

Спиричев Илья Игоревич

аспирант, Российский новый университет, Москва.

SPIN-код: 8565-1252, AuthorID: 1321573

Электронный адрес: ispirichev@internet.ru

Цуа I. Spirichev

Postgraduate, Russian New University, Moscow.

SPIN-code: 8565-1252, AuthorID: 1321573

E-mail address: ispirichev@internet.ru

Лабунец Леонид Витальевич

доктор технических наук, старший научный сотрудник, профессор кафедры автономных информационных и управляющих устройств, Московский государственный технический университет имени Н.Э. Баумана (национальный исследовательский университет), Москва; профессор кафедры информационных систем в экономике и управлении, Российский новый университет, Москва.

SPIN-код: 7021-3977, AuthorID: 363688

Электронный адрес: labunets@bmstu.ru; llv_rosnou@labnet.ru

Leonid V. Labunets

Doctor of Technical Sciences, Senior Researcher, Professor at the Department of autonomous information and control devices, Bauman Moscow State Technical University (National Research University), Moscow; Professor at the Department of information systems in economics and management, Russian New University, Moscow.

SPIN-code: 7021-3977, AuthorID: 363688

E-mail address: labunets@bmstu.ru; llv_rosnou@labnet.ru

АРХИТЕКТУРА СИСТЕМ РАСПОЗНАВАНИЯ ЛИЦ: ЭТАПЫ ОБРАБОТКИ И КЛАССИФИКАЦИЯ МЕТОДОВ

Аннотация. Рассматриваются архитектурные основы современных систем распознавания лиц. Проанализированы ключевые этапы обработки: обнаружение, предварительная обработка, выравнивание, извлечение признаков и сравнение. Представлена детальная классификация методов – от традиционных подходов и методов на основе локальных признаков (локальные бинарные шаблоны, гистограммы направленных градиентов, масштабно-инвариантное преобразование признаков, ускоренные устойчивые признаки) до современных архитектур глубокого обучения (сверточные нейронные сети, включая подходы FaceNet и ArcFace). Показана взаимосвязь качества всех этапов конвейера обработки и её влияние на конечную точность распознавания.

Ключевые слова: распознавание лиц, этапы обработки, локальные признаки, глубокое обучение, сверточные нейронные сети, биометрия.

Для цитирования: Спиричев И.И., Лабунец Л.В. Архитектура систем распознавания лиц: этапы обработки и классификация методов // Вестник Российского нового университета. Серия: Сложные системы: модели, анализ, управление. 2026. № 1. С. 146–159. DOI: 10.18137/RNU.V9187.26.01.P.146

ARCHITECTURE OF FACE RECOGNITION SYSTEMS: PROCESSING
STAGES AND CLASSIFICATION OF METHODS

Abstract. The article examines the architectural foundations of modern face recognition systems. Key processing stages are analyzed: detection, preprocessing, alignment, feature extraction, and matching. The authors provide a detailed classification of methods – from traditional approaches and local-feature-based methods (Local Binary Patterns, Histogram of Oriented Gradients, Scale-Invariant Feature Transform, Speeded-Up Robust Features) to modern deep learning architectures (Convolutional Neural Networks, including FaceNet and ArcFace approaches). The article demonstrates the interdependence of all pipeline stages and its impact on final recognition accuracy.

Keywords: face recognition, processing stages, local features, deep learning, convolutional neural networks, biometrics.

For citation: Spirichev I.I., Labunets L.V. (2026) Architecture of face recognition systems: Processing stages and classification of methods. *Vestnik of Russian New University. Series: Complex Systems: Models, analysis, management.* No. 1. Pp. 146–159. DOI: 10.18137/RNUV9187.26.01.P.146 (In Russian).

Введение

Современные технологии цифровой идентификации личности находятся на этапе интенсивного развития, а распознавание лиц как один из наиболее перспективных биометрических методов занимает особое место в этой трансформации. В условиях стремительного роста объёмов цифровых изображений и видео в повседневной жизни человека, а также возрастающих требований к безопасности и персонализации сервисов задача автоматического распознавания лиц приобретает стратегическое значение для множества сфер деятельности – от систем контроля доступа и правоохранительных структур до развлекательных приложений и социальных сетей.

Распознавание лиц представляет собой сложную систему, объединяющую методы компьютерного зрения, машинного обучения и обработки изображений. Исторически развитие этой области прошло путь от простых геометрических методов до современных архитектур глубокого обучения, демонстрирующих точность, превосходящую человеческую в контролируемых условиях. Ключевым фактором успеха современных систем является корректная организация вычислительного конвейера – последовательности этапов обработки, каждый из которых критически влияет на конечный результат.

Несмотря на значительные достижения последнего десятилетия, проблема остаётся актуальной из-за сохраняющихся вызовов: вариативности условий регистрации изображений, необходимости работы в реальном времени при ограниченных вычислительных ресурсах и требований к устойчивости алгоритмов к геометрическим искажениям.

Целью данной статьи является систематизация архитектурных решений современных систем распознавания лиц: анализ ключевых этапов обработки (детекция, предварительная обработка, выравнивание, извлечение признаков, сравнение) и детальная классификация методов – от традиционных подходов и методов на основе локальных признаков до современных архитектур глубокого обучения.

Работа представляет собой теоретико-методологический фундамент для понимания принципов построения эффективных систем распознавания лиц.

Основные этапы системы распознавания лиц

Успешное распознавание лица – это многоступенчатый процесс, где каждый этап критически важен для конечного результата. Ошибки на ранних стадиях (например, некорректное обнаружение или выравнивание лица) могут существенно снизить точность работы всей системы даже при использовании совершенных алгоритмов извлечения признаков.

Рассмотрим ключевые этапы, составляющие основу современных систем распознавания лиц.

1. Обнаружение лица (детекция лиц). Первоначальным и фундаментальным этапом любой системы распознавания лиц является обнаружение и локализация лиц на входном изображении или видеопоследовательности. Основная задача данного этапа – определить координаты расположения лиц с помощью ограничивающих прямоугольников. Без надежного выполнения этой операции все последующие этапы обработки становятся невозможными [1].

Исторически значимым прорывом стал алгоритм Виолы – Джонса, представленный в 2001 году [1]. Данный метод характеризуется следующими ключевыми особенностями:

- использование признаков Хаара в виде прямоугольных шаблонов для выявления контрастных областей лица;
- применение интегрального изображения для эффективного вычисления сумм пикселей;
- последовательное усиление адаптивного метода для агрегирования слабых классификаторов в сильный;
- высокая скорость работы за счет последовательного применения каскадной структуры быстрых классификаторов.

Несмотря на эффективность, классический метод Виолы – Джонса демонстрирует ограниченную устойчивость к поворотам лица, изменчивым условиям освещения и сложному фону [2].

В современных условиях доминирующее положение заняли детекторы на основе сверточных нейронных сетей, способные автоматически формировать иерархию признаков [3]. Их преимущества включают способность обнаруживать лица при значительных углах поворота и в сложных условиях съемки [4], а также обеспечивать высокую точность локализации и устойчивость работы.

Ключевыми требованиями к современным детекторам лиц являются [5]:

- точность локализации;
- устойчивость к изменяющимся условиям съемки;
- высокая скорость обработки;
- способность обнаруживать множественные лица на изображении.

Качественное выполнение этапа обнаружения лица является критически важным фундаментом всей системы, поскольку ошибки на этой стадии необратимо влияют на результат последующего распознавания.

Современные подходы. В современных системах распознавания лиц доминирующее положение занимают детекторы на основе сверточных нейронных сетей. Среди наиболее эффективных современных решений следует выделить многоцелевую каскадную сверточную нейронную сеть (MTCNN), которая представляет собой каскад из трех последова-

Архитектура систем распознавания лиц:
этапы обработки и классификация методов

тельных сетей и обеспечивает не только обнаружение лиц, но и одновременное определение ключевых точек. Другим значимым достижением является детектор RetinaFace, относящийся к классу одноступенчатых детекторов. Его отличительной особенностью является предсказание трехмерной ориентации лица и карты видимости его участков, что значительно повышает точность локализации в сложных условиях.

Современные требования, предъявляемые к детекторам лиц, обеспечивают высокую точность локализации при сохранении приемлемой производительности, устойчивость к изменениям условий съемки, способность обрабатывать изображения в реальном времени, эффективное обнаружение множественных лиц различного масштаба. Особое значение имеет способность детектора сохранять работоспособность при наличии помех и дефектов сжатия, что особенно актуально для систем видеонаблюдения.

2. Предварительная обработка. Успешное обнаружение лиц на изображении позволяет на этапе предварительной обработки выполнить подготовку изображений к последующему анализу. Целью данного этапа является повышение информативности визуальных данных и их стандартизация для обеспечения стабильной работы алгоритмов распознавания.

Реальные изображения лиц часто содержат дефекты, обусловленные неидеальными условиями съемки: вариациями освещенности, неравномерностью теней, недостаточным контрастом, различиями в масштабе и цветопередаче. Указанные факторы не несут информации об идентификационных характеристиках лица, но, как правило, существенно затрудняют корректное выделение значимых признаков.

Основные направления предварительной обработки включают несколько подходов.

Масштабирование и приведение к единому размеру обеспечивает стандартизацию входных данных для алгоритмов извлечения признаков, требующих изображения фиксированного формата. Для сохранения информативности при изменении размера применяются специализированные алгоритмы интерполяции.

Коррекция проблем освещенности и контраста представляет наиболее сложную задачу. Для ее решения применяются: метод выравнивания гистограммы, перераспределяющий яркости пикселей для повышения общей контрастности; адаптивное выравнивание гистограммы, работающее с локальными участками изображения и эффективно подавляющее шумы; гамма-коррекция для нелинейного преобразования яркости; сглаживающая фильтрация для подавления шумов цифровой съемки.

Особое значение имеет *работа с цветовыми характеристиками изображения*. Для многих алгоритмов распознавания цветовая информация является избыточной и может снижать надежность работы. Преобразование в градации серого позволяет сфокусироваться на текстуре и форме лица. Альтернативным подходом является использование цветных пространств с разделенными компонентами яркости и цвета, что обеспечивает независимую коррекцию освещенности.

Качественная предварительная обработка обеспечивает снижение зависимости от условий съемки, выделение значимых идентификационных признаков и повышение точности последующих этапов обработки.

3. Выравнивание лица (нормализация позы). Вариации позы лица представляют собой один из наиболее значимых факторов, снижающих эффективность систем распознавания. Лица на изображениях редко имеют идеально фронтальное положение, поскольку люди естественным образом поворачивают, наклоняют или приподнимают голову. Эти

вариации приводят к изменению пространственного расположения анатомических ориентиров (глаз, носа, рта) относительно границ изображения и друг друга, что затрудняет последующий анализ. Этап выравнивания лица, также известный как нормализация позы, направлен на преобразование изображения таким образом, чтобы ключевые анатомические точки занимали строго определенные позиции, формируя стандартизованную систему координат для последующих этапов обработки.

Алгоритм выравнивания обеспечивает детекцию ключевых точек лица, что является фундаментальной операцией всего процесса. Минимально необходимый набор включает центры зрачков обоих глаз, однако для повышения точности обычно определяются дополнительные ориентиры (см. Рисунок): кончик носа, углы рта, контур подбородка, точки бровей и контур лица (височные, скуловые области).

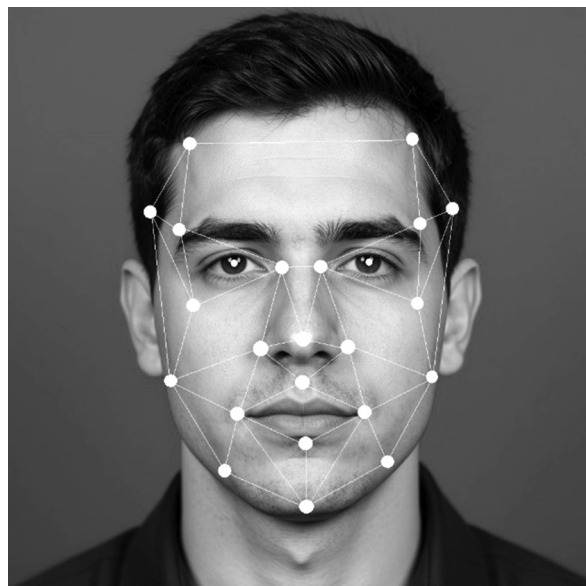


Рисунок. Детекция ключевых точек лица

Источник: рисунок выполнен авторами

Современные методы, основанные на глубоком обучении (например, реализованные в детекторе MTCNN или специализированных моделях), обеспечивают высокую точность и скорость определения ключевых точек даже при небольших поворотах и наклонах головы. После обнаружения ключевых точек задается целевое положение для фронтального, выровненного лица, где центры глаз находятся на одной горизонтальной линии, расстояние между ними составляет фиксированную долю от ширины итогового изображения, а соединяющая их линия параллельна верхнему и нижнему краям изображения [6].

На основе вычисленной разницы между текущим и целевым положением ключевых точек определяют параметры геометрического преобразования. Наиболее широко применяется аффинное преобразование, представляющее собой комбинацию поворота, масштабирования и сдвига. Это преобразование сохраняет параллельность прямых линий и эффективно корректирует наклоны и небольшие повороты головы. Для случаев со значи-

Архитектура систем распознавания лиц:
этапы обработки и классификация методов

тельными перспективными искажениями (сильные наклоны головы вперед или назад) применяют проективное преобразование (гомографию), которое позволяет более точно проецировать лицо на фронтальную плоскость, однако требует больших вычислительных ресурсов.

Выравнивание лица критически важно для эффективности всей системы распознавания по нескольким причинам:

- 1) устраняет один из основных источников вариативности входных данных;
- 2) повышает точность извлечения признаков, поскольку алгоритмы работают с согласованными пространственными областями лица;
- 3) улучшает качество сравнения векторов признаков за счет их пространственной согласованности;
- 4) обеспечивает устойчивость системы к небольшим отклонениям позы при распознавании новых изображений.

Необходимо отметить ограничения методов выравнивания: они эффективны преимущественно для умеренных поворотов и наклонов головы (обычно в пределах ± 30 градусов по каждому из углов). При очень сильных поворотах (профильное изображение) или значительных перекрытиях лица (руками, шарфом, маской) точность детекции ключевых точек резко снижается, что ограничивает применимость методов нормализации позы.

4. Извлечение признаков. Этап извлечения признаков представляет собой центральный компонент системы распознавания лиц, определяющий ее эффективность и надежность. Если предыдущие этапы (обнаружение, предварительная обработка, выравнивание) выполняют подготовительные функции, направленные на локализацию лица и приведение изображения к стандартизированному виду, то извлечение признаков непосредственно решает ключевую задачу – формирование уникального числового представления, отражающего индивидуальные характеристики лица.

Основная цель данного этапа – преобразование подготовленного изображения лица в компактный вектор признаков, удовлетворяющий трем критическим требованиям:

- 1) вектор должен обеспечивать высокую способность к распознаванию объектов, позволяя надежно идентифицировать человека среди множества других людей;
- 2) представление должно обладать инвариантностью к внешним факторам, проявляя минимальную изменчивость при вариациях освещения, незначительных изменениях выражения лица, небольших отклонениях позы и естественных возрастных трансформациях;
- 3) вектор признаков должен быть компактным для обеспечения эффективного хранения и быстрого сравнения в больших базах данных.

Существует два принципиально различных подхода к формированию признаков. Первый подход, характерный для традиционных методов, предполагает ручное конструирование признаков на основе экспертных знаний. В рамках геометрических методов производят измерение расстояний между анатомическими ориентирами (межзрачковое расстояние, пропорции лицевых частей), углов и их соотношений. В методах, основанных на анализе текстуры, применяются алгоритмы, такие как локальные бинарные шаблоны (LBP), разбивающие лицо на локальные области и описывающие текстурные признаки в каждой из них. Преимущества традиционных методов заключаются в интерпретируемости и относительной вычислительной простоте, однако они демонстрируют ограниченную устойчивость к сложным условиям регистрации изображений и требуют значительных экспертных усилий при разработке.

Современный стандарт в области извлечения признаков формируют методы глубокого обучения, в частности сверточные нейронные сети. В отличие от традиционных подходов, нейронные сети автоматически извлекают иерархические признаки непосредственно из обучающих данных, без необходимости ручного проектирования. Архитектура сети строится таким образом, что начальные слои выявляют простые элементы (края, текстуры), промежуточные слои комбинируют их в сложные структуры (черты лица), а «глубокие» слои формируют высокоуровневое семантическое представление, специфичное для конкретного лица. Результат работы сети представлен эмбедингом – вектором числовых значений (обычно размером 128...512 элементов), который служит цифровым отпечатком лица. Лица одного человека должны иметь близкие эмбединги в пространстве признаков, тогда как эмбединги разных людей должны быть значительно удалены друг от друга [7].

Этап извлечения признаков имеет фундаментальное значение для всей системы распознавания по нескольким причинам:

1) качество извлеченных признаков непосредственно определяет точность и надежность системы в целом, поскольку последующие этапы сравнения работают именно с этим представлением;

2) компактная форма вектора признаков обеспечивает возможность эффективного сравнения: вычисление меры сходства между двумя векторами значительно быстрее и проще, чем пиксельное сравнение изображений;

3) использование векторных представлений обеспечивает масштабируемость системы, позволяя работать с базами данных, содержащими миллионы и миллиарды записей, при относительно скромных вычислительных ресурсах.

Современные достижения в области глубокого обучения позволили создать методы автоматического формирования признаков, обеспечивающие беспрецедентную точность и устойчивость, что обуславливает их доминирующее положение в современных системах распознавания лиц.

5. Классификация / Сравнение (сопоставление). Векторное представление лица позволяет реализовать заключительный этап процесса распознавания – классификацию или сравнение (сопоставление). Данный этап решает ключевую задачу идентификации личности путем анализа сходства между полученным вектором признаков и эталонными образцами, хранящимися в базе данных. В зависимости от прикладной задачи выделяют два основных режима работы системы: верификация по правилу «один к одному» и идентификация «один ко многим».

Верификация представляет собой процесс подтверждения предполагаемой личности. Система сравнивает вектор признаков входного изображения с единственным эталонным вектором, соответствующим заявленному идентификатору. Результатом является бинарное решение: лицо совпадает с заявленным или не совпадает. Данная задача широко применяется в системах контроля доступа, разблокировки мобильных устройств и аутентификации в финансовых приложениях.

Идентификация представляет собой поиск неизвестного лица в базе данных, содержащей записи множества персон. Вектор признаков анализируемого лица последовательно сравнивается со всеми эталонными векторами в базе (которая может содержать от тысяч до миллиардов записей). Результатом работы является либо идентификатор наиболее похожего лица, либо ранжированный список кандидатов, упорядоченный по степени сход-

Архитектура систем распознавания лиц:
этапы обработки и классификация методов

ства. Эта задача находит применение в правоохранительной деятельности, системах видеонаблюдения и автоматической маркировке изображений.

Сравнение векторов признаков осуществляют, например, по принципу минимума расстояния или максимума математических мер сходства в многомерном пространстве признаков. Наиболее распространенными являются евклидово расстояние и сходство по критерию нормированного скалярного произведения векторов признаков. Евклидово расстояние определяет прямолинейную дистанцию между двумя точками в пространстве признаков, где каждая координата соответствует одному признаку. Малые значения расстояния свидетельствуют о высоком сходстве лиц. Мера сходства на основе скалярного произведения измеряет косинус угла между векторами признаков, не учитывая их абсолютную длину. Значение данного показателя находится в диапазоне от -1 до 1 , где значения, близкие к 1 , указывают на высокую степень сходства. Этот критерий является робастным по отношению к вариациям освещения и контраста, поскольку игнорирует общую яркость изображения.

Для традиционных методов распознавания, основанных на ручном конструировании признаков, широко применяются алгоритмы машинного обучения. Метод k ближайших соседей (k -NN) классифицирует входное лицо на основе голосования среди k наиболее похожих эталонов. Метод опорных векторов (SVM) формирует оптимальную разделяющую гиперплоскость в признаковом пространстве, максимизирующую зазор между различными классами. Для задач верификации часто используют модификацию SVM-классификатора, построенного по правилам «один против всех», когда синтезируется дискриминантная поверхность между векторами признаков конкретного человека и всех остальных индивидов или «один против одного», когда формируются дискриминантные поверхности для каждой пары индивидов.

Ключевым моментом процесса принятия решения является выбор порогового значения. Для верификации результат сравнения (расстояние или сходство) сопоставляется с заранее определенным порогом. Значение порога напрямую влияет на соотношение между двумя типами ошибок: частотой ложных срабатываний (FAR – разрешение доступа посторонним) и частотой ложных отказов (FRR – отказ в доступе зарегистрированным пользователям). Снижение порога повышает безопасность системы, но ухудшает удобство использования, и наоборот. Оптимизация порогового значения обычно осуществляется методом скользящей проверки на этапе калибровки системы с помощью репрезентативных тестовых выборок.

Значение этапа сравнения определяется несколькими факторами:

1) именно на этом этапе принимается окончательное решение о принадлежности лица к определенной личности, что непосредственно влияет на практическую применимость системы;

2) алгоритмы сравнения векторов обеспечивают вычислительную эффективность, позволяя осуществлять поиск в масштабных базах данных в режиме реального времени;

3) гибкость подхода позволяет настраивать систему под различные требования безопасности и удобства использования путем изменения пороговых значений и методов сравнения.

Эффективность данного этапа напрямую зависит от качества извлеченных признаков и корректного выбора методов сравнения, адаптированных к конкретным условиям применения системы.

Классификация методов распознавания лиц

История развития технологий распознавания лиц – это путь от простых, интуитивно понятных подходов к мощным, но сложным алгоритмам на основе искусственного интеллекта. Все многообразие существующих методов можно разделить на три большие категории, каждая из которых отражает определенный этап эволюции и свой взгляд на то, как «увидеть» уникальность человеческого лица. В этом разделе представлен анализ методов распознавания лиц, начиная с самых ранних традиционных методов.

Методы на основе локальных признаков. Осознавая ограничения традиционных подходов, особенно их уязвимость к вариациям освещения, позы и мимики, исследователи разработали методы на основе локальных признаков, которые стали значительным прорывом в области распознавания лиц. Данный подход основан на фундаментальной идее: вместо анализа лица как единого целого сосредоточиться на его устойчивых локальных деталях, которые сохраняются даже при изменении условий регистрации изображения.

Суть методологии заключается в декомпозиции изображения выровненного лица на множество локальных областей (блоков), каждая из которых анализируется независимо для извлечения характерных текстурных или структурных особенностей. Для каждой такой области вычисляется компактный локальный описатель, отражающий ключевые визуальные шаблоны. Затем полученные описатели объединяются в единое глобальное представление лица с помощью статистических методов, таких как построение гистограмм или применение метода коллекции визуальных слов. Такой подход обеспечивает значительную устойчивость к помехам, поскольку локальные детали менее подвержены влиянию глобальных изменений освещения или геометрии лица.

Среди ключевых методов локального анализа выделяют *метод локальных бинарных шаблонов* (LBP). Его принцип основан на сравнении яркости центрального пикселя с яркостью окружающих соседей в заданной окрестности. Результат сравнения кодируется в бинарную последовательность, которая преобразуется в десятичное число – код LBP для данного пикселя. Данный метод демонстрирует выдающуюся устойчивость к изменениям освещения, поскольку кодирует относительные, а не абсолютные значения яркости. Для формирования глобального описания лица применяют гистограммы локальных бинарных шаблонов (LBPН), которые агрегируют распределение кодов LBP по всем блокам изображения.

Гистограммы направленных градиентов (HOG) представляют собой другой важный подход, фокусирующийся на анализе градиентов яркости изображения. Данный метод вычисляет направление и силу перепадов яркости в каждой локальной области, делит направления на определенное количество угловых секторов и строит гистограмму распределения градиентов по этим направлениям. Метод HOG особенно эффективен для описания формы и контуров основных черт лица, таких как глаза, нос и рот, и обладает устойчивостью к небольшим смещениям и поворотам в пределах анализируемого блока.

Более сложные и мощные методы включают *масштабно-инвариантное преобразование признаков* (SIFT) и *ускоренные устойчивые признаки* (SURF). Эти алгоритмы автоматически обнаруживают ключевые точки на лице – области с высокой контрастностью и уникальной локальной структурой, не обязательно совпадающие с анатомическими ориентирами. Для каждой ключевой точки строится многомерный описатель, учитывающий

Архитектура систем распознавания лиц:
этапы обработки и классификация методов

распределение градиентов в ее окрестности с нормализацией по масштабу и ориентации. Такой подход обеспечивает инвариантность к изменению размера изображения, поворотам и частичной устойчивости к вариациям освещения.

Преимущества методов на основе локальных признаков многообразны:

- 1) они демонстрируют высокую устойчивость к изменениям освещения, особенно LBP, что позволяет работать в условиях, недоступных для традиционных методов;
- 2) система сохраняет работоспособность при частичных перекрытиях лица, поскольку информация из визуально видимых областей остается полезной для распознавания;
- 3) данные методы лучше справляются с вариациями позы и мимики, так как локальные свойства текстуры и градиентов в стабильных областях лица изменяются меньше, чем глобальная геометрия.

Несмотря на существенные преимущества, методы локальных признаков имеют определенные ограничения. Они требуют тщательной ручной настройки параметров, таких как размер блоков, радиус анализа ближайших соседей, параметры деления, что существенно влияет на результативность системы. Проблема оптимального объединения информации из тысяч локальных описателей в единое глобальное представление остается нетривиальной задачей, так как простые методы агрегации могут терять важные пространственные взаимосвязи между областями. Кроме того, устойчивость данных методов к сильным поворотам головы (более 30 градусов), экстремальным условиям освещения или значительным перекрытиям (капюшоны, медицинские маски) остается ограниченной.

Методы LBP, HOG, SIFT и SURF стали основой для многих практических приложений в 2000–2010-х годах и заложили концептуальные основы для последующего развития методов глубокого обучения. Хотя сегодня эти методы в чистом виде используются реже, их принципы продолжают влиять на архитектуры современных нейронных сетей и находят применение в гибридных системах распознавания лиц.

Историческое значение и современное применение. Традиционные методы распознавания лиц, в частности алгоритм «собственных лиц» (Eigenfaces), имеют фундаментальное историческое значение для развития компьютерного зрения и биометрических технологий. Внедренные в начале 1990-х годов, эти методы впервые продемонстрировали принципиальную возможность автоматической идентификации личности по изображению лица, заложив теоретические основы для последующих исследований в этой области. Метод главных компонент (РСА), примененный к задаче распознавания лиц в работах Терка и Пентлаанда [8], стал первым математически строгим подходом, позволившим формализовать проблему в терминах многомерного анализа и статистической обработки изображений.

Теоретическая ценность этих методов заключается в разработке концептуального аппарата для представления лиц в многомерном пространстве признаков и формализации задачи классификации. Практическое значение проявилось в создании первых прототипов систем контроля доступа и демонстрации потенциала биометрической идентификации в контролируемых условиях. Тем не менее фундаментальные ограничения традиционных методов, обусловленные их высокой чувствительностью к вариациям освещения, позы головы и выражения лица, препятствовали их широкому практическому применению в реальных условиях.

В современной практике непосредственные реализации геометрических методов и подходов на основе глобальных шаблонов, таких как Eigenfaces и Fisherfaces, практиче-

ски не применяются в промышленных высокоточных системах распознавания лиц. Это связано с появлением более устойчивых методов на основе локальных признаков, а затем – с революцией, вызванной методами глубокого обучения, которые продемонстрировали существенно более высокую точность и робастность. Тем не менее традиционные методы сохраняют определенную значимость в образовательном контексте как иллюстрация эволюции подходов к решению задачи распознавания лиц.

Методы на основе глубокого обучения. Наступление эры глубокого обучения, в частности сверточных нейронных сетей (СНС), ознаменовало собой качественный прорыв в области распознавания лиц, позволив достичь точности, превосходящей человеческую в контролируемых условиях, обеспечив работоспособность систем в реальных неконтролируемых средах. В отличие от предыдущих подходов методы глубокого обучения принципиально изменяют парадигму формирования признаков, автоматизируя процесс извлечения репрезентативных характеристик непосредственно из «сырых» пиксельных данных, без необходимости ручного конструирования и экспертного проектирования.

Фундаментальное отличие методов глубокого обучения заключается в их способности к автоматическому извлечению иерархических признаков через многослойную обработку входных данных. Сверточные нейронные сети последовательно формируют представления возрастающей сложности в начальных, промежуточных и глубоких слоях, формируя высокоуровневые семантические представления, специфичные для идентификации личности. Данный процесс обеспечивается архитектурой сети, состоящей из чередующихся сверточных слоев, применяющих фильтры для выявления локальных особенностей, слоев подвыборки (пулинга), обеспечивающих инвариантность к небольшим сдвигам и снижающих пространственное разрешение, и полносвязных слоев, формирующих окончательное векторное представление.

Ключевым концептуальным достижением в области распознавания лиц стал переход от задачи классификации к задаче обучения метрического пространства признаков. Современные архитектуры не только оптимизируют корректную классификацию изображений по идентификаторам, но и формируют такое признаковое пространство, где векторные представления лиц одного человека максимально компактны, а представления разных людей значительно удалены друг от друга в метрическом пространстве. Данная теория реализована через специализированные функции потерь, среди которых выделяются следующие подходы.

Подход FaceNet, разработанный в Google, ввел функцию потерь на основе троек (TripletLoss), использующую обучающие шаблоны изображений: «якорь» – изображение человека, «позитив» – другое изображение того же человека, «негатив» – изображение другого человека. Функция потерь оптимизируется таким образом, чтобы евклидово расстояние между эмбедингами якоря и позитива было меньше расстояния между якорем и негативом на заданную границу. Данный подход обеспечивает формирование плотных кластеров векторных представлений для каждого идентификатора в признаковом пространстве.

ArcFace представляет собой более современный подход, добавляющий угловую границу (angular margin) в процесс обучения. В отличие от FaceNet, который работает в евклидовом пространстве расстояний, ArcFace модифицирует функцию потерь, добавляя суммирующую угловую границу непосредственно в целевой угол между эмбедингом и вектором весов соответствующего класса. Данная модификация создает дополнительный

Архитектура систем распознавания лиц:
этапы обработки и классификация методов

«зазор» между классами в угловом пространстве, обеспечивая более четкое разделение кластеров разных личностей и повышенную компактность представлений внутри классов. Аналогичные принципы лежат в основе методов SphereFace и CosFace, различающихся способом введения границы в метрическое пространство.

Эволюция архитектур сверточных сетей для распознавания лиц прошла путь от относительно простых моделей DeepFace и DeepID, демонстрировавших потенциал подхода в 2014 году, до современных высокопроизводительных решений. VGG-Face, основанная на архитектуре VGG-16, стала популярным базисом благодаря своей структурной простоте и высокой точности. Революционным прорывом стали остаточные сети (ResNet), благодаря которым была введена концепция «сквозных» соединений (skip connections) для решения проблемы исчезающих градиентов в глубоких архитектурах. Это позволило создавать сети с сотнями слоев, обеспечив беспрецедентную точность распознавания. Современные системы часто используют гибридные архитектуры, такие как Inception-ResNet, сочетающие преимущества различных подходов к организации вычислений.

Преимущества методов глубокого обучения многогранны:

1) обеспечивают выдающуюся точность на стандартных тестовых наборах, значительно превосходя все предыдущие подходы, особенно в сложных условиях регистрации изображений;

2) демонстрируют высокую устойчивость к вариациям освещения, частичным перекрытиям, изменениям мимики и умеренным отклонениям позы головы;

3) автоматизация процесса извлечения признаков устраняет субъективность и трудоемкость этапа ручного конструирования описателей, что значительно ускоряет разработку и внедрение систем.

Несмотря на значительные достижения, методы глубокого обучения сталкиваются с рядом существенных ограничений. Проблема интерпретируемости («эффект черного ящика») затрудняет анализ причин ошибок и диагностику систем, что особенно критично в высокорискованных приложениях безопасности и правоохранительной деятельности. Требовательность к объему и качеству обучающих данных создает барьер для разработки специализированных систем, тогда как несбалансированность данных приводит к систематической ошибке в распознавании для определенных демографических групп. Вычислительная сложность обучения требует значительных аппаратных ресурсов (мощных GPU/TPU), что ограничивает доступность технологии для небольших организаций и исследователей. Кроме того, уязвимость к адверсарияльным атакам и спуфингу (предъявление фотографий, видео, 3D-масок) представляет серьезную угрозу для безопасности систем, основанных на биометрической аутентификации.

В современной практике методы глубокого обучения, особенно архитектуры на основе ResNet с функциями потерь типа ArcFace, составляют основу промышленных систем распознавания лиц. Они нашли широкое применение в мобильных устройствах (разблокировка по лицу), социальных сетях (автоматическая маркировка пользователей), системах видеонаблюдения (поиск подозреваемых), биометрических паспортах и контроля доступа в критически важные объекты. Текущие исследования фокусируются на повышении устойчивости к атакам, разработке компактных моделей для устройств с ограниченными ресурсами, устранении систематической ошибки и адаптации к экстремальным условиям регистрации (сильные повороты, низкое освещение, наличие медицинских масок).

Несмотря на существующие проблемы интерпретируемости, требовательности к данным и уязвимости к атакам, глубокое обучение остается наиболее перспективным направлением для дальнейшего развития технологий распознавания лиц, обеспечивая баланс между теоретической строгостью и практической применимостью.

Заключение

В ходе исследования проведена систематизация архитектурных решений современных систем распознавания лиц. Показано, что эффективная работа системы требует корректной организации всего вычислительного конвейера – от этапа обнаружения лица с применением современных детекторов на основе сверточных нейронных сетей (MTCNN, RetinaFace) до этапов предварительной обработки, выравнивания и извлечения признаков.

Эволюция методов распознавания лиц прошла путь от традиционных геометрических и статистических подходов (метод «собственных лиц») к методам на основе локальных признаков (LBP, HOG, SIFT, SURF), а затем – к современным архитектурам глубокого обучения. Сверточные нейронные сети, в частности решения на основе архитектур ResNet с функциями потерь ArcFace и подходом FaceNet, обеспечивают беспрецедентную точность извлечения признаков за счёт автоматического формирования иерархических представлений непосредственно из пиксельных данных.

Ключевым выводом работы является взаимосвязь качества всех этапов обработки: даже самые совершенные методы глубокого обучения не способны компенсировать ошибки, возникшие на ранних стадиях (некорректная детекция или выравнивание лица). Поэтому проектирование надёжных систем распознавания требует комплексного подхода к оптимизации всего конвейера обработки, а не только к выбору алгоритма извлечения признаков.

Перспективными направлениями дальнейших исследований в области архитектур и методов представляются следующие: разработка компактных моделей для устройств с ограниченными вычислительными ресурсами; повышение устойчивости к экстремальным ракурсам и условиям освещения; интеграция этапов детекции и выравнивания в единую архитектуру, а также совершенствование функций потерь для формирования более дискриминативных признаков пространств.

Литература

1. Кухарев Г.А., Каменская Е.И., Матвеев Ю.Н., Щеголева Н.А. Методы обработки и распознавания изображений лиц в задачах биометрии. СПб. : Политехника, 2013. 387 с. ISBN 978-5-7325-1028-7.
2. Николенко С.И., Кадурич А.А., Архангельская Е.О. Глубокое обучение: погружение в мир нейронных сетей. СПб. : Питер, 2021. 476 с. ISBN 978-5-4461-1537-2.
3. Охотников К.А., Мартыненко Б.В. Сравнительный анализ методов аутентификации: от традиционных паролей до биометрических решений // Динамика сложных систем. 2025. Т. 19. № 3. С. 52–13. DOI: 10.18127/j19997493-202503-05. EDN YUGYDI.
4. Соколова А.Д., Савченко А.В., Николенко С.И. Поиск редких данных в задаче распознавания лиц на изображениях // Компьютерная оптика. 2022. Т. 46. № 5. С. 801–807. DOI: 10.18287/2412-6179-CO-1061. EDN NIOWXV.
5. Визильтер Ю.В., Горбачевич В.С., Моисеенко А.С. Одноэтапный детектор лиц и особых точек на цифровых изображениях // Компьютерная оптика. 2020. Т. 44. № 4. С. 589–595. DOI: 10.18287/2412-6179-CO-674. EDN USSCTT.

Архитектура систем распознавания лиц:
этапы обработки и классификация методов

6. Кухарев Г.А. Биометрические системы: Методы и средства идентификации человека. СПб. : Политехника, 2001. 239 с. ISBN 5-7325-0623-3.
7. Бобков А.В., Хтет А. Алгоритм распознавания лиц при их частичном закрытии // Автоматизация. Современные технологии. 2024. Т. 78. № 2. С. 77–87. DOI: 10.36652/0869-4931-2024-78-2-77-87. EDN AWQSLN.
8. Turk M., Pentland A. Eigenfaces for Recognition // Journal of Cognitive Neuroscience. 1991. Vol. 3. No. 1. P. 71–86. DOI: 10.1162/jocn.1991.3.1.71.

References

1. Kukharev G.A., Kamenskaya E.I. (2013) *Metody obrabotki i raspoznavaniya izobrazheniy lits v zadachakh biometrii* [Methods of Face Image Processing and Recognition in Biometrics]. St. Petersburg : Politekhnik Publ. 387 p. ISBN 978-5-7325-1028-7. (In Russian).
2. Nikolenko S.I., Kadurin A.A., Arkhangelskaya E.O. (2021) *Glubokoe obuchenie: pogruzhenie v mir neironnykh setey* [Deep Learning: Immersion in the World of Neural Networks]. St. Petersburg : Piter. 476 p. ISBN 978-5-4461-1537-2. (In Russian).
3. Okhotnikov K.A., Martynenkov B.V. (2025) Promising Authentication methods: From traditional passwords to biometric solutions. *Dynamics of complex systems*. Vol. 19. No. 3. Pp. 52–13. DOI: 10.18127/j19997493-202503-05 (In Russian).
4. Sokolova A.D., Savchenko A.V., Nikolenko S.I. (2022) Search for Rare Data in the Face Recognition Task on Images. *Computer Optics*. Vol. 46. No. 5. Pp. 801–807. DOI: 10.18287/2412-6179-CO-1061 (In Russian).
5. Vizilter Yu.V., Gorbatshevich V.S., Moiseenko A.S. (2020) Single-shot face and landmarks detector. *Computer Optics*. Vol. 44. No. 4. Pp. 589–595. DOI: 10.18287/2412-6179-CO-674. (In Russian).
6. Kukharev G.A. (2021) *Biometricheskie sistemy: Metody i sredstva identifikatsii lichnosti chheloveka* [Biometric Systems: Methods and Means of Human Identification]. St. Petersburg : Politekhnik Publ. 239 p. ISBN 5-7325-0623-3. (In Russian).
7. Bobkov A.V., Htet A. (2024) Algorithm for Face Recognition When They Are Partially Closed. *Automation. Modern Technologies*. Vol. 78. No. 2. Pp. 77–87. DOI: 10.36652/0869-4931-2024-78-2-77-87. (In Russian).
8. Turk M., Pentland A. (1991) Eigenfaces for Recognition. *Journal of Cognitive Neuroscience*. Vol. 3. No. 1. Pp. 71–86. DOI: 10.1162/jocn.1991.3.1.71

Поступила в редакцию: 14.01.2026

Received: 14.01.2026

Поступила после рецензирования: 07.02.2026

Revised: 07.02.2026

Принята к публикации: 26.02.2026

Accepted: 26.02.2026