**Аннотация.** В данной статье рассматриваются современные методы детекции лиц, основанные на различных алгоритмах и архитектурах нейронных сетей. Описаны популярные детекторы, включая классический метод каскадных классификаторов Хаара, реализованный в OpenCV, а также современные нейросетевые подходы: SSD, MTCNN, RetinaFace, YUNet и Dlib. Рассмотрены их особенности, принципы работы и области применения. Особое внимание уделено эффективности и точности обнаружения лиц в сложных условиях, таких как различные цветовые искажения, изменяющееся освещение, окклюзия. В статье приведено сравнение представленных методов по скорости работы, вычислительным требованиям и качеству детекции. Проведён анализ преимуществ и недостатков различных подходов, что позволяет определить наиболее подходящий метод в зависимости от требований к точности, скорости и вычислительным ресурсам.

**Введение**

Распознавание лица по видео и фотоизображению является одной из наиболее значимых задач в области компьютерного зрения и искусственного интеллекта. Она охватывает широкий спектр применений – от систем безопасности и биометрической аутентификации до улучшения взаимодействия человека с машинами и создания умных устройств. С развитием технологий машинного обучения, особенно нейросетевых моделей, распознавание лиц достигло значительных успехов и стало частью повседневной жизни [1].

Одним из главных достижений в области распознавания лиц стали глубокие нейронные сети, которые позволяют достигать высокой точности в решении этой задачи. Модели на основе сверточных нейронных сетей (англ. Convolutional Neural Network, CNN) продемонстрировали впечатляющие результаты в различных условиях: освещенности и даже частичной маскировке лица [2]. Применение таких моделей открывает новые возможности в области идентификации людей, повышения безопасности и удобства пользовательского опыта.

Однако, несмотря на значительные успехи, задача распознавания лиц остаётся актуальной и требует дальнейших исследований. Проблемы, связанные с различиями во внешности людей, изменяющимися условиями среды, а также с необходимостью обработки большого объёма данных, требуют разработки более эффективных и адаптивных методов. В своей работе [3] авторы обсуждают важность адаптивных методов для улучшения устойчивости детекторов лиц. Таким образом, данная работа направлена на всестороннее исследование нейросетевых подходов к распознаванию лиц, анализ их преимуществ и недостатков, а также перспектив применения в реальных системах.

**Задача детекторов лиц**

Процесс распознавания лиц включает несколько ключевых этапов. На первом этапе происходит получение изображения или видеокадра с камеры. Далее выполняется предварительная обработка, включающая коррекцию освещённости, шумоподавление и нормализацию изображения для приведения его к единому формату. После этого применяется детектор лиц, задача которого – обнаружение области, содержащей лицо (рис. 1).

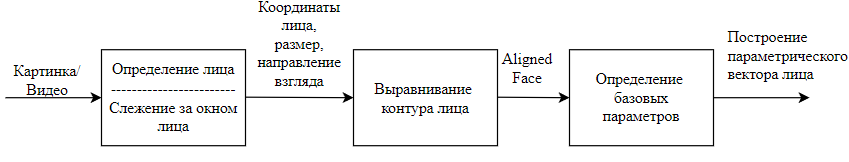


Рисунок 1 – Общая схема механизма работы детектора

В статье [4] обсуждается, как современные методы детекции могут быть адаптированы для работы в различных условиях. Подробнее о механизмах работы и архитектурах детекторов рассмотрено в следующем разделе, включая использование сверточных нейронных сетей, которые показывают высокую точность в детекции лиц [2], а также более сложные подходы, такие как метод многозадачной каскадной сети, о котором говорится в статье [5].

Как следует из вышесказанного, в контексте построения систем видео-распознавания личности особое внимание стоит уделять алгоритмам детекции, в частности, их производительности и необходимости в вычислительных ресурсах. Так как различные системы, оснащенные функцией распознавания лиц, имеют различные технические и вычислительные характеристики, которые могут оказать влияние на результат работы. Так же не мало важно учитывать качество входной информации, в том числе случае, когда представлено видео или изображение низкого качества или на лице имеются перекрывающие его элементы (маска, очки), что препятствует получению точного результата. Для этих целей был проведен предложенный анализ систем распознавания личности на основе разных детекторов и при наличие разной входной информации.

**Исследуемые детекторы**

Для исследования в рамках данной работы были выбраны шесть наиболее популярных детекторов:

* OpenCV;
* SSD;
* MTCNN;
* RetinaFace;
* YUNet;
* Dlib.

OpenCV (англ. Open Source Computer Vision Library – рус. библиотека компьютерного зрения с открытым исходным кодом) – это широко используемая библиотека для компьютерного зрения, которая предоставляет множество инструментов для обработки изображений и видео, включая детекторы лиц. В OpenCV классическим методом для детекции лиц является использование каскадных классификаторов Хаара, основанных на алгоритме AdaBoost. Этот метод в первую очередь применяется для распознавания лиц в статических изображениях. В статье [6] рассматривается применение каскадных классификаторов Хаара для детекции лиц, а в статье [4] приводится анализ их эффективности в сравнении с более современными методами, такими как нейронные сети.

Поиск лица на изображении в OpenCV реализован [методом Виолы-Джонса](http://habrahabr.ru/post/133826/). Метод подставляет к изображению так называемые примитивы Хаара, которые являются набором элементарных сочетаний тёмных и светлых областей, и если на изображении найдена область, в которой подходит достаточное количество примитивов, то объект найден. Для того чтобы понять, светлая эта область или тёмная, необходимо просуммировать значение соседних пикселей. Для того чтобы не делать это много раз в процессе поиска объекта, изображение переводится в [интегральное представление](http://robocraft.ru/blog/computervision/536.html).

Ключевые характеристики OpenCV:

1. Каскадные классификаторы Хаара основаны на обучении слабых классификаторов для выделения признаков (например, контуры и текстуры) лиц.

2.Преимущества – высокоскоростная обработка и относительно низкие вычислительные затраты.

3. Недостатки – чувствительность к условиям освещения и положению лица (например, наклон головы).

SSD (англ. Single Shot MultiBox Detector – рус. мультибоксовый детектор с одиночным выстрелом) – это современная архитектура для обнаружения объектов, использующая сверточные нейронные сети. SSD предназначен для одновременного обнаружения множества объектов на изображении, включая лица, с высокой точностью и эффективностью. В отличие от традиционных методов, SSD работает в реальном времени и подходит для обработки видео. В статье [1] обсуждается, как SSD используется для обнаружения лиц в реальном времени, а также приводится сравнительный анализ его производительности с другими методами, такими как YOLO и Faster R-CNN.

На начальном этапе входное изображение проходит через несколько сверточных слоев, которые постепенно уменьшают его размер, при этом извлекая важные характеристики. Эти слои создают карты признаков, которые содержат информацию о различных аспектах изображения, таких как контуры, текстуры и цветовые паттерны. На разных уровнях карты признаков нейросеть генерирует различные объектные ящики (англ. bounding boxes) и предсказывает классы для объектов внутри этих ящиков. Таким образом крупные объекты могут быть распознаны на более низких уровнях сети, а мелкие – на более высоких.

Каждый объектный ящик имеет свои координаты, а также вероятность принадлежности к определённому классу. После этого происходит этап постобработки, на котором применяется алгоритм нелокального подавления (англ. Non-Maximum Suppression, NMS) для удаления дублирующихся предсказаний и выбора наиболее уверенных результатов.

Ключевые характеристики SSD:

1. Многоуровневое предсказание – используется несколько уровней сетки для обнаружения объектов на разных масштабах.

2. Преимущества – высокая производительность, подходит для реального времени, точность на уровне современных систем.

3. Недостатки – сложность в настройке и необходимость большого объёма данных для обучения.

MTCNN (англ. Multi-task Cascaded Convolutional Networks – рус. многозадачные каскадные сверточные сети) – это сеть, которая решает несколько задач одновременно: детектирование лиц, выравнивание лиц и обнаружение ключевых точек (например, глаз, носа, рта). Эта архитектура применяет каскадные сверточные сети, что позволяет эффективно обрабатывать сложные условия, такие как различные углы наклона или частичная маскировка. В статье [5] подробно рассматриваются особенности работы MTCNN для детекции и выравнивания лиц, а также оценка его эффективности в различных сценариях, включая влияние углов наклона и маскировки.

На первом этапе сеть анализирует изображение и создает карту вероятностей, указывая на возможные местоположения лиц. Этот этап выполняется с помощью легкой сверточной сети, которая быстро обрабатывает изображение, выделяя области, где могут находиться лица. Затем MTCNN применяет более глубокую сверточную сеть для уточнения координат найденных лиц. Здесь происходит более детальная обработка, которая позволяет точно определить границы лицевых областей и улучшить качество обнаружения. На финальном этапе сеть определяет ключевые точки на лице, такие как глаза, нос и уголки рта. Это важно для дальнейших задач, таких как выравнивание лиц или распознавание эмоций.

Ключевые характеристики MTCNN:

1. Многозадачность – помимо нахождения лиц, MTCNN также выполняет выравнивание лиц и определяет ключевые точки.

2. Многокаскадная структура – использует три этапа (прежде чем перейти к следующему каскаду, необходимо пройти предыдущий).

3. Преимущества – высокая точность при разнообразных углах наклона, эффективен в условиях ограниченных вычислительных ресурсов.

4. Недостатки – требует относительно большего времени для обработки по сравнению с другими детекторами.

RetinaFace (рус. сетчатка лица) – это современная архитектура для детектирования, которая использует глубокие сверточные нейронные сети для обнаружения лица с высокой точностью и детализированием ключевых точек. RetinaFace работает с изображениями в реальном времени и эффективно обрабатывает лица в сложных условиях, таких как плохое освещение и частичные маскировки. В статье [6] описывается эффективность RetinaFace в реальных условиях, включая применение в различных сценариях с низким качеством освещения и на изображениях с частичными маскировками, а также его возможности для точного выявления ключевых точек.

На каждом уровне сети генерируются объектные ящики, которые предсказывают возможные местоположения лиц. Для каждого ящика также вычисляется вероятность принадлежности к классу «лицо», что позволяет точно определять, присутствует ли лицо в заданной области. Для повышения точности предсказаний применяется алгоритм нелокального подавления (NMS), который убирает дублирующиеся предсказания и оставляет только наиболее уверенные результаты.

Ключевые характеристики RetinaFace:

1. Глубокие нейронные сети – использует остаточные сети (ResNet) для обучения признаков.

2. Обнаружение ключевых точек – наряду с определением лиц, модель также определяет координаты ключевых точек, таких как глаза, нос и рот.

3. Преимущества – высокая точность при различных условиях, включая углы наклона и освещенность.

4. Недостатки – требует значительных вычислительных ресурсов для обучения и инференса.

YUNet – это быстрое и точное решение для детектирования лиц, которое стало популярным благодаря своей производительности и способности работать на мобильных устройствах с ограниченными вычислительными мощностями. YUNet использует эффективные нейросетевые структуры (архитектура представлена на рисунке 6), что позволяет достичь хороших результатов при ограниченных ресурсах. В статье [7] обсуждаются преимущества YUNet в контексте мобильных устройств, где ограниченные вычислительные ресурсы требуют оптимизации моделей, что делает эту архитектуру эффективным выбором для приложений в реальном времени на устройствах с ограниченными мощностями.

YUNet использует каскадный подход, который позволяет обрабатывать изображения на нескольких уровнях. Это означает, что на каждом этапе анализа сеть уточняет и улучшает результаты, обеспечивая более глубокое понимание структуры лиц на изображении. Применяются многоуровневые сверточные слои для извлечения признаков на различных масштабах. Сеть также определяет ключевые точки на лицах, что позволяет улучшить качество распознавания

Ключевые характеристики YUNet:

1. Высокая производительность – быстрое детектирование лиц, особенно на устройствах с низкими вычислительными мощностями.

2. Простота использования – хорошо работает в реальных условиях, таких как мобильные приложения.

3. Преимущества – быстрое детектирование, хорошо работает в реальном времени.

4. Недостатки – может уступать в точности на сложных изображениях.

Dlib – это известная библиотека для машинного обучения, которая включает высококачественные инструменты для обработки изображений и видео. Один из основных методов для детектирования лиц в Dlib – это использование каскадных классификаторов Хаара, а также алгоритмов на основе поэтапного обнаружения ключевых точек лица, что делает её удобной для более точного выравнивания лиц и анализа выражений. После обнаружения лиц Dlib применяет технику, известную как «выравнивание лиц», которая позволяет стандартизировать положение и размер лиц в изображении. Затем используется CNN для извлечения векторных признаков, которые представляют собой уникальные характеристики каждого лица. Эти векторы позволяют сравнивать и идентифицировать лица с высокой степенью точности.

В статье [8] подробно описывается использование Dlib для задач распознавания лиц, где акцент сделан на точности обнаружения и выравнивания лиц с использованием каскадных классификаторов и алгоритмов для детектирования ключевых точек.

Ключевые характеристики Dlib:

1. Высокая точность – особенно полезна для выравнивания лиц и обнаружения ключевых точек.

2. Модели для распознавания – предоставляет как детекторы для нахождения лиц, так и алгоритмы для анализа выражений и поз.

3. Преимущества – хорошая точность и гибкость.

4. Недостатки – относительная медленность по сравнению с более современными моделями.

**Эксперименты**

В рамках данного исследования рассматриваются два этапа тестирования детекторов лиц.

Первый этап призван оценить производительности детекторов, затрачиваемых на распознавание потока лиц. В качестве входных данных для каждой сети использовался набор из 300 последовательных изображениях различных людей, взятый из открытого датасета с сайта www.kaggle.com.

В ходе этого блока эксперимента анализировались следующие показатели:

1. Процент правильно предсказанных лиц из 300 изображений. Этот показатель оценивает, насколько точно детектор идентифицирует лица без ложных срабатываний и пропусков.

2. Время работы алгоритма. Замерялось время выполнения детекции для всех изображений, что позволяет оценить эффективность алгоритма в режиме реального времени и переделить среднее время работы каждого из исследуемых детекторов.

3. Степень загрузки CPU (англ. Central Processing Unit – рус. центральный процессов) во время работы детектора. Этот параметр критически важен при выборе детектора для задач, требующих высокой производительности на ограниченных вычислительных ресурсах.

Второй этап эксперимента – оценка устойчивости детекторов к искажениям. Проводилась проверка качества работы детекторов на изображениях, подвергнутых различным видам искажений. Это позволяло сравнить их устойчивость к реальным условиям, таким как:

1. Черно-белый фильтр (англ. grayscale conversion) – проверка работы детекторов в условиях отсутствия цветовой информации. Некоторые модели могут зависеть от цветовых признаков, поэтому важно оценить их способность обнаруживать лица в черно-белых изображениях, таких как снимки с камер видеонаблюдения.

2. Изменение контрастности и освещённости – проверка детектора на способность работать при слишком ярком освещении. В реальных условиях лица могут быть освещены неравномерно, поэтому важно оценить устойчивость алгоритмов к этим факторам.

3. Размытие (англ. gaussian blur) – имитация движения или плохого фокуса камеры. Детекторы должны уметь находить лица даже при нечетких изображениях, например, при съемке в движении.

4. Шум (noise: salt & pepper, gaussian noise) – добавление случайных пиксельных артефактов, моделирующих помехи при передаче или сохранении изображения. Этот тест оценивает способность детекторов работать с изображениями низкого качества.

5. Частичная окклюзия (заслонение части лица) – тестирование детекторов в условиях, когда лицо частично закрыто (например, очками, маской, руками, капюшоном). Такой сценарий актуален для реальных ситуаций, например, в системах безопасности.

Для проверки каждого из перечисленных критериев был подготовлен набор исходных данные, предварительно обработанных, в соответствие с исследуемым параметром (рис.2).



а б в г д е ж

Рисунок 2 – Пример исходных данных для второго этапа эксперимента:

а – без изменений, б – черно-белый фильтр, в – фильтр повышенной яркости, г – фильтр Гауссово шума, д – окклюзия (маска), е – окклюзия (очки), ж – окклюзия (маска и очки)

В качестве устройства проведения исследования использовался персональный компьютер с GPU (NVIDIA GeForce RTX 3050) и CPU (11th Gen Intel(R) Core(TM) i5-11400F @ 2.60GHz).

**Результаты первого блока эксперимента**

Точность детекции лиц – показывает процент правильно обнаруженных лиц из 300 потоковых изображений (рис. 3).

Рисунок 3 – Т**очность детекции лиц**

Анализируя результаты детекторов, можно отметить, что RetinaFace показал наилучшие показатели с 282 положительными и всего 18 отрицательными срабатываниями, что свидетельствует о его высокой точности. MTCNN (266/34) и Dlib (218/82) также продемонстрировали хорошие результаты, обеспечивая баланс между успешными детекциями и количеством ошибок. В то же время SSD (123/177) и Yunet (160/140) показали наибольшее число ложных отрицаний, что делает их менее надежными. В целом, RetinaFace и MTCNN можно считать наиболее точными и стабильными для задач детекции лиц.

Время обработки изображений – показывает среднее время обработки потока из 300 изображений каждого алгоритма детекции (рис. 4).

**Рисунок 4 – Среднее время обработки изображений**

Анализируя среднее время работы, можно отметить, что SSD (100 мс), Dlib (105 мс) и Yunet (110 мс) являются самыми быстрыми, что делает их подходящими для задач, требующих высокой скорости обработки. OpenCV (175 мс) также демонстрирует относительно быстрое время работы, однако MTCNN (501 мс) и RetinaFace (818 мс) значительно уступают по скорости, что может быть критично для реального времени, несмотря на их высокую точность. Таким образом, выбор детектора зависит от компромисса между точностью и производительностью.

Загрузка процессора (рис. 5) – показывает среднюю загрузку CPU во время работы детекторов.

**Рисунок 5 – Загрузка процессора**

Анализ средней загрузки CPU показывает, что RetinaFace (100%) является самым ресурсоемким детектором, что может негативно сказаться на работе системы в условиях ограниченных вычислительных мощностей. OpenCV (70%) также потребляет значительный объем ресурсов, в то время как SSD (50%) и MTCNN (50%) демонстрируют более сбалансированную нагрузку. Yunet (40%) и Dlib (20%) являются наименее требовательными к процессору, что делает их хорошими кандидатами для работы на слабых устройствах или при многозадачности. Таким образом, выбор детектора зависит от доступных вычислительных ресурсов и требований к скорости работы.

**Результаты второго блока эксперимента**

Первостепенно была осуществлена проверка детекции изображения без каких-либо искажений (рис. 2а), что позволило провести базовую и стабильную проверку их работы (рис. 6).

**Рисунок 6 –** **Детекция изображения без искажения**

Анализ показателей точности относительно детекции изображения без искажения, отражающих степень неразличимости между детекцией и фоновыми объектами, свидетельствует, что SSD, RetinaFace и Dlib имеют наивысший показатель (100%), т.е. имеют высокую уверенность алгоритма в принятии решений. MTCNN немного уступает (99%), но остается на высоком уровне. OpenCV и Yunet (94%) демонстрируют наименьший показатель, что может указывать на чуть более низкую точность в отделении лиц от фона. В целом, разница между методами минимальна, но SSD, RetinaFace и Dlib могут быть предпочтительнее в задачах, где важна уверенность модели.

При применении черно-белого (ЧБ) фильтра (рис. 2б) точность детекции остается высокой для большинства алгоритмов (рис. 7). SSD, MTCNN, RetinaFace и Dlib показывают максимальную точность (100%), что свидетельствует об их устойчивости к изменению цветовой информации. OpenCV и Yunet демонстрируют небольшое снижение точности (94% и 93% соответственно), что может указывать на их зависимость от цветовых характеристик изображения. В целом, фильтр ЧБ практически не влияет на работу детекторов, за исключением незначительных колебаний для отдельных моделей.

**Рисунок 7 – Детекция черно-белого изображения**

Фильтр повышенной яркости (рис. 2в) на изображении создает яркие засветы, которые могут усложнить задачу распознавания лиц. Применение такого фильтра к изображениям для различных детекторов лиц приводит к ухудшению их эффективности, так как яркие области на изображении могут закрывать или искажать важные черты лица, затрудняя точную детекцию. Несмотря на это, все детекторы показали высокую точность, варьирующуюся от 91% до 100%, что указывает на то, что фильтр повышенной яркости воздействует на производительность, но не вызывает значительных потерь в точности распознавания (рис. 8).

**Рисунок 8 – Детекция изображения с фильтром повышенной яркости**

Фильтр Гауссова шума (рис. 2г) добавляет случайные искажения на изображение, что может значительно ухудшить качество распознавания лиц, снижая четкость и контрастность важных признаков. Применение этого фильтра к изображениям для различных детекторов лиц (OpenCV, SSD, MTCNN, RetinaFace, Yunet, Dlib) приводит к значительным ухудшениям в точности распознавания, особенно для SSD, MTCNN и RetinaFace, где точность равна 0% (рис. 9). Однако OpenCV и Dlib показывают более высокие результаты (95% и 100% соответственно), что указывает на разную степень устойчивости детекторов к добавленному шуму, но в целом фильтр существенно снижает эффективность распознавания.

**Рисунок 9 – Детекция изображения с фильтром Гауссово шума**

При применении окклюзия (маска на лице) изображение (рис. 2д) усложняется за счет присутствия инородного объекта, который может скрывать часть лицевых признаков, что затрудняет распознавание. В результате применения маски для различных детекторов лиц наблюдается незначительное снижение точности распознавания (рис. 10). Несмотря на это, все детекторы остаются достаточно точными, с результатами в пределах от 90% до 100%, что свидетельствует о высокой устойчивости большинства моделей к наличию маски на лице. Однако, маска все же создает дополнительные сложности для распознавания, особенно для Yunet, где точность ниже по сравнению с другими детекторами.

**Рисунок 10 – Детекция изображения** **с окклюзией (маска)**

Наличие очков на лице (рис. 2е) усложняет задачу распознавания, так как они могут закрывать часть глаз и других важных черт лица, что влияет на точность детекции. В результате наблюдается понижение точности детекции (рис. 11). Детекторы, такие как SSD, MTCNN,RetinaFace и YUnet, показывают отличные результаты с точностью 100%, в то время как OpenCV имеет немного более низкий результат (96%). При этом Dlib сталкивается с серьезными трудностями в распознавании, показав точность 0%, что указывает на низкую устойчивость этой модели к наличию очков на лице.

**Рисунок 11 – Детекция изображения с окклюзией (очки)**

Когда на изображении присутствуют одновременно очки и маска (рис. 2ж), это сильно усложняет задачу распознавания лиц, поскольку оба предмета могут закрывать наибольшее количество важных частей лица – глаза, нос и рот. Результаты для различных детекторов лиц показывают резкое ухудшение точности (рис. 12). Только детектор RetinaFace демонстрирует отличную точность (70%), в то время как SSD достигает 56%. Остальные детекторы, такие как OpenCV, MTCNN, Yunet и Dlib, не смогли успешно распознать лица, показывая результаты 0%. Это указывает на значительную сложность распознавания при наличии этих двух предметов одновременно.

**Рисунок 12 – Детекция изображения с окклюзией (маска и очки)**

**Заключение**

Результаты тестирования показывают, что выбор детектора лиц зависит от требований к точности, скорости обработки и загрузке вычислительных ресурсов. RetinaFace (282 из 300) и MTCNN (266 из 300) продемонстрировали наивысшую точность, но их скорость оставляет желать лучшего — 818 мс и 501 мс соответственно, а RetinaFace также потребляет 100% CPU, что делает его требовательным к вычислительным ресурсам. Dlib (218 из 300) и OpenCV (200 из 300) показали сбалансированные результаты, предлагая умеренную точность и приемлемую скорость (105 мс и 175 мс), при этом Dlib имеет минимальную загрузку процессора (20%), что делает его подходящим для встраиваемых решений. SSD (123 из 300) и YuNet (160 из 300) работают быстрее всех (100–110 мс) и требуют меньше ресурсов (SSD — 50%, YuNet — 40%), но их точность оказалась наихудшей, что ограничивает их применение в сложных условиях.

Дополнительный анализ устойчивости детекторов к различным искажениям показал, что:

1. При обработке изображений черно-белым фильтром, повышением яркости или добавлением маски точность детекторов практически не изменилась, за исключением YuNet, который показал небольшое снижение устойчивости.

2. В условиях значительного шумового загрязнения (Гауссов шум) эффективность большинства детекторов резко снизилась. Например, SSD, MTCNN, RetinaFace и YuNet полностью провалили тест, тогда как OpenCV и Dlib продолжили демонстрировать приемлемую устойчивость.

3. При наличии очков на лице точность RetinaFace осталась максимальной, а большинство других детекторов показали значительное снижение эффективности.

4. При сочетании маски и очков все детекторы, кроме RetinaFace, полностью потеряли способность обнаруживать лица.

Таким образом, RetinaFace и MTCNN являются лучшими вариантами для задач, требующих высокой точности, но их высокая вычислительная сложность делает их менее применимыми для работы в реальном времени. Dlib и YuNet показали сбалансированные результаты, особенно в условиях ограниченных ресурсов. SSD и YuNet продемонстрировали высокую скорость, но низкую точность, что делает их подходящими для потоковой обработки видео, где критична скорость работы. Выбор детектора должен основываться на специфике задачи: для высокоточных решений лучше использовать RetinaFace и MTCNN, для встраиваемых систем предпочтительнее Dlib, а для быстрой обработки видео SSD и YuNet.

Исследование подтвердило, что нейронные сети являются мощным инструментом для обработки изображений и анализа данных, однако их эффективность зависит от баланса между точностью, скоростью обработки и потреблением вычислительных ресурсов.

**Список источников**

1. Hu, W., & Tan, Y. "Fast Face Detection Using Multi-scale Convolutional Neural Networks." IEEE Transactions on Image Processing, 2016.
2. Ren, S., He, K., & Sun, J. "Region-based Convolutional Networks for Accurate Object Detection and Segmentation." IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2015.
3. Liao, S., Zhu, X., & Lei, Z. "A Fast and Accurate Face Detector for Real-Time Applications." IEEE Transactions on Image Processing, 2015.
4. Sun, Y., Wang, X., & Tang, X. "Deep Learning Face Representation by Joint Identification-Verification." IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2014.
5. Zhang, Z., & Liu, Y. "Face Detection in the Wild with Deep Convolutional Neural Networks." European Conference on Computer Vision (ECCV), 2016.
6. Liu, X., Han, H., Jain, A. K., & Tong, S. "Deep Sparse Representation for Robust Face Recognition." IEEE Transactions on Image Processing, 2016.
7. Xie, Z., & Zhang, Z. "Deep Learning-Based Face Detection: A Survey." Journal of Visual Communication and Image Representation, 2017
8. Yu, C., & Liu, W. "Face Detection with Deep Convolutional Neural Networks." Proceedings of the IEEE International Conference on Computer Vision (ICCV), 2015.