Министерство науки и высшего образования РФ

Федеральное государственное бюджетное образовательное учреждение высшего образования

«Рыбинский государственный авиационный технический университет

имени П.А. Соловьева»

Институт информационных технологий и систем управления

Кафедра Математическое и программное обеспечение ЭВС

ОТЧЕТ

по учебной, ознакомительной практике

РАЗРАБОТКА И ИССЛЕДОВАНИЕ ЭЛЕМЕНТОВ СИСТЕМЫ РАСПОЗНАВАНИЯ ОБРАЗОВ В ПОТОКОВОМ ВИДЕО

по профилю – Разработка программно-информационных систем

по направлению – 09.04.04 Программная инженерия

Соискатель, студент группы ПИМ-23

*(Код) (Подпись, дата) (Фамилия И.О.)*

Руководитель к.т.н., доцент

*(Уч. степень, звание) (Подпись, дата) (Фамилия И.О.)*

##### Рыбинск 2024

Оглавление

[Введение 3](#_Toc171602375)

[Актуальность темы. 3](#_Toc171602376)

[Характеристика проблемы. 3](#_Toc171602377)

[Объект исследования. 3](#_Toc171602378)

[Предмет исследования. 3](#_Toc171602379)

[Цель исследования. 3](#_Toc171602380)

[Задачи исследования. 3](#_Toc171602381)

[1 Аналитический обзор методов распознавания лиц в видеопотоке 4](#_Toc171602382)

[1.1.1 Классификация по ближайшему среднему значению 7](#_Toc171602383)

[1.1.2 Классификация по расстоянию до ближайшего соседа 7](#_Toc171602384)

[1.1.3 Структурный или синтаксический 8](#_Toc171602385)

[1.1.4 Байесовский подход к принятию решений 8](#_Toc171602386)

[1.1.5 Нейронные сети 8](#_Toc171602387)

[1.1 Методы и алгоритмы распознавания лиц 10](#_Toc171602388)

[1.2.1 Алгоритм Виолы - Джонсона 12](#_Toc171602389)

[1.2.2 Метод главных компонентов 14](#_Toc171602390)

[1.2.3 3D Распознавание лиц 15](#_Toc171602391)

[1.2 Особенности распознавания лиц в видеопотоке 15](#_Toc171602392)

[1.3.1 Проблема освещенности 16](#_Toc171602393)

[1.3.2 Проблема изменение лица 17](#_Toc171602394)

[1.3.3 Обучение нейронной сети 17](#_Toc171602395)

[1.3.4 Распознавание по изображениям низкого качества 18](#_Toc171602399)

[1.3 Анализ систем и методов алгоритмов распознавания 19](#_Toc171602400)

[1.4.1 Статические методы распознавания 19](#_Toc171602401)

[1.4.2 Достоинства и недостатки основных методов распознавания 20](#_Toc171602402)

[1.4.3 Системы распознавания 22](#_Toc171602403)

[Заключение 23](#_Toc171602404)

[Список использованных источников 24](#_Toc171602405)

Введение

Актуальность темы.

Характеристика проблемы.

Объект исследования.

Предмет исследования.

Цель исследования.

Задачи исследования.

1 Аналитический обзор методов распознавания лиц в видеопотоке

В настоящее время рост распознавания лиц обусловлен параллельными достижениями в области обработки компьютерного зрения [1]. Это далеко идущая технология. Стремительно увеличивается количество и качество камер видеонаблюдения.

Несмотря на значительные успехи в разработке и реализации систем для распознавания образов в потоковом видео на данный момент нет ни одной «идеальной» системы. Это связано с постоянным развитием математического аппарата, аппаратного обеспечения и со сложностью программной реализации подобных систем. Появление новых камер и увеличение емкости систем хранения данных открывает возможность использования принципиально новых алгоритмов обработки потокового видео. Постоянно развиваются программные библиотеки, содержащие низкоуровневые подпрограммы для обработки видео.

Особым случаем является задача распознавания образов в потоковом видео. В рамках реализации программной системы распознавания образов в потоковом видео необходимо решать, как научные (сравнение различных алгоритмов распознавания образов в зависимости от качества изображения, его разрешения, степени сжатия и т. д.), так и крайне сложные инженерные задачи (эффективная реализация алгоритмов распознавания, хранение и обработка больших объемов данных). Решение задачи построения элементов программной системы распознавания образов в потоковом видео имеет как практическое, так и значительное теоретическое значение.

1.1 Методы и алгоритмы распознавания объектов в видеопотоке

Несмотря на большой прогресс в исследованиях компьютерного зрения, идентификация лица все еще создает много проблем из-за некоторых типичных проблем, таких как выражения лица, изменения позы, изменения освещения, частичные затенения, старение и изменение внешнего вида. На данный момент существует множество методов распознавания образов. Необходимо выбрать метод или группу методов, которые предоставят наилучшие показатели.

В диссертационной работе [2] предлагается, разделение методов выделения объекта на изображении в видеопотоке на группы:

1. Комбинированные методы;
2. Вероятностные методы;
3. Детерминированные методы;
4. Нейросетевые методы;

Было выявлено, что в группе детерминированных методов основанных на поиске точек объекта являются боле устойчивыми к искажениям. Эти методы инвариантны к проективным преобразованиям. Они более устойчивы к шумовым искажениям изображения и изменению яркости света.

По результатам работы были представлены ограничения к целевому объекту для его поиска на изображении.

Таблица 1.1 – Ограничения к целевому объекту для поиска

|  |  |
| --- | --- |
| Ограничения | Значение |
| Минимальный размер изображения объекта | 20 х 20 точек |
| Максимальное изменение (увеличение/уменьшение) размера изображения объекта в кадре | До 20 раз |
| Цветовая характеристика | Цветовые гистограммы изображений образца и объекта в кадре должны соответствовать по мере Бхаттачария > 80% |
| Изменение яркости | Значение детерминанта матрицы Гессе для ключевой точки должно быть не меньше 300 |

В статье [3] представлен эффективный подход для распознавания лиц и сходства изображений. Подход основан на теоретико-информационной мере подобия, полученной с использованием энтропии одномерной версии двухмерной совместной гистограммы между двумя изображениями. Были использованы две энтропии, Шеннон и Реньи, что привело к двум показателям: сходство Шеннона-гистограммы (SHS) и мера сходства Реньи (RSM). Работоспособность RSM и SHS была проверена на основе эффективного существующего сходства на основе признаков метрик сходства (FSIM), структурного сходства (SSIM). Сравнение с недавним теоретико-информационным ISSIM также рассматривалось. Экспериментальные результаты показали превосходную производительность для предложенных мер с точки зрения правильных решений с минимальной путаницей в распознавании лиц и сходстве изображений, используя базы данных AT & T, FEI, TID2008 и базы данных изображений IVC. Путаница в распознавании вводится как фактор производительности, измеряемый как разница между сходством, полученным в лучшем совпадении, и сопоставлением, полученным в результате второго наилучшего совпадения. В работе был применен глобальный анализ лица, где обрабатывается все изображение сразу. Хотя хорошие результаты были получены при использовании стандартной базы данных, на практике могут возникнуть трудности. Алгоритм обнаружения лиц Виола-Джонса и локальный анализ изображений лиц сыграли значительную роль в улучшении распознавания лиц.

### Классификация по ближайшему среднему значению

В этом методе вектор признаков для классификации известен заранее (получается в процессе обучения системы) или предсказывается в режиме реального времени. Следовательно, некий объект будет относится к классу A, если он ближе к вектору математического ожидания класса A, чем к другим классам.

Такой метод применяется при кучном расположении точек признаков и находятся на большем расстоянии от точек других классов.

### Классификация по расстоянию до ближайшего соседа

Метод основан на оценке сходства объектов. Некий вектор признаков относится к классу А, которому принадлежат ближайшие к нему объекты обучающей выборки.

Преимущества подхода заключаются в добавлении новых образцов в базу данных, а также использование древовидных и сеточных структур сокращают количество вычисляемых расстояний.

В статье [4] используют строки в качестве нового локального дескриптора и классификатора ближайшего соседа с расстоянием Хеллингера для идентификации людей по их изображениям. Строковый дескриптор обеспечивает эффективное описание краев и форм в анализе изображений и прост в реализации; однако, чтобы построить вектор признаков лица, нужно разделить его на непересекающиеся блоки, где описание каждой области дается набором слов, которые присваиваются словам словаря с использованием расстояния Левенштейна.  Экспериментальные результаты продемонстрировали надежность подхода с точки зрения скорости распознавания с подходящими параметрами.

### Структурный или синтаксический

Метод, при котором объект может быть разделен на составляющие. Так же не имеет значение, при котором система детерминировано может определить класс объекта. Используют этот метод при распознавании символов или фигур. Еще этот метод предполагает использование языка, с помощью которого можно структурно элементы и подэлементы. В таком случае, метод необходим для распознавания сложных образов.

### Байесовский подход к принятию решений

Метод в основе которого лежит теорема Байеса и определении априорных вероятностей. Вероятность исходов (принадлежность объекта к классу) будет изменятся после подтверждения новых признаков. Появление образа на изображении содержит случайный характер и вероятность события можно получить с помощью закона распределения вероятностей многомерной случайной величины. Байесовский классификатор определит объект к такому классу, к которому распознанный объект принадлежит с высокой степенью вероятности.

Отрицательной стороной является то, что размерность признаков может быть очень большой. В результате оценки вероятности случайных величин может не хватить количеств наблюдений.

### Нейронные сети

С помощью нейронных сетей можно решить широкий круг задач и представляют структуру из слоев – искусственных нейронов и связей между ними. На вход поступают сигналы и на выходе выдает ответ, выходные сигналы, которые являются решением задачи. На нейроны входного слоя поступают сигналы, которые умножаются на веса нейронов. Далее функция активации (представляющее собой пороговое значение) приводит сумму всех сигналов в число от 0 до 1. Следовательно, сеть выдает вероятностное значение, как результат метода Байеса. В отличии от линейных методов, нейронная сеть точно описывает набор данных и строит с высокой степенью эффективности нелинейные зависимости. Нейросеть обучается на полной выборке данных, не разделяя ее, это и приводит к более точной настройки сети.

В статье [5] предлагается метод распознавания лиц (FRMCD) основанный на центрально-симметричной локальной двоичной структуре (CS-LBP) для извлечения признаков текстуры CS-LBP, затем эти свойства текстуры изучаются глубокой сетью доверия (DBN) для завершения классификации и распознавания. Приведенный метод позволяет DBN получать локальные характеристики изображений лица, поэтому полезной информации о входных изображениях лиц в DBN больше, чтобы обеспечить высокую скорость распознавания. Как показано в экспериментальных результатах, скорость распознавания FRMCD превосходит другие традиционные методы распознавания лиц, а FRMCD также обладает высоким эффектом распознавания и надежностью.  Между тем FRMCD имеет низкую сложность вычислений и занимает мало времени.

В работе [6] обосновывается, что обычные методы, такие как использование нескольких извлеченных функций вручную и основанные на взаимодействии нескольких методов, могут решить задачу для приложений "неподвижное видео", в то время как методы глубокого обучения предлагают более универсальные решения для различных приложений распознавания лиц. В настоящее время глубокие архитектуры CNN (сверточная нейронная сеть) обеспечивают сложные решения, которые менее подходят для приложений видеонаблюдения в реальном времени.

В статье [7] приведен пример использования таких средств как: Amazon Rekognition и Amazon Simple Storage Service (Amazon S3), оба из Amazon Web Services (AWS). Amazon Rekognition - это облачное программное обеспечение, обеспечивающее возможности компьютерного зрения, в то время как Amazon S3 обеспечивает хранение объектов. Также рассмотрели вопрос об использовании Google Cloud Platform (GCP), платформы, которая предоставляла услуги, аналогичные тем, что делал AWS.

## Методы и алгоритмы распознавания лиц

Существует множество алгоритмов распознавания, такие как: метод главных компонентов, метод Виолы – Джонса, метод сравнения эластичных графов, сверточная нейронная сеть, метод главных компонентов, скрытые Марковские модели. Важные отличия всех алгоритмов заключаются в вычислении признаков и в сравнении их с эталоном в базе данных.

Например, метод сравнения эластичных графов. Он заключается в сопоставлении эластичных графов. Граф со взвешенными вершинами и ребрами представляет лицо. Вершины графа расположены на ключевых точках лица. В них вычисляются значения признаков, используя для этого фильтры Габора. Вычисление происходит локальным путем свертки значений яркости пикселей с фильтрами Габора. Разница между графами вычисляется с помощью функции деформации. Корректировка графа происходит путем смещения вершин в позицию, в которой разница между значениями признаков в вершине деформируемого и эталонного графа будет минимальной. Выполнение происходит до тех пор, пока не будет найдена наименьшее суммарное различие между признаками графов. Процедура деформации должна выполнится для всех эталонных лиц в базе данных. Результатом будет являться эталон с лучшим значением функции деформации.

Основные недостатки: низкая эффективность при запоминании новых эталонов в базу данных. Сложность вычислительной процедуры распознавания. Линейная зависимость времени распознавания лица от размера базы данных эталонов.

В результате теоретического и эмпирического метода [8] было выявлено, что метод классификации Виолы – Джонса и Сверточная нейронная сеть работают эффективней других с высокой степенью распознавания.

Таблица 2 – Сравнительная характеристика алгоритмов(картинка)



Результат выполнения алгоритмов по времени представлены в таблица – 3.

Таблица 3 – Сравнительная характеристика по точности определения

|  |  |  |
| --- | --- | --- |
| Алгоритм  Свойство | Метод Виолы – Джонсона совместно с использованием ЛБШ | Сверточная нейронная сеть |
| Точность определения на изображении | 100 % | 62.5 % |
| Точность определения в видео потоке | от 99.43 %  до 90.35 % | от 50 %  до 53 % |
| Общее время работы | 3 минуты 1.4 секунды | 10 минут 13 секунд |

Алгоритм не специализируется только на выделении лиц, его можно использовать для других объектов, но необходимо применить нужный для этого классификатор. Но наиболее широко его применяют для обнаружения лиц на растровом изображении.

Поиск лиц использует несколько необходимых методов:

1. свертка функций Хаара с исходным изображением;
2. машинное обучение AdaBoost;
3. каскадный классификатор;

### Алгоритм Виолы - Джонсона

Алгоритм Виолы – Джонса основан на вейвлетах Хаара. Прямоугольный признак Хаара (Рисунок - 1) определяется как разность сумм пикселей двух смежный областей внутри прямоугольника. Он может занимать разные положения и менять масштаб на изображении. Этот вид признаков Хаара называется 2-прямоугольным. Так же Виола и Джонс смогли определить 3-прямоугольные и 4-прямоугольные признаки. Любой признак может показать существование какой-либо характеристики изображения, например, изменение текстур или границы между темным и светлым регионами.

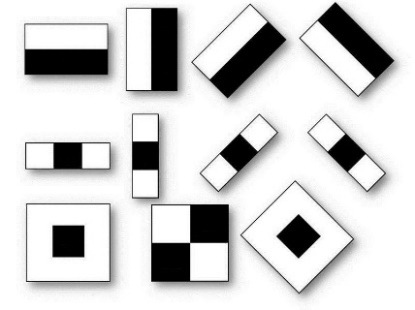


Рисунок 1.1 - Примеры признаков Хаара, белый 1, черный -1

При результате сверки больше заданного порога, который был определен в процессе обучения классификатора, принято считать, что на данном участке изображения присутствует функция Хаара. Ниже (Рисунок 2) представлены примеры размещения признаков на изображении.

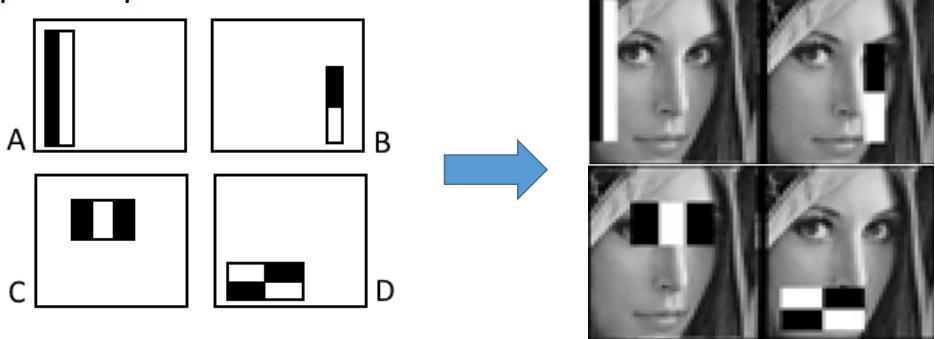


Рисунок 1.2 - Размещение признаков Хаара на изображении лица

После сверки выполняется подбор вейвлетов Хаара с сопоставлением пороговых величин. Их сочетание характеризует на изображении одно или несколько лиц. Так как размер признака Хаара задают размер искомого лица, то для поиска лиц разного масштаба используются признак – пирамида. Стоит заметить, что при использовании алгоритма Виолы – Джонса для поставленной задачи, то на каждом кадре видеопотока будут найдены прямоугольные области. Эти области соответствуют найденным лицам.

Кроме свойств, представленных в таблицах (Таблица 1, Таблица -2) основным показателем распознавания может служить изменение ракурс объекта. Для понимания ракурса объекта использовался алгоритм разбиения классов на подклассы при нечетких ракурсах наблюдения. Эти алгоритмы, основанные на использование показателей сопряженности в качестве сходства и обеспечивающие повышение точности распознавания объекта до 97%.

В статье [9] представили метод обнаружения глаз, основанный на детекторе Виолы-Джонса, цвете кожи и применении шаблона глаза. При любых условиях, выражения или наличия затенения метод обеспечивает обнаружение глаз. Полученные результаты показывают, что этот метод имеет много преимуществ с точки зрения качества и скорости обнаружения.

Таблица 4 – Результаты обнаружения глаз

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Количество всего глаз | Количество обнаруженных глаз | | Количество ложных глаз | |
| 336 | 310 | 92,26% | 31 | 9,09% |

В статье [9] приведен способ для улучшения метода распознавания лица, с помощью компонентов Виолы – Джонса. Этот метод начинается с обнаружения и обрезки компонентов Виолы – Джонса. Лицевые компоненты обнаруживаются и точно обрезаются при любых изменениях ракурса. Обрезанные компоненты представлены гистограммой ориентированных градиентов (HOG). Вес каждого компонента должен быть определен с использованием процесса проверки. Объединение этих весов было сделано простой техникой голосования.

Алгоритм выбора области состоит из этапов:

1. Первым компонентом, который нужно искать, является грань.
2. Обнаружение пары глаз на обрезанном изображении лица.
3. Область под парой глаз внутри обрезанного изображения лица будет областью поиска для носа.
4. Определенная область используется для обнаружения рта. В случае обнаружения нескольких устьев объект с более значительным значением оси y (самый низкий объект) выбирается для использования в качестве компонента устья.

### Метод главных компонентов

В статье [11] предложен метод анализа главных компонентов (PCA) на основе исправлений для распознавания лиц. Многие основанные на PCA методы распознавания лиц используют корреляцию между пикселями, столбцами или строками. Но локальная пространственная информация не используется или используется не полностью в этих методах. Исправления являются более значимыми базовыми единицами для распознавания лиц, чем пиксели, столбцы или строки, поскольку лица различаются по участкам с глазами и носами. Чтобы вычислить корреляцию, изображения лица делятся на участки, а затем эти участки преобразуются в векторы столбцов, которые будут объединены в новую «матрицу изображений». Заменяя изображения на новую «матрицу изображений» в двумерной структуре PCA, можно напрямую вычисляем корреляцию разделенных участков путем вычисления полного разброса. Оптимизируя общий разброс проецируемых образцов, получается проекционная матрицу для выделения признаков. Также, используется классификатор ближайшего соседа. Было выявлено, что этот метод повышает точность по сравнению с одномерными и двумерными PCA.

### 3D Распознавание лиц

В статье [12] предлагается метод трехмерного распознавания лиц, основанный на локальной конформной параметризации и анализе изо-геодезических полос. Используя локальную конформную параметризацию, достигается двумерного представления лица, которое устраняет влияние позы головы. На основе изо-геодезических полос извлекается черты лица из двумерного представления лица. Преимущества предложенного метода заключается к устойчивости разных поз головы, выражений лица. Однако ограничением метода является чувствительность к окклюзиям в области носа.

## Особенности распознавания лиц в видеопотоке

Особенностью распознавания объектов в видеопотоке заключается в том, что изображения предоставляют низкую информационную емкость. Наибольшее количество точек изображения содержит нейтральную информацию, которая может повлиять на распознавание.

Основная проблемой, которую необходимо решить для точного распознавания это изменчивость визуальных образов. Изменчивость заключается в изменении выражения лица, дополнительной одежды или аксессуаров, поворота или наклона головы. Решение этой проблемы требует высокой производительности вычислительных средств и необходимой этому большого объема памяти.

### Проблема освещенности

С целью оценки эффективности предложенных алгоритмов распознавания лиц агентство DARPA и исследовательская лаборатория армии США разработали программу FERET (face recognition technology).  
В масштабных тестах программы FERET принимали участие алгоритмы, основанные на гибком сравнении на графах и всевозможные модификации метода главных компонент (PCA). Эффективность всех алгоритмов была примерно одинаковой [13]. В этой связи трудно или даже невозможно провести четкие различия между ними (особенно если согласовать даты тестирования). Для фронтальных изображений, сделанных в один и тот же день, приемлемая точность распознавания, как правило, составляет 95%. Для изображений, сделанных разными аппаратами и при разном освещении, точность, как правило, падает до 80%. Для изображений, сделанных с разницей в год, точность распознавания составило примерно 50%. При этом стоит заметить, что даже 50 % — это более чем приемлемая точность работы системы подобного рода.

Для решения проблем с освещенностью необходимо проводить предварительную обработку [14]. Необходимо применить метод выделения инварианта освещения на основе фазы Габора. Сначала нормализуется изображения лица, используя метод предварительной обработки, основанный на гомоморфном фильтре, для предварительного устранения последствий изменений освещенности. Затем для преобразования изображения используется набор двумерного реального вейвлета Габора с различными направлениями, и несколько коэффициентов Габора объединяются в одно целое при рассмотрении как спектра, так и фазы. В итоге, инвариант освещения получается путем извлечения фазового признака из объединенных коэффициентов. Этот метод не требует 3D-информации о форме лица или данных для обучения. И извлеченный инвариант освещения содержит более существенную информацию о различении, в то же время значительно уменьшая эффект изменений освещения в то же время.

### Проблема изменение лица

В статье [15] предложен подход модифицированной масштабно-инвариантной трансформации признаков SIFT (MSIFT) для повышения эффективности распознавания SIFT. В статье работа выполняется в три этапа. Во-первых, сглаживание изображения было выполнено с использованием DWT. Во-вторых, вычислительная сложность SIFT при вычислении дескриптора уменьшается путем вычитания среднего значения из каждого дескриптора вместо нормализации. В-третьих, алгоритм сделан автоматическим с использованием коэффициента корреляции (CoC) вместо использования отношения расстояний (которое требует взаимодействия с пользователем). Основным достижением этого метода является уменьшение размера базы данных, поскольку для его хранения требуется хранить только нейтральные изображения, а не все выражения одного и того же изображения лица.

### Обучение нейронной сети

Распознавание лиц с одним образцом является проблемой в компьютерном зрении. Поскольку для обучения требуется только один образец, это затрудняет распознавание при изменении лица, наклона головы и освещенности. Для преодоления проблемы, предлагается схема комбинированного метода традиционного и глубокого обучения (TDL) [16].

1. Предлагает метод расширенной выборки, основанный на традиционном подходе. Кроме того, он может генерировать образцы, такие как маскировка, выражение и смешанные вариации.
2. Использует трансферное обучение, а затем выбирает несколько расширяющихся выборок для точной настройки модели.
3. Для реализации эксперимента используется тонко настроенная модель.

Изучение глубокого классификатора требует построения набора данных. Наборы данных в средствах массовой информации часто зависят от ситуации, с разными наборами или ландшафтами. Становится актуальным использовать поисковые системы для создания набора данных или сортировки больших наборов изображений на основе метаданных и эвристики. Эти методы не идеальны и тогда появляется шум. В статье [17] рассмотрены подходы, охватывающие широкий спектр вариантов использования, в зависимости от набора данных: имеет ли он проверенные или исправленные метки или нет, а также предполагаемую долю шумных меток. Все они имеют различные характеристики устойчивости: некоторые могут работать хорошо при низком коэффициенте шума, но быстро ухудшаться, в то время как другие могут иметь немного более низкую оптимальную точность, но не ухудшаться так сильно при высоком коэффициенте шума.



### Распознавание по изображениям низкого качества

В статье [18] обобщаются и анализируются работы по проблеме LQFR. Распределяя по категориям работы, авторы предоставили обзор многих современных представлений с последующей оценкой эффективности с использованием известных наборов данных исследований. Среди этих методов, глубокое обучение очень устойчиво к некоторым проблемам: освещенность, размытость, отклонение головы и т. д., но в некоторых случаях оно дает очень плохие результаты. В научном мире существует высокий уровень интереса к распознаванию лиц на изображениях низкого качества из-за многообещающих применений (криминалистика, наблюдение и т. д.).

## Анализ систем и методов алгоритмов распознавания

### Статические методы распознавания

Подходы статистического распознавания образов широко применяются для распознавания образов и целей классификации. В современной литературе предлагаются различные методы статистического распознавания образов. Для выбора более подходящего статистического распознавания образов метод требует комплексной оценки эффективности современных методов. В статье [] предоставляется анализ производительности статических методов распознавания.

Таблица 4 – Сравнительные результаты эффективности статистических методов распознавания образов (точность в%)

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Data Sets | Kernel-C | LDA-C |  | Logistic-C | PolQuadLMS-С | PSOLDA-С | NB-С |  |
| LinearLMSC |  |
|  |  |  |  |  |  |  |  |
| Bupa | 57,98 | 69,82 | 69,53 | 69,53 | 70,12 | 72,76 | 67,25 |  |
|  |  |  |  |  |  |  |  |
| Ecoli | 42,56 | 72,41 | 72,66 | 76,79 | 75,98 | 80,65 | 82,15 |  |
|  |  |  |  |  |  |  |  |
| Glass | 39,46 | 61,33 | 32,90 | 67,69 | 17,72 | 75,68 | 70,61 |  |
|  |  |  |  |  |  |  |  |
| Heberman | 73,19 | 75,78 | 74,49 | 74,82 | 75,14 | 76,43 | 72,20 |  |
|  |  |  |  |  |  |  |  |
| Iris | 93,33 | 96,00 | 88,00 | 95,33 | 90,00 | 98,00 | 94,67 |  |
|  |  |  |  |  |  |  |  |
| Monks | 52,66 | 47,34 | 47,34 | 97,27 | 47,34 | 97,27 | 97,27 |  |
|  |  |  |  |  |  |  |  |
| Pima | 65,11 | 76,19 | 76,45 | 77,36 | 76,31 | 80,87 | 74,88 |  |
|  |  |  |  |  |  |  |  |

Наиболее успешным методом обнаружения шаблонов из стандартных наборов эталонных данных является PSOLDA-C по сравнению с другими конкурентными методами статистического распознавания образов. Метод Logistic-C находится на второй позиции с учетом средней эффективности методов для всех наборов данных.

### Достоинства и недостатки основных методов распознавания

В статье [20] проводится обзор и анализ с выделением преимуществ и недостатков некоторых методов. Методы, указанные в таблице 1, иллюстрируют эволюцию технологии распознавания лиц. Следует подчеркнуть огромный потенциал методов, основанных на дескрипторе лица, в связи с тем фактом, что идея локального дескриптора была признана в качестве наиболее важной основы проектирования для задач идентификации и проверки лица.

Таблица 5 – Преимущества и недостатки

|  |  |  |  |
| --- | --- | --- | --- |
| № | Метод | Преимущества | Недостатки |
| 1. | Классические алгоритмы распознавания лиц | Проецируют лицо на линейное подпространство, охватываемое изображениями собственного лица. | Методы не могут адекватно представлять лица, когда происходят большие изменения в выражении лица освещения и других факторах. |
| 2. | Нейронные сети | Радиальная базисная функция интегрируется с неотрицательной матричной факторизацией. Идеальное решение, для распознавания изображений лица с частичным искажением. | Требование большего количества обучающих образцов. |
| 3. | Габор вейвлеты | Характеристики захвата характерных визуальных свойств: селективность ориентации пространственной локализации и пространственная частота. | Значительно большая размерность пространства признаков Габора. Подход требует большого объема вычислений и нецелесообразен для приложений реального времени. |
| 4. | Методы, основанные на дескрипторе лица | Устойчивы к изменениям освещения и изменениям положения головы. | Требует большой объем вычислений. |
| 5. | 3D распознавание лиц | Распознавание не зависит от освещения. | Вычислительно дорогой процесс, а также данные должны быть максимально точными. |
| 6. | Распознавание на основе видео | Возможность использования избыточности, присутствующей в видео, для улучшения систем неподвижных изображений. | Измерение нескольких изображений и мало исследовательских данных. |

В статье [21] провели работу над модифицированным подходом LBP, называемым расширенным локальным двоичным шаблоном (A-LBP), который использует неоднородные шаблоны в процессе представления. Разработанный метод учитывает неоднородные структуры и извлекает доступную им дискриминационную информацию, чтобы доказать их полезность. Они используются в сочетании с соседними единообразными шаблонами и извлекают полезную информацию из локальных дескрипторов. Среди четырех мультиалгоритмических подходов три подхода показывают лучшее значение вероятность справедливого уровня соответствия TAR на 97,79%, 97,05% и 97,95%.

В статье [22] предлагается новая структура извлечения признаков DCFB для надежной оценки позы головы и распознавания лиц. Эффективный 1-мерный оптимальный компромиссный фильтр 1D-OTF, разработан в DCFB с использованием частотных представлений 1D-признаков в линейном подпространстве, полученных анализом главных компонентов. Экспериментальные результаты продемонстрировали эффективность DCFB для оценки позы головы, распознавания лица и оценки окклюзии.

### Системы распознавания

В статье [23] мы представляют систему учета посещаемости с использованием распознавания лиц. Она может быть использована в колледжах или офисах с использованием распознавания лиц в реальном времени с помощью Convolution Neural Network (CNN). Для обнаружения лиц использовали классификатор Хаара, предложенный Виолой и Джонсом. Распознавания лиц происходило в результате применения сиамской сети. Сиамская сеть - это архитектура нейронной сети, которая содержит две или более идентичных (т.е. одинаковых конфигурации с одинаковыми параметрами и весами) подсетей. Поскольку распознавали лица, то использовали сверточную сиамскую сеть. Эта система обучалась один раз на всех участниках. Но у этой системы есть недостаток в плане большой выборки при обучении. Так как при изменении позы головы более 20 градусов человек не будет идентефицирован.

Решение этой проблемы описано в статье [24]. Описывается новая модель, которая генерирует изображения в новых позах, например изменяя выражение лица и ориентации. В отличие от предыдущего подхода, который требует большего набора данных конкретного человека для обучения, данный подход может работать в небольшом наборе изображений. С этой целью вводится адаптивный характер GAN (CainGAN) нормализации идентичности, который использует пространственные характеристики, которые объединяются с исходным изображением. Идентификация распространяется по всей сети, путем применения условной нормализации. Экспериментальные результаты показывают, что качество генерируемых изображений масштабируется с размер входного набора, используемого при выводе.

Распознавание лица с помощью рекуррентных нейронных сетей рассмотрено в статье [25]. Дезинформация в онлайн контенте растет и необходимость в обнаружении такого контента необходимо. В этой работе показано, что сочетание рекуррентных сверточных сетей улучшает подход к выравниванию моделей и граней. Было проведено исследование различных стратегий как для выравнивания, так и для комбинирования функций CNN через повторение. Был найден ориентир на основе выравнивания лица для достижения наилучших результатов в момент изменения лица в видео.

# Заключение

# Список использованных источников

1. M. Andrejevic, N. Selwyn, Facial recognition technology in schools: critical questions and concerns. – 2019. - [Электронный ресурс]. Режим доступа:https://www.tandfonline.com/doi/full/10.1080/17439884.2020.1686014?af=R&
2. Пастушков А. В. Методы и алгоритмы поиска объекта в видеопотоке: диссертация Национальный исследовательский Томский гос. университет, Томск, 2017. – Режим доступа: https://www.dissercat.com/content/metod-i-algoritmy-poiska-obekta-v-videopotoke/read
3. M. A. Alijanabi, Z. M. Hussain, S. F. Lu, Entropy-Histogram Approach for Image Similarity and Face Recognition. – 2018. [Электронный ресурс]. Режим доступа: https://www.hindawi.com/journals/mpe/2018/9801308/
4. A. Saaidi, R. El Alami,  M. Abarkan. A New Local Descriptor Based on Strings for Face Recognition. — 2020. [Электронный ресурс]. Режим доступа: http://downloads.hindawi.com/journals/jece/2020/3451808.pdf
5. Y. Xin, Y. Mingxin, W. Yang, F. Jianying, The Face Recognition Method Based on CS-LBP and DBN, 2018 - [Электронный ресурс]. Режим доступа: https://www.hindawi.com/journals/mpe/2018/3620491/
6. S. Bashbaghi, Face Recognition in Video Surveillance from a Single Reference Sample Through Domain Adaptation. – 2017. - [Электронный ресурс]. Режим доступа: https://www.etsmtl.ca/getattachment/Unites-de-recherche/LIVIA/Recherche-et-innovation/Theses/Saman\_PhD\_2017.pdf
7. N. Tkauc, T. Tran, K. Hernandez-Diaz, F. Alonso-Fernandez, Cloud-Based Face and Speech Recognition for Access Control Applications. – 2020. - [Электронный ресурс]. Режим доступа: https://arxiv.org/abs/2004.11168
8. Упская О. К. Исследование и программная реализация алгоритмов распознавания лиц: диссертация Комсомольск-на-Амуре гос. университет, Комсомольский-на-Амуре, 2018. - Режим доступа: https://knastu.ru/media/files/page\_files/page\_391/magistr\_referat/Avtoreferat.\_Upskaya\_O.K.\_6VSm-1.pdf
9. S. El Kaddouhi, A. Saaidi и M. Abarkan, Eye detection based on the Viola-Jones method and corners points, *Multimedia Tools and Applications* , vol. 76, нет. 21, с. 23077–23097, 2017. http://article.nadiapub.com/IJCA/vol11\_no5/6.pdf
10. D. Isaam, Al-B. Hussein, Improving the Component-Based Face Recognition Using Enhanced Viola–Jones and Weighted Voting Technique. – 2019. - [Электронный ресурс]. Режим доступа: https://www.hindawi.com/journals/mse/2019/8234124/
11. T.-X. Jiang, T.-Z. Huang, X.-L. Zhao, T.-H. Ma, Patch-Based Principal Component Analysis for Face Recognition. – 2017. - [Электронный ресурс]. Режим доступа: https://www.hindawi.com/journals/cin/2017/5317850/
12. L. Chenlei, Z. Junli, 3D Face Recognition based on Local Conformal Parameterization and Iso-Geodesic Stripes Analysis. – 2018. - [Электронный ресурс]. Режим доступа: https://www.hindawi.com/journals/mpe/2018/4707954/
13. Гришин А.И. Исследование эффективности методов обработки изображений в системах распознавания лиц. - [Электронный ресурс]. Режим доступа: https://dspace.susu.ru/xmlui/bitstream/handle/0001.74/8344/2016\_344\_grishin.pdf?sequence=1?sequence=1
14. W. Shuping, Z. Hao, Efficient Gabor Phase Based Illumination Invariant for Face Recognition – 2017. - [Электронный ресурс]. Режим доступа: https://www.hindawi.com/journals/am/2017/1356385/
15. N. Nirvair, K. Lakhwinder, Modified SIFT Descriptors for Face Recognition under Different Emotions. – 2016. - [Электронный ресурс]. Режим доступа: https://www.hindawi.com/journals/je/2016/9387545/
16. Z. Junying, Z. Xiaoxoxia, G. Junying, M. Chaoyun, Z. Yikui, W. Fan, Deep Convolutional Neural Network Used in Single Sample per Person Face Recognition. – 2018. - [Электронный ресурс]. Режим доступа: https://www.hindawi.com/journals/cin/2018/3803627/
17. G. Sanchez, V. Guis, R. Marxer, F. Bouchara, Deep learning classification with noisy labels. – 2020. - [Электронный ресурс]. Режим доступа: https://arxiv.org/abs/2004.11116
18. P. Li, L. Prieto, D. Mery, P. Flynn, Face Recognition in Low Quality Images: A Survey. - 2018. - [Электронный ресурс]. Режим доступа: https://www.researchgate.net/publication/325439782\_Face\_Recognition\_in\_Low\_Quality\_Images\_A\_Survey
19. A. Zulfqiar, S. Syed Khurram, S. Waseem, Performance Analysis of Statistical Pattern Recognition Methods in KEEL. – 2017. - [Электронный ресурс]. Режим доступа: https://www.sciencedirect.com/science/article/pii/S1877050917315016
20. W. Wokcik, K. Gromaszek, M. Junisbekov, Face Recognition: Issues, Methods and Alternative Applications. – 2016. - [Электронный ресурс]. Режим доступа: https://api.intechopen.com/chapter/pdf-preview/51031
21. R. Shyam, Y.N. Singh, Multialgorithmic, Frameworks for Human Face Recognition. – 2016. - [Электронный ресурс]. Режим доступа: https://www.hindawi.com/journals/jece/2016/4645971/
22. S. Chen, D. Yan, Y. Yan, Directional Correlation Filter Bank for Robust Head Pose Estimation and Face Recognition. – 2018. [Электронный ресурс]. Режим доступа: https://www.hindawi.com/journals/mpe/2018/1923063/
23. A. Shailesh, M. Hrithink, P. Vishal, Smart Attendance System Using CNN. – 2020. [Электронный ресурс]. Режим доступа: https://arxiv.org/abs/2004.14289
24. A. T. Ardelean, L.M. Sasu, Pose Manipulation with Inentity Preservation. – 2020. [Электронный ресурс]. Режим доступа: https://arxiv.org/abs/2004.09169
25. E. Sabir, J. Cheng, A. Jaiswal, W. AbdAlmageed, I. Masi, Recurrent Convolutional strategies for face manipulation detection in videos. Электронный ресурс]. Режим доступа: https://arxiv.org/abs/1905.00582