

СРАВНЕНИЕ СОВРЕМЕННЫХ ТЕХНОЛОГИЙ РАСПОЗНАВАНИЯ ОБРАЗОВ

Т. С. Акобян

Сургутский государственный университет, г. Сургут, Российская Федерация
ORCID: <https://orcid.org/0009-0000-6207-2374>, ✉ ackobyantigran2018@yandex.ru

Аннотация: в представленной работе выполнен обзор методов и технологий, применяемых при построении искусственных нейронных сетей. Биометрические технологии, основанные на уникальных физических и поведенческих характеристиках человека, становятся ключевым инструментом для идентификации личности. Настоящее исследование посвящено сравнительному анализу различных моделей распознавания лиц, включая алгоритмы каскадов Хаара, dlib, MTCNN и FaceNet. Проведен анализ по девяти критериям, включая точность, полноту, время обработки и потребление вычислительных ресурсов. Для оценки эффективности алгоритмов использовался набор из 250 изображений, классифицированных по различным условиям. Полученные результаты демонстрируют, что каждый алгоритм обладает своими преимуществами и недостатками в зависимости от конкретных задач и условий эксплуатации.

Ключевые слова: нейронные сети, распознавание образов, сравнение технологий, искусственные нейронные сети, каскады Хаара.

Для цитирования: Акобян Т. С. Сравнение современных технологий распознавания образов. *Успехи кибернетики*. 2025;6(1):175–179.

Поступила в редакцию: 12.02.2025.

В окончательном варианте: 18.03.2025.

MODERN IMAGE RECOGNITION TECHNOLOGIES: A REVIEW

T. S. Akobyan

Surgut State University, Surgut, Russian Federation
ORCID: <https://orcid.org/0009-0000-6207-2374>, ✉ ackobyantigran2018@yandex.ru

Abstract: in this paper, we provided an overview of the methods and technologies used in the construction of artificial neural networks (ANN). Biometric technologies based on a person's unique physical and behavioral characteristics have become a key tool for personal identification. We conducted a comparative analysis of various face recognition models, including Haar cascades, dlib, MTCNN, and FaceNet algorithms. We based the analysis on nine criteria, including accuracy, completeness, processing time, and computing resource consumption. To evaluate the effectiveness of the algorithms, we used a set of 250 images classified according to various conditions. The results demonstrated that each algorithm has its own advantages and disadvantages, depending on the specific tasks and operating conditions.

Keywords: neural networks, pattern recognition, technology comparison, artificial neural networks, Haar cascades.

Cite this article: Akobyan T. S. Modern Image Recognition Technologies: A Review. *Russian Journal of Cybernetics*. 2025;6(1):175–179.

Original article submitted: 12.02.2025.

Revision submitted: 18.03.2025.

Биометрия базируется на свойствах человеческого организма, которые не подвержены изменениям (в отличие от физических паролей, пин-кодов, карт). Биометрические системы используют данные о последовательном поведении человека (речь, походка), его физических параметрах (палец, глаз, лицо), биометрических особенностях (отпечаток пальца, подпись) для идентификации и проверки личности [1].

Распознавание лица — это один из наиболее прогрессивных способов авторизации личности на сегодняшний день. Он позволяет быстро и точно идентифицировать человека, необходимого для выполнения определенных функций, например, распознавание и доступ к контролируемым зонам. Эти технологии стали неотъемлемой частью безопасности в ряде отраслей, включая государственную и корпоративную безопасность, медицинскую и банковскую сферы, аэропорты и другие крупные объекты, где безопасность является первоочередной задачей [2].

Цель исследования состоит в том, чтобы провести сравнительный анализ различных моделей распознавания лиц с целью выявления их преимуществ и недостатков в различных условиях и задачах, а также реализовать новый метод на базе комбинирования лучших технологий. Это поможет как исследователям, так и практикам в выборе наиболее подходящей модели для конкретных приложений.

Для достижения этой цели необходимо:

1. Найти, собрать и проанализировать данные для стартового набора.
2. Выбрать модели и методы.
3. Провести оценки производительности.
4. Проанализировать результат эксперимента.
5. Реализовать на базе лучших технологий комбинированный метод распознавания.

Исследование имеет практическую значимость для разработчиков систем безопасности, развития биометрических технологий, а также для обеспечения приватности и конфиденциальности личных данных.

Одним из основных методов распознавания лиц является биометрическое сопоставление, которое основано на уникальных характеристиках лица человека, таких как расстояние между глазами, форма и размер носа и т. д. Другие методы включают в себя использование глубокого обучения и нейронных сетей для анализа и сопоставления изображений лиц [3].

Для выбора оптимального алгоритма были выполнены тесты с набором данных.

Сам набор представляет из себя 250 изображений, разделенных на 5 групп:

1. Манекены и рисунки из мультфильмов с изображением лиц.
2. Лица людей в масках.
3. Лица людей с разных ракурсов.
4. Белокожие люди.
5. Темнокожие люди.

В каждой группе по 50 изображений и на каждой картинке есть как минимум одно лицо. Все изображения были подготовлены и размечены, то есть лица на изображениях были отмечены человеком — это было сделано для сравнения результата конкретного метода с результатом, который сделал человек.

Алгоритмы будут сравниваться по девяти следующим критериям:

1. Количество найденных лиц (может быть больше, чем лиц на изображении).
2. Количество ненайденных лиц.
3. Количество ложных срабатываний.
4. Точность (Precision) — процент найденных и ненайденных лиц, т.е. полнота работы системы без ложных срабатываний.
5. Полнота (Recall) — процент найденных лиц из общего количества лиц в системе.
6. Пересечение над объединением (IOU) — критерий, предполагающий сравнение области, найденной методом, с тем, что отметил человек на этапе формирования набора данных.
7. Время работы алгоритма — это требуемое время для распознавания лица в изображении.
8. CPU — процент производительности, который требуется от центрального процессора для выполнения задачи.
9. GPU — процент производительности, который требуется от графического процессора для выполнения задачи.

В ходе проведения тестов были получены следующие результаты (см. рис. 1): каскады Хаара распознали лишь треть результатов, возможно, это связано с тем, что находить людей в масках и с разным профилем является тяжелой задачей для данного алгоритма, но нужно учесть, что он является лучшим в производительности алгоритмом. DBFace же, наоборот, оказался лучшим по распознаванию, но худшим в производительности. dlib показал неплохой результат в распознавании, а также хороший результат в производительности. FaceNet показал результат, аналогичный DBFace, т.е. хорошее распознавание, но для данных действий требуется много ресурсов.

Выбрать, какой алгоритм является лучшим, достаточно тяжело, ведь критериев много, и каждый алгоритм показывает как хорошие, так и плохие результаты. В таком случае стоит воспользоваться комбинированным методом принятия решений.

Определим веса критериев для данного метода, для этого нужны оценки экспертов по ранее

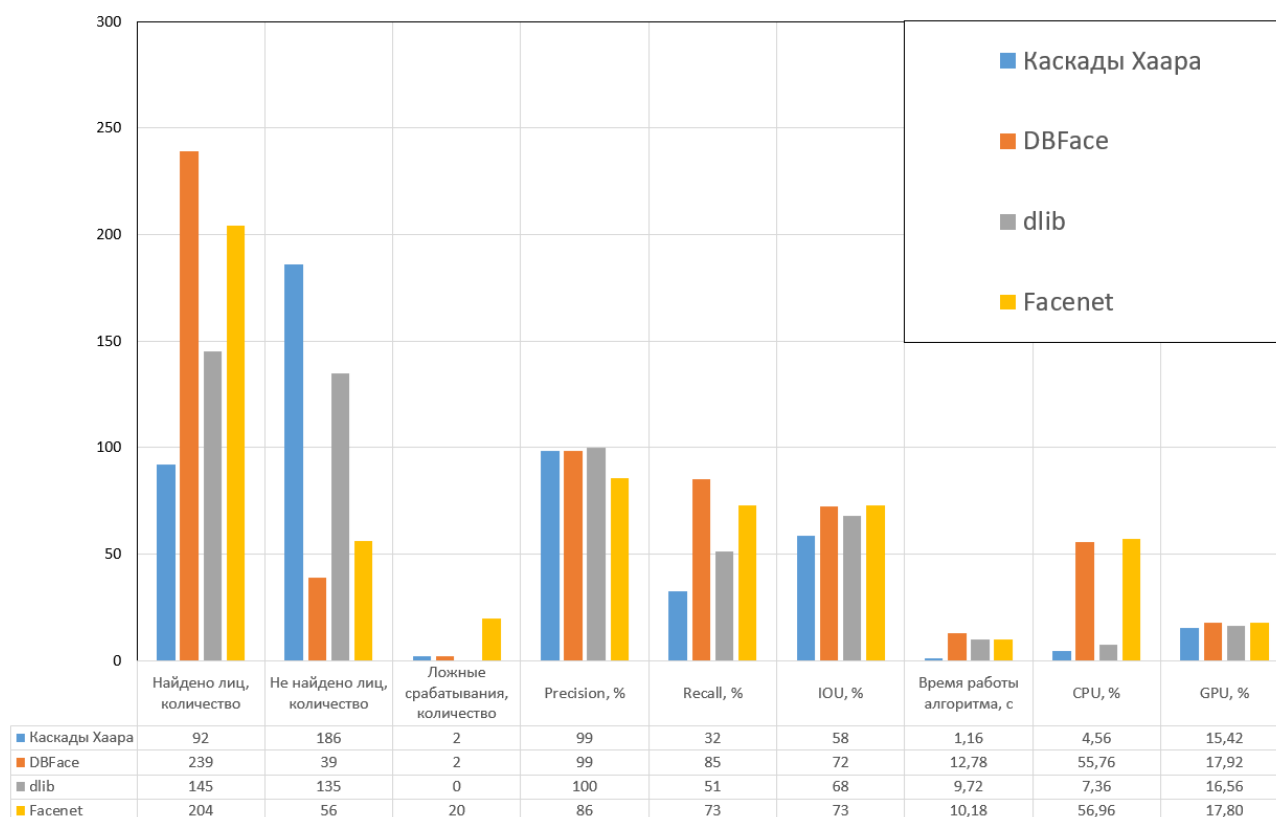


Рис. 1. Результат применения методов после проведения тестов

озвученным 9 критериям, по этому поводу был проведен опрос шести независимых экспертов (см. таблицу 1). Эти оценки были нормализованы, а потом вычислены веса критериев. Далее были определены суммы рангов у всех экспертов, они были равными, что говорило о согласованности между экспертами.

Таблица 1

Оценка независимых экспертов

Эксперт	Критерии									Сумма рангов
	Y1	Y2	Y3	Y4	Y5	Y6	Y7	Y8	Y9	
Э1	6	6	4	4	5	5	2	2	2	36
Э2	5	5	5	3	5	6	2	4	1	36
Э3	7	6	6	1	6	5	3	1	1	36
Э4	5	5	6	5	6	5	1	2	1	36
Э5	4	5	5	6	5	6	2	2	1	36
Э6	7	5	4	5	5	4	4	1	1	36

В комбинированном методе использовалось ограничение по максимуму (для критериев: ложные срабатывания, точность, полнота и время работы алгоритма), а для выбора оптимального алгоритма критерий относительной уступки (для критериев: найдено лиц, не найдено лиц, пересечение над объединением, CPU, GPU). Нормализуем данные с метрик, используя метод нормализации критерием Сэвиджа для минимизации максимально возможных потерь, а потом относительную нормализацию для более точного сравнения и оценки альтернатив. Далее использовали ограничение максимина для исключения решения, которое не защищает от потенциальных неблагоприятных ситуаций. В ходе расчета данным ограничением были исключены алгоритмы DBFace и FaceNet. В конечном итоге, используя критерий относительной уступки, мы определили, что лучшей альтернативой будет являться алгоритм dlib. Хотелось скомбинировать возможности двух лучших методов — dlib и каскады Хаара — для реализации нового гибридного метода распознавания образов.

Обоснование выбора между последовательным и параллельным методами программирования

для распознавания лиц важно по нескольким причинам, поскольку каждый из этих методов имеет свои преимущества и недостатки. Чтобы понять, какой подход будет более эффективным, определим их основные характеристики и сравним.

Последовательный метод:

Преимущества:

1. Последовательная проверка результатов с использованием dlib повышает точность. Если каскад Хаара ошибается, dlib может помочь откорректировать его ошибки.
2. Меньше параллельных вычислений, что снижает вероятность ошибок из-за несовпадений в результатах.

Недостатки: потери по времени из-за двух этапов обработки: сначала каскады Хаара, затем dlib.

Параллельный метод:

Преимущества:

1. Может быть более быстрым, так как оба метода выполняются одновременно, что снижает общее время работы программы.
2. Потенциально более эффективен для обработки нескольких изображений.

Недостатки:

1. Большие требования к ресурсам (память и процессор).
2. Высокая вероятность ошибок при некорректной синхронизации или пересечении результатов двух методов.

Если говорить о математическом обосновании, то можно рассматривать задачи с точки зрения скорости (время обработки) и точности (вероятность ошибки).

Время обработки:

Параллельный метод будет эффективнее с точки зрения времени, если $T_{\text{послед}} > T_{\text{параллель}}$, где:

$$T_{\text{послед}} = T_{\text{Хаара}} + T_{\text{dlib}}$$

$$T_{\text{параллель}} = \max(T_{\text{Хаара}}, T_{\text{dlib}}).$$

Так как оба метода работают одновременно, то время обработки ограничено временем самого долгого метода, а не суммой времени обоих.

Точность:

1. В случае с параллельным методом вероятность ошибки (например, ложные положительные срабатывания) может быть выше, так как два метода могут не всегда точно совпадать в результатах.
2. В случае последовательного метода dlib будет служить фильтром для корректировки результатов, что снижает количество ложных срабатываний [4].

Вывод: так как приоритетом является точность, а не скорость, то будет реализован последовательный метод.

Для распознавания лиц с помощью комбинированного метода каскадов Хаара и библиотеки dlib необходимо использовать два этапа. На первом этапе, с помощью каскадов Хаара, будет происходить быстрое и предварительное распознавание лиц в изображении, а на втором этапе метод из библиотеки dlib уточнит результаты, проверив найденные области.

Шаги решения:

1. Применение каскадов Хаара: каскады Хаара являются быстрым методом для обнаружения объектов на изображении. Это основано на обучении классификаторов, которые с помощью простых признаков (например, разница в яркости пикселей в разных областях изображения) определяют, является ли область изображения лицом. Каскады Хаара используют набор слабых классификаторов, которые комбинируются в более мощный классификатор [5].
2. Использование библиотеки dlib: после того как каскады Хаара распознают лица, мы применяем dlib для уточнения положения лиц. Это происходит путем применения метода поиска лица с помощью модели HOG (Histogram of Oriented Gradients), который дает более точные результаты в случае, когда каскады Хаара могут ошибиться.

Алгоритм:

1. Каскады Хаара: алгоритм Хаара использует признаки, которые вычисляются через суммы значений пикселей в прямоугольных областях. Признаки выглядят как простые «окна» для каждого объекта, которые обучаются на наборе лиц и фона. С помощью классификаторов, обученных на этих признаках, каскад Хаара может быстро искать объекты на изображении.

Если рассматривать это как математическое преобразование, то задача заключается в том, чтобы для каждого окна вычислить сумму пикселей внутри и вокруг него, а затем принять решение, что это лицо (или нет), используя классификатор.

2. Обнаружение лиц с помощью dlib: метод из библиотеки dlib основан на использовании градиентов изображений, которые вычисляются как:

$$G_x = \frac{\partial I}{\partial x}, G_y = \frac{\partial I}{\partial y},$$

где I — это изображение, а G_x и G_y — градиенты изображения по осям x и y соответственно. Эти градиенты характеризуют изменения яркости в разных направлениях и помогают выделить контуры, которые могут быть признаками лиц.

3. Комбинированный метод: идея комбинированного подхода заключается в том, что каскады Хаара дают быстрый, но не всегда точный результат, а метод dlib помогает уточнить положение лиц, уменьшая количество ложных срабатываний.

Заключение

В ходе исследования был проведен сравнительный анализ алгоритмов распознавания лиц: каскадов Хаара, dlib, MTCNN и FaceNet. Эксперименты показали, что точность и полнота выше у нейронных сетей (MTCNN, FaceNet), однако они требуют значительных вычислительных ресурсов. В то же время каскады Хаара демонстрируют высокую скорость работы, но уступают в качестве распознавания в сложных условиях. Алгоритм dlib показал сбалансированные результаты по точности и производительности, что делает его универсальным решением.

Полученные результаты подтверждают, что выбор алгоритма зависит от конкретных задач и условий эксплуатации, а также требований к скорости обработки и точности. Настоящее исследование может быть полезно разработчикам систем безопасности и биометрических технологий при выборе оптимального решения.

ЛИТЕРАТУРА

1. Частикова В. А., Титова А. А., Войлова Д. О. Аналитический обзор методов идентификации личности на основе биометрических характеристик. *Вестник Адыгейского государственного университета. Серия: Естественно-математические и технические науки*. 2022;1(296):92–112. DOI: 10.53598/2410-3225-2022-1-296-92-112.
2. Ковшов Е. Е., Завистовская Т. А. Разработка автоматизированной системы контроля и управления доступом на основе анализа динамики носогубных мышц лица человека. *Современные проблемы науки и образования*. 2013;2:185.
3. Иванова Е. В., Струева А. Ю. Система учета посещаемости студентов на основе распознавания лиц. *Вестник Южно-Уральского государственного университета. Серия: Вычислительная математика и информатика*. 2021;4:60–73. DOI: 10.14529/cmse210404.
4. Калинин Н. Н., Скурыдин Ю. Г., Скурыдина Е. М. Оценка производительности моделей машинного обучения, основанных на библиотеках tensorflow и dlib при определении состояния сонливости на Raspberry Pi. *Высокопроизводительные вычислительные системы и технологии*. 2023;1:106–114.
5. Хорошева А. А., Федько М. Д., Юсупов А. Р. Применение каскада Хаара для разработки системы распознавания птиц. *Молодой исследователь Дона*. 2020;2(23):166–170.