Санкт-Петербургский государственный университет

**Кафедра компьютерного моделирования**

**и многопроцессорных систем**

**Асташенкова Лада Константиновна**

**Выпускная квалификационная работа бакалавра**

**Сегментация изображений человеческих лиц**

Направление 010302

Прикладная математика и информатика

Научный руководитель,  
кандидат тех. наук,  
доцент  
Гришкин В. М.

Санкт-Петербург

2019

# Содержание

Введение 3

Постановка задачи 5

Обзор литературы 6

Глава 1. Обзор алгоритмов обнаружения и сегментации лиц 7

* 1. Локализация областей лиц на изображении 7
     1. Метод Виолы-Джонса 7
     2. Нейронные сети 9
     3. HOG и SVM 11
  2. Обнаружение ключевых точек лица 14
     1. Понятие ключевых точек 14
     2. Дескрипторы ключевых точек 14
     3. Поиск ключевых точек лица 17

Глава 2. Этапы работы алгоритма сегментации 19

* 1. Описание выбранных методов 19
  2. Описание алгоритма сегментации 21

Глава 3. Практическая реализация 23

* 1. Описание инструментов 23
  2. Результаты 24

Выводы 28

Заключение 29

Список литературы 30

**Введение**

В настоящее время современный человек не представляет свою жизнь без информационных технологий и электроники последнего поколения, которые уже тесно сплелись воедино с его каждодневными делами. Это обосновывает тот огромный интерес, с которым разработчики подходят к созданию все более совершенных технических устройств и программных продуктов, призванных сделать жизнь пользователей в разы комфортнее.

Согласно [1] уже 7 лет назад количество фотографий, загруженных пользователями в социальную сеть для обмена фотографиями и видеозаписями Instagram с момента ее запуска, превысило 5 миллиардов. Настолько огромные числа свидетельствуют о том, что частота использования мобильных камер в повседневных ситуациях с каждым днем только набирает обороты. Однако люди все чаще желают не просто сохранить значимые моменты в памяти, но и привнести в них что-то необычное и поделиться в социальных сетях. По этой причине многочисленные приложения для обработки фотографий пользователей становятся все популярнее. Вместе с этим развиваются и методы, используемые в этих системах. Одним из вспомогательных этапов таких алгоритмов можно выделить и процесс сегментации изображений лица человека.

Сегментация изображений человеческих лиц является задачей в области компьютерного зрения и искусственного интеллекта. Результатом сегментации является выделение черт лица человека. В частности, из-за особенностей строения лица чаще всего производят определение наиболее характерных частей, к которым относятся области глаз, бровей, носа, рта и подбородка. Как уже упоминалось ранее, сегментация данных областей может быть как конечной целью, так и промежуточным шагом для широкого круга задач, начиная с систем безопасности и заканчивая мультипликацией. А именно, примерами таких задач можно выделить:

* определение эмоционального состояния и мимики;
* оценивание положения головы;
* выравнивание изображения лица;
* обнаружение моргания;
* вычисление меры сходства по изображениям людей;
* автоматическое редактирование изображений лица;
* компьютерная анимация.

Поскольку отличительными характеристиками внешности каждого человека могут служить ключевые точки лица, то именно на их основе в данной работе и предлагается проводить сегментацию изображения человеческого лица. На данный момент существует большое количество встроенных функций и методов для поиска ключевых точек, например, в библиотеки OpenCV и DLib, с помощью которых предлагается реализовать алгоритм сегментации.

Таким образом, можно убедиться, что компьютерное зрение играет далеко не последнюю роль во многих сферах человеческой деятельности и значительно повышает качество жизни, требуя от пользователя все меньше действий и времени. Это и создает необходимость совершенствования существующих методов и создания новых подходов к решению подобных задач.

**Постановка задачи**

Основной целью данной работы является разработка программной системы для сегментации лица человека на цифровом изображении для выделения таких областей, как:

* правый глаз;
* левый глаз;
* правая бровь;
* левая бровь;
* нос;
* рот;
* линия подбородка.

Задачу сегментации в данном случае можно разделить на следующие этапы: локализация всех лиц на изображении, детектирование ключевых точек на этих лицах и выявление основных частей каждого лица. На каждом этапе проводится исследование существующих алгоритмов и выбор наиболее эффективного для конкретной ситуации решения.

**Обзор литературы**

На данный момент существует множество книг, статей и Интернет-ресурсов, по которым можно ознакомиться и разобраться с темами из области компьютерного зрения и обработки изображений.

Из книг [2, 3] можно приобрести достаточное количество информации для погружения в данную сферу. Данная литература позволяет узнать основные понятия компьютерного зрения, ознакомиться с распространенными задачами в этой области, а также возможными алгоритмами их решения.

Одним из этапов сегментации можно выделить локализацию области лица. По этой причине были рассмотрены существующие способы локализации лиц, описанные в материале [4]. Статья [5] и электронный ресурс [6] дали представление о методе Виолы-Джонса. Теоретический аспект нейронных сетей был раскрыт в материалах [7, 8]. Описание метода детектирования лиц с помощью гистограммы направленных градиентов представлено в [9]. Про метод опорных векторов информация была почерпнута из [10].

Другим этапом сегментации изображения лица человека является поиск ключевых точек, определение которых можно найти в материале [11]. Различные методы их обнаружения рассмотрены в статьях [12, 13]. В работе [14] описан процесс выделения ключевых точек на изображении лица при помощи деревьев регрессии, информация о которых была добыта из электронных ресурсов [15, 16].

**Глава 1. Обзор алгоритмов обнаружения и сегментации лиц**

В данной главе приведены этапы сегментации изображений лиц человека и основные методы для их выполнения.

Задача сегментации изображения лица является двухэтапным процессом:

* этап №1. Локализация всех лиц на изображении;
* этап №2. Определение ключевых точек и частей каждого лица.

Рассмотрим каждый этап сегментации подробнее, разобрав основные способы их реализации.

* 1. **Локализация областей лиц на изображении**

Первым этапом сегментации изображения лица человека является локализация всех лиц на изображении. Обнаружение области лица может быть выполнено с использованием ряда различных методов, рассмотрим некоторые из них.

* + 1. **Метод Виолы-Джонса**

Одним из самых известных методов обнаружения и детектирования лица на изображении является метод Виолы–Джонса, предложенный в 2001 году Полом Виолой и Майклом Джонсом [5]. Данный метод для поиска лица и его черт использует принцип сканирующего окна. Согласно этому принципу изображение сканируется окном поиска с последующим применением к каждому положению классификатора. Основные моменты данного алгоритма, описанные в [6], следующие:

1. Интегральное представление изображения, что представляет собой матрицу, размер которой совпадает с размером исходного изображения. В каждом элементе этой матрицы хранится сумма интенсивностей пикселей, которые находятся левее и выше данного элемента. Данный принцип позволяет быстро вычислять необходимые объекты;
2. Признаки Хаара для поиска нужного объекта, например, лица и его черт. В стандартном методе Виолы-Джонса принято использовать примитивы Хаара, т.е. признаки прямоугольной формы, вид которых представлен на рис. 1. Вычислением значения такого признака является разность суммы значений яркостей, закрываемых светлой частью признака, и суммы значений яркостей, закрываемых темной частью;

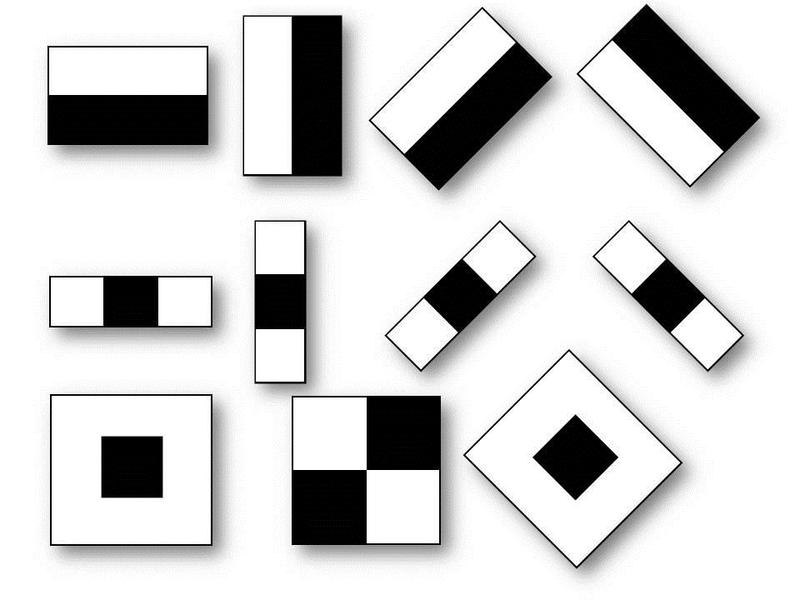


Рис. 1. Примитивы Хаара

1. Бустинг для выбора наиболее подходящих признаков искомого объекта на рассматриваемой области изображения. Бустинг (от англ. boost – улучшение, усиление) – метод усиления слабых классификаторов, где под «слабым» подразумевается классификатор, вероятность которого получить правильный ответ практически такая же, как при случайном угадывании. А под «усилением», что путем комбинирования таких слабых классификаторов с присвоенными им весами можно достигнуть большей эффективности и, следовательно, корректный результат;
2. Простые бинарные классификаторы после принятия на вход признаков отсеивают их, давая результат «истина» или «ложь»;
3. Каскады признаков для быстрого исключения окон, где лицо не найдено. Каскады признаков представляет собой дерево принятия решений, узлы которого детектируют практически все интересующие образы и отклоняют регионы, которые точно не являются образами.

Можно сделать вывод, что метод Виолы-Джонса возможно использовать не только для поиска области лица, но и для обнаружения частей лица, т.е. решения задачи сегментации. Но, несмотря на то, что классификатор обладает высокой скоростью и точностью обнаружения лиц на изображениях, он имеет и недостатки: медленное обучение классификатора и использование большого количества данных для этого. Кроме того, качество работы алгоритма заметно снижается при недостаточном освещении и затемнении. По этим причинам метод Виолы-Джонса следует считать в данной работе не самым эффективным средством для решения поставленной задачи.

* + 1. **Нейронные сети**

Детекторы лиц также основываются на нейронных сетях (НС) [7, 8]. Построены такие алгоритмы на последовательном преобразовании сигнала нейронами – вычислительными единицами, работающими параллельно. Связь между двумя нейронами называется синапсом и имеет один параметр – вес. Благодаря такой связи в виде синапса входная информация меняется и передается от нейрона к нейрону, в конечном итоге достигая результата. Процентная величина, отражающая расхождение между ожидаемым и полученным ответами, считается за ошибку, которую алгоритм с каждой итерацией во время обучения стремится уменьшить.

Предполагается, что обученная НС сможет за счет обобщающих способностей применить полученный в процессе обучения опыт на неизвестные объекты. Идея обучения заключается в настройке весов межнейронных связей в процессе решения задачи оптимизации методом градиентного спуска. В процессе обучения НС автоматически извлекаются ключевые признаки, определяется их важность и строится взаимосвязь между ними.

Основанные на НС системы обнаружения объектов на изображениях имеют иерархическую структуру. Данная структура проводит первичную обработку вектора признаков грубой сетью с высоким уровнем ошибок второго рода, т.е. часто не находит объект, присутствующий на изображении. В случае если ошибка второго рода произошла, и объект не был классифицирован, решение корректируется более точной сетью, которая, однако, уже уступает по скорости. Точность алгоритмов, основанных на НС, повышают за счет настройки параметров, что заметно влияет на результат.

Тем не менее, НС считается в данной работе неприемлемой для решения поставленной задачи по причине того, что обладает рядом недостатков, к которым относятся:

* необходимость полного переобучения сети при внесении нового эталона в базу данных, что является затратным по времени процессом;
* возможность возникновения переобучения – явления, при котором алгоритм обучения вырабатывает предсказания, слишком близко соответствующие конкретному набору данных, что делает алгоритм менее точным уже в процессе использования на дополнительных данных;
* трудно формализуемый этап выбора архитектуры сети (количество нейронов и характер связей между ними).
  + 1. **HOG и SVM**

Данный алгоритм детектирования лиц основывается на классификаторе, обученном на гистограмме направленных градиентов (Histogram of Oriented Gradients, HOG) и методе опорных векторов (Support Vector Machine, SVM).

Использование HOG широко распространено в области компьютерного зрения. Данный метод изначально был предложен для обнаружения пешеходов на изображениях в 2005 году [9]. В последующем он был модифицирован для обнаружения людей и других объектов и на видео.

Дескриптор HOG основывается на предположении, что внешний вид и форма объекта в рассматриваемой области изображения могут быть описаны распределением градиентов интенсивности пикселей, соответствующих всей этой области, а не каждой точке конкретно [12]. Описать этапы работы дескриптора HOG можно следующим образом:

1. Разбиение на ячейки. Исходное изображение разбивается на небольшие связные области (ячейки);
2. Расчет градиента. С помощью фильтрации с ядрами и ячеек вычисляются величины и . Другими словами, с такими ядрами производится свертка изображения – операция вычисления нового значения выбранного пикселя с учетом значений пикселей, окружающих его. Во время вычисления нового значения пикселя центр ядра свертки как бы «прикладывается» к данному пикселю так, что окружающие его пиксели тоже накрываются ядром. Далее вычисляется сумма, слагаемые которой являются произведениями значений интенсивностей пикселей и значений ячейки ядра, накрывшей этот пиксель, затем вычисленная сумма делится на сумму всех элементов ядра свертки. Полученные значения и являются и , на их основе вычисляется величина и направление градиента ячеек

, ;

1. Объединение в блоки. Ячейки группируются в более крупные блоки для последующего локального нормирования;
2. Расчет гистограммы градиентов в блоках. Теперь каждому пикселю в блоке соответствует градиент, т.е. его величина и направление. Для обеспечения более компактного представления значений градиенты преобразуются в гистограмму. Интервал гистограммы определяется направлением градиента, например, 0, 20, 40 градусов и т.д. Таким образом, каналы гистограммы равномерно распределяются от 0 до 180 или же от 0 до 360 градусов, в зависимости от того, вычисляется «знаковый» или «беззнаковый» градиент, основываясь на принадлежности угла градиента к тому или иному интервалу;
3. Нормализация блоков. По причине того, что рассчитанная гистограмма не очень устойчива к изменениям освещения, градиенты локально нормируются. Каждый элемент вектора гистограммы делится на величину вектора.

Как правило, блоки перекрываются, т.е. каждая ячейка входит более чем в один конечный дескриптор. Таким образом, дескриптор HOG является вектором компонент нормированных гистограмм ячеек из всех областей блока. При анализе некоторого объема изображений, каждое из них описывается таким вектором значений, который можно рассматривать как точку в многомерном пространстве. Для разделения этих точек на классы при детектировании объектов используется SVM [10].

Метод SVM применяется для классификации объектов, относя каждый из них к одному определенному классу из нескольких (два и более), заранее известных. Идея метода можно проиллюстрировать на простом примере классификации точек на плоскости, разбитых на два класса. В данном случае два класса можно разделить линией, которую называют разделяющей прямой. В пространствах большей размерности роль такой линии играет гиперплоскость – пространство, размерность которого на единицу меньше размерности исходного изображения. Если искать гиперплоскость, обладающей наибольшей точностью классификации, то ей окажется гиперплоскость, расстояние от которой до каждого класса максимально, тогда она носит название оптимальной разделяющей гиперплоскости. Данный метод классификации устойчив к переобучению и обладает высокой скоростью работы, хотя по точности уступает некоторым другим методам.

Детектирование алгоритмом HOG и SVM производится по принципу скользящего окна, полностью его описать можно следующим образом. Сначала считаются гистограммы направленных градиентов для всех изображений объектов обучающей выборки, которые необходимо детектировать. Затем на этих данных тренируется SVM-классификатор. В процессе детектирования используются группа окон, размеры которых фиксированы. Каждое такое окно проходит по изображению и вычисляет HOG заданного участка, а натренированный классификатор SVM решает, относить ли найденный объект к искомому. В конце этого процесса получается группа окон с вероятными объектами, из которой лишние исключаются путем объединения окон, имеющих наибольшее пересечение, и удаления остальных.

Так как HOG работает локально, то данный метод инвариантен к геометрическим (смещение, поворот, масштаб) и фотометрическим (изменение яркости изображения) преобразованиям, кроме ориентации объекта. Благодаря преимуществам HOG данный дескриптор можно считать приемлемым средством детектирования областей лиц людей на изображении.

* 1. **Обнаружение ключевых точек лица**

Вторым этапом сегментации изображения лица человека является детектирование ключевых точек на каждом из лиц, локализованных на предыдущем этапе. С помощью найденных ключевых точек происходит последующее обнаружение частей лица. В данном параграфе рассматривается само понятие ключевых точек и способы их обнаружения и описания.

* + 1. **Понятие ключевых точек**

Ключевые или особые точки это такие точки изображения, у которых локальные окрестности обладают некоторыми отличительными особенностями по сравнению с окрестностями других точек изображения [11]. Примерами ключевых точек лица являются уголки глаз, рта, точки вдоль дуг бровей и линии подбородка, кончик носа, и так далее. Проблема автоматического определения ключевых точек лица осложняется следующими естественными факторами:

* масштаб изображения;
* наличие шумов и качество освещения;
* поза и выражение;
* внутренняя изменчивость лица человека.
  + 1. **Дескрипторы ключевых точек**

Существует множество методов поиска и описания ключевых точек лица, но все эти методы в основном пытаются локализовать и отметить ключевые точки вдоль таких областей лица, как рот, глаза, брови, нос, подбородок.

**Использование HOG-дескрипторов**

Так как гистограммы ориентированных градиентов являются известным дескриптором и имеют много реализаций, они могут использоваться и для обнаружения ключевых точек лица. Дескриптор функции — это упрощенное представление изображения или его части, в котором отброшена посторонняя информация. Как правило, дескриптор функции преобразует изображение в вектор. В дескрипторе HOG в качестве функции используется распределение (гистограмма) направлений градиентов. Суть состоит в том, что большая величина градиентов соответствует краям и углам объекта, т.е. областям с резким изменением интенсивностей. Необходимые этапы для расчета дескриптора HOG были описаны выше в 1.1.3. На рис. 2 показан результат описания возможного входного изображения лица человека дескриптором HOG.

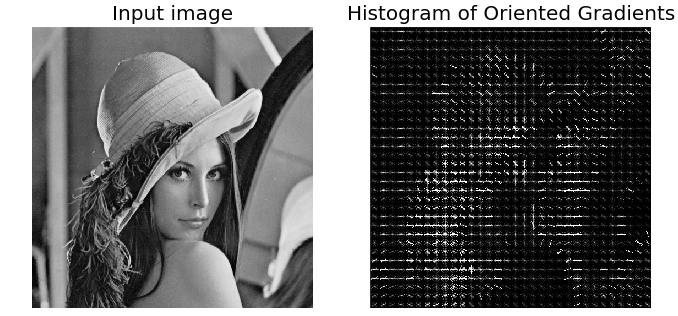


Рис. 2. HOG-дескриптор изображения человека

**Использование LBP-дескрипторов**

Существует метод описания ключевых точек лица на основе локальных бинарных паттернов (Local Binary Pattern, LBP). Оператор впервые был предложен для распознавания текстур в 1996 году Т. Ояла [13]. Метод с LBP описывает локальный текстурный паттерн с помощью бинарного кода, зависящего от интенсивностей соседних пикселей изображения. Классический LBP, применяемый к пикселю изображения, использует его окрестность из восьми пикселей. Он принимает яркость центрального пикселя за порог и сравнивает с ним значения яркости каждого пикселя из окрестности с ним. Если значение больше или равно пороговому значению, то пикселю присваивается значение 1, иначе – 0. В итоге получается 8-битное число, характеризующее окрестность пикселя. Пример работы оператора LBP на полутоновом изображении показан на рис. 3.

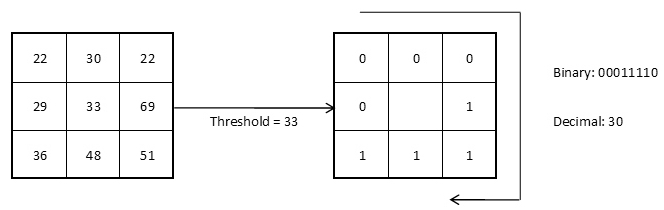


Рис. 3. Пример работы LBP-оператора

В модификациях LBP радиус окрестности может меняться так, что для описания используются пиксели, находящиеся на некотором расстоянии от центрального пикселя. В этом случае пиксели окрестности лежат на круге некоторого радиуса, а количество точек на этом круге задается произвольно. В работе [13] было предложено использовать для анализа изображения только те LPB, в которых присутствует не более двух переходов с «0» на «1» или наоборот. Такой подход, во-первых, требует меньше памяти, так как использует только часть LBP. И во-вторых, определяет только важные локальные текстуры, к которым можно отнести концы линий, края и углы, пятна на изображении. На рис. 4 приведен пример LBP, детектирующих некоторые особенности изображенных объектов, где пиксели со значением «0» помечены светлыми кружками, а со значением «1» - темными.

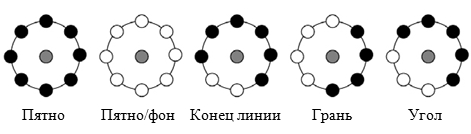


Рис. 4. Примеры локальных особенностей, детектируемых LBP

Применяя оператор LBP к каждому пикселю изображения, можно построить гистограмму, в которой каждому равномерному коду LBP соответствует отдельный столбец. Изображения лиц могут представляться набором всевозможных локальных особенностей, которые описываются LBP. А для учета информации о расположении найденных особенностей изображение разбивается на подобласти, в каждой из которых вычисляется своя гистограмма LBP. В результате объединения этих гистограмм получается общая гистограмма, учитывающая и локальные, и глобальные особенности изображения. Благодаря работе только с целочисленной арифметикой LBP эффективен в вычислительном плане, также оператор инвариантен к изменениям яркости изображения, связанными с условиями освещения при съемке. Однако вычислительная сложность LBP при описании ключевых точек лица является одним из недостатков, из-за которого этот дескриптор в данной работе отвергнут.

* + 1. **Поиск ключевых точек лица**

Задача поиска ключевых точек лица решается и при выравнивании изображений лиц. В качестве примера работы с такой задачей можно выделить статью В. Каземи и Дж. Салливана [14] 2014 года. Основная идея алгоритма, описанного в статье, заключается в выделении 68 ключевых точек, имеющихся на каждом лице, которые описывают края глаз, положение носа, выступающую часть подбородка и т. п. На поиск данных точек авторы статьи предлагают обучать ансамбль деревьев регрессии.

Деревом принятия решения называют дерево, в листьях которого содержатся значения целевой функции, а в остальных узлах – условия перехода, определяющие по какому из ребер двигаться [15]. Для наблюдения, условие для которого принимает значение «истина», осуществляется переход по левому ребру, иначе – по правому. При классификации в листьях дерева стоят результирующие классы. Отличие деревьев регрессии от них в том, что в листьях деревьев регрессии содержится уже какое-то определенное значение целевой функции.

Построение из деревьев регрессии ансамбля предлагается из-за соображения, что усреднение результатов работы деревьев регрессии может обеспечить более устойчивую и надежную оценку [16]. Это объясняется тем, что влияние случайных отклонений полученного деревом ответа от правильного ответа ослабляется. Поэтому комбинирование моделей зачастую ведет к повышению качества предсказаний.

В алгоритме поиска ключевых точек с помощью ансамбля деревьев регрессии возможно использование меньшего количества ключевых точек лица, чем 68, например 5. Это улучшает быстродействие алгоритма, но такое изменение может привести к потере точности, поэтому в данной работе процесс сегментации изображения лица проводится на основе информации о 68 точках.

**Глава 2. Этапы работы алгоритма сегментации**

В предыдущей главе были рассмотрены методы локализации областей лиц и способы описания ключевых точек. В данной главе перейдем к более детальному рассмотрению алгоритма, разработанного для решения поставленной задачи.

**2.1. Описание выбранных методов**

На этапе локализации областей лиц требуется получить -координаты всех лиц на изображении. В данной работе для локализации лиц предлагается остановиться на варианте с использованием предварительно обученного детектора HOG в связке с линейным методом опорных векторов. Найдя таким образом области лиц, можно переходить к следующему этапу – детектированию ключевых точек и частей на каждом локализованном лице.

Детектирование ключевых точек в разработанном алгоритме выполнено с помощью метода, основанного на ансамбле деревьев регрессии. При обучении детектора ключевых точек, описанного в статье [14], необходимы следующие предварительные данные:

1. Обучающий набор помеченных вручную ключевых точек на изображении, т.е. конкретный набор -координаты областей, окружающих каждую часть лица (рот, глаз, бровь и т.д.);
2. Заданные вероятности расстояния между парами входных пикселей.

Эти данные являются частью 68-точечного набора данных iBUG 300-W, на котором и обучался предиктор ключевых точек лица DLib. Можно отметить, что существуют и другие разновидности детекторов ключевых точек лица, включая модель из 194 точек, которую можно обучить на наборе данных HELEN. Однако 68 ключевых точек вполне достаточно для обнаружения искомых частей лица.

Учитывая эти обучающие данные, ансамбль деревьев регрессии обучается оценивать положения ключевых точек лица непосредственно по самим интенсивностям пикселей, т.е. извлечение признаков не происходит. Конечным результатом является детектор ключевых точек лица, который можно использовать для обнаружения ключевых точек лица в режиме реального времени с высококачественными предсказаниями.

Реализация выбранного метода детектирования ключевых точек лица из статьи [14] содержится в библиотеке DLib и рассчитана на поиск 68 ключевых точек, представленных на рис. 5.

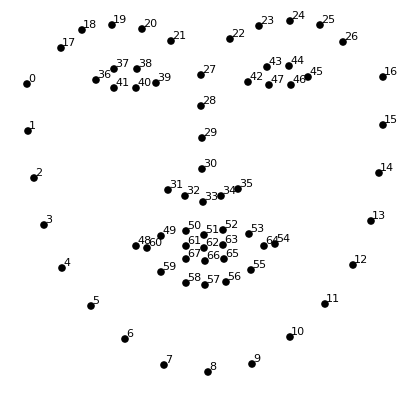


Рис. 5. 68 ключевых точек лица

Как видно на рис. 5, ключевым точкам можно присвоить определенные индексы, по которым впоследствии возможно их соотносить с определенными чертами лица. Соответствие каждой части лица человека и индексов ключевых точек в массиве представлены в таблице 1.

|  |  |
| --- | --- |
| **Часть лица** | **Индексы ключевых точек** |
| Линия подбородка | 0 – 16 |
| Правая бровь | 17 – 21 |
| Левая бровь | 22 – 26 |
| Нос | 27 – 35 |
| Правый глаз | 36 – 41 |
| Левый глаз | 42 – 47 |
| Рот | 48 – 68 |

Табл. 1. Соответствие ключевых точек частям лица

* 1. **Описание процесса сегментации**

В итоге предлагается следующий алгоритм сегментации изображения лица человека, представленный в виде блок-схемы на рис. 6.

Начинается алгоритм сегментации с того, что на вход подается изображение, которое затем обрабатывается. В процессе обработки меняются размеры изображения до ширины 500 пикселей, что является приемлемой шириной изображения для данного детектора лиц.

После этого на изображении локализуются все области лиц с помощью предварительно обученного детектора лиц библиотеки DLib, который является модификацией стандартной гистограммы ориентированных градиентов (HOG) и линейного метода опорных векторов (SVM) для обнаружения объектов. В результате этого получаются -координаты прямоугольников, ограничивающих области лиц.

Затем, учитывая найденные области лиц, происходит переход к обнаружению на каждом лице ключевых точек с помощью предиктора ключевых точек лица библиотеки DLib. Данный предиктор для каждого лица выявляет 68 -координат ключевых точек и формирует из них массив.

Для того чтобы фактически найти каждую часть лица, вычисляется прямоугольник, ограничивающий соответствующие точки. Вокруг каждой части лица алгоритм отображает рамку, координируясь индексами ключевых точек в массиве, представленными в табл. 1 параграфа 2.1.

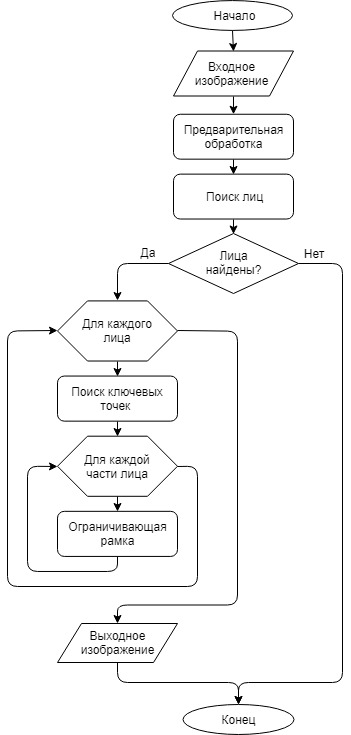


Рис. 6. Процесс сегментации изображения лица человека

Таким образом, как результат сегментации алгоритм отображает выходное изображение. Также по индексам ключевых точек можно извлечь каждую часть лица изображения человека по отдельности.

**Глава 3. Практическая реализация**

В данной главе будут описаны инструменты, используемые для решения задачи сегментации изображения лица человека, подробно описана реализация разработанного алгоритма и представлены результаты его работы.

**3.1. Инструменты разработки**

Для работы с изображениями была выбрана библиотека OpenCV по причине того, что она содержит большое количество алгоритмов компьютерного зрения и обработки изображений. А также используется библиотека машинного обучения DLib, включающая в себя реализации многих функций, среди которых обученный HOG-каскад для поиска лиц и готовое решение для поиска ключевых точек.

Для разработки программы выбор был сделан в пользу языка программирования Python, являющегося одним из самых популярных языков программирования в мире из-за ряда его очевидных преимущества, среди которых простой синтаксис и легкая читаемость кода. А множество встроенных библиотек и встроенных функций для обработки изображений значительно упрощают процесс разработки на данном языке программирования.

Запуск программы осуществлялся в виртуальной среде Python - Jupiter Notebook, которая представляет собой веб-приложение с открытым исходным кодом. Данная среда позволяет создавать и обмениваться документами, которые содержат код, формулы, изображения и сопровождающее описание. Функционал этого веб-приложения включает в себя преобразование данных, численное и статическое моделирование, визуализацию данных, машинное обучение и многое другое. Notebook поддерживает не только Python, но и другие языки программирования, в числе которых R, Julia и Scala.

**3.2. Результаты**

В данном параграфе продемонстрированы результаты разработанного алгоритма сегментации изображения лица с помощью описанных выше инструментов для случаев присутствия на изображении одного человека и нескольких.

**Случай изображения лица одного человека**

Промежуточный результат сегментации с найденными ключевыми точками лица на изображении одного человека показан на рис. 7.

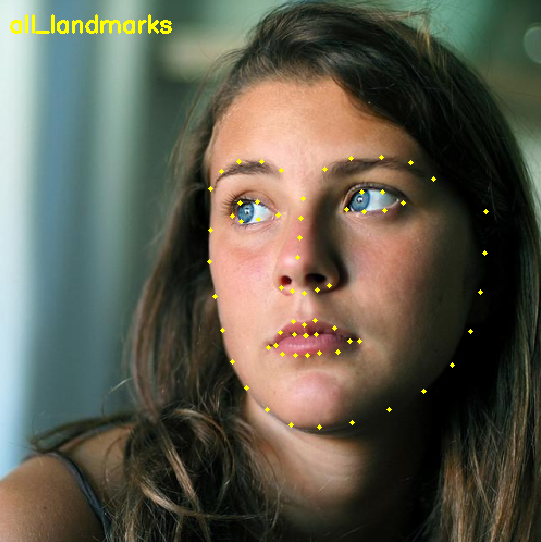


Рис. 7. Ключевые точки лица одного человека

На рис. 8 показан конечный результат сегментации изображения лица одного человека.

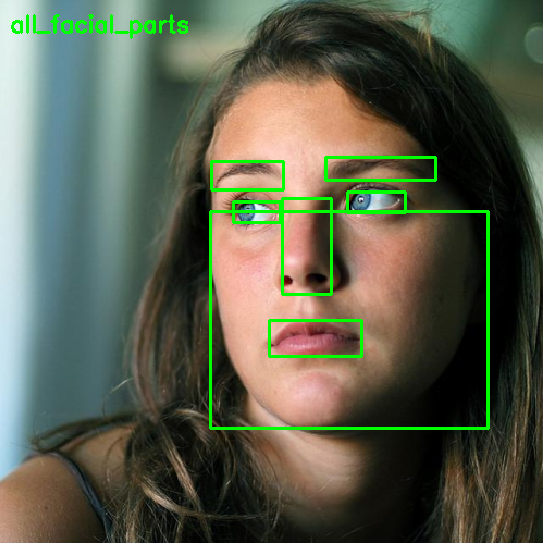


Рис. 8. Результат сегментации изображения лица одного человека

**Случай изображения лиц нескольких человек**

Промежуточный результат сегментации с найденными ключевыми точками лиц на изображении нескольких человек показан на рис. 9.

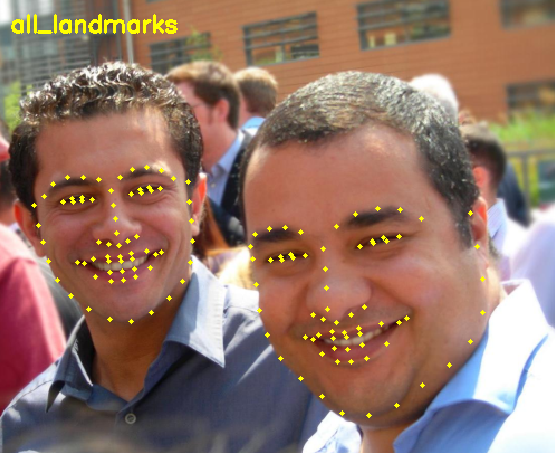


Рис. 9. Ключевые точки лиц нескольких человек

На рис. 10 показан конечный результат сегментации изображений лиц нескольких человек.

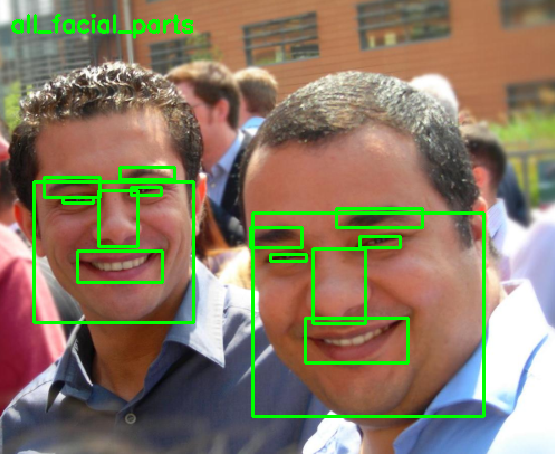


Рис. 10. Результат сегментации изображения лиц нескольких человек

На данных визуализациях видно, что выбранный алгоритм успешно справился с задачей сегментации. Также можно заметить, что ограничивающие части лица рамки отображаются вокруг глаз, бровей, носа и рта, а рамка вокруг линии подбородка ограничивает все лицо за исключением области лба.

**Оценка точности**

Точность разработанного алгоритма сегментации изображений лиц человека проверена на тестовом наборе из 130 изображений из набора данных [17]. Точность сегментации лица на выборке оказалась 0,85.

Еще одна оценка точности сегментации на тестовом наборе изображений может быть вычислена по следующей формуле

,

где точность сегментации равна отношению пересечения найденной алгоритмом области с областью, выявленной на глаз человеком, к объединению этих же областей. На основании этой формулы выявлена примерная величина, равная 0,8.

На основании полученных данных можно сделать вывод, что реализованный алгоритм работает успешно и результат сегментации не зависит от количества лиц людей на изображении.

**Выводы**

В данной работе реализован алгоритм сегментации изображения лица человека. В первой главе рассмотрены методы локализации лиц и описания ключевых точек. В первом параграфе первой главы в результате оценки методов локализации областей лиц было выбрано решение, основанное на гистограмме направленных градиентов и методе опорных векторов. Во втором параграфе первой главы приведены способы описания ключевых точек и предложено проводить поиск ключевых точек лица с использованием ансамбля деревьев регрессии, реализация которого присутствует в библиотеке машинного обучения DLib. Во второй главе подробнее рассмотрен разработанный алгоритм сегментации изображений лиц и приведена его блок-схема. В третьей главе продемонстрированы результаты метода и инструменты разработки, с помощью которых они были получены.

**Заключение**

Таким образом, в данной работе проведен анализ существующих алгоритмов, которые могут использоваться в процессе сегментации изображения лица человека. При анализе эффективности каждого метода были выбраны те из них, которые можно считать наиболее приемлемыми для решения поставленной задачи. На основе выбранных средств разработана программная реализация, осуществляющая сегментацию изображения лица одного человека или лиц нескольких человек с использованием языка программирования Python и библиотек OpenCV и DLib. Тестирование разработанной системы показало результаты, которые можно считать успешными для поставленной цели. В частности, с помощью разработанного алгоритма возможно обнаружить и извлечь следующие части лица: глаза, брови, носи, рот и линию подбородка. Данный алгоритм впоследствии может применяться в задачах распознавания эмоций, определения моргания, анимации и автоматического редактирования изображений лиц людей.

**Список литературы**

1. Интернет в цифрах: мировая статистика за 2012 год. https://habr.com/ru/company/webnames/blog/166457/
2. Шапиро Л. Компьютерное зрение/ под ред.: С. М. Соколов, пер.: А. А. Богуславский, Л. Шапиро. М.: БИНОМ. Лаборатория знаний, 2013. 762 с.
3. David A. Forsyth, Jean P. Computer vision: a modern approach. Pearson. 2011. 792 p.
4. Анализ существующих подходов к распознаванию лиц. https://habr.com/ru/company/synesis/blog/238129/
5. Viola P., Jones M. Robust real-time face detection // International Journal of Computer Vision. 2004. Vol. 57. No 2. P. 137–154.
6. Метод Виолы-Джонса (Viola-Jones) как основа для распознавания лиц. https://habr.com/ru/post/133826/
7. Хайкин С. Нейронные сети: полный курс // Издательство «Вильямс». 2006.
8. Нейронные сети для начинающих. Часть 1. https://habr.com/ru/post/312450/
9. Dalal N., Triggs B. Histograms of oriented gradients for human detection. https://ieeexplore. ieee.org/document/1467360 (дата обращения: 25.07.05).
10. Воронцов К. В. Лекции по методу опорных векторов. http://www.ccas.ru/voron/download/SVM.pdf
11. Dhananjay R., Vinay A, Shylaja S. Facial landmark localization. // International Journal of Current Engineering and Technology. 2014. Vol. 4. No 3. P. 1901-1907.
12. Learn OpenCV. Histogram of oriented gradients. https://www.learnopencv.com/histogram-of-oriented-gradients/
13. Ojala T., Pietikäinen M., Harwood D. A comparative study of texture measures with classiﬁcation based on feature distributions // Pattern Recognition. 1996. Vol. 29. No 1. P. 51–59.
14. Kazemi V., Sullivan J. One millisecond face alignment with an ensemble of regression trees // The IEEE Xplore. 2014. P. 1867–1874.
15. Классификация и регрессия с помощью деревьев принятия решений. https://habr.com/ru/post/116385/
16. Шитиков В. К., Мастицкий С. Э. Классификация, регрессия и другие алгоритмы Data Mining с использованием R. https://github.com/ranalytics/data-mining
17. Localizing Parts of Faces Using a Consensus of Exemplars / Proceedings of the 24th IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Peter N. Belhumeur, David W. Jacobs, David J. Kriegman, Neeraj Kumar. June 2011.