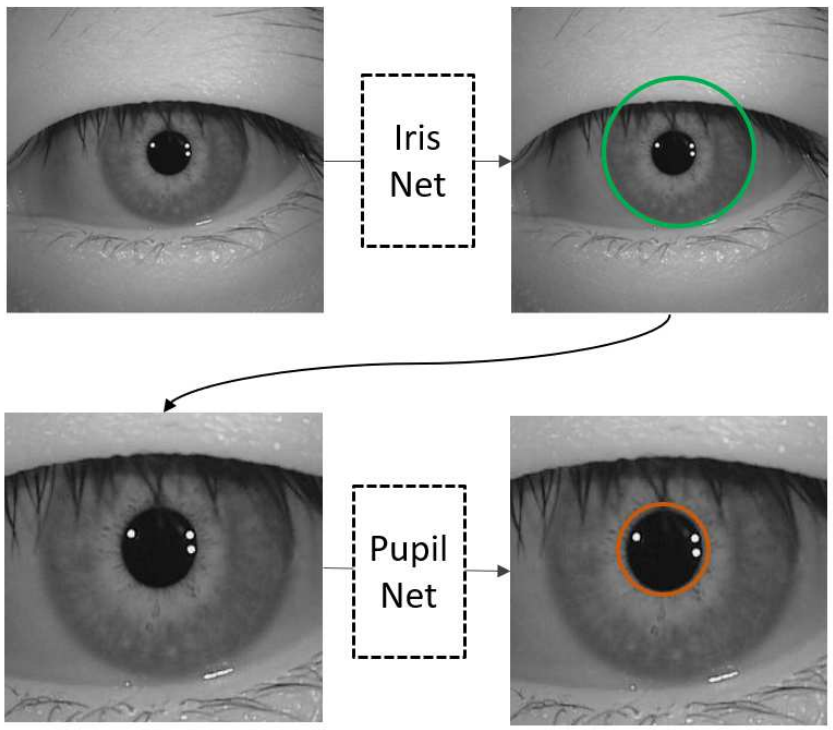


Рис. 4.3. Схема применения комбинации сетей для поиска границ радужки

Входными данными для нейронной сети первого этапа является изображение глаза, результатом — параметры окружности «радужка-склера» (x_i, y_i, d_i), полученные как решения трех задач классификации в соответствии с разд. 4.2

Этап 2. Осуществляется поиск параметров внутренней окружности (x_p, y_p, d_p). Исходя из особенностей строения человеческого глаза и статистических исследований [2], можно ввести ограничения, связывающие между собой упомянутые выше параметры:

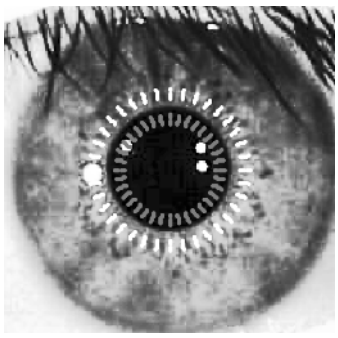
$$D_p \in \left[\frac{1}{7}D_i; \frac{3}{4}D_i\right] . \tag{4.7}$$

Исходя из (4.7), имеет смысл осуществлять поиск зрачка внутри области изображения, заданного как квадратная область, содержащаяся между точками p_1 и p_2:

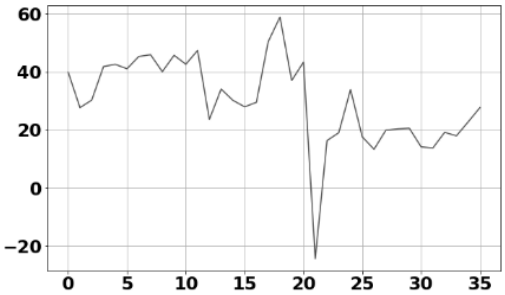
$$p_{1,2} = \left(x_i \mp \frac{3}{8}d_i, y_i \mp \frac{3}{8}d_i\right) .$$

Полученная область изображения масштабируется к размеру $W \times W$ с учетом 4.6 и подается в сверточную нейронную сеть, которая осуществляет поиск (x_p, y_p, d_p) как решение трех задач классификации в соответствии с разд.4.2

Оценка качества детектирования.



(а)



(б)

Рис. 4.4. Оценка качества детектирования зрачка: а) – области извлечения признаков; б) – признаковое описание контрастности границы

Предлагаемый метод детектирования границ радужки, как и ряд других, по построению не позволяет оценить точность получаемой аппроксимации. На

практике наличие подобной оценки позволило бы снизить долю ложных срабатываний метода и, следовательно, не допустить некорректно обработанные входные данные на последующие этапы работы системы распознавания.

Здесь предлагается простой эвристический подход к оценке качества аппроксимации границ. Пограничные зоны радужной оболочки характеризует присутствие контраста для областей «радужка-зрачок» и «радужка-склера», степень выраженности которых может отличаться из-за особенностей строения глаза или изменчивости условий регистрации изображения. Анализ контрастности пограничных областей позволит получить оценку качества аппроксимации. Граница «радужка-склера» редко видима полностью вследствие перекрытия веками и ресницами, в отличие от области «радужка-зрачок». При наличии данных о расположении век и ресниц было бы возможно произвести анализ контрастности внешней границы, однако это выходит за рамки предлагаемого метода. Поэтому оценивается только аппроксимация внутренней границы, как изображено на рис. 4.4, *а*.

Построим признаковое описание для контрастности внутренней границы радужки $\theta_p = (x_p, y_p, d_p)$. Выделим на аппроксимирующей окружности 36 точек с шагом в $\phi = 10°$:

$$V = \left\{ \left(x_p + \frac{d_p}{2}\cos(k\phi), y_p + \frac{d_p}{2}\sin(k\phi)), k = \overline{0, 35} \right) \right\} .$$

Для каждого пикселя $(x, y) \in V$ с помощью билинейной интерполяции можно оценить средние значения яркости изображения снаружи и внутри данной окружности для соседних с (x, y) пикселей. Оценкой контрастности в окрестности (x, y) будет разность этих средних значений. Совокупность этих оценок позволяет построить признаковое описание $f = F(V; I, \theta_p)$ для аппроксимирующей окружности. Работа данного алгоритма проиллюстрирована на рис. 4.4.

Пусть $q \in [0; 1]$ — величина, принимающая значение 0 для идеальной аппроксимации окружностями и 1 — для ошибки сегментации. Построим алго-

ритм, вычисляющий:

$$q = a_p(f; \theta_p) = a_p(I; \theta_p). \tag{4.8}$$

Рассмотрим задачу построения такого алгоритма как задачу бинарной классификации и воспользуемся методом логистической регрессии. Обучающую выборку можно сформировать, используя истинные параметры аппроксимации границы «радужка-зрачок» для объектов нулевого класса. Признаковое описание для объектов положительного класса можно получить, используя заведомо некорректные значения координаты центра и диаметра зрачка. Для повышения робастности подобной оценки качества также предлагается применять эквализацию гистограммы входного изображения I перед построением признакового описания f.

4.3. Экспериментальные результаты.

Для обучения и тестирования моделей глубокого обучения, осуществляющих поиск параметров границ радужки, применялись данные открытых баз изображений. Использованные базы можно разделить на две группы. Первую составляют базы, содержащие изображения высокого разрешения CASIA 2[30], ES, ICE[105] и MMU[62], а также UBIRIS v.1[106], имеющая примеры изображений низкого качества. Вторую группу составляли изображения низкого разрешения, характерные для растров, которые получаются с небольших камер, встраиваемых в мобильные устройства: CASIA Mobile [29] и собранная вручную база изображений Raspberry DB. База изображений CASIA Mobile составлена из трех частей: M1, M2 и M3. В данной работе, однако, были использованы только последние две части, поскольку первая содержит изображения недопустимо низкого в соответствии с (4.6) разрешения. База изображений Raspberry DB была получена вручную при помощи одноименного микрокомпьютера, оборудованного совместимой инфракрасной камерой с активной инфракрасной подсветкой. Более подробная информация об упомянутых базах изображений при-

ведена в табл. 4.3, а примеры изображений из каждой использованной базы — на Рис. 4.5.

Название	Код	Кол-во изображений, тыс.	Разрешение
CASIA2-Iris-Lamp	CAS	5	640×480
ICE Database	ICE	3	640×480
CASIA-Iris-M1-S2	CM2	5	1968×1024
CASIA-Iris-M1-S3	CM3	1	1920×1920
ES dataset	ES	25	640×480
MMU GASPFA	MMU	2	320×286
Raspberry DB	RAS	1	640×480

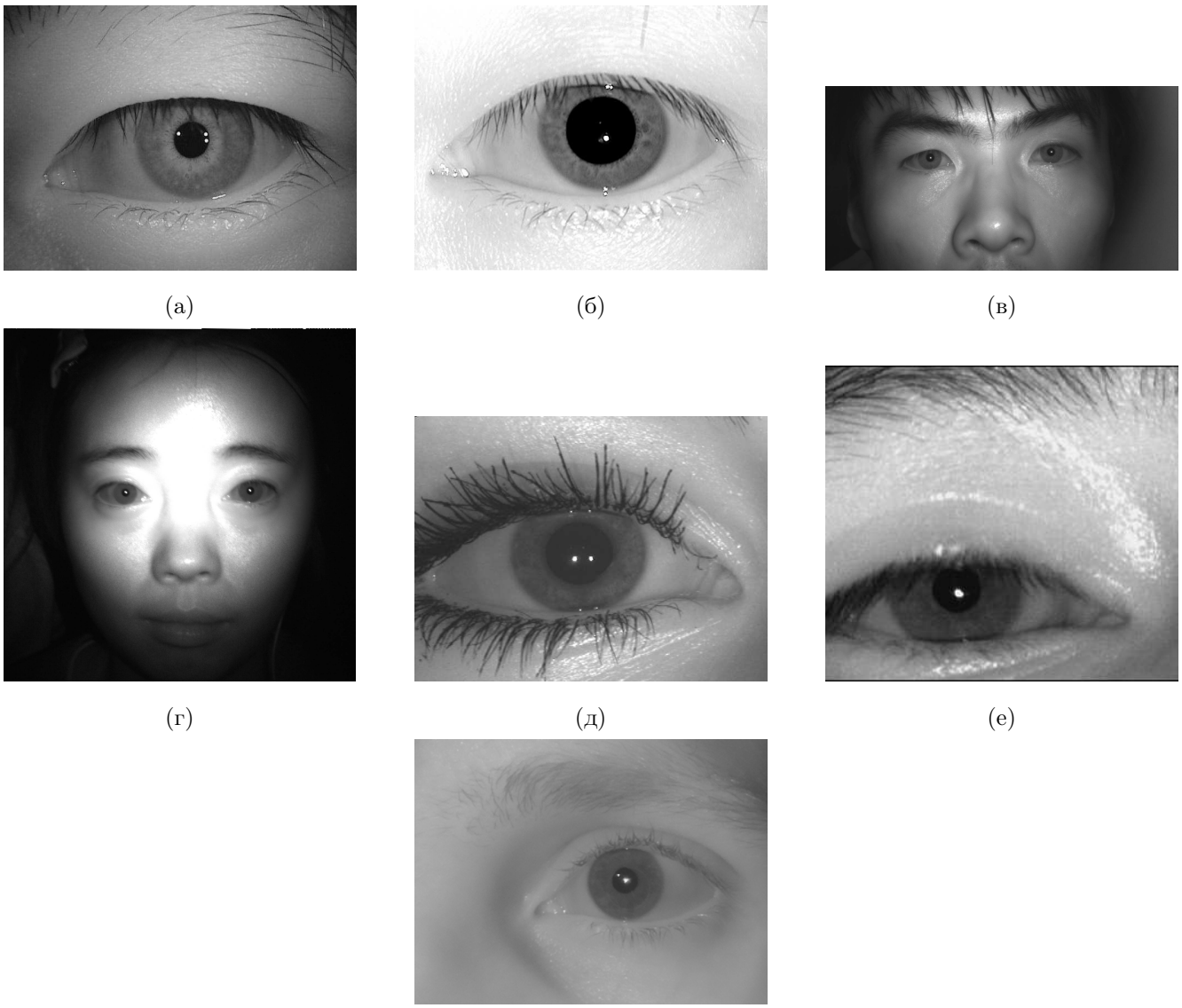
Таблица 4.3. Использованные базы изображений радужки

Целью поставленного вычислительного эксперимента была проверка качества работы предлагаемого метода на изображениях разного качества и из разных доменов. Параметры истинных аппроксимирующих границы радужки окружностей (X_i, Y_i, D_i) и (X_p, Y_p, D_p) для баз были получены с помощью разметки человека-эксперта.

Подготовка данных.

Все эксперименты с обучением нейронных сетей проводились на обучающей выборке, составленной из изображений баз CASIA2, CASIA Mobile S2 и ES. В качестве валидационной выборки были применены базы изображений MMU, CASIA Mobile S3, а также половина изображений из базы ICE. В качестве отложенной тестовой выборки использовалась другая половина изображений базы ICE и база изображений низкого качества UBIRIS v.1.

Во время обучения для повышения обобщаемости модели применялись аугментации случайной яркостной коррекции, наложения случайного пуассоновского шума, случайного аффинного поворота, случайного отображения по горизонтали и извлечения случайной области внутри изображения. Масштаби-



(а) (б) (в)

(г) (д) (е)

(ж)

Рис. 4.5. Примеры изображений используемых баз: (а) CASIA; (б) ICE; (в) CASIA Mobile v2; (г) CASIA MOBILE v3; (д) ES; (е) MMU; (ж) RAS

рование изображений к целевому разрешению (4.6) производилось после описанных преобразований. Обучение моделей осуществлялось методом стохастического градиентного спуска с адаптивным моментом [76]. Первоначальное значение темпа обучения составляло 0.001 и уменьшалось экспоненциально в 0.9 раз каждые 10 эпох.

Процедура обучения и подбор гиперпараметров.

Архитектура предлагаемых нейронных сетей такова, что их выходной слой может содержать произвольное количество нейронов, а значит, осуществлять

предсказание произвольного количества классов для координат центра и диаметра. В случае, если количество предсказываемых классов будет превосходить размеры изображения, можно утверждать, что модель будет обучена находить параметры границ радужки с субпиксельной точностью. Однако в данной работе подобный эксперимент не был произведен, поскольку постановка задачи допускает некоторую степень неточности решения задачи детектирования при условии, что применяемая сверточная нейронная сеть будет сравнительно небольшой. Для поиска оптимального размера выходных слоев произведен вычислительный эксперимент. Осуществлялось сравнение моделей, предсказывающей количество классов, соразмерное разрешению входного изображения «original» и уменьшенное вдвое «reduced». Постановка этого эксперимента имела смысл ввиду большого количества возможных искажений видимого качества изображений и неизбежных ошибок экспертной разметки. Каждая модель была обучена на протяжении 50 эпох, после чего проводилось сравнение по мере качества (4.4) со значением $\alpha = 0.05$. Большей точности на валидационной выборке достигли модели с размерами выходных слоев «original». В дальнейших экспериментах были использованы соответствующие их версии. Отличия по качеству решения задачи для IrisModel составили 85.3% против — 60.9%, для PupilModel — 99.7% против — 99.1%. Заметим, что разные версии моделей PupilModel практически не отличаются по точности решения задачи аппроксимации границы зрачка, что можно объяснить тем, что поиск производится уже внутри выделенной области радужки, тогда как IrisModel вынуждена искать положение центра глаза на изображении периокулярной области, которая, как правило, более вариабельна.

После выбора размеров выходных слоев обе модели обучались на протяжении 128 эпох на независимых выборках. Полученные значения мер качества (4.3) — (4.5) приведены в табл. 4.4.

Применение моделей в комбинации.

Детектирование границ радужной оболочки на изображении глаза осу-

Модель	Выборка	$H(0.05)$	$H(0.1)$	$H(0.2)$	MAE
		%			
IrisModel	Валидация	88.03	99.76	99.9	0.54
	Тест	90.1	99.5	99.9	0.43
PupilModel	Валидация	99.76	99.4	99.9	0.7
	Тест	95.48	99.6	99.9	0.9

Таблица 4.4. Качество моделей на валидационной и тестовой выборках

ществляется в два этапа, как описано в разд. 4.2. Качество работы комбинации двух моделей оценивается при помощи меры качества (4.5). В данном эксперименте применялись следующие базы изображений: CASIA Mobile S3, Raspberri DB, UBIRIS v.1, MMU и ICE как составляющие валидационной и тестовой выборок.

Определение точности комбинации моделей IrisNet и PupilNet было проведено как совместно с применением алгоритма оценки качества аппроксимации границы $a_p(I;\theta_p)$ (4.8), так и без него (табл. 4.5).

База изображений	$Q_p(0.05)$	$Q_i(0.05)$	$Q_p^a(0.05)$	$Q_i^a(0.05)$	FRR
	%				
Raspberry DB	95.5	98.34	97.4	98.34	0.2
CASIA Mobile S3	97.9	98.5	98.1	98.6	0.9
UBIRIS v.1	85.8	95.3	98.7	98.1	9.1
ICE	97.6	95.8	98.6	95.8	0.7
MMU	88.3	98.5	90.2	98.8	0.2

Таблица 4.5. Результаты применения комбинации моделей

Для изображений, где найденная окружность «радужка-зрачок» была помечена $a_p(I;\theta_p)$ как ошибочная, изображение не учитывалось при расчете меры качества 4.8. Скорректированные таким образом значения данной меры каче-

ства приведены в столбцах Q_p^a и Q_i^a. Доля объектов с неверными метками ошибочной аппроксимации внутренней границы радужки рассмотрена в столбце FRR.

Для оценки точности метода на мобильной базе изображений CASIA Mobile S3 требовалось бы применение постороннего алгоритма выделения области глаза на входном изображении, поскольку это изображение содержит обширную область лица участника. В данной работе предложен иной способ оценки точности, использующий искусственное создание выборки изображений лица с применением экспертной разметки. Для каждого изображения лица из данной базы извлекается пять случайных прямоугольных областей глаза следующим образом:

1. размер области глаза выбирается случайно из равномерного распределения $U(2D_i, 4D_i)$;

2. координаты центра области (x, y) выбираются случайно из равномерного распределения $U(-D_i, D_i)$.

Таким образом удается оценить точность детектирования радужки в условиях ошибочной работы детектора области глаза.

Приведенные результаты вычислительного эксперимента демонстрируют снижение качества работы предлагаемого метода на базах изображений низкого качества, таких, как UBIRIS. Данная база содержит изображения с низкой контрастностью на границе «радужка-зрачок», что также подтверждает высокий уровень ложных срабатываний предложенного алгоритма $a_p(I; \theta_p)$ в столбце FRR, что, несомненно, усложняет задачу аппроксимации границ радужки. Обучающая выборка предлагаемого метода имеет сравнительно малое количество примеров изображений глаз, видимое качество которых подобно изображениям UBIRIS. Наконец, база изображений UBIRIS отличается не только низким контрастом границы «радужка-зрачок», но и распределением размеров зрачка относительно радужки: большая часть изображений содержит в себе зрачки

малого размера, сопоставимого по величине с бликом от осветителя. Низкая точность работы детектора зрачка на базе MMU вызвана низким разрешением исходных изображений и аналогично UBIRIS — большим количеством изображений с малым видимым размером зрачка.

Применение моделей с последующим уточнением.

Помимо описанного в предыдущем разделе способа применения комбинации моделей, было проанализировано использование последующего уточнения получаемой аппроксимации границ с помощью метода [35] со значительно суженной областью поиска. По построению предлагаемый метод имеет ошибку детектирования не выше 5% от истинного диаметра радужки. Таким образом, область поиска параметров границы $\hat{\theta} = (\hat{x}, \hat{y}, \hat{d})$ при уточнении методом [35] можно определить следующим образом в зависимости от первичной аппроксимации $\theta = (x, y, d)$:

$$\hat{x} \in [x - 0.05D_i; x + 0.05D_i] \ ,$$

$$\hat{y} \in [y - 0.05D_i; y + 0.05D_i] \ ,$$

$$\hat{d} \in [d - 0.05D_i; d + 0.05D_i] \ .$$

Вычислительная стоимость поиска внутри такой области пространства параметров будет сравнительно невысокой. Результаты применения предлагаемого метода с уточнениями даны в табл. 4.6.

Уточнение границ радужной оболочки при помощи классического метода [35] позволяет повысить качество работы комбинации методов и снизить долю ложных срабатываний алгоритма $a_p(I; \theta_p)$ на базе UBIRIS, вызванных низкой контрастностью на границе «радужка-зрачок» и ошибками детектирования зрачка. Однако прирост точности детектирования для границы зрачка нельзя назвать существенным. Аналогично для базы изображений MMU применение метода [35] для уточнения границы зрачка не привело к значительным

База изображений	$Q_p(0.05)$	$Q_i(0.05)$	$Q_p^a(0.05)$	$Q_i^a(0.05)$	FRR
	%				
Raspberri DB	96.1	98.3	97.9	98.9	0.1
CASIA Mobile S3	98.3	98.7	98.7	99.0	0.3
UBIRIS v.1	84.5	97.5	92.7	98.9	5.9
ICE	97.6	95.8	98.6	95.8	0
MMU	89.2	98.8	92.3	99.2	0.1

Таблица 4.6. Результаты применения комбинации моделей с последующим уточнением

улучшениям, поскольку данный метод работает неустойчиво на изображениях с небольшими по величине зрачками.

Сравнение с существующими методами.

Предлагаемый метод детектирования границ радужной оболочки на изображении глаза также сравнивался с иными методами, описанными в [35, 89, 92, 140, 2].

Мера качества	Метод детектирования						
	Уильдс	Дугман	Мазек	Ма	Ганькин	CNN	CNN и уточнение
ε_c	3.15	2.61	4.98	3.92	0.97	1.4	1.3
ε_r	6.12	4.39	5.15	5.39	1.13	1.9	1.7
t_{c+r} (ms)	379.61	523.14	97.52	363.64	106.60	8	10

Таблица 4.7. Результаты сравнения с существующими методами

Сравнение производилось с применением базы изображений MMU при помощи следующих мер качества:

1. относительная ошибка детектирования центров:

$$\epsilon_c = \frac{1}{N}\sum_{k=1}^{N}\sqrt{(y^k - Y^k)^2 + (x^k - X^k)^2};$$

2. относительная ошибка детектирования радиусов:

$$\epsilon_r = \frac{1}{N}\sum_{k=1}^{N}|r^k - R^k|.$$

Помимо точности детектирования в сравнение было включено медианное время выполнения (t_{c+r}) на одном ядре процессора Qualcomm Snapdragon 845. Результаты приведены в табл. 4.7.

4.4. Выводы к четвертой главе

Предложен метод детектирования параметров, аппроксимирующих границы радужки окружностей для мобильных биометрических систем. Метод состоит из двух сверточных нейронных сетей небольшого размера. Предлагаемый подход достигает показателей точности детектирования, сравнимых с описанными в современной литературе, однако обладает существенно меньшим временем выполнения на современных мобильных процессорах. Метод может быть применен для получения грубой оценки границ, которая может быть использована в дальнейшем для ее улучшения иными методами.

Глава 5

Определение живости радужки

Радужная оболочка глаза как биометрическая модальность является перспективной технологией, применяемой в современных мобильных устройствах 1.1. К таковым относится несколько моделей известных производителей [37, 95, 116]. Как правило, методы биометрического распознавания применяются при разблокировке устройства для повышения общего уровня безопасности. Допускается и применение упрощении доступа к личной информации: например, к платежным системам.

Система распознавания по РОГ как правило использует изображения, регистрируемые в ближнем инфракрасном спектре излучения с активной подсветкой одним или несколькими диодами. Взлом системы путем предъявления искусственно созданного биометрического образца (спуфинг, spoofing) затрудняется ввиду того, что требуется воссоздать видимые в ИК-диапазоне характеристики глаза жертвы. Более того, наблюдаемые характерные особенности подделки будут более заметны в условиях активной подсветки. Тем не менее, были выявлены факты успешного взлома или спуфинг-атаки нескольких оснащенных сканером радужки мобильных устройств. Авторами выступили несколько независимых групп профессионалов, чьей специализацией является взлом и демонстрация уязвимостей систем безопасности, в т.ч. и биометрических [22, 28].

Настоящая работа экспериментально подтверждает возможность успешного обмана сканера радужки путем спуфинга, однако, лишь при выполнения ряда важных условий. Во-первых, растр радужной оболочки глаза должен быть создан при помощи инфракрасной камеры высокого разрешения, без размытия и ошибок экспонирования кадра. Во-вторых, глаза должны быть в достаточной мере открыты и направлены взглядом в камеру. Стоит отметить, что соблюдение этих условий требует съемки объекта спуфинг-атаки с очень короткого

расстояния или с применением телеобъектива высокого разрешения и трудно реализуется на практике без кооперации или тесного контакта с жертвой. Наконец, поддельное изображение РОГ должно быть напечатано на бумаге с разрешением не менее 600 dpi и иметь диаметр не менее 250-300 точек. Тем не менее, можно сделать вывод, что задача детектирования подделок является актуальной и для рассматриваемых мобильных биометрических систем.

5.1. Обзор способов выявления подделок радужки

В современной литературе описаны несколько эффективных способов подлога изображения РОГ при распознавании [33, 44, 60]. Наиболее распространенным видом спуфинга является демонстрация фотографии изображения радужки жертвы, напечатанная на принтере высокого разрешения (свыше 600 dpi). Аналогичным способом является использование экранов других устройств для показа изображения или видеозаписи глаза пользователя в случае, если система распознавания оперирует в видимом спектре. Допускаются также более экзотичные и порой трудно обнаруживаемые виды атак: демонстрация искусственно созданного из стекла или пластика глазного протеза или нанесение текстуры РОГ жертвы на прозрачную контактную линзу, а также иные варианты, имитирующие подлинный человеческий глаз для сенсоров системы распознавания.

В литературе описано несколько групп методов определения подделок РОГ:

- Требующие и не требующие внедрение в биометрическую систему дополнительной аппаратуры, специально предназначенной для детектирования физиологических свойств подлинной радужки. К таковым можно отнести глазной гиппус или изменение размера зрачка после неожиданной смены освещения, получаемое в результате подсветки дополнительным диодом [33, 44] в процессе регистрации изображения РОГ сенсором.

- Кооперативные и некооперативные по отношению к пользователю систе-

мы. Взаимодействие с таковым осуществляется, к примеру, путем вывода подсказок с просьбой закрыть/открыть веки, совершить определенное движение глазами и др.

Применение дополнительных аппаратных средств как и усложнение процедуры распознавания путем взаимодействия с ее участником как правило затруднительно для мобильных устройствах потому, что такой класс методов детектирования подделок способен заметно уменьшить удобство использования и при этом увеличить стоимость технологии в целом [98]. По этой причине фокус внимания исследователей в основном направлен на полностью автоматические методы, которые наиболее выгодны при коммерческом применении. В то же время к таковым предъявляются повышенные требования к их устойчивости к вариабельности входных данных, универсальности и удобства поддержки.

Возможность детектирования подделок в системе распознавания по радужной оболочке была исследована Джоном Дугманом. В его работе [36] рассматриваются атаки на биометрическую систему при помощи распечаток фотографий глаз, сделанных в инфракрасном диапазоне. Для печати были использованы лазерные принтеры, созданные по устаревшей на данный момент технологии, вследствие чего имеющие характерные артефакты в виде видимой невооруженным глазом на распечатке сетки точек. Дугман предложил искать побочные максимумы в частотной области результата двухмерного дискретного преобразования Фурье входного изображения. Данный подход показал высокую точность детектирования распечаток глаз, однако оказался практически бесполезным против подделок, изготовленных на современных лазерных принтерах, обладающих высокой плотностью печати. Несмотря на это, идея применения частотного анализа для детектирования изображений неживых глаз получила развитие в работах [32, 58, 112].

Другим подходом к решению задачи определения живости глаза является использование информации об особенностях текстуры объектов, представлен-

ных на изображении глаза. Физические свойства подделки человеческого глаза отличаются от таковых у биологических тканей, что вносит изменения в распределение интенсивности отраженного света и порождает видимые артефакты на изображении глаза-подделки.

Ряд работ предлагает использовать локальные дескрипторы для описания и анализа текстуры РОГ с целью детектирования спуфинга. К примеру, модификации текстурного дескриптора LBP [100] (local binary patterns, локальные бинарные шаблоны) могут эффективно применяться [56, 60] против нескольких известных типов атак: искусственные глаза из пластика и стекла, контактные линзы с рисунком радужки, бумажные распечатки и т.д. Анализ [110] бинаризованных локальных статистик изображения (BSIF, binarized statistical image features) также продемонстрировал практическую применимость к решению задачи определения живости на примере нескольких специализированных баз растров. Комбинированное решение с применением LBP, BSIF и частотного анализа (LPQ, local phase quantization) для извлечения разнородных текстурных признаков предложено в объемном исследовании [54], посвященном детектированию подделок для нескольких биометрических модальностей: РОГ, отпечаток пальца и лицо.

Меры качества изображения применяться [44] для обнаружения подделок в биометрической системе. Обоснованием такого подхода служит гипотеза о том, что запечатленный при помощи камеры растр спуфинг-атаки будет качественно отличаться от снимка живого глаза человека в нормальных (фиксированных) условий распознавания. Как правило, корректная работа сенсора регистрации изображения на поддельных образцах не гарантируется производителем. Несколько специфичных для изображений РОГ мер качества [44] было изучено в контексте детектирования атак вида бумажных распечаток.

Перспективным подходом к обнаружению спуфинг-атак в биометрических системах сегодня считается применение методов глубокого обучения. Они демонстрируют высокую производительность при сравнению с иными уже суще-

ствующими подходами. Первой работой на тему нейросетевых решений задачи анти-спуфинга для модальностей радужки, лица и отпечатков пальцев считается [94]. Объемные сравнительные исследования различных подходов к построению подобных классификаторов «живости» биометрических образцов проводятся в рамках LivDet соревнований, демонстрируя при этом превосходство методов глубокого обучения [143—145].

Часть вышеупомянутых методик были исследованы в контексте применения в контексте мобильных приложениях. При таком сценарии на решение задачи накладываются достаточно жесткие ограничения по времени обработки входных данных и потреблению вычислительных ресурсов вычислительного устройства. Среди описанных в литературе были выбраны несколько подходов [110, 122, 123], соответствующих упомянутым ограничениям и демонстрирующих перспективные результаты. Методики определения живости РОГ при помощи набора мер качества изображений описаны в [122, 123], а в [110] — применение BSIF [70] разных масштабов.

5.2. Детектирование подделок радужки

Предложен основанный на глубокой сверточной нейронной сети (CNN) метод определения живости глаза против спуфинг-атак разных типов. Применение метода требует информации о положении и размерах зрачка и радужки, которые предлагается аппроксимировать параметрическими окружностями. Входными данными являются пара изображений: квадратная область глаза $\mathbf{I_{ER}}$, центр которой совпадает с центром окружности зрачка, и развернутая в прямоугольник кольцевая область радужной оболочки $\mathbf{I_{NI}}$, которую принято называть нормализованной радужкой.

Процедура нормализации была впервые предложена в работе [34] и подразумевает трансформацию растра с изображением радужки $I(x, y)$ за счет смены системы координат (Рис. 5.1) с Декартовой (x, y) на полярную (r, θ) (5.1):

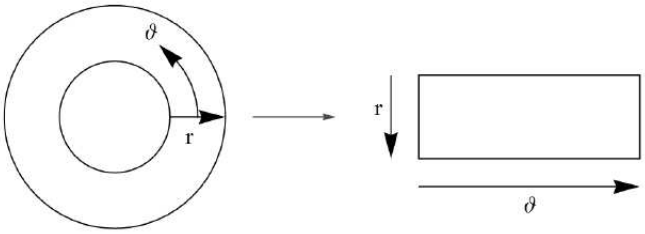


Рис. 5.1. Схема нормализации радужки

$$I(x(r,\theta), y(r,\theta)) \rightarrow I(r,\theta) \tag{5.1}$$

где (r, q) - соответствующие нормализованные координаты в полярной. $x(r,\theta)$ и $y(r,\theta)$ заданы в виде линейных комбинаций групп точек границ зрачка $(x_p(\theta), y_p(\theta))$ и радужки $(x_i(\theta), y_i(\theta))$:

$$\begin{aligned} x(r,\theta) &= (1-r)\cdot x_p(\theta) + r\cdot x_i(\theta) \\ y(r,\theta) &= (1-r)\cdot y_p(\theta) + r\cdot y_i(\theta) \end{aligned} \tag{5.2}$$

Алгоритм обнаружения потенциального подделывания применяется сразу после этапа сегментации радужки на входном изображении глаза (Глава 4). Перед процедурой вычисления результата предсказания нейронной сети, входные данные масштабируются к заранее определенным размерам, Рис. 5.2. Регион глаза $I_{ER}(M_{ER}, N_{ER})$ извлекается с параметрами $M_{ER} = N_{ER} = 3R_i$, где R_i — радиус окружности внешней границы радужки. Центр изображения $I_E R$ совмещен с центром окружности зрачка. Размер входных растров был определен как оптимальный для выбранной архитектуры с учетом требований точности получаемого решения и скорости обработки.

Архитектура решения

В нейросетевой модели применяются идеи работы MobileNets [63] как обладающие необходимыми для мобильных приложений характеристиками быстродействия при сохранении описательной способности. Архитектуры такого виде имеют особую конструкция сверточных блоков, в которых достигается сокращение вычислительной сложности по сравнению с классической при сохранении

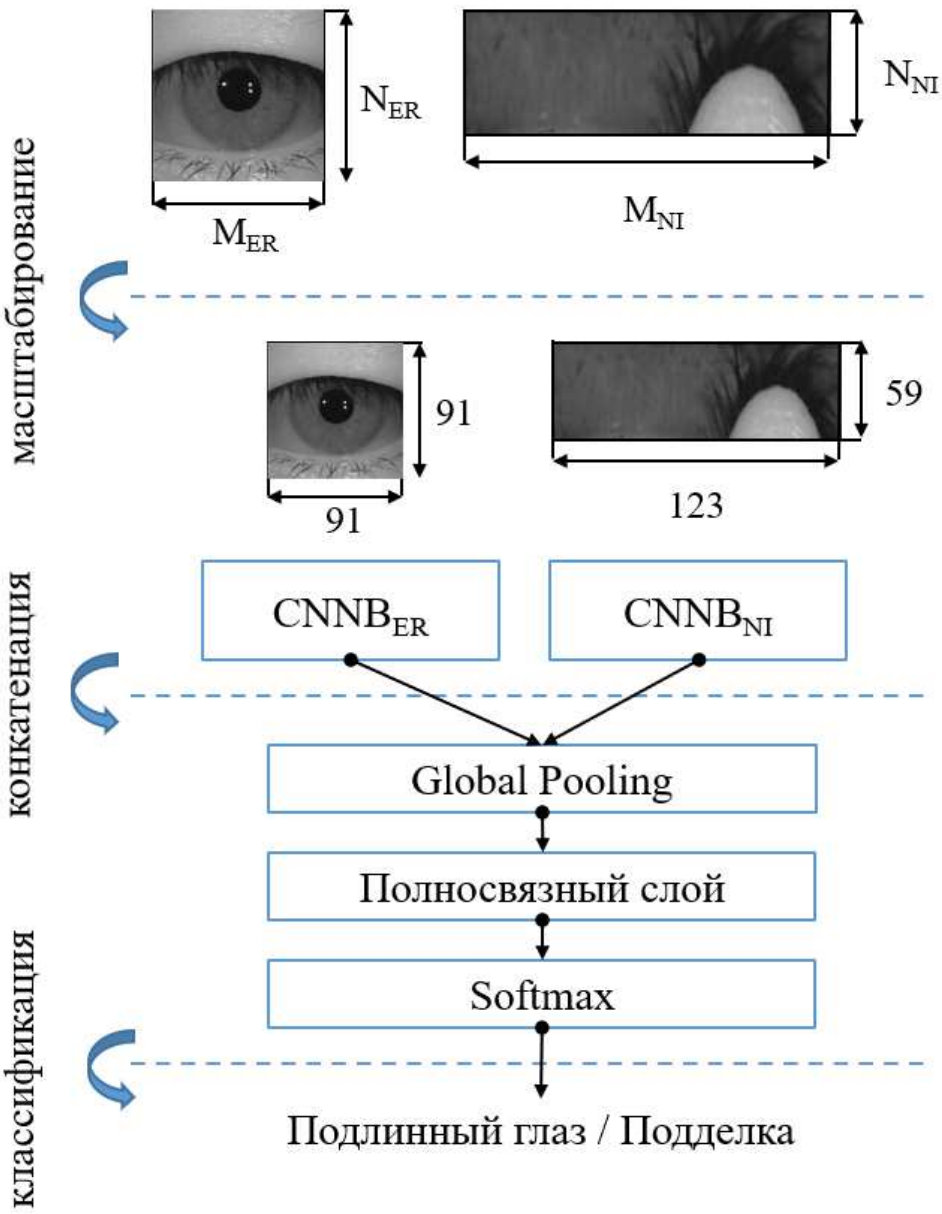


Рис. 5.2. Строение нейросетевого классификатора для обнаружения подделок радужки

емкостных характеристик за счет использования сепарабельных ядер. Трехмерной свертка раскладывается в композицию операций «по глубине» (depthwise), применяющую одно ядро для всех каналов тензора, и «точечную» с ядром размера 1×1, позволяющей скомбинировать результаты «предыдущей» и сформировать более сложное признаковое описание, Табл. 4.1. Такое разбиение позволяет значительно сократить количество атомарных операций (сложение или умножение) и уменьшить число параметров в сверточных блоках.

Сверточные блоки $CNNB_{ER}$ и $CNNB_{NI}$ принимают на вход пару растров

I_{ER} и I_{NI} соответственно, Рис. 5.2, полученные из одного исходного изображения глаза. Строение $CNNB_{ER}$ и $CNNB_{NI}$ дано в Таб. 5.1. Блоки обладают схожей структурой, составными элементами являются сверточные блоки семейства архитектур MobileNet [63], обозначенные как MCB_I. Их структура описана в 4.1.

Элемент архитектуры сети	Размер входного тензора	
	$\mathbf{CNNB_{ER}}$	**$\mathbf{CNNB_{NI}}$**
Сверточный слой ($k_h = k_w = 3, s' = 2$)	$1 \times 91 \times 91$	$1 \times 59 \times 123$
Блок $\mathrm{MCB_I}(k_h = k_w = 3, s' = 2)$	$8 \times 45 \times 45$	$8 \times 29 \times 61$
Блок $\mathrm{MCB_I}(k_h = k_w = 3, s' = 2)$	$16 \times 43 \times 43$	$16 \times 27 \times 59$
Блок $\mathrm{MCB_I}(k_h = k_w = 3, s' = 1)$	$32 \times 21 \times 21$	$32 \times 13 \times 29$
Блок $\mathrm{MCB_I}(k_h = k_w = 3, s' = 2)$	$64 \times 19 \times 19$	$64 \times 11 \times 29$
Блок $\mathrm{MCB_I}(k_h = k_w = 3, s' = 1)$	$64 \times 9 \times 9$	$64 \times 5 \times 13$
Глобальный усредняющий пулинг	$64 \times 7 \times 7$	$64 \times 3 \times 11$

Таблица 5.1. Структура блоков $CNNB_{ER}$ и $CNNB_{NI}$: k_h, k_w — размеры ядер свертки по вертикали и горизонтали соответственно, s' — шаг (stride) применения операции свертки.

Выбранная версия архитектуры модели имеет меньшую вычислительную сложность прямого прохода (forward pass) сети по сравнению с оригинальной [63], несмотря на это позволяя при этом решать поставленную задачу с высокой точностью. Промежуточные признаковые описания изображений I_{ER} и I_{NI}, полученные на выходе соответствующих блоков, обрабатываются операцией глобального усредняющего пулинга (global average pooling) и комбинируются в общий вектор признаков путем конкатенации. Результат подается на вход полносвязному (fully-connected,linear) слою. Вероятности принадлежности P_{live} и P_{spoof} изображения глаза к одному из двух классов («живой» или «подделка» соответственно) оцениваются при помощи softmax классификатора, обучение которого осуществляется при помощи логистической функции потерь.

Описание базы данных подделок В открытом доступе присутству-

ют несколько баз данных, содержащих изображения как подлинных (живых), так и поддельных радужек. Такие базы можно разделить на две группы: полученные в видимом и ближнем инфракрасном (БИК) спектрах. Использование при распознавании по радужной оболочке глаза БИК спектра считается более надежным решением за счет ряда преимуществ [35]. По этой причине современные производители мобильных устройств применяют инфракрасные камеры и активную ИК-подсветку.

Кроме того, описанные в литературе виды подделок радужки в БИК спектре можно разбить на две смысловые группы: созданные с целью имитировать биометрическую характеристику жертвы и, наоборот, нацеленные на сокрытие личности участника процедуры идентификации. К первой группе можно отнести следующие способы спуфинга: распечатка области глаза на бумаге; живая радужка, покрытая полупрозрачной линзой с воспроизведенным на ней рисунком РОГ другого человека; глазные протезы с текстурой РОГ. К последней относятся изображения живых глаз, покрытых текстурированными (узорчатыми) контактными линзами и глазные протезы различной степени реалистичности. Для случая обхода мобильных систем распознавания характерны именно приемы из первой группы, поскольку мобильное устройство, как правило, имеет одного или нескольких пользователей среди узкого круга лиц, например, членов семьи. Подделывание рисунка радужки на полупрозрачной линзе или глазном протезе не рассматривается в данной работе ввиду его высокой сложности реализации, особенно для второго типа подделок. В некоторых работах предлагаются способы детектирования и таких подделок, но случай реализации решения для мобильных устройств не рассматривается. Печать изображения глаза жертвы спуфинг-атаки на бумаге является более интуитивным и простым.

На данный момент в открытом доступе не представлены базы изображений подделок, полученных при помощи мобильного устройства в БИК диапазоне. Поэтому подобная выборка данных была предварительно собрана вручную. В нее были включены следующие типы искусственных образцов: (i) распечатка

качественной фотографии глаза человека (PR), (ii) распечатка с покрытой прозрачной контактной линзой областью радужной оболочки (PWL) и (iii) распечатка с нанесенным на регион РОГ прозрачным канцелярским клеем (PWG). Выбор подобных типов подделывания был сделан с учетом успешного их применения при обходе мобильных биометрических систем [22, 28]. Для регистрации качественных изображений-подделок применялась NIR-камера высокого разрешения с диапазоном расстояний от 20 до 40 (см). Создание подделок производилось путем печати на белой бумаге плотности 80 г/м^3. Для искусственных примеров в равной пропорции применялась плотность печатных элементов в 600 и 1200 (dpi).

Для формирования выборки подлинных примеров были выбраны два вида условий освещенности: (i) нормальный уровень освещенности внутри жилого помещения (IN); (ii) повышенный уровень освещения в солнечную погоду на улице (OUT). Выбор таких категорий был продиктован необходимостью рассмотреть изменения условий окружения, характерным для применения систем распознавания в мобильных устройства.

Для получения изображений в ближнем инфракрасном диапазоне излучения при создании собственной выборки примеров применялся портативный маломощный компьютер Raspberry Pi. В качестве сенсора захвата области глаза применялась совместимая камера (PiCamera v2.1) с заменой фильтра видимого света на полосно-пропускающий в диапазоне 850 $\pm$ 20нм, а для активной подсветки при съемке в помещении был задействован светодиод с соответствующей пиковой частотой излучения. Подробное описание набора данных дано в Таб. 5.2. Разбиение базы на подвыборки для обучения и тестирования учитывало требования непересечения их по субъектам. Рис. 5.3 содержит примеры растров I_{ER}.

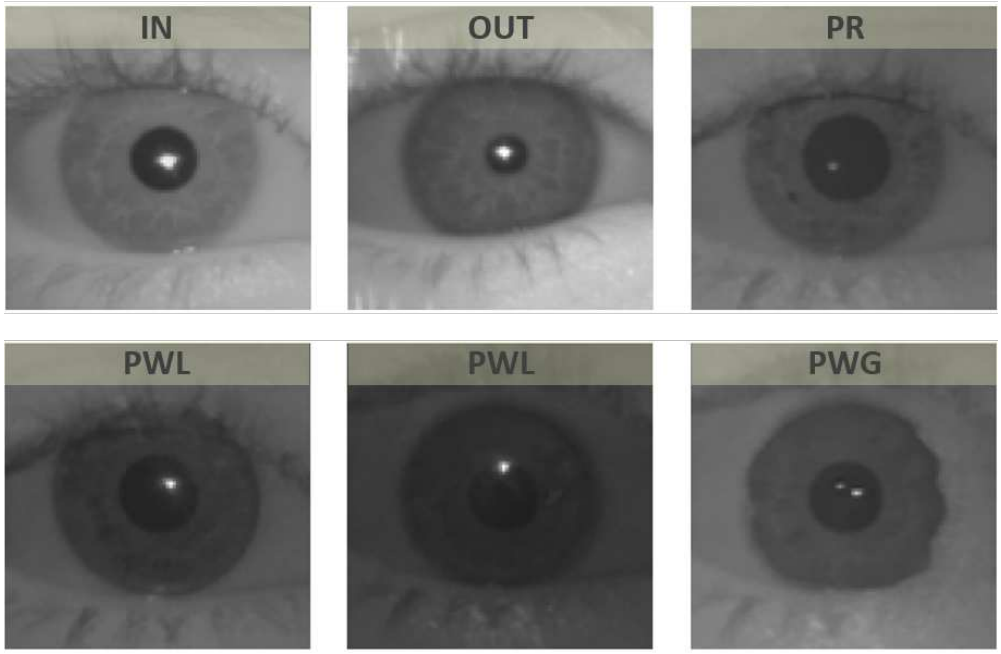


Рис. 5.3. Примеры изображений подделок

5.3. Экспериментальные результаты

Для оценки производительности предлагаемого алгоритма было проведено сравнение по точности с несколькими описанными в современной литературе подходами, зарекомендовавшими себя как демонстрирующие наивысшую производительность на наборах данных растров, полученных в БИК диапазоне согласно обзору, приведенному в работе [43]. При сравнении использовалась собранная база изображений.

Используемые для сравнения методы были реализованы и обучены на ее

Параметр	**Значение**
Разрешение	320×240
Персон/Глаз	23/46
Подделки/Настоящие радужка	18548/18031
IN/OUT/PR/PWL/PWG (всего)	10679/7869/6233/5907/5891
IN/OUT/PR/PWL/PWG (тест)	2534/2006/1436/1452/1568

Таблица 5.2. Описание собранной базы изображений

тренировочной подвыборке. К таковым относятся следующие известные работы: построенные с применением частотного анализа методы [32, 58]; опирающиеся на признаки, извлекаемые при помощи текстурных дескрипторов LBP [56] и BSIF [110] решения, а также подход [122], использующий численные показатели качества изображения. Нейросетевой метод на основе пары CNN в комбинации с набором эвристик, предложенный группой исследователей из CASIA в [145] был исключен из рассмотрения по причине его высокой вычислительной сложности и, как следствие, неприменимости в мобильных биометрических системах, работающих в режиме реального времени. Быстродействие прямого прохода нейронной сети для этого подхода превышает время обработки предлагаемым в данной работе методом на два порядка.

При оценке качества методов применялись следующие показатели:

- *APCER (attack presentation classification error rate)* — доля изображений подделок, ошибочно классифицированных как живые;
- *BPCER (bona fide presentation classification error rate)* — доля изображений живых образцов, ошибочно классифицированных как подделки;
- *CCR (correct classification rate)* — доля правильно классифицированных на всей выборке.

Результаты тестирования упомянутых ранее методов и предлагаемого подхода, включая медианное время выполнения на мобильном процессоре t_{run}, представлены в Таб. 5.3. Стоит упомянуть, что только два упомянутых решения ([122] и [110]) представлялись авторами как допускающие возможность применения в биометрических системах на мобильных устройствах. Метод [122] действительно обладает простотой и относительным быстродействием алгоритма на фоне аналогов. В то же время подход [110] мало применим для применения на видеопотоке в режиме реального времени в мобильном устройстве, поскольку требует произведения вычислительно сложных операций свертки с

Метод	BPCER	APCER	CCR	IN	OUT	PR	PWL	PWG	t_{run}, мс
Czajka [32]	0.505	0.207	0.661	0.614	0.959	0.121	0.598	0.885	64
He *et al.* [58]	0.370	0.739	0.442	0.782	0.919	0.513	0.465	0.530	42
Gupta *et al.* [56]	0.294	0.251	0.749	0.871	0.968	0.049	0.644	0.684	51
Raghavendra [110]	0.076	0.128	0.897	0.916	0.980	0.867	0.836	0.901	900
Sequeira [122]	0.320	0.293	0.694	0.461	0.932	0.644	0.542	0.834	126
Предлагаемый	**0.038**	**0.034**	**0.969**	**0.983**	**0.961**	**0.981**	**0.97**	**0.986**	**5**

Таблица 5.3. Сравнительный анализ методов детектирования подделок радужки

фильтрами большой размерности: от 7×7 до 17×17.

Тестирование методов осуществлялось при помощи мобильного устройства с операционной системой Android. Медианное время классификации нейронной сетью t_{run} с применением одного ядра процессора Qualcomm Snapdragon 835 CPU (2.45 GHz) составляло 4-6 миллисекунд.

5.4. Выводы к пятой главе

Рассмотрены принципы построения систем защиты от взлома при помощи подделок в применении к системам распознавания по радужке на мобильном устройстве. С учетом опыта исследовательских групп профессионалов в данной области, изучена и воспроизведена процедура подделывания современных мобильных систем идентификации по радужке. Произведена классификация общих подходов к защите от подделывания, а также обзор известных из литературы методов, их преимуществ и недостатков. Предложены и исследованы новые пути создания и предъявления подделок системе, ранее не рассматриваемые в литературе: (i) распечатка с покрытой прозрачной контактной линзой областью радужной оболочки; (ii) распечатка с нанесенным в область радужной оболочки прозрачным клеем. Предложена и реализована методология сбора мобильным устройством базы изображений подделок с учетом изменчивости условий окружения, содержащей в том числе новые виды спуфинг-атак. Разработан, протестирован и внедрен новый метод определения живости радужки

при помощи классификатора на основе глубокой сверточной нейронной сети. Предложенный подход показал высокую точность решения задачи, превосходящую таковую для описанных в литературе аналогов, а также достаточную для применения на мобильном устройстве в режиме реального времени скорость обработки.

Заключение

1. Исследованы особенности построения алгоритмов противодействия взлому при помощи подделок для методов биометрического распознавания по видеообразу лица в приложениях мобильных устройств. Исследованы зависимости и причины изменения видимого образа лица с учетом специфики поведения пользователя устройства и характерных нестандартных и изменчивых условий окружения, присущих сценариям регистрации изображений объектов в мобильных приложениях. Разработан, предложен и внедрен метод детектирования подделок, допускающий применение в режиме реального времени в мобильных устройствах.

2. Исследованы методы и алгоритмы извлечения характеристик изображения лица применительно к решению задачи детектирования подделок. Разработан и внедрен метод раннего обнаружения спуфинг-атак для мобильных приложений, позволяющий до применения вычислительно сложных алгоритмов обнаруживать неестественные артефакты и атрибуты, присущие попыткам взлома, учитывать и использовать данные экспозиции камеры и вспомогательных сенсоров устройства с целью получения дополнительной информации об окружении.

3. Исследованы особенности обнаружения попыток подделывания лица при распознавании с мобильного устройства, оборудованного стереокамерой с малым стереобазисом. Разработан и протестирован новый метод определения живости лица при помощи классификатора в виде сверточной нейронной сети. Предложенное решение показало высокую точность и быстродействие детектирования подделок, в том числе тестировании на отложенной выборке данных открытой базы мобильных стереофотографий, содержащей изображения лиц, полученных в широком диапазоне условий окружения.

4. Разработаны, исследованы и внедрены алгоритмы аппроксимации окружностями границ радужной оболочки на изображении глаза основанные на применении методологии глубокого обучения. Предложенные подходы позволяют осуществлять оценку положений границ радужки в режиме реального времени для растров как высокого, так и низкого качества.

5. Изучена специфика построения систем обнаружения попыток взлома мобильных систем распознавания по радужке и новые способы подделывания этой БХЧ. Разработан, протестирован и внедрен новый метод определения живости в виде классификатора в виде сверточной нейронной сети. Предложенное решение показало высокий уровень производительности и быстродействия при детектировании подделок, значительно превышающий таковой для описанных в литературе аналогичных методов.

6. Собраны, обработаны и размечены следующие базы данных: наборы изображений сниженного качества для подлинных лиц и распространенных типов подделок, содержащих более 1000 уникальных личностей и извлеченных при помощи мобильного устройства с имитацией реальных сценариев повседневного использования в изменчивых условиях окружения и применения, набор данных стереоизображений лица (более 90000), набор данных изображений подлинных и поддельных радужек, содержащий как известные, так и новые виды атак (более 160000).

7. Созданы программные средства для проведения вычислительных экспериментов по оценке качества разработанных алгоритмов.

8. Созданы библиотека и демо-приложения для апробации реализованных методов и алгоритмов на мобильном устройстве.

Список литературы

1. Биометрия в СберБанке. — 01.2022. — URL: https://www.sberbank.ru/ru/person/dist_services/bio ; Accessed: 2022-01-30.
2. *Ганькин К.* [и др.]. Сегментация изображения радужки глаза, основанная на приближенных методах с последующими уточнениями // Известия РАН. Теория и системы управления. — 2014. — Т. 53, № 2. — С. 224—238.
3. *Гринчук О.* Методы определения подлинности изображений лиц. Диссертация на соискание ученой степени кандидата технических наук. — 2020.
4. *Ефимов Ю.* [и др.]. Выделение точных границ радужки на изображении глаза // Информационные Технологии. — 2017. — Т. 23, № 4. — С. 300—309.
5. *Лобанов А.* Фотограмметрия. М.: Недра. — 1984.
6. *Матвеев И.* Методы и алгоритмы автоматической обработки изображений радужной оболочки глаза. Диссертация на соискание учёной степени доктора технических наук. — 2014.
7. *Одиноких Г.* Методы и алгоритмы биометрического распознавания человека по радужной оболочке глаза на мобильном устройстве. Диссертация на соискание ученой степени кандидата технических наук. — 2019.
8. *Прэтт У.* Цифровая обработка изображений. — Мир, 1982.
9. *Р.М. Болл Р.* [и др.]. Руководство по биометрии. — Техносфера, 2007. — Пер. с англ.
10. *Чигринский В.* [и др.]. Быстрый алгоритм поиска границ зрачка и радужной оболочки глаза // Машинное обучение и анализ данных. — 2016. — Т. 2, № 2. — С. 159—172.
11. *Шапиро Л. С. Д.* Компьютерное зрение. — Бином. Лаборатория знаний., 2006.

12. Android Ice Cream Sandwich to Feature Face Unlock. — 2011. — URL: https://www.theverge.com/2011/10/18/android-ice-cream-sandwich-feature-face-unlock ; Accessed: 2022-01-30.

13. Apple Face ID Technology. — 2017. — URL: https://support.apple.com/en-us/HT208108 ; Accessed: 2022-01-30.

14. Apple Touch ID Technology. — 2014. — URL: https://en.wikipedia.org/wiki/Touch_ID ; Accessed: 2022-01-30.

15. ARM Security Technology. Building a Secure System using TrustZone Technology. — 2009. — URL: https://documentation-service.arm.com/static/5f212796500e883ab8e74531?token= ; Accessed: 2022-01-30.

16. *Arsalan M.* [et al.]. Deep Learning-Based Iris Segmentation for Iris Recognition in Visible Light Environment // Symmetry. — 2017. — Vol. 9. — P. 263.

17. *Atoum Y.* [et al.]. Face anti-spoofing using patch and depth-based CNNs // 2017 IEEE International Joint Conference on Biometrics (IJCB). — 2017. — P. 319–328.

18. *Badrinarayanan V.* [et al.]. SegNet: A Deep Convolutional Encoder-Decoder Architecture for Image Segmentation // IEEE Transactions on Pattern Analysis and Machine Intelligence. — 2017. — Vol. 39. — P. 2481–2495.

19. *Bao W.* [et al.]. A liveness detection method for face recognition based on optical flow field // 2009 International Conference on Image Analysis and Signal Processing. — 2009. — P. 233–236.

20. *Basit A.* [et al.]. Localization of iris in gray scale image using intensity gradient // Optics and Lasers in Engineering. — 2007. — Vol. 45. — P. 1107–1114.

21. *Bhalgat Y.* [et al.]. LSQ+: Improving low-bit quantization through learnable offsets and better initialization //. — 06/2020. — P. 2978–2985.

22. Bkav Corporation: Galaxy S8 Iris Scanner bypassed by glue. — 2017. — URL: http://www.bkav.com/top-news/-/view_content/content/94273/galaxy-s8-iris-scanner-bypassed-by-gl-1 ; Accessed: 2018-10-01.

23. *Boulkenafet Z.* [et al.]. On the generalization of color texture-based face anti-spoofing // Image and Vision Computing. — 2018. — Vol. 77. — P. 1–9.

24. *Bowyer K.* [et al.]. Image Understanding for Iris Biometrics: A Survey // Comput. Vis. Image Underst. — New York, NY, USA, 2008. — Vol. 110, no. 2. — P. 281–307. — URL: http://dx.doi.org/10.1016/j.cviu.2007.08.005.

25. *Boyd M.* [et al.]. MSc Computing Science Group Project Iris Recognition : Master's thesis / Boyd M. — Imperial College, London, 2010.

26. CASIA-SURF: A Large-Scale Multi-Modal Benchmark for Face Anti-Spoofing / S. Zhang [et al.] // IEEE Transactions on Biometrics, Behavior, and Identity Science. — 2020. — Vol. 2, no. 2. — P. 182–193.

27. *Chang J.* [et al.]. Pyramid Stereo Matching Network // Proc. IEEE/CVF Conf. Computer Vision and Pattern Recognition. Salt Lake City, Utah, USA. — 2018. — P. 5410–5418.

28. Chaos Computer Club (CCC): Chaos Computer Club breaks iris recognition system of the Samsung Galaxy S8. — 2017. — URL: https://www.ccc.de/en/updates/2017/iriden ; Accessed: 2018-10-01.

29. Chinese Academy of Sciences Institute of Automation (CASIA), CASIA-Iris-Mobile-V1.0. — 2015. — URL: http://biometrics.idealtest.org/.

30. Chinese Academy of Sciences Institute of Automation. Iris image database, ver. 2. — 2004. — URL: http://www.cbsr.ia.ac.cn/IrisDatabase.htm.

31. Chinese Academy of Sciences Institute of Automation. Iris image database, ver. 3. — 2005. — URL: http://www.cbsr.ia.ac.cn/IrisDatabase.htm.

32. *Czajka A.* Database of iris printouts and its application: Development of liveness detection method for iris recognition // Proceedings of 18th International Conference on Methods Models in Automation Robotics (MMAR'13). — 2013. — P. 28–33.

33. *Czajka A.* [et al.]. Presentation Attack Detection for Iris Recognition: An Assessment of the State-of-the-Art // ACM Comput. Surv. — New York, NY, USA, 2018. — Vol. 51, no. 4. — 86:1–86:35. — URL: http://doi.acm.org/10.1145/3232849.

34. *Daugman J.* High confidence personal identification by rapid video analysis of iris texture // Proc. IEEE Internat. Carnahan conf. on security technology. — 1992. — P. 50–60.

35. *Daugman J.* How iris recognition works // IEEE Transactions on Circuits and Systems for Video Technology. — 2004. — Vol. 14, no. 1. — P. 21–30.

36. *Daugman J.* Iris recognition and anti-spoofing countermeasures // Proc. of 7-th International Biometrics conference. — 2004.

37. Delta ID Inc.: Fujitsu smartphone powered by Delta ID iris recognition. — 2017. — URL: http://www.deltaid.com/ ; Accessed: 2019-10-01.

38. *DiVerdi S.* [et al.]. Geometric calibration for mobile, stereo, autofocus cameras // 2016 IEEE Winter Conference on Applications of Computer Vision (WACV). — 2016. — P. 1–8.

39. Facial Recognition Market Forecast to 2028 - COVID-19 Impact and Global Analysis By Component, Application, and Vertical. — URL: https://www.researchandmarkets.com/reports/5557882/facial-recognition-market-forecast-to-2028 ; Accessed: 2022-01-30.

40. *Fuhl W.* [et al.]. PupilNet: Convolutional Neural Networks for Robust Pupil Detection. — 2016. — URL: https://arxiv.org/abs/1601.04902.

41. Fujitsu Limited: Fujitsu Develops Prototype Smartphone with Iris Authentication. — 2015. — URL: http://www.fujitsu.com/global/about/resources/news/press-releases/2015/0302-03.html ; Accessed: 2019-09-01.

42. Full frontal vulnerability: Photos can still trick, unlock Android mobes via facial recognition. — 2019. — URL: https://www.theregister.com/2019/01/04/photos_trick_smartphones ; Accessed: 2020-01-30.

43. *Galbally J.* [et al.]. A review of iris anti-spoofing // Proc. of the 4th International Conference on Biometrics and Forensics (IWBF). — 2016. — P. 1–6.

44. *Galbally J.* [et al.]. Iris liveness detection based on quality related features // Proc. 5th IAPR Int. Conf. Biometrics. — 2012. — P. 271.

45. *Gangwar A.* [et al.]. DeepIrisNet2: Learning Deep-IrisCodes from Scratch for Segmentation-Robust Visible Wavelength and Near Infrared Iris Recognition. — 2019. — URL: https://arxiv.org/abs/1902.05390.

46. *Ge H.* [et al.]. Face Anti-Spoofing by the Enhancement of Temporal Motion // 2020 2nd International Conference on Advances in Computer Technology, Information Science and Communications (CTISC). — 2020. — P. 106–111.

47. *George A.* [et al.]. Deep Pixel-wise Binary Supervision for Face Presentation Attack Detection // 2019 International Conference on Biometrics (ICB). — 2019. — P. 1–8.

48. *George A.* [et al.]. Deep Pixel-wise Binary Supervision for Face Presentation Attack Detection //. — 2019. — P. 1–8.

49. *Girshick R.* [et al.]. Rich Feature Hierarchies for Accurate Object Detection and Semantic Segmentation // Proc. of IEEE Conference on Computer Vision and Pattern Recognition. — Washington, DC, USA : IEEE Computer Society, 2014. — P. 580–587. — (CVPR '14). — URL: https://doi.org/10.1109/CVPR.2014.81.

50. Google Android Camera API. — 2021. — URL: https://developers.google.com/ar/reference/java/com/google/ar/core/CameraIntrinsics ; Accessed: 2022-01-30.

51. Google Face Unlock Technology. — 2019. — URL: https://support.google.com/pixelphone/answer/9517039 ; Accessed: 2022-01-30.

52. Google Pixel 3 Device Description. — 2021. — URL: https://en.wikipedia.org/wiki/Pixel_3 ; Accessed: 2021-10-01.

53. Google Pixel 4 Face Unlock works if eyes are shut. — 2019. — URL: https://www.bbc.com/news/technology-50085630 ; Accessed: 2021-10-01.

54. *Gragnaniello D.* [et al.]. An Investigation of Local Descriptors for Biometric Spoofing Detection // IEEE Transactions on Information Forensics and Security. — 2015. — Vol. 10, no. 4. — P. 849–863.

55. *Guo J.* [et al.]. Improving Face Anti-Spoofing by 3D Virtual Synthesis //. — 06/2019. — P. 1–8.

56. *Gupta P.* [et al.]. On Iris Spoofing Using Print Attack // Proc. of 22nd International Conference on Pattern Recognition. — 2014. — P. 1681–1686.

57. *He K.* [et al.]. Deep Residual Learning for Image Recognition // Proc. of IEEE Conference on Computer Vision and Pattern Recognition (CVPR). — 2016. — P. 770–778.

58. *He X.* [et al.]. A fake iris detection method based on FFT and quality assessment // Proc. of Chinese Conference on Pattern Recognition (CCPR'08). — 2008. — P. 1–4.

59. *He X.* [et al.]. A New Fake Iris Detection Method // Advances in Biometrics. — 2009. — Vol. abs/1412.6980.

60. *He Z.* [et al.]. Efficient Iris Spoof Detection via Boosted Local Binary Patterns // Proc. of Advances in Biometrics: Third International Conference. — 2009. — P. 1080–1090.

61. History of Iris Recognition. — 2016. — URL: https://www.cl.cam.ac.uk/~jgd1000/history.html ; Accessed: 2019-06-01.

62. *Ho C. C.* [et al.]. MMU GASPFA: A COTS Multimodal Biometric database // Pattern Recognition Letters. — 2013. — Vol. 34, no. 15. — P. 2043–2050.

63. *Howard A. G.* [et al.]. MobileNets: Efficient Convolutional Neural Networks for Mobile Vision Applications. — 2017. — URL: https://arxiv.org/abs/1704.04861.

64. *Hua Y.* [et al.]. Holopix50k: A Large-Scale In-the-wild Stereo Image Dataset. — 2020. — URL: https://arxiv.org/abs/2003.11172.

65. *ICAO.* ICAO Document 9303: Machine Readable Travel Documents, Part 9: Deployment of Biometric Identification and Electronic Storage of Data in eMRTDs. — 2015. — URL: https://www.icao.int/publications/Documents/9303_p9_cons_en.pdf.

66. *Jain A. K.* [et al.]. Handbook of Biometrics. — 1st. — Springer Publishing Company, Incorporated, 2010.

67. *Jalilian E.* [et al.]. Iris Segmentation Using Fully Convolutional Encoder–Decode Networks // Advances in Computer Vision and Pattern Recognition. — 2017. — P. 133–155.

68. *Jee H.-K.* [et al.]. Liveness Detection for Embedded Face Recognition System // International Journal of Computer and Information Engineering. — 2008. — Vol. 2, no. 6. — P. 2142–2145.

69. *Jourabloo A.* [et al.]. Face De-spoofing: Anti-spoofing via Noise Modeling // Computer Vision – ECCV 2018. — Cham : Springer International Publishing, 2018. — P. 297–315.

70. *Kannala J.* [et al.]. BSIF: Binarized statistical image features // Proceedings of the 21st International Conference on Pattern Recognition (ICPR2012). — 2012. — P. 1363–1366.

71. *Karpathy A.* [et al.]. Large-scale Video Classification with Convolutional Neural Networks // Proceedings of International Computer Vision and Pattern Recognition (CVPR 2014). — 2014.

72. *Khamis S.* [et al.]. StereoNet: Guided Hierarchical Refinement for Real-Time Edge-Aware Depth Prediction // Proc. 15th Europ. Conf. Computer Vision. Munich, Germany. — 2018. — P. 596–613.

73. *Kim G.* [et al.]. Face liveness detection based on texture and frequency analysis // 2012 5th IAPR International Conference on Biometrics (ICB). — 2012. — P. 67–72.

74. *Kim S.* [et al.]. Face liveness detection using variable focusing // 2013 International Conference on Biometrics (ICB). — 2013. — P. 1–6.

75. *Kim T.* [et al.]. BASN: Enriching Feature Representation Using Bipartite Auxiliary Supervisions for Face Anti-Spoofing // 2019 IEEE/CVF International Conference on Computer Vision Workshop (ICCVW). — 2019. — P. 494–503.

76. *Kingma D. P.* [et al.]. Adam: A Method for Stochastic Optimization. — 2014. — URL: https://arxiv.org/abs/1412.6980.

77. *Kollreider K.* [et al.]. Evaluating liveness by face images and the structure tensor // Fourth IEEE Workshop on Automatic Identification Advanced Technologies (AutoID'05). — 2005. — P. 75–80.

78. *Korobkin M.* [et al.]. Iris Segmentation in Challenging Conditions // Proceedings of International Conference on Pattern Recognition and Artificial Intelligence (ICPRAI). — Pattern Recognition, Image Analysis. Springer., 2018. — P. 652–657.

79. *Krizhevsky A.* [et al.]. ImageNet Classification with Deep Convolutional Neural Networks // Advances in Neural Information Processing Systems 25 / ed. by F. Pereira [et al.]. — Curran Associates, Inc., 2012. — P. 1097–1105. — URL: http://papers.nips.cc/paper/4824-imagenet-classification-with-deep-convolutional-neural-networks.pdf.

80. *Lagorio A.* [et al.]. Liveness detection based on 3D face shape analysis // 2013 International Workshop on Biometrics and Forensics (IWBF). — 2013. — P. 1–4.

81. LG's Palm-reading G8 Has a Unique Vision of the Future. — 2019. — URL: https://www.theverge.com/circuitbreaker/2019/2/24/18235249/lg-g8-z-camera-hand-id-palm-scanner-ir-infrared-veins-unlock-mwc-2019 ; Accessed: 2022-01-30.

82. *Li J.* [et al.]. Live face detection based on the analysis of Fourier spectra // Biometric Technology for Human Identification. Vol. 5404 / ed. by A. K.

Jain [et al.]. — International Society for Optics, Photonics. SPIE, 2004. — P. 296–303. — URL: https://doi.org/10.1117/12.541955.

83. *Li L.* [et al.]. An original face anti-spoofing approach using partial convolutional neural network // 2016 Sixth International Conference on Image Processing Theory, Tools and Applications (IPTA). — 2016. — P. 1–6.

84. *Li Z.* [et al.]. An Effective Face Anti-Spoofing Method via Stereo Matching // IEEE Signal Processing Letters. — 2021. — Vol. 28. — P. 847–851.

85. *Lin B.* [et al.]. Face Liveness Detection by RPPG Features and Contextual Patch-Based CNN // Proceedings of the 2019 3rd International Conference on Biometric Engineering and Applications. — Stockholm, Sweden : Association for Computing Machinery, 2019. — P. 61–68. — (ICBEA 2019). — URL: https://doi.org/10.1145/3345336.3345345.

86. *Liu N.* [et al.]. Accurate iris segmentation in non-cooperative environments using fully convolutional networks // International Conference on Biometrics (ICB). — 2016. — P. 1–8.

87. *Liu Y.* [et al.]. Learning Deep Models for Face Anti-Spoofing: Binary or Auxiliary Supervision // 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition. — 2018. — P. 389–398.

88. *Liu C.* Beyond Pixels: Exploring New Representations and Applications for Motion Analysis. Doctoral Thesis. — 2009. — URL: https://people.csail.mit.edu/celiu/OpticalFlow/.

89. *Ma L.* [et al.]. Efficient iris recognition by characterizing key local variations // IEEE Transactions on Image Processing. — 2004. — Vol. 13. — P. 739–750.

90. *Määttä J.* [et al.]. Face spoofing detection from single images using micro-texture analysis // 2011 International Joint Conference on Biometrics (IJCB). — 2011. — P. 1–7.

91. *Marcel S.* [и др.]. Handbook of Biometric Anti-Spoofing: Trusted Biometrics under Spoofing Attacks. — Springer Publishing Company, Incorporated, 2014.

92. *Masek L.* [et al.]. MATLAB Source Code for a Biometric Identification System Based on Iris Patterns. Master Thesis. : Master's thesis / Masek L. — The School of Computer Science, Software Engineering, 2003.

93. Measuring Biometric Unlock Security. — 2019. — URL: https://source.android.com/security/biometric/measure ; Accessed: 2022-01-30.

94. *Menotti D.* [et al.]. Deep Representations for Iris, Face, and Fingerprint Spoofing Detection // IEEE Transactions on Information Forensics and Security. — 2015. — Vol. 10, no. 4. — P. 864–879.

95. Microsoft Corporation: Unlock your Lumia 950 or Lumia 950 XL with a look. — 2017. — URL: https://support.microsoft.com/en-us/instantanswers/4ea145a3-b98e-f8ed-a262-055ec78cdb80/unlock-your-lumia-950-or-lumia-950-xl-with-a-look ; Accessed: 2019-10-01.

96. *Moravcik T.* An Approach to Iris and Pupil Detection in Eye Image : Master's thesis / Moravcik T. — University of Zilina, 2010.

97. Motorola Atrix Device Specifications. — 2011. — URL: https://www.gsmarena.com/motorola_atrix-3709.php ; Accessed: 2022-01-30.

98. *Odinokikh G.* [et al.]. Iris Anti-spoofing Solution for Mobile Biometric Applications // Proceedings of International Conference on Pattern Recognition and Artificial Intelligence. — 2018. — P. 666–671.

99. *Odinokikh G. A.* [et al.]. High-Performance Iris Recognition for Mobile Platforms // Pattern Recognition and Image Analysis. — 2018. — Vol. 28, no. 3. — P. 516–524. — URL: https://doi.org/10.1134/S105466181803015X.

100. *Ojala T.* [et al.]. A comparative study of texture measures with classification based on featured distributions // Pattern Recognition. — 1999. — Vol. 29, no. 1. — P. 51–59.

101. *Pan G.* [et al.]. Monocular camera-based face liveness detection by combining eyeblink and scene context // Telecommunication Systems. — 2011. — Vol. 47. — P. 215–225.

102. *Pan G.* [et al.]. Eyeblink-based Anti-Spoofing in Face Recognition from a Generic Webcamera // 2007 IEEE 11th International Conference on Computer Vision. — 2007. — P. 1–8.

103. *Pan L.* [et al.]. Iris Localization based on Multi-resolution Analysis // Proc. 19th Intern. Conf. Pattern Recognition. — 2008. — P. 1–4.

104. *Parkhi O.* [et al.]. Deep Face Recognition // Proceedings of the British Machine Vision Conference (BMVC). — 09/2015. — P. 41.1–41.12.

105. *Phillips P.* [et al.]. Frvt2006 and ice2006 large-scale experimental results // IEEE PAMI. — 2010. — Vol. 5. — P. 831–846.

106. *Proenca H.* [et al.]. A noisy iris image database // Proc. 13th Int. Conf. Image Analysis and Processing. — 2005. — P. 970–976.

107. *Proenca H.* [et al.]. Iris segmentation methodology for non-cooperative recognition // IEEE Proc. Vision, Image and Signal Processing. Vol. 153. — 2006. — P. 199–205.

108. *Proença H.* [et al.]. IRINA: Iris Recognition (Even) in Inaccurately Segmented Data // Proc. of IEEE Conference on Computer Vision and Pattern

Recognition. — 2017. — P. 6747–6756. — URL: https://doi.org/10.1109/CVPR.2017.714.

109. PSA: Your Note 8's Face Unlock can easily be fooled. — 2017. — URL: https://www.cnet.com/tech/mobile/samsung-note-8-fooled-face-unlock-not-secure ; Accessed: 2022-01-30.

110. *Raghavendra R.* [et al.]. Robust Scheme for Iris Presentation Attack Detection Using Multiscale Binarized Statistical Image Features // IEEE Transactions on Information Forensics and Security. — 2015. — Vol. 10, no. 4. — P. 703–715.

111. *Raja K. B.* [et al.]. Smartphone based robust iris recognition in visible spectrum using clustered K-means features // Proc. of 2014 IEEE Workshop on Biometric Measurements and Systems for Security and Medical Applications. — 2014. — P. 15–21.

112. *Raja K. B.* [et al.]. Presentation attack detection using Laplacian decomposed frequency response for visible spectrum and Near-Infra-Red iris systems // Proc. of IEEE 7th International Conference on Biometrics Theory, Applications and Systems (BTAS). — 2015. — P. 1–8.

113. *Rehman Y. A. U.* [et al.]. SLNet: Stereo face liveness detection via dynamic disparity-maps and convolutional neural network // Expert Systems with Applications. — 2020. — Vol. 142. — P. 113002. — URL: https://www.sciencedirect.com/science/article/pii/S0957417419307195.

114. *Ronneberger. O.* [et al.]. U-Net: Convolutional Networks for Biomedical Image Segmentation // Medical Image Computing and Computer-Assisted Intervention – MICCAI 2015. — Cham : Springer International Publishing, 2015. — P. 234–241.

115. *Ruder S.* An Overview of Multi-Task Learning in Deep Neural Networks. — 2017. — URL: https://arxiv.org/abs/1706.05098.

116. Samsung Electronics Inc.: Security. — 2017. — URL: http://www.samsung.com/global/galaxy/galaxy-s8/security/ ; Accessed: 2018-10-01.

117. Samsung Galaxy S10 Unlock Hack (Youtube Video). — 2019. — URL: https://www.youtube.com/watch?v=BGgQ9woZQOg ; Accessed: 2022-01-30.

118. Samsung Galaxy S10+ Device Description. — 2021. — URL: https://en.wikipedia.org/wiki/Samsung_Galaxy_S10 ; Accessed: 2021-10-01.

119. *Sang J.* [et al.]. Face Image Quality Evaluation for ISO/IEC Standards 19794-5 and 29794-5 // Advances in Biometrics. — Berlin, Heidelberg : Springer Berlin Heidelberg, 2009. — P. 229–238.

120. *Schuckers S.* Spoofing and Anti-Spoofing Measures // Information Security Technical Report. — 2002. — Vol. 7, no. 4. — P. 56–62.

121. *Seo J.* [et al.]. Face Liveness Detection Using Thermal Face-CNN with External Knowledge // Symmetry. — 2019. — Vol. 11, no. 3. — P. 360. — URL: https://doi.org/10.3390/sym11030360.

122. *Sequeira A.* [et al.]. MobiLive 2014 - Mobile Iris Liveness Detection Competition // Proc. of IEEE International Joint Conference on Biometrics. — 2014. — P. 1–6.

123. *Sequeira A. F.* [et al.]. Iris liveness detection methods in mobile applications // Proc. of International Conference on Computer Vision Theory and Applications. Vol. 3. — 2014. — P. 22–33.

124. *Shelhamer E.* [et al.]. Fully Convolutional Networks for Semantic Segmentation // IEEE Trans. Pattern Anal. Mach. Intell. — 2017. — Vol. 39, no. 4. — P. 640–651. — URL: https://doi.org/10.1109/TPAMI.2016.2572683.

125. *Song L.* [et al.]. Face Liveness Detection Based on Joint Analysis of RGB and Near-Infrared Image of Faces // Electronic Imaging. — 2018. — Vol. 6, no. 1. — P. 3731–3736.

126. *Sun L.* [et al.]. Blinking-Based Live Face Detection Using Conditional Random Fields // Advances in Biometrics. — Berlin, Heidelberg : Springer Berlin Heidelberg, 2007. — P. 252–260.

127. *Sun W.* [et al.]. Face Spoofing Detection Based on Local Ternary Label Supervision in Fully Convolutional Networks // IEEE Transactions on Information Forensics and Security. — 2020. — Vol. 15. — P. 3181–3196.

128. *Sun X.* [et al.]. Dual Camera Based Feature for Face Spoofing Detection // Communications in Computer and Information Science. — 2016. — Vol. 662. — P. 332–344. — URL: https://doi.org/10.1007/978-981-10-3002-4_28.

129. *Sun X.* [et al.]. Multimodal Face Spoofing Detection via RGBD Images // Proceedings of International Conference on Pattern Recognition. Beijing, China. — 2018. — P. 2221–2226.

130. *Tan X.* [et al.]. Face Liveness Detection from a Single Image with Sparse Low Rank Bilinear Discriminative Model // Computer Vision – ECCV 2010 / ed. by K. Daniilidis [et al.]. — Springer Berlin Heidelberg, 2010. — P. 504–517.

131. *Thavalengal S.* [et al.]. User Authentication on Smartphones: Focusing on iris biometrics // IEEE Consumer Electronics Magazine. — 2016. — Vol. 5, no. 2. — P. 87–93.

132. *Toshev A.* [et al.]. DeepPose: Human Pose Estimation via Deep Neural Networks // Proc. of the IEEE Conference on Computer Vision and Pattern Recognition. — Washington, DC, USA : IEEE Computer Society, 2014. —

P. 1653–1660. — (CVPR '14). — URL: https://doi.org/10.1109/CVPR.2014.214.

133. Unique Identification Authority of India: What is Aadhaar. — 2021. — URL: https://uidai.gov.in/my-aadhaar/about-your-aadhaar.html ; Accessed: 2022-01-30.

134. University of Bath. Iris Image Database. (2005). — 2005. — URL: http://www.bath.ac.uk/elec-eng/research/sipg/irisweb/.

135. *Vergne C.* [et al.]. World Payments Report (WPR 2021). — 2021. — URL: https://www.paymentscardsandmobile.com/world-payment-report-non-cash-transaction-growth-hit-hard-by-covid ; Accessed: 2022-01-30.

136. *Wang S.* [et al.]. Fast dropout training // Proc. of the 30th International Conference on Machine Learning. Vol. 28–2 / ed. by S. Dasgupta [et al.]. — Atlanta, Georgia, USA : PMLR, 2013. — P. 118–126. — (Proceedings of Machine Learning Research). — URL: http://proceedings.mlr.press/v28/wang13a.html.

137. *Wang T.* [et al.]. Face liveness detection using 3D structure recovered from a single camera // 2013 International Conference on Biometrics (ICB). — 2013. — P. 1–6.

138. *Wang Z.* [et al.]. Deep Spatial Gradient and Temporal Depth Learning for Face Anti-Spoofing // 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). — 2020. — P. 5041–5050.

139. Why face unlock is a bad idea. — 2020. — URL: https://www.kaspersky.com/blog/face-unlock-insecurity/21618 ; Accessed: 2022-01-30.

140. *Wildes R.* Iris Recognition: An Emerging Biometric Technology // Proc. IEEE. Vol. 85. — 1997. — P. 1348–1363.

141. *Xiao J.* [et al.]. SUN database: Large-scale scene recognition from abbey to zoo // 2010 IEEE Computer Society Conference on Computer Vision and Pattern Recognition. — 2010. — P. 3485–3492.

142. *Xu Z.* [et al.]. Learning temporal features using LSTM-CNN architecture for face anti-spoofing // 2015 3rd IAPR Asian Conference on Pattern Recognition (ACPR). — 2015. — P. 141–145.

143. *Yambay D.* [et al.]. LivDet-iris 2013 - Iris Liveness Detection Competition 2013 // Proc. of IEEE International Joint Conference on Biometrics. — 2013. — P. 1–8.

144. *Yambay D.* [et al.]. LivDet-Iris 2015 - Iris Liveness Detection Competition 2015 // Proc. of IEEE International Conference on Identity, Security and Behavior Analysis (ISBA). — 2015. — P. 1–6.

145. *Yambay D.* [et al.]. LivDet-Iris 2017 - Iris Liveness Detection Competition 2017 // Proc. of IEEE International Joint Conference on Biometrics (IJCB). — 2017.

146. *Yan J.* [et al.]. Face liveness detection by exploring multiple scenic clues // 2012 12th International Conference on Control, Automation, Robotics and Vision (ICARCV). — 2012. — P. 188–193.

147. *Yang X.* [et al.]. Face Anti-Spoofing: Model Matters, so Does Data // 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). — 2019. — P. 3502–3511.

148. *Yu Z.* [et al.]. Face Anti-Spoofing with Human Material Perception // Computer Vision – ECCV 2020. — Cham : Springer International Publishing, 2020. — P. 557–575.

149. *Yu Z.* [et al.]. Deep Learning for Face Anti-Spoofing: A Survey. — 2021. — URL: https://arxiv.org/abs/2106.14948.

150. *Yu Z.* [et al.]. TransRPPG: Remote Photoplethysmography Transformer for 3D Mask Face Presentation Attack Detection // IEEE Signal Processing Letters. — 2021. — Vol. 28. — P. 1290–1294.

151. *Yuan W.* [et al.]. A rapid iris location method based on the structure of human eyes // Proc. of 27th Annual Conf. on Engineering in Medicine and Biology. Vol. 45. — 2005. — P. 3020–3023.

152. *Zhang F.* [et al.]. GA-Net: Guided Aggregation Net for End-to-End Stereo Matching // Proc. IEEE/CVF Conf. Computer Vision and Pattern Recognition. Long Beach, CA, USA. — 2019. — P. 185–194.

153. *Zhang K.* [et al.]. Joint Face Detection and Alignment Using Multitask Cascaded Convolutional Networks // IEEE Signal Processing Letters. — 2016. — Vol. 23, no. 10. — P. 1499–1503.

154. *Zhang M.* [et al.]. The BTAS Competition on Mobile Iris Recognition // Proceedings of 8th International Conference on Biometrics Theory, Applications and Systems. — 2016. — URL: http://gen.lib.rus.ec/scimag/index.php?s=10.1109/BTAS.2016.7791191.

155. *Zhang Y.* [et al.]. Du2Net: Learning Depth Estimation from Dual-Cameras and Dual-Pixels // Proc. ECCV 2020, 16th European Conference on Computer Vision, Glasgow, UK. — 2020. — P. 582–598.

156. *Zhou Z.-H.* [et al.]. Projection functions for eye detection // Pattern Recognition. — 2004. — Vol. 37. — P. 1049–1056.