

МИНИСТЕРСТВО НАУКИ И ВЫСШЕГО ОБРАЗОВАНИЯ РОССИЙСКОЙ ФЕДЕРАЦИИ  
федеральное государственное бюджетное образовательное учреждение  
высшего образования  
«Тольяттинский государственный университет»

«Институт математики физики и информационных технологий»  
(наименование института полностью)

---

Кафедра «Прикладная математика и информатика»  
(наименование)

02.03.03 Математическое обеспечение и администрирование информационных систем  
(код и наименование направления подготовки, специальности)

---

Web-дизайн и мультимедиа  
(направленность (профиль) / специализация)

---

## **ВЫПУСКНАЯ КВАЛИФИКАЦИОННАЯ РАБОТА (БАКАЛАВРСКАЯ РАБОТА)**

на тему Разработка алгоритма обнаружения лица человека в видеопотоке

Студент Е.А. Горбулёв \_\_\_\_\_  
(И.О. Фамилия) (личная подпись)

Руководитель кандидат технических наук, доцент, А.Б. Кузьмичев \_\_\_\_\_  
(ученая степень, звание, И.О. Фамилия)

Консультант \_\_\_\_\_  
М.В. Дайнеко

## Аннотация

Выпускная квалификационная работа по теме «Разработка алгоритма обнаружения лица человека в видеопотоке» содержит 56 страниц текстового документа, 25 рисунков, 24 использованных источника, 4 таблицы и 1 приложение, которое представлено на диске.

Работа состоит из введения, 3 глав, заключения и приложения.

Во введении приводится актуальность проводимого исследования, формулируется цель и ставятся задачи, которые необходимо решить.

В первом разделе производится анализ существующих методов обнаружения лица человека на цифровых изображениях.

Во втором разделе описывается разработка алгоритма обнаружения лица человека в видеопотоке.

В третьем разделе приводятся результаты тестирования разработанного алгоритма и демонстрируются результаты работы.

В заключении представлены выводы по проделанной работе.

Объектом работы является алгоритм обнаружения лиц в видеопотоке.

Целью ВКР является создание программного алгоритма, позволяющего обнаруживать лица на кадрах транслируемого видеопотока в реальном времени на основе метода Виолы-Джонса.

Актуальность данной работы заключается в том, что в последние годы технологии компьютерного зрения начали применять не только в охранных системах, но и в сфере кино, искусства, видеоигр и дополненной реальности. И сфера влияния продолжает расти, затрагивая всё новые аспекты жизни человека.

Данный проект реализован при помощи метода Виолы-Джонса для поиска лиц на изображении. Код программы написан в среде Matlab, использовалась бесплатная профессиональная пробная версия.

## Abstract

This graduation work is devoted to developing an algorithm for detecting faces in video streams in real time.

The aim of the graduation work is to develop an algorithm for detecting faces in a video stream.

The object of the research is the methods and algorithms for detecting a person's face in a video stream.

The subject of the investigation is the Viola-Jones method.

The graduation work describes in details the implementation of this method in Matlab based on the OpenCV library, as well as covers its algorithm based on Haar cascades and AdaBoost adaptive learning.

To achieve the aim set, the following tasks have to be solved:

1. to analyze the algorithms for detecting faces;
2. to select the most appropriate method to develop an algorithm;
3. to ensure the software implementation of the suggested algorithm.

The graduation work may be divided into several logically connected parts.

In the first chapter, much attention is paid to the methods for detecting and recognizing a person's face. The considered algorithms are based on contour models, comparison with a template and training. This chapter also presents their mathematical implementation, as well as explains the advantages and disadvantages.

The second chapter is devoted to developing a detection algorithm based on the Viola-Jones method. The chapter also describes the applied image preprocessing filters, such as the median filter, the bilateral filter, and the grayscale image transfer filter.

The third chapter dwells on the work and analysis of the developed algorithm. This chapter also discusses the results of the occurrence probability of FN and FP errors, the percentage being 2% and 6%. Testing the preprocessing filters shows a significant increase in the the algorithm effectiveness.

In conclusion, we would like to highlight that a comparison of the numerical test results confirms that the Viola-Jones method is one of the most effective modern methods, and the applied preprocessing methods help to increase its effectiveness.

## Содержание

Введение.....	6
1 Анализ методов обнаружения лица человека на цифровых изображениях .....	7
1.1 Обзор существующих методов обнаружения лиц на изображениях .....	7
1.2 Анализ и выбор метода для обнаружения лица .....	19
2 Разработка алгоритма обнаружения лица человека в видеопотоке .....	26
2.1 Разработка алгоритма обучения .....	26
2.2 Интегральное представление изображения.....	28
2.3 Алгоритм работы сканирующего окна .....	30
2.4 Разработка алгоритма обучения классификаторов AdaBoost.....	34
2.5 Разработка алгоритма обнаружения.....	38
3 Тестирование разработанного алгоритма .....	45
3.1 Оценка разработанного алгоритма обнаружения лица в видеопотоке.....	45
3.2 Оценка работы алгоритма с применением фильтров предобработки .....	48
Заключение .....	52
Список используемой литературы.....	54

## Введение

В мире быстро развивающихся технологий не найдётся такого предприятия, на котором не установлены в качестве охранных средств безопасности системы контроля и управления доступом. И если прежде для идентификации сотрудников на пропускном пункте предприятия размещали турникеты с устройством сканирования карт или пропусков, то на сегодняшний день, благодаря активному развитию нейросетей и биометрических систем, компании всё больше устанавливают иные, более современные методы обнаружения и распознавания, которые точнее, удобнее и, главное, надёжнее[4].

В данной работе рассматриваются методы обнаружения лица человека на цифровых изображениях и видеопотоке в реальном времени.

Область применения задачи обнаружения имеет широкий спектр, затрагивающий не только системы безопасности, но и социальные сети и даже видеоигры с дополненной реальностью. Задача обнаружения лиц становится одним из главных направлений развития технологий машинного обучения и компьютерного зрения[1].

Предметом дипломной работы является обработка видеопотока для обнаружения лица человека.

Целью работы была установлена разработка алгоритма для выделения лица человека на кадрах видеопотока, основываясь на указанном ранее методе, а также увеличение эффективности работы алгоритма путём добавления методов предобработки изображений.

Для реализации работы требуется выполнить следующие задачи:

1. Проанализировать методы обнаружения лиц;
2. Определить наиболее подходящий метод для реализации;
3. Разработать алгоритм обнаружения лица на видеопотоке;
4. Провести тестирование алгоритма.

# 1 Анализ методов обнаружения лица человека на цифровых изображениях

## 1.1 Обзор существующих методов обнаружения лиц на изображениях

Создано огромное множество разнообразных алгоритмов обнаружения лиц, некоторые из которых содержат в себе другие алгоритмы. Некоторые доступны в виде математических моделей и их описания, а другие в виде программной реализации. Например, библиотека алгоритмов компьютерного зрения OpenCV – библиотека с открытым кодом.

В целом, методы обнаружения можно разделить на две категории: методы на основе знаний, которые основополагаются на опыт человека (эмпирические методы) и методы по внешним признакам. На рисунке 1 представлена классификация данных методов[3].



Рисунок 1 – Методы обнаружения лиц

Эмпирические методы обнаружения строятся на правилах, которые основываются на известных знаниях человека о лице, таких как его черты, форма, цвет кожи, размер и форма глаз, носа, рта. Чтобы фрагмент анализируемого изображения мог считаться человеческим лицом, он должен удовлетворять некоему набору правил, который составлен на основе выше указанных принципов. Определить набор правил не составит труда[2].

Правила – это формальные знания, полученные на опыте, которым руководствуется человек для определения, находится лицо перед ним или нет. Например:

- Глаза всегда располагаются симметрично относительно друг друга, между ними нос, а линия рта горизонтальна;
- Цвет этих частей лица сильно отличается по яркости относительно остальной области.
- Яркость и цвет центральной части лица однородные;
- Соотношение размеров частей лица между собой.

Полагаясь на эти и другие характеристики, составляются алгоритмы, выявляющие на полученном изображении наличие данных правил.

Например, учёные Yang и Huang составили систему из трёх правил на основе многомасштабной иерархии, которая создаётся благодаря усреднению и субсэмплированию изображения[24]. Пример иерархии приведён на рисунке 2.



Рисунок 2 – Многомасштабная иерархия



Изначально изображение проверяется на первом уровне с самым низким разрешением на наличие подходящих для дальнейшей работы кандидатов. Затем, на втором уровне, выравниваются гистограммы и выделяются области лица изображений на основе полученных сведений. На третьем уровне сканирующее окно на основе полученных границ проверяет возможные лица. Сканирующее окно следует известным правилам, описывающим лицо.

Данный метод сильно уменьшает вычислительные операции за счёт уменьшения изображений.

Ещё один эмпирический метод разработали ученые Kotropoulos и Pitas.

В данном методе строятся вертикальные и горизонтальные гистограммы в предполагаемой области, содержащей лицо. Построение осуществляется на основе формул:

$$HI(x, y) = \sum_{y=1}^n I(x, y), \quad (1)$$

$$VI(x, y) = \sum_{x=1}^m I(x, y), \quad (2)$$

где  $I(x, y)$  – интенсивность изображения в точке  $(x, y)$ .

Строится горизонтальная гистограмма изображения, определяются локальные минимумы на основе резких изменений в  $HI(x, y)$ , которые и показывают края лица. Таким же образом строятся вертикальные гистограммы со своими локальными минимумами, которые показывают области расположения рта, носа и глаз. На основе всех локальных минимумов и высчитывается возможная область, содержащая лицо [18]. На рисунке 3 приведён пример работы метода:

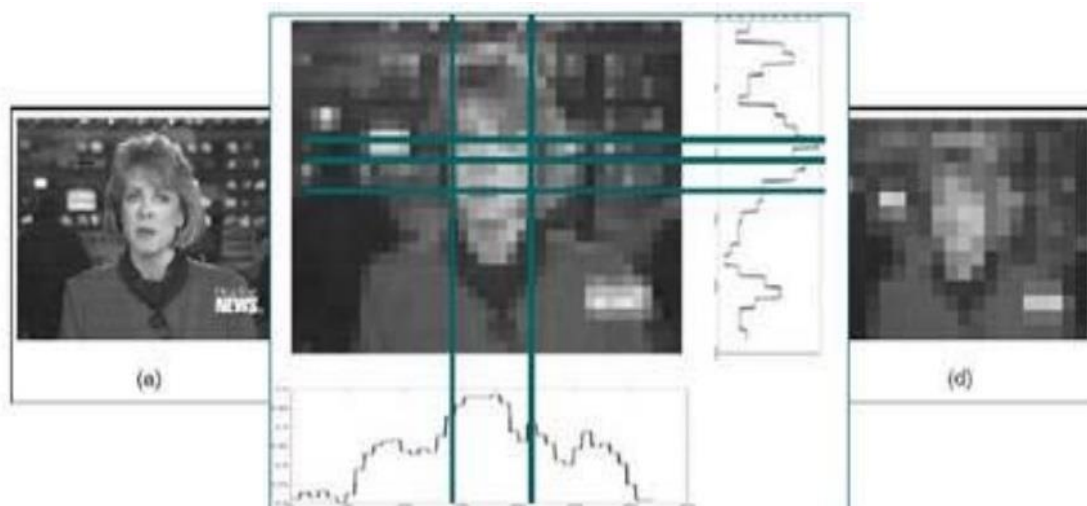


Рисунок 3 – Построенные вертикальные и горизонтальные гистограммы

Методы сравнения с шаблоном также относятся к методам, основанным на знаниях и являются наиболее простыми для обнаружения лица на получаемом изображении. Они основываются на анализе всего изображения локальной маской, и на её основе образ в поле маски сопоставляется с шаблоном, хранящимся в базе. В свою очередь шаблон- это изображение установленного размера с размещённым по центру лицом человека в анфас.

Шаблоны бывают двух видов: недеформируемые и деформируемые. В то время как шаблоны первого типа неизменны (за исключением масштаба), шаблоны второго типа подстраиваются под входные данные, выделяя характерные части лица человека.

Для определения соответствия маски шаблону используют так называемую среднеквадратичную ошибку – СКО. Если этот показатель минимален, то достигнуто наибольшее соответствие области изображения под маской эталону.

Пусть функция  $f(x,y)$  обозначает изображение размером  $M$  на  $N$ , тогда функция  $w(x,y)$  описывает шаблон размера  $H$  на  $W$ , размеры которых меньше или равны размерам изображения. По формуле

$$c(x, y) = \sum_s \sum_t f(s, t)w(x + s, y + t), \quad (3)$$

где  $x = 0, 1, 2, \dots, M-1$ ,  $y = 0, 1, 2, \dots, N-1$ , находится корреляция в той области, где шаблон совпадает с изображением.

Рисунок 4 демонстрирует процесс корреляции. Окно сканирования (шаблон) располагается в левой верхней части полученного на входе изображения и двигается вниз до тех пор, пока находится в области изображения.

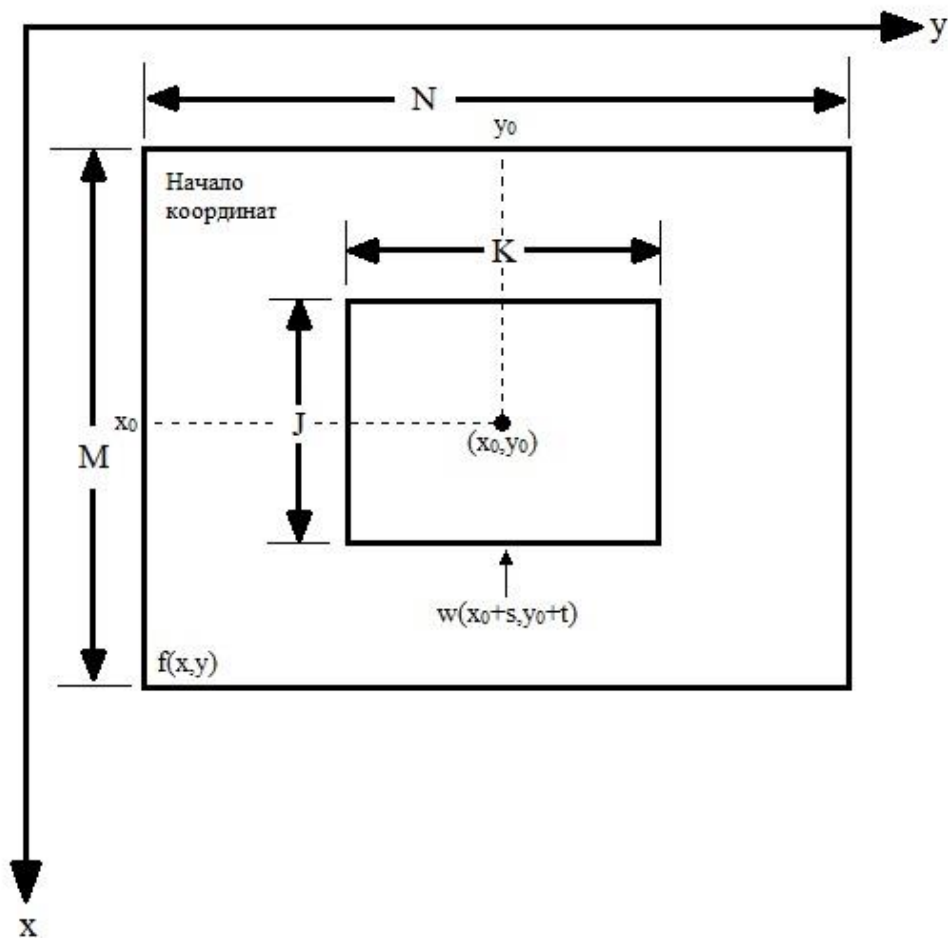


Рисунок 4 – Процесс вычисления корреляции

Вычислив корреляцию для точки  $(x_0, y_0)$ , получим единственное значение  $c(x_0, y_0)$ . Когда изменяются значения  $x$  и  $y$ , сканирующее окно перемещается по изображению, получая функцию  $c(x, y)$ , максимумы которой обозначают точки

наибольшего совпадения шаблона и исходного изображения. Минусом является то, что на краях анализируемой области изображения эффективность снижается, так как чем больше область шаблона, вышедшая за границы исходного изображения, тем меньше вероятность корреляции[4].

Так же ещё одним минусом является уменьшение эффективности при изменениях амплитуды шаблона и самого изображения. Для того, чтобы избавиться от этой проблемы используют НККФ – нормированную кросскорреляционную функцию, которая является независимой к изменениям изображения. НККФ выглядит следующим образом:

$$\gamma(x, y) = \frac{\sum_s \sum_t [f(s,t) - \bar{f}(s,t)] [w(x+s,y+t) - \bar{w}]}{\sqrt{\sum_s \sum_t [f(s,t) - \bar{f}(s,t)]^2 \sum_s \sum_t [w(x+s,y+t) - \bar{w}]^2}}, \quad (4)$$

где  $x = 0, 1, 2, \dots, M-1$ ,  $y = 0, 1, 2, \dots, N-1$ ,  $\bar{w}$ - это среднее значение пикселей в шаблоне,  $\bar{f}$ - среднее значение элементов в области текущего положения шаблона.

Хоть данный метод и показывает отличные результаты, но существенным недостатком является необходимость равенства размеров эталона и искомого объекта. Данная проблема решается тем, что процесс сканирования изображения проводится некоторое количество раз, при каждом сканировании уменьшая размер анализируемого изображения, что демонстрируется на рисунке 5.

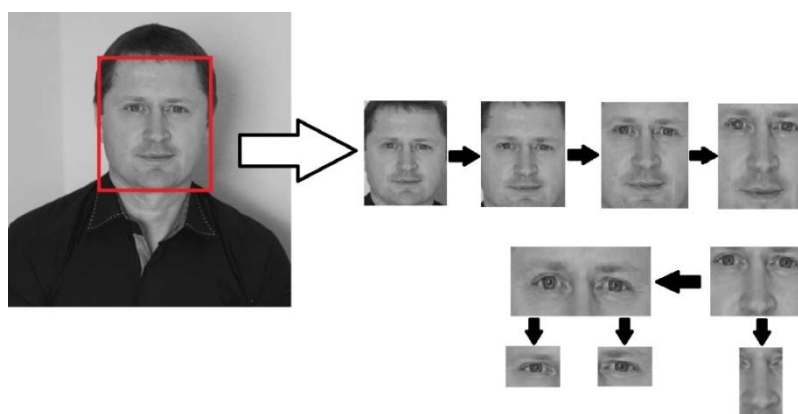


Рисунок 5 – Эталоны, вырезанные из исходного изображения

Из исходного изображения были получены эталоны, описывающие характерные неизменяемые черты лица, такие как глаза, но и брови.

На рисунке 6 представлен фрагмент изображения с лишним шумом на фоне, шаблон и результат выделения лица, а также график НККФ с наблюдаемым максимумом корреляции.

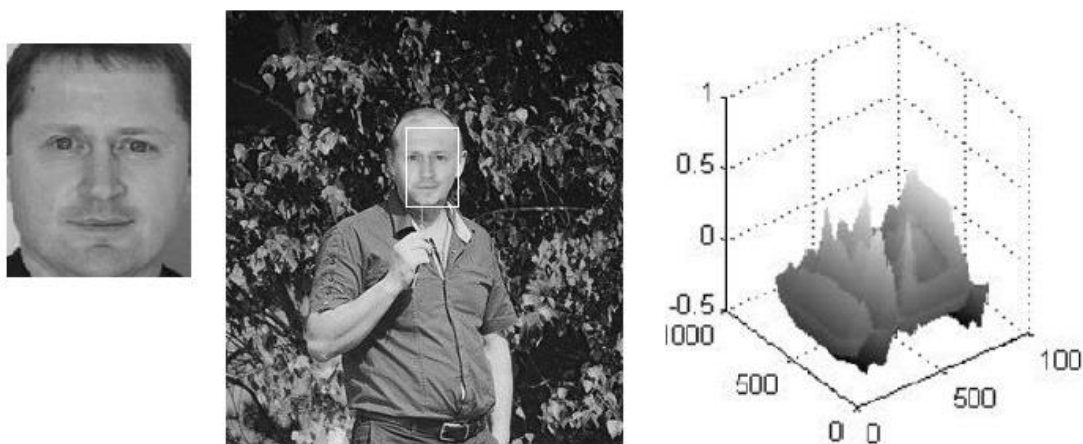


Рисунок 6 – Работа алгоритма сравнения с шаблоном

Как уже указывалось выше, чем ближе значение НККФ к 1, тем точнее выделяется лицо на изображении. Но чем более неоднороден фон и освещение, в сравнении с шаблоном, тем более вероятно возникновение ошибки работы метода.

Метод довольно широко используется, но характеризуется высокой вычислительной сложностью из-за минусов, приведённых ранее.

К сожалению, при всей своей работоспособности и низких требованиях, данные методы не очень эффективны, так как показывают хороший результат только на изображениях без лишнего загрязняющего шума, с благоприятным освещением, высоким качеством и фоном, который не меняется. В общественных местах, где не будет фиксированного ракурса и поворота лица, а также освещение и задний план часто меняются, вероятность ошибок и ложного распознавания увеличивается во много раз. Благодаря этому они не страдают от

проблем методов, основанных на знаниях, из-за чего надёжны и более распространены охранных системах видеонаблюдения[1].

Анализируемые далее методы на основе характерных внешних признаков не формализуют в явном виде происходящие в человеческом мозге процессы, а, наоборот, выделяют зависимости и характеристики лица неявно на основе математических методов и методов машинного обучения[15].

Рассмотрев эмпирические методы обнаружения лица, можно прийти к следующим основным минусам:

1. Методы, основанные на знаниях, трудно реализуемы из-за того, что сложно составлять работающие правила на основе неких знаний человека об устройстве лица;
2. Задача ещё более усложняется, если анализируемое изображение имеет шум, сложный фон, плохое качество или немонотонное освещение;
3. Методы сравнения с шаблоном лучше всего работают, если фон исходного изображения однородный, что не всегда выполнимо в условиях реального использования.

Эффективность рассматриваемых далее методов на основе внешних признаков не зависит от условий, при которых страдают предыдущие методы.

Для методов обнаружения по характерным внешним признакам необходимо предварительное обучение системы путём тестирования на предварительно подготовленных наборах образцов. Методы обладают избыточностью и большой вычислительной сложностью за счёт большого объёма работы. Чтобы ускорить процесс нахождения лиц и упростить процесс работы, применяются различные дополнительные вспомогательные методы для уменьшения числа рассматриваемых фрагментов изображения, например - каскадные классификаторы[22].

Одним из самых эффективных методов на основе обучения является метод Виолы-Джонса. Алгоритм показывает отличные результаты и высокую работоспособность, а также верно обнаруживает лица, повернутые вплоть до

угла в 30 градусов[8]. Именно этот метод используется в основе выполненной работы.

Метод Виолы-Джонса работает на следующих алгоритмах:

1. Обнаружение изображения человека на основе признаков Хаара;
2. Каскадная классификация для моментального исключения фрагментов изображения, не содержащих лицо;
3. Приведение полученного кадра к интегральному виду для вычисления характерных точек лица;
4. Классификатор признаков сразу даёт результат «верно» либо «ложь»;
5. Использование метода обучения – бустинга, для выбора признаков.

Для большей точности при работе с видеопотоком в данном методе используется алгоритм приведения изображения к интегральному виду и расчёту суммы интенсивности всех пикселей при помощи так называемого сканирующего окна. Это позволяет создать матрицу, равную по размерам полученному кадру-изображению и содержащую в себе элементы, представляющие собой сумму яркости пикселя и яркости пикселей слева от него и над ним[5].

Так как алгоритм основан на обучении классификатора, то нуждается в обучающем наборе, состоящем из изображений с лицом и без лица. Причём в обучающей выборке изображения, содержащие лица, маленького разрешения. При работе алгоритма это поможет верно обрабатывать изображения с низким разрешением. Также изображения обучающей выборки разнообразны по освещению, фону, выражению лица, содержат людей разного гендера и цвета кожи. Изображения без человека имеют тот же размер, что и с человеком и их количество в выборке многократно больше[19].

Преимущества метода:

- Скорость распознавания;

— Метод показывает большую эффективность даже при повороте исследуемого лица (эффективность сильно снижается при увеличении угла поворота более 30 градусов).

Недостатки:

- Длительное время на анализ тестовых изображений;
- Имеются ограничения на положение обнаруженного лица.

*Метод гибкого сравнения на графах (Elastic graph matching).* Метод относится к методам 2D моделирования и его алгоритм работает, основываясь на сравнении графов, характеризующих лица: область лица представляется в виде сетки с персональным размещением вершин и рёбер. Создаётся также шаблонный граф, определяющий основную характеристику распознавания. Он не изменяется, в то время как остальные созданные графы искажаются под влиянием структуры лица, но при этом сохраняют свою привязку к главным точкам строения лица, например: промежуток между ушами, расстояние между глазами, положение линии носа, ширина и длина губ и т.п. Точность распознавания объекта увеличивается с увеличением антропометрических точек, но также увеличивается и время на обработку одного объекта.

Недостатки метода:

- Требуется значительных вычислительных мощностей из-за сложности алгоритма распознавания;
- Введение новых шаблонов в базу данных довольно трудозатратно;
- Быстродействие на прямую зависит от размера зарегистрированной базы.

*Метод главных компонент.* Задача данного метода заключается в уменьшении количества признаков без существенной потери данных и эффективности, причем так, чтобы признаки наиболее точно описывали самые характерные черты, свойственные множеству лиц. Данный метод применяется главным образом для представления лица как вектора малой размерности для дальнейшего сравнения с векторами, хранящимися в базе[9].



Для анализа новых изображений лиц, которые можно представить в виде взвешенной комбинации собственных векторов, применяется выборка векторов, вычисленных однажды на обучающем наборе. Сжатые и трансформированные изображения лица можно хранить в БД в виде вектора, по значению которого, в случае необходимости, можно найти конкретное изображение лица и реконструировать в нормальный вид.

Преимущества:

- Хранение и поиск изображений в больших БД;
- Точность распознавания более 90 %.

Недостатки метода:

- Показатели эффективности падают при изменении мимики лица;
- Сильные отличия в освещении и фоне обрабатываемой фотографии от хранящейся в базе приводят к снижению эффективности;
- Вычислительная сложность.

*Нейросетевые методы.* Набирающие популярность методы, которые включают в себя множество других различных алгоритмов. Характерной особенностью нейросетей можно назвать их удивительную обучаемость на заранее занесённых в базу данных наборах с примерами. Обучаясь, нейронная сеть сама выделяет ключевые характеристики и строит связи между признаками. Для распознавания нового объекта уже обученная нейросеть использует полученный на примерах опыт[20].

Нейросети показывают наиболее эффективные результаты в области распознавания объектов. Ранее они считались самыми сложными в реализации, что справедливо и сейчас, но успехи в данной области делают большие шаги вперёд в последние годы.

Преимущества данного метода:

- Очень высокий показатель обнаружения и распознавания.

Недостатки:

- Трудная архитектура сети, включающая множество нейронов, слоев и различные характеры связи;
- Сложность вычислений;
- Сложная процедура внесения изменений, требующая переобучения сети.

*Метод опорных векторов.* Метод предназначен для решения многих задач регрессионного анализа и классификации.[10] Это набор алгоритмов обучения с учителем, предназначенных для нахождения гиперплоскости в пространстве признаков, которая отделяет изображения на классы с наличием лиц и без лиц. Из всех найденных гиперплоскостей между этими двумя классами выбирается та, которая максимально удалена от каждого из классов одновременно. Суть алгоритма – найти гиперплоскость, разделяющую данные на классы.

Метод предназначен для снижения размерности пространства признаков, при этом не приводя к заметной потере информативности обучающей выборки.

Преимущества метода:

- В сравнении с нейросетями работает быстрее;
- Большой уровень устойчивости к переобучению;
- Вариативность между повышенной точностью и уровнем снижения шума.
- Недостатки:
- Небольшая точность (относительно других методов);
- Требовательность к вычислительной мощности.

*Скрытые Марковские модели.* Статистическое сравнение объектов с имеющейся базой данных. Задачей является разгадывание неизвестных ранее характеристик на основе уже наблюдаемых. Метод основывается на статистических свойствах сигналов и их пространственных характеристиках. Работа СММ состоит из: множества состояний, начального распределения вероятностей, скрытых состояний, матрицы вероятности генерации наблюдения и матрицы вероятностей переходов между состояниями[13].

Общий алгоритм работы заключается в следующем: вначале строится Марковская модель для всех элементов. При поиске лица на изображении

проверяются все полученные Марковские модели для имеющейся выборки. Далее все построенные модели проверяются и вычисляется самая большая из наблюдаемых вероятностей того, что набор наблюдений для анализируемого объекта создан соответствующей моделью и совпадает с ней.

Можно выделить следующие минусы в работе модели:

- Не самый эффективный алгоритм обучения;
- Низкая скорость срабатывания и точность распознавания;

## 1.2 Анализ и выбор метода для обнаружения лица

Таким образом, были выявлены достоинства и недостатки каждого алгоритма, которые приведены в таблице 1 [14].

Таблица 1 – Достоинства и недостатки методов обнаружения лиц

Метод	Достоинства	Недостатки
Эмпирические методы	Относительная простота; Низкие вычислительные требования;	Зависимость от таких факторов на изображении, как освещение, сложный фон, шум, мимика лица.
Метод Виолы-Джонса	Точность распознавания более 90 %; Низкий процент ложных срабатываний; Скорость распознавания; Низкое влияние мимики лица.	Длительное время на анализ тестовых изображений; Имеются ограничения на положение обнаруженного лица.
Метод гибкого сравнения на графах	Устойчивость к смене ракурса.	Введение новых шаблонов в базу данных довольно трудозатратно; Вычислительная сложность;

		Зависимость времени работы от размера базы данных.
Метод главных компонент	Точность распознавания более 90 %; Хранение и поиск изображений в больших базах данных.	Показатели эффективности падают при изменении мимики лица; Сильные отличия в освещении и фоне обрабатываемой фотографии от хранящейся в базе приводят к снижению эффективности; Вычислительная сложность
Метод опорных векторов	Большой уровень устойчивости к переобучению; Вариативность между повышенной точностью и уровнем снижения шума.	Чувствителен к шумам и стандартизации данных Небольшая точность (относительно других методов); Требовательность к вычислительной мощности.
Скрытые Марковские модели	Точность распознавания 90-95 %;	Не самый эффективный алгоритм обучения; Низкая скорость срабатывания и точность распознавания;
Нейросетевые методы	Очень высокий показатель обнаружения и распознавания.	Трудная архитектура сети, включающая множество нейронов, слоев и различные характеры связи; Сложность вычислений; Сложная процедура внесения изменений, требующая переобучения сети.

Также была составлена таблица, показывающая зависимость методов от характеристик обрабатываемого изображения[14].

Таблица 2 – Зависимость методов обнаружения от характеристик изображения

Метод	Осве- ще- ние	Ра- кур с	Выраж е-ние лица	Акс ессу ары	Шу м	Разр еше ние
Эмпирически методы	+	+	+	+	+	-
Метод Виолы - Джонс а	-	+	-	-	-	?
Метод гибкого сравнения на графах	?	-	-	?	+	?
Метод главных компонент	+	+	+	+	-	+
Метод	+	-	-	?	+	?

опорных векторов						
Скрытые Марковские модели	?	?	+	+	-	-
Нейронные сети	-	-	+	-	-	+

Для таблицы установлены обозначения:

- «+» –зависит;
- «-» – не зависит;
- «?» – зависимость не определена.

Выбирать метод для решения поставленной задачи исходя из цифровых показателей качества не целесообразно, так как эффективность зависит от условий, в которых должен работать алгоритм.

В качестве условий, влияющих на выбор метода, использовались следующие параметры:

- Расположение лица на видеопотоке: угол наклона и поворота;
- Количество лиц: одно или множество;
- Разнообразие людей: тип лица, цвет кожи, наличие растительности на лице;

- Масштаб, разрешение и качество изображения;
- Освещение: монотонное, резко меняющееся, фиксированное;
- Наличие фона: фиксированный, контрастный, зашумлённый.

Для работоспособности алгоритма был составлен список требований, влияющих на его эффективность:

1. Желательно фронтальное расположение лица в кадре;
2. Отсутствие факторов, перекрывающих лицо;
3. Отсутствие излишних эмоций;
4. Одинаковое освещение и фон для изображений тестовой выборки.

В связи с чем для реализации данной работы было принято решение использовать наиболее подходящий и эффективный в большинстве условий метод Виолы-Джонса для обнаружения лица в видеопотоке.

Метод и по сей день показывает хорошую эффективность, скорость и точность, а также относительно низкую требовательность к вычислительной мощности.

Помимо этих правил, было принято решение реализовать в разрабатываемой системе для повышения эффективности работы дополнительные алгоритмы обработки лиц, помогающие алгоритмам обнаружения и распознавания.

Для оценки работы алгоритма были применены критерии оценки - ошибки ложного отказа и ложного принятия (FN и FP). Подробнее о критериях оценки и результата тестирования в 3 главе.

Алгоритм показал отличные результаты на изображениях с лишним шумом, но были выявлены ошибки ложного обнаружения (FP), вероятность которых составила 11,91%. Результат обнаружения лица продемонстрирован на рисунке 7.

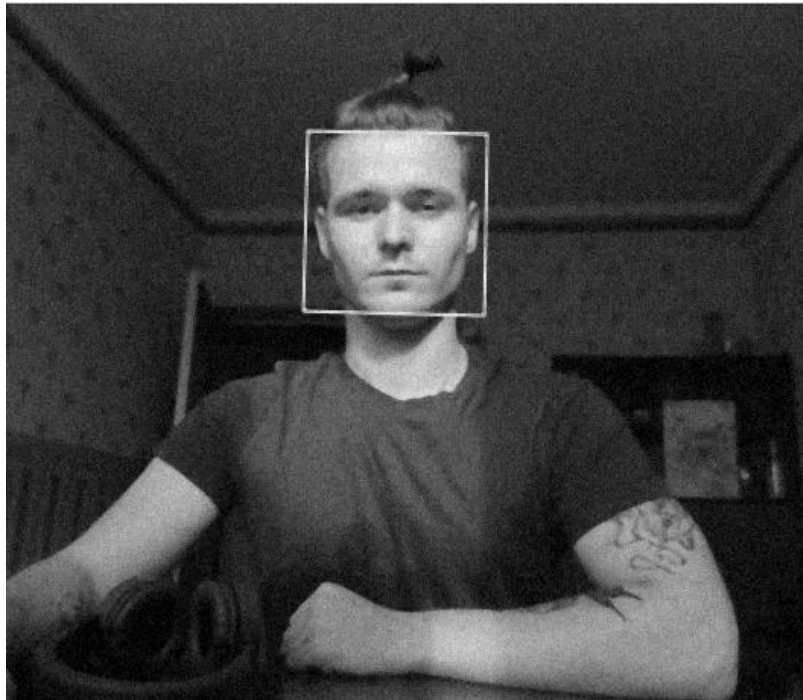


Рисунок 7 – Результат работы метода обнаружения

Для того, что бы снизить вероятность появления ошибок на зашумлённых кадрах видеопотока и для повышения эффективности работы алгоритма, было принято решение включить в программу дополнительные алгоритмы предобработки получаемых из видеопотока кадров.

Таким образом, общий алгоритм обнаружения лица на видеопотоке в реальном времени можно представить в следующем виде (рисунок 8):

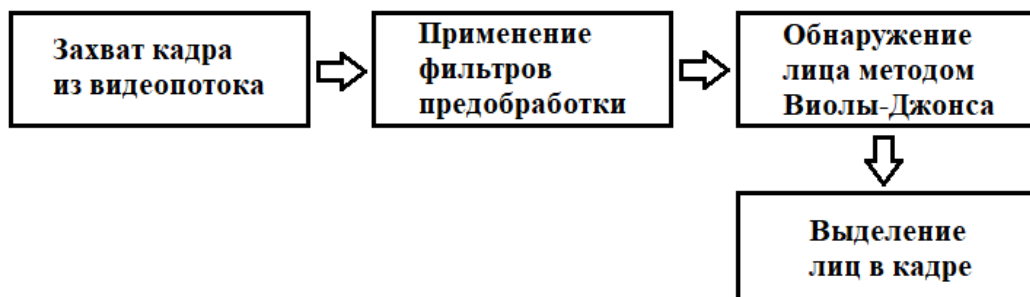


Рисунок 8 – Общее представление алгоритма обнаружения лиц



По обобщенному алгоритму обработки видеопотока, были установлены требуемые от разрабатываемого алгоритма функциональные возможности:

- Отображение на экране видеопотока в реальном времени с подключённой к компьютеру камеры;
- Отображение графического выделения обнаруженного лица;
- Возможность обучения алгоритма на увеличенных выборках с целью увеличения эффективности работы.

## **2 Разработка алгоритма обнаружения лица человека в видеопотоке**

### **2.1 Разработка алгоритма обучения**

При анализе метода Виолы-Джонса в 1 разделе уже указывалось, что данному методу нужно предварительное обучение на тестовой выборке изображений. В связи с чем разработку можно разбить на два этапа: разработка алгоритма обучения и разработка алгоритма обнаружения лица на видеопотоке.

Обучение классификаторов проводилось на двух тестовых выборках, содержащих изображения с лицом и без лица. Для корректной работы алгоритма требуется обучение на выборке из положительных и отрицательных элементов[21]. Для этого была загружена открытая база данных «NIST Mugshot Identification Database», из которой было взято 2430 изображений с изображением людей и 4550, не содержащих лица, все изображения. Важно, чтобы все изображения были одинакового размера, в данной выборке он был равен 300 на 400 пикселей, формат «.pgm». Обучение на описанной выборке заняло достаточно много времени – приблизительно 23 часа на четырёхядерном процессоре intel core i7-3770 с частотой 3.40GHz.

Алгоритм обучения продемонстрирован на рисунке 9.

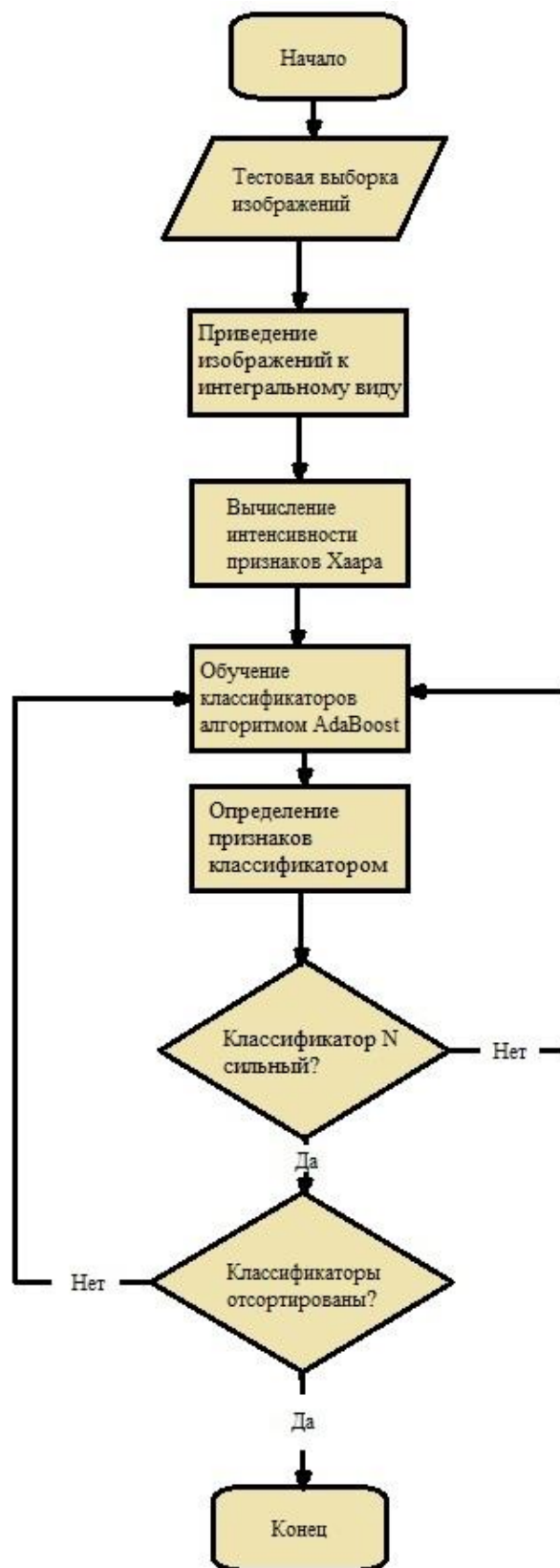


Рисунок 9 – Блок-схема алгоритма обучения классификаторов

## 2.2 Интегральное представление изображения

Для работы с изображениями используется их интегральное представление, что позволяет быстро рассчитать интенсивность пикселей анализируемой области. Создаётся матрица, равная по размерам исходному изображению, каждая ячейка которой содержит значение суммы интенсивности пикселя, а также соседнего левого и верхнего пикселей[22]. Это значение вычисляется по формуле:

$$ii(x', y') = \sum_{x \leq x', y \leq y'} i(x, y), \quad (5)$$

где  $i(x, y)$  – это интенсивность пикселя с координатами  $x, y$  на исходном изображении, а  $ii(x', y')$  интенсивность пикселя, имеющего координаты  $x', y'$  на интегральном изображении. Пример на рисунке 10.

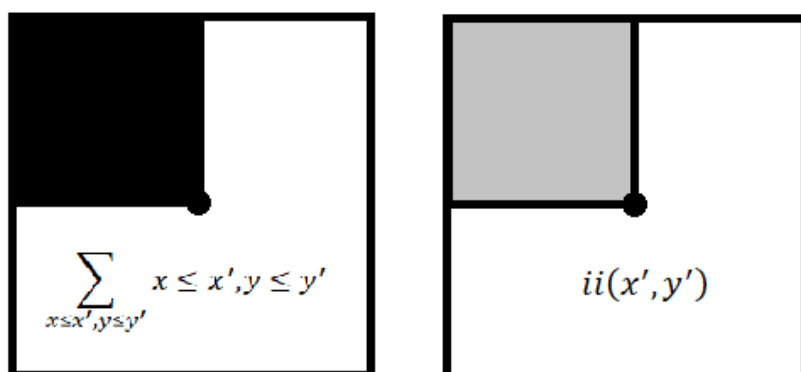


Рисунок 10 – Исходное и интегральное изображения

За один проход сканирующего окна по изображению рассчитывается одно такое интегральное выражение. Для примера возьмём пиксель, координаты которого  $i(x, y)$ . Для наглядности рассмотрим рисунок 11.

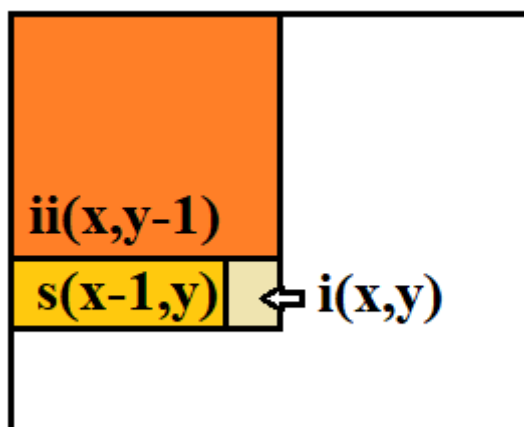


Рисунок 11 – Общий вид вычисление интегрального изображения

Сумма интенсивностей пикселей вычисляется по формуле

$$s(x, y) = s(x - 1, y) + i(x, y), \quad (6)$$

$$ii(x, y) = ii(x, y - 1) + s(x, y), \quad (7)$$

где  $s(x, y)$  – это накопленная сумма строки, тогда следует, что  $s(x, -1) = 0$ ,  $ii(-1, y) = 0$ .

Когда интегральное изображение построено, сумма яркости пикселей области любого размера легко вычисляется. Возьмём произвольный прямоугольник ABCD, где A, B, C, D – значения интегрального выражения в углах. Пример изображён на рисунке 12.

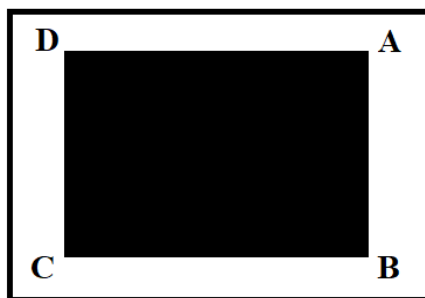


Рисунок 12 – Область изображения для вычисления интенсивности

Тогда сумма значений яркости пикселей исходного изображения считается по формуле

$$S(ABCD) = A - B - C + D. \quad (8)$$

Из формулы видно, что расчёт матрицы линейный и выполняется всегда за константное время, зависящее от числа пикселей в изображении, благодаря чему можно быстро обрабатывать получаемые из видеопотока кадры, не уменьшая производительность.

Имея интегральное выражение, для подсчёта суммы яркостей пикселей в любой области исходного изображения достаточно трёх сложений, но при этом количество выявленных характерных признаков получается очень большим- 16000 признаков для окна 24 на 24 пикселя. Поэтому следует из общего числа создать выборку таких наиболее подходящих признаков, которые позволят эффективно разбить обучающий набор на изображения, содержащие лицо и на изображения без лица.

### 2.3 Алгоритм работы сканирующего окна

Алгоритм работы сканирующего окна можно представить в общем виде на рисунке 13.



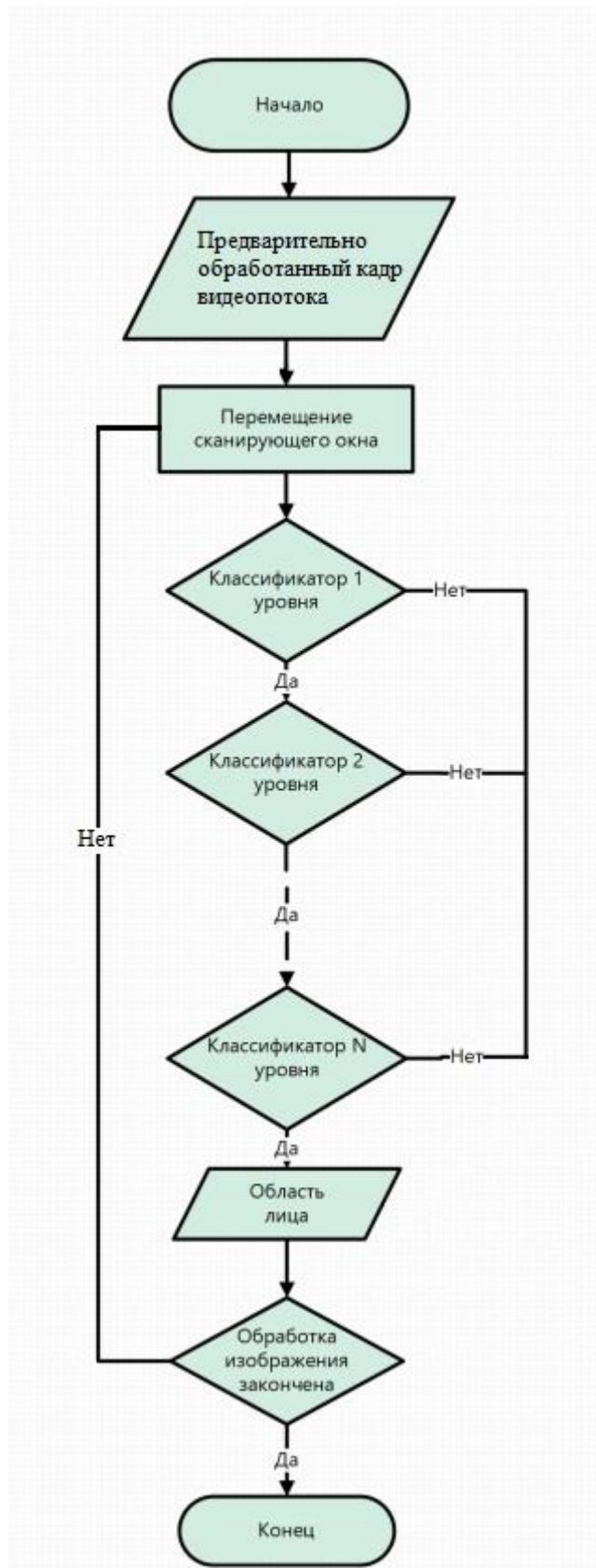


Рисунок 13 – Анализ кадра видеопотока сканирующим окном



По итогу своей работы алгоритм сканирующего окна должен определить возможные черты лица, присутствующие на кадре видеопотока.

Для извлечения ключевых признаков лиц из обучающего набора используют признаки Хаара. Они выглядят как некие прямоугольные фрагменты, объединённые белые и чёрные области. Условное их изображение приведено на рисунке 14.

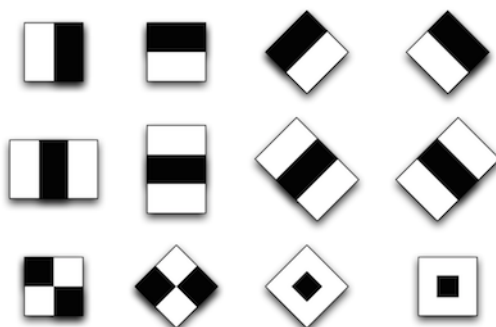


Рисунок 14 – Признаки Хаара

Закрывая фрагменты изображения белыми и чёрными областями, каскады Хаара суммируют яркость пикселей под разными областями и позволяют получить разность сумм яркости пикселей между ними по формуле:

$$f = \sum IS_1 - \sum IS_2, \quad (9)$$

где  $I$  – это значение яркости пикселей,  $S_1$ - это пиксели, попадающие под белую область, а  $S_2$  – под чёрную.

Значение получаемой разности присваивается конкретным признакам и сравнивается с неким порогом, полученным при обучении: превышение порога означает, что был обнаружен искомый объект.

Окно сканирования каждый раз перемещаясь по кадру применяет к пикселям области под собой каскад классификаторов. Каждый новый применённый классификатор проверяет признаки Хаара анализируемой области.

Когда программа удачно применяет все классификаторы каскада, то считает, что возможное лицо было обнаружено. В случае, когда один из классификаторов каскада возвращает отрицательное значение, алгоритм решает, что под данной областью сканирования лицо не обнаружено, тогда окно сканирования делает шаг дальше, проходя таким образом всю область изображения[16].

## 2.4 Разработка алгоритма обучения классификаторов AdaBoost

Классификатор в задачах классификации – это функция, определяющая принадлежность объекта к тому или иному классу. Классификация выглядит следующим образом: есть зависимость  $Y^*:X \Rightarrow Y$ , где  $X$  – множество описаний объекта,  $Y$  – множество номеров классов[17]. Обучающая выборка описывается формулой:

$$X_m = \{(x_1, y_1), \dots, (x_m, y_m)\}. \quad (10)$$

Классификаторы можно разделить на сильные и слабые. Сильные показывают высокую точность работы, тогда как точность слабых 51% или менее. Что бы выбрать верное множество признаков, при обучении используют метаалгоритм машинного обучения - бустинг. Это широко распространённый в современных системах метод классификации, преобразующий слабые обучающие алгоритмы в сильные, так как каждый последующий алгоритм обучения стремится решить недостатки предыдущего алгоритма[17].

На всех новых итерациях алгоритм бустинга формирует простые классификаторы:

$$h_j(z) = \begin{cases} 1, & \text{если } p_j f_j(z) < p_j \theta_j, \\ 0 & \end{cases} \quad (11)$$

где  $z$ - это сканирующее окно,  $j$ - номер итерации,  $h_j(z)$ - классификатор,  $p_j$ - направление знака неравенства для всех признаков,  $\theta_j$ - порог классификации,  $f_j(z)$ - вычисленное значение признака.

В данной работе применяется адаптивный алгоритм бустинга - AdaBoost, который отличается тем, что образует взвешенную группу из простых классификаторов, создавая, таким образом, один «сильный» классификатор – каскад классификаторов.

Алгоритм бустинга представлен в следующем виде:

1. Определяются слабые классификаторы по признакам Хаара
2. Окно сканирования перемещается, на каждом шаге вычисляя характерные признаки.
3. Определяется пороговое значение для каждого признака.
4. Выявляются лучшие признаки и пороговое значение.
5. Классификаторам присваивается новый вес.

Применение каскадного классификатора имеет большую эффективность при работе с изображениями и видео, содержащими небольшое количество лиц, так как в таком случае классификаторам проще сразу отбросить области, не содержащие лицо, тем самым повысив скорость обнаружения[22]. Пример каскада классификаторов представлен на рисунке 15.



Рисунок 15 – Каскадная модель сильных классификаторов

Таким образом общий вид разработанного алгоритма обучения классификаторов AdaBoost имеет следующий вид (рисунок 16):

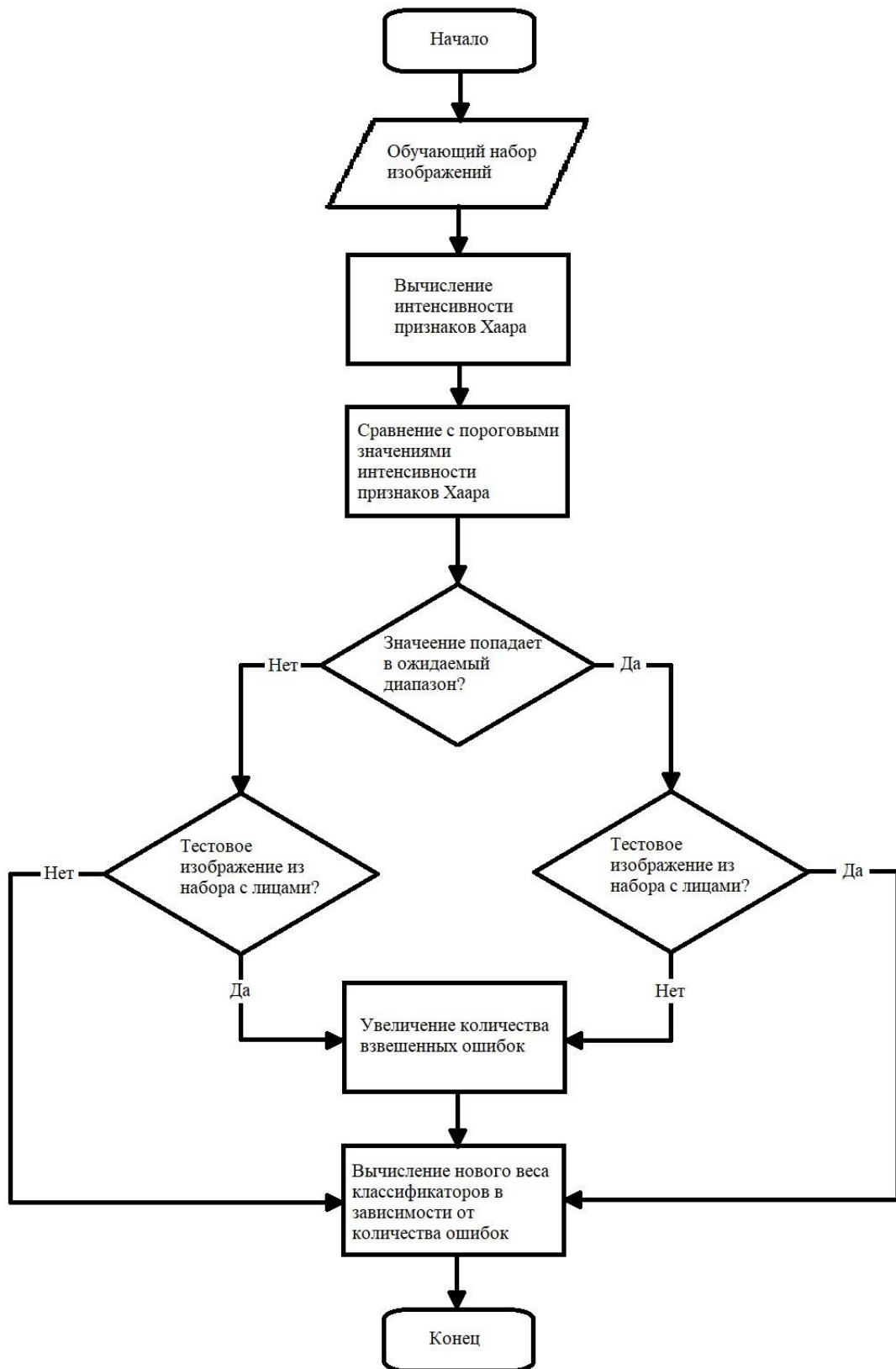


Рисунок 16 – Алгоритм обучения AdaBoost

## 2.5 Разработка алгоритма обнаружения

Разработанный алгоритм обнаружения лиц из видеопотока представлен в виде блок-схемы на рисунке 17.

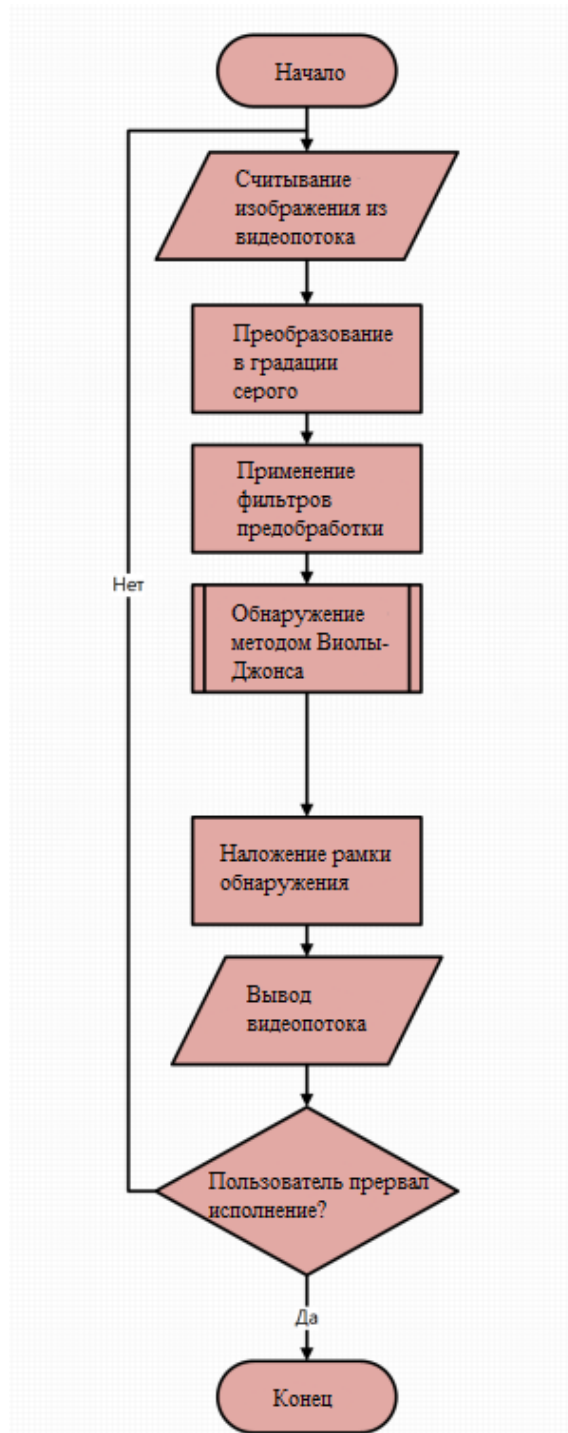


Рисунок 17 – Блок-схема алгоритма распознавания лиц в видеопотоке

Метод обнаружения Виолы-Джонса работает с изображениями, поэтому для его реализации каждый новый кадр из видеопотока сохраняется во временный массив и подвергается преобразованиям для дальнейшей работы: осуществляется преобразование в интегральный вид и перевод изображения в полутона серого, так как цвет не имеет значения при обнаружении.

Поступающий из видеопотока кадр представлен в цветах формата RGB, но наличие цвета не несёт полезной информации для алгоритма обнаружения, поэтому было принято решение перевести изображения в формат YUV, то есть в градации серого[11].

Перевод изображения осуществляется по формулам:

$$Y = 0.299R + 0.587G + 0.114B \quad (12)$$

$$U = -0.14713R - 0.28886G + 0.436B \quad (13)$$

$$V = 0.615R - 0.51499G - 0.10001B \quad (14)$$

Для полутонового изображения имеет значение только Y- яркостная составляющая. Именно она содержит изображение в оттенках серого.

Но для повышения эффективности работы алгоритма представления изображения в полутоновом спектре недостаточно.

Применяются фильтры предобработки для подавления шума. Наиболее распространённым фильтром является фильтр Гаусса, но в результате работы он сильно размывает изображение, что негативно влияет на работу алгоритма Виолы-Джонса. По этому было принято решение применить комбинированный метод, включающий билатеральный фильтр и медианный фильтр[12].

*Билатеральный фильтр* - это метод адаптивной фильтрации, сохраняющий чёткость крупных объектов, но подавляющий шумы и сглаживающий небольшие объекты.

Фильтр основывается на вычислении интенсивности пикселей, представляя их как взвешенное среднее значение яркости пикселей окрестности. Формула фильтра выглядит следующим образом:

$$y_{i,j} = \sum_{k,m \in Q} x_{i+k,j+m} W(i, j, k, m), \quad (15)$$

где  $y_{i,j}$  – пиксель на выходе, а  $W(i, j, k, m)$  – взвешенная сумма интенсивностей соседних пикселей. Веса пикселей не имеют фиксированного значения, а зависят от расположения анализируемого пикселя. Зависимость описывает следующая формула:

$$W(i, j, k, m) \approx \exp\left(-\frac{(x_{i,j}-x_{i+k,j+m})^2}{h^2}\right) \exp\left(-\frac{k^2+m^2}{p^2}\right). \quad (16)$$

Первая часть формулы отображает фотометрическую близость. Она изучает яркость пикселей путём оценки яркости центрального пикселя и пикселей-соседей. При схожей интенсивности яркости значения пикселей усредняются, иначе усреднение не применяется для предотвращения размытия границ. Таким образом происходит подавление лишнего шума и сохранение чёткости контрастных границ[6].

Вторая часть формулы оценивает расстояние пикселя в окне до центра и является фильтром Гаусса.

Фильтр называется билатеральным благодаря тому, что оценивается близость пикселя со стороны интенсивности и со стороны расстояния.

*Медианная фильтрация* - адаптивный фильтр, показывающий хорошие результаты шумоподавления. Его работа основывается на том, что производится последовательная оценка каждой точки изображениядвигающимся окном –



апертурой фильтра. Центральный элемент окна располагается над пикселем фильтрации, а остальные элементы симметричны относительно него. Чаще всего оцениваются данные текущей строки и столбца, а также соседние им, поэтому фильтр имеет форму квадрата или крестообразный вид[7]. Рисунок 18 демонстрирует примеры сканирующего окна фильтра.

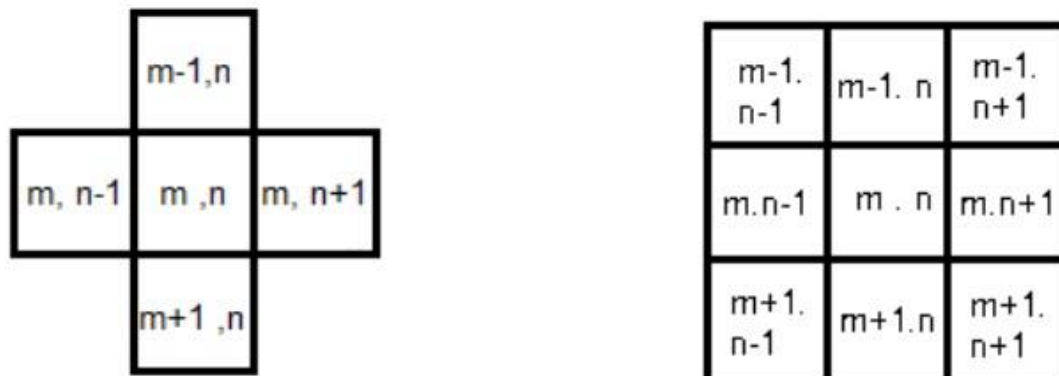


Рисунок 18– Окно сканирования медианного фильтра

Из пикселей, попавших под окно сканирования, формируется выборка для дальнейшей работы. Если данную выборку представить в виде упорядоченной по возрастанию последовательности значений, то её центральный элемент будет служить медианой. Медиана, являющаяся результатом применения фильтра, считается по формуле:

$$\text{med} = \arg \min \sum_{f_i \in W, f_j \in W} |f_i - f_j|. \quad (17)$$

В данной формуле  $W$  - это выборка пикселей, среди которых вычисляется медиана, а  $f_i$  и  $f_j$  – значения пикселей в выборке.

Тогда формула двумерного медианного фильтра выглядит следующим образом:

$$G_{i,j} = \text{med}[E_{i+s,j+1}; (s, t) \in W], \quad (18)$$

$$i, j \in Z^2. \quad (19)$$

В формуле  $G_{i,j}$  отображает элемент матрицы изображения после применения фильтра,  $E_{i,j}$  – элемент матрицы до применения фильтра.

Данный фильтр имеет существенный недостаток – если величина мелких фрагментов изображения меньше величины половины размера окна сканирования, то эти детали размываются. Поэтому область окна фильтра увеличивают.

В данной работе используется комбинация рассмотренных выше фильтров с апертурой 3. Благодаря использованию комбинированного метода шум и лишние элементы фона успешно сглаживаются, и на первой итерации каскада классификаторов алгоритм успешно выделяет лица[7].

Затем при помощи метода Виолы-Джонса на кадре обнаруживается лицо и накладывается рамка выделения.

В итоге работы алгоритма, обработанный кадр с выделенным лицом показывается пользователю, затем считывается и обрабатывается следующий кадр.

По причине того, что приходится обрабатывать каждый кадр видеопотока, сильно снижается плавность отображаемого для пользователя результата. В связи с чем было принято решение разбить программу на два параллельных потока, в один из которых вынесена функция распознавания, которая отдаёт в первый поток только области обнаружения. Таким образом значительно увеличивается плавность и производительность.

Блок-схема, описывающая принцип работы, приведена на рисунке 19.

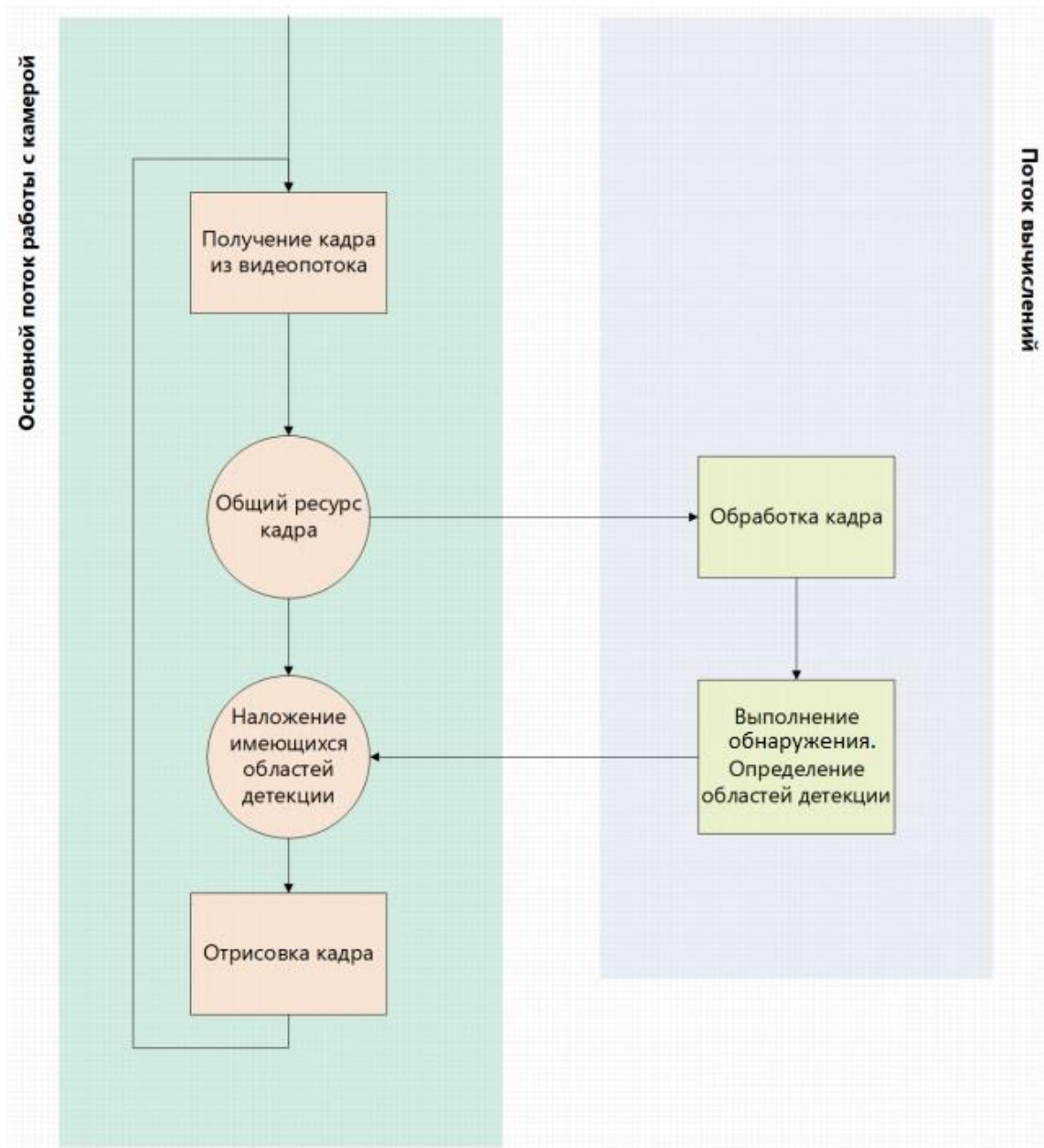


Рисунок 19 – Блок-схема распараллеливания алгоритма

Сначала с видеопотока получается изображение и инициализируется функция обнаружения лиц, затем в цикле из видеопотока получают кадры, передаются в функцию обнаружения и отображаются пользователю. Параллельно в функции обнаружения анализируются получаемые кадры и возвращаются зоны обнаружения, которые так же отображаются пользователю[23].

Таким образом, весь алгоритм распознавания лица на видеопотоке можно представить в виде следующей диаграммы декомпозиции (рисунок 20):

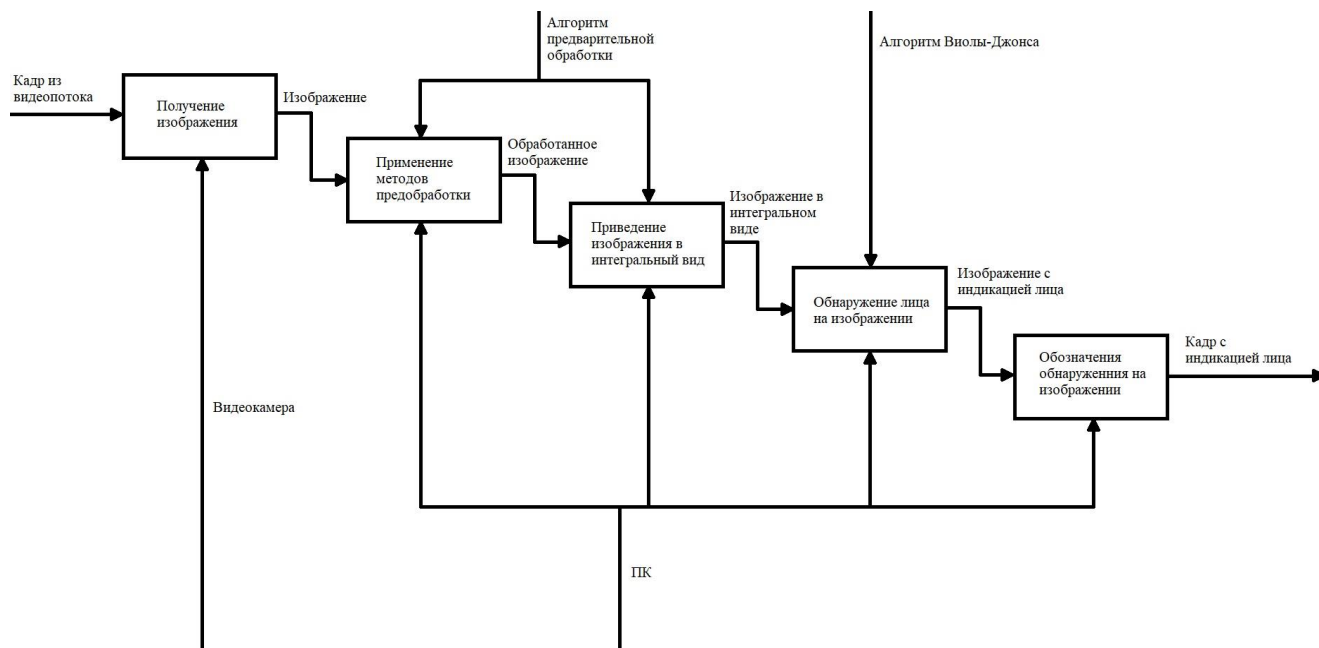


Рисунок 20 – Диаграмма декомпозиции разработанного алгоритма обнаружения лиц на видеопотоке

Во втором разделе был приведён разработанный алгоритм обнаружения лица на видеопотоке на основе метода Виолы-Джонса.

В целях повышения эффективности алгоритма обнаружения был включён дополнительный этап предобработки кадров видеопотока перед обнаружением лица.

### 3 Тестирование разработанного алгоритма

#### 3.1 Оценка разработанного алгоритма обнаружения лица в видеопотоке

Выделяют два основных параметра, по которым оценивают методы обнаружения: Коэффициент ложного доступа FAR — False Acceptance Rate и Коэффициент ложного отказа FRR — False Rejection Rate.

Коэффициент ложного отказа обозначает, насколько часто происходит ошибочное отвержение изображений, содержащих потенциальное лицо. Точность обнаружения зависит от того, насколько мал коэффициент. Коэффициент ложного допуска наоборот показывает вероятность, насколько часто система может ошибочно признать лицо на изображении без лица.

Существует также коэффициент EER, показывающий равную вероятность первых и вторых типов ошибок, то есть вероятность, при которой и ошибка ложного приёма и ошибка ложного отказа равновероятны. Это позволяет провести сравнение тестируемых систем и выбрать ту, у которой показатель коэффициента ERR ниже, что свидетельствует о большей точности и эффективности системы. Рисунок 21 демонстрирует взаимосвязь всех трёх коэффициентов FAR, FRR и EER:

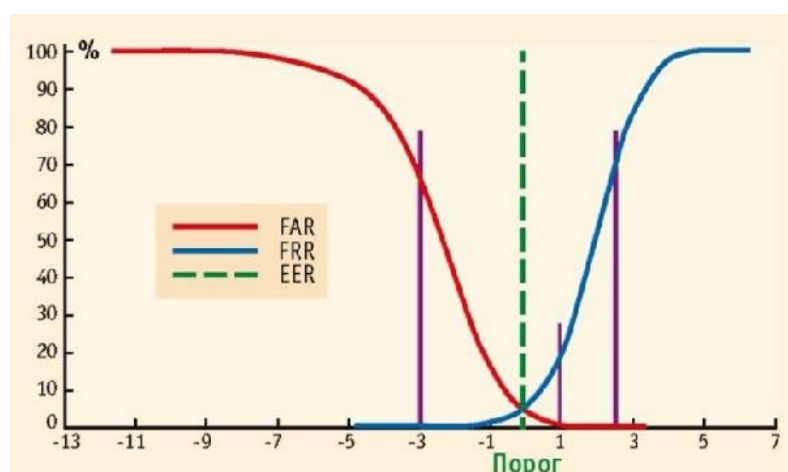


Рисунок 21 – Взаимосвязь характеристик FAR, FRR и EER

Приведённый рисунок демонстрирует график ROC- Receiver Operating Characteristic, визуализирующий сравнение систем в зависимости типов ошибок друг от друга.

Для проверки работоспособности и эффективности разработанного алгоритма был выполнен ряд тестовых испытаний. Для тестирования был создан набор из 25 кадров изображений. В серии тестовых видеотрансляций использованные для теста изображения были сняты на подключенную видеокамеру в процессе выполнения программы. Программа выделяет обнаруженное лицо жёлтым прямоугольником. Пример изображений тестовой выборки продемонстрирован на рисунке 22.



Рисунок 22 – Набор изображений для тестирования

Изображения отличаются разрешением, цветовой гаммой, освещением и положением людей, наличием аксессуаров и мимики.

Для наглядного сравнения эффективности разработанного алгоритма было проведено тестирование всех выше описанных методов.

Оценка эффективности проводилась путём определения ошибки ложного отказа (FN) и ложного обнаружения (FP) на изображениях с различными условиями по формуле:

$$FN = \frac{N_{обн}}{N_{общ}}, \quad (20)$$

$$FP = \frac{N_{необн}}{N_{общ}}, \quad (21)$$

где FN – ошибки первого рода; FP – ошибки второго рода;  $N_{обн}$  – число распознанных объектов;  $N_{необн}$  – число нераспознанных объектов;  $N_{общ}$  – общее число изображений;

В первом случае лицо на изображении не определяется программой и изображение пропускается, а во втором случае алгоритм неверно определяет объект на изображении без лица и вызывает ложное обнаружение.

На следующей таблице продемонстрированы вероятности возникновения ошибок обоих типов на тестируемом наборе для методов, рассмотренных в 1 разделе:

Таблица 3 – Вероятность ошибки методов обнаружения

Метод	Метод гибкого сравнения на графах	Метод главных компонент	Метод опорных векторов	Скрытые Марковские модели	Виолы-Джонса
FN, %	1,97	1,97	1,97	0	0
FP, %	50,96	70,61	47,04	74,57	11,91

Метод Виолы-Джонса продемонстрировал наилучшие результаты работы, а вероятность возникновения ошибок обоих типов значительно ниже, чем у остальных методов. Худшие результаты показал метод корреляции и методы сегментации.

### 3.2 Оценка работы алгоритма с применением фильтров предобработки

Однако в некоторых случаях всё же возникали ошибки ложного обнаружения. Пример работы алгоритма Виолы-Джонса продемонстрирован на рисунках 23 и 24.

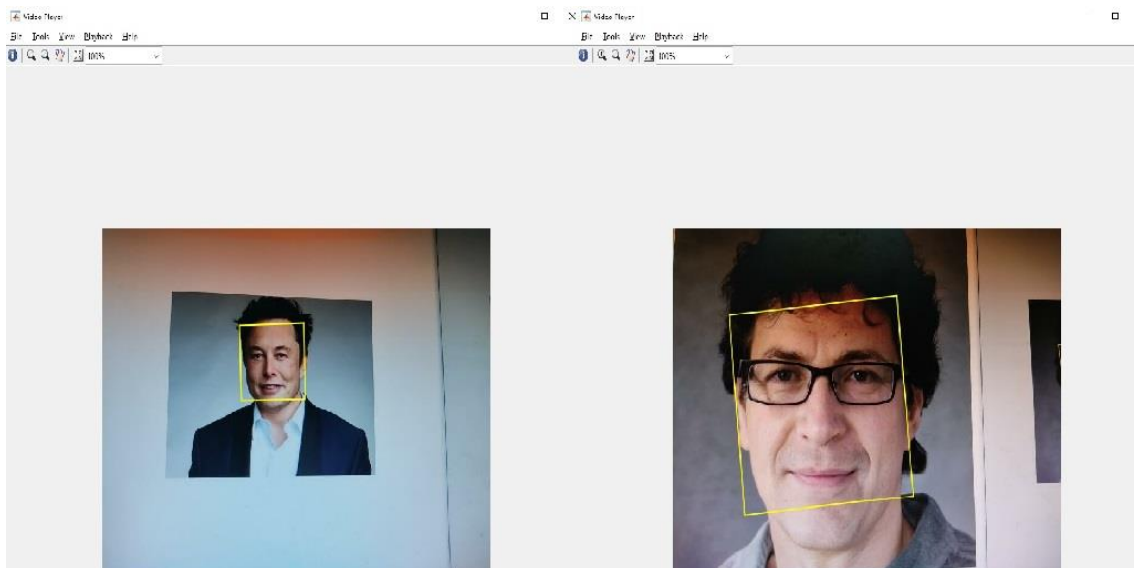


Рисунок 23 – Пример корректной работы алгоритма

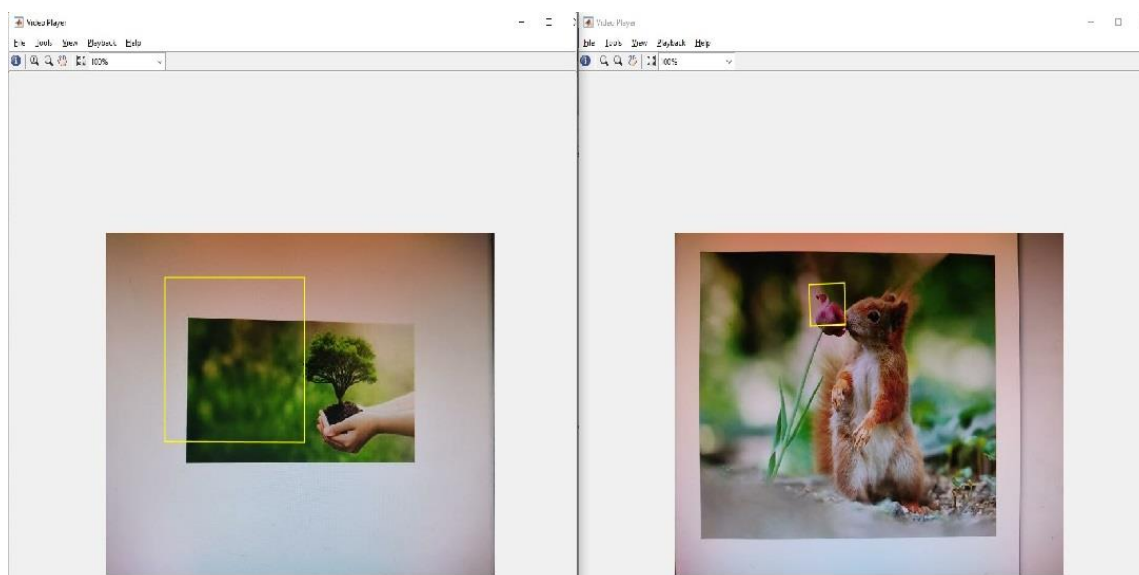


Рисунок 24 – Пример возникновения ошибки ложного допуска



Получаемые на входе изображения из видеопотока зачастую не удовлетворяют требуемому качеству для дальнейшей работы алгоритма, поэтому проводится их предварительная обработка. Как правило, изображение подвергается так называемой нормализации: изменяется масштаб, изображение кодируется. Существуют разнообразные методы, фильтры шумоподавления, например фильтр размытия Гаусса, медианный и подобные им. Выбор применяемого метода или же их совокупности зависит от цели исследования.

Для уменьшения вероятности возникновения ошибок при работе алгоритма обнаружения было принято решение включить в работу вспомогательные алгоритмы предобработки:

- Применение чёрно-белого фильтра;
- Снижение шума;
- Применение фильтра сглаживания.

Как уже указывалось ранее, для снижения вероятности возникновения ошибок в методе Виолы-Джонса был применён алгоритм предварительной обработки, включающий в себя билатеральный и медианный метод с апертурой 3. В таблице 4 приведены вероятности возникновения ошибок на той же тестовой выборке при использовании данных фильтров по отдельности и вместе.

Таблица 4 – Вероятность возникновения ошибок при использовании фильтров

Фильтр	Медианный		Билатеральный	Комбинированный
	[3,3]	[5,5]		
FN,%	0	0	0	0
FP,%	4,97	0	2,35	0

В результате проделанной работы был разработан алгоритм на основе метода Виолы-Джонса и комбинированного метода предварительной обработки изображений. Программная часть была реализована в среде Matlab.

В результате анализа работы алгоритма была выявлена близкая к нулю вероятность возникновения ошибок ложного допуска и ложного обнаружения, что свидетельствует о высокой работоспособности данного алгоритма в различных условиях освещения, при наличии мимики или при повороте лица. На рисунке 25 приведён пример испытания программы.

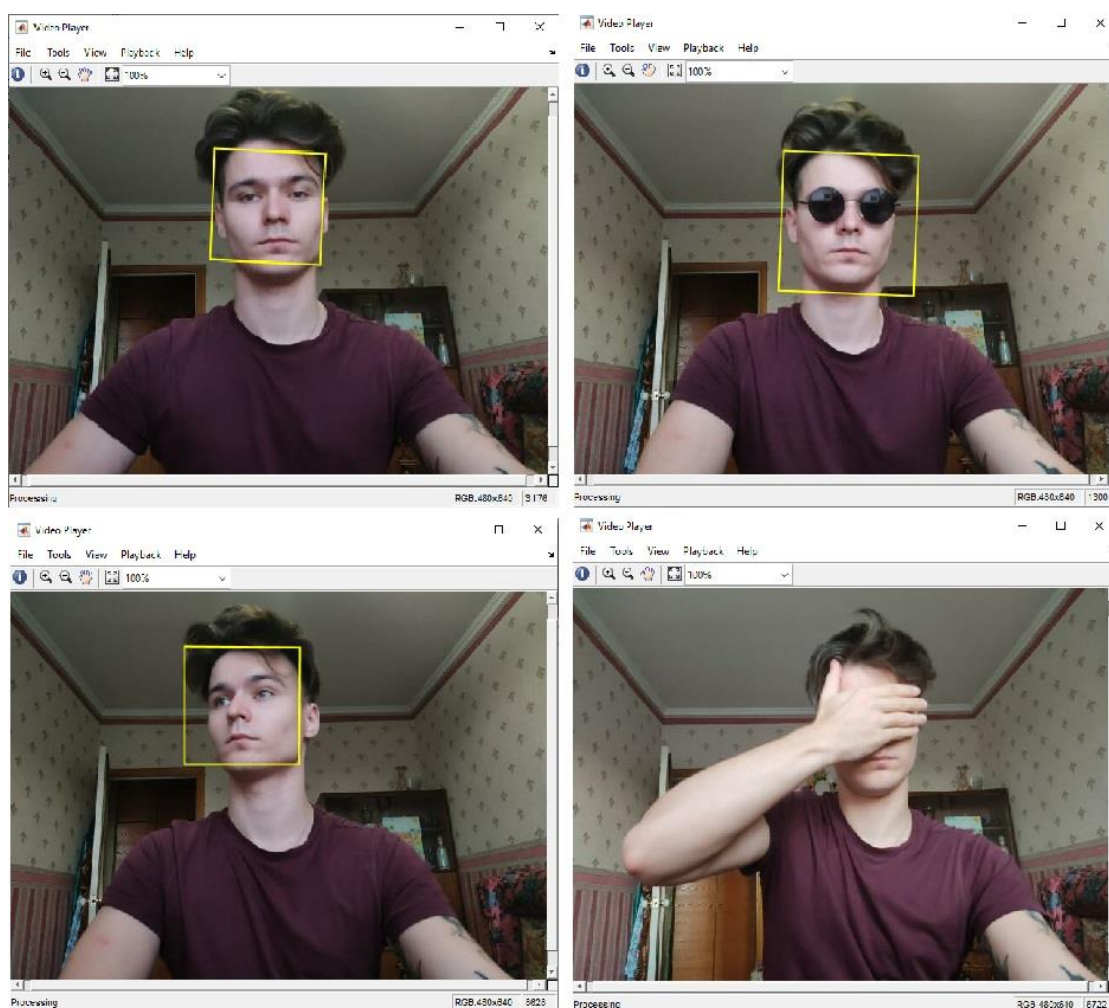


Рисунок 25 – Результаты работы алгоритма

При тестировании на видеопотоке через подключённую камеру алгоритм показал хорошую эффективность при различных условиях освещения, наличии аксессуаров и небольшом повороте головы до 30 градусов. Как правило,

алгоритм работает во всех случаях, но были выявлены исключения, когда на видео присутствует лишь часть лица, видео размыто или лицо слишком далеко. В иных случаях отсутствие обнаружения является не ошибкой алгоритма, а недостаточной натренированностью классификатора.

## Заключение

По итогу сделанной выпускной квалификационной работы были выполнены все поставленные задачи:

1. Был проведён анализ методов и алгоритмов обнаружения лиц на цифровых изображениях и кадрах видеопотока, были выявлены достоинства и недостатки. На основании чего можно сделать вывод, что основной проблемой методов обнаружения можно назвать высокую зависимость от внешних факторов, таких как: монотонность освещения, разрешение исходного изображения и положение объекта на изображении. Алгоритмы, зависящие в меньшей степени от воздействия подобных факторов, как правило, требовательны к вычислительной мощности и сложны в обслуживании, что усложняет их применение для обнаружения в реальном времени;
2. На основе проведённого анализа и наблюдений был определён наиболее подходящий метод для основы создания алгоритма обнаружения лиц на видеопотоке – метод Виолы-Джонса на основе признаков Хаара и адаптивного обучения классификаторов AdaBoost. За счёт адаптивного обучения классификаторов на тестовой выборке, удалось добиться создания каскада сильных классификаторов, что позволило получить высокую эффективность обнаружения лиц;
3. Разработан и программно осуществлён алгоритм обнаружения лиц на кадрах видеопотока в реальном времени на основе алгоритма Виолы-Джонса, который был модифицирован алгоритмом предобработки, который включает в себя перевод изображения в градации серого, применение билатерального и медианного фильтров шумоподавления. Это позволило увеличить его эффективность. В результате тестирования доказано, что при применении комбинированного фильтра предобработки эффективность

алгоритма достигает 97%. Подавляющая часть объектов была успешно распознана.

Таким образом, разработанный в результате работы алгоритм в дальнейшем может быть обучен на большей обучающей выборке изображений, доработан алгоритмом распознавания и адаптивной системой предобработки, что позволит использоваться его при решении задач как личного характера, так и для прямого применения в охранных системах контроля допуска и распознавания людей.

## Список используемой литературы

1. Актуальность 2D алгоритмов в определенных задачах автоматического распознавания [Электронный ресурс]. // Реноме. – Режим доступа: <https://moluch.ru/conf/tech/archive/2/138/> (дата обращения: 21.05.2021).
2. Броневи́ч А. Н. Лекции по методам машинного обучения [Электронный ресурс]. // Режим доступа: [http://window.edu.ru/resource/files/lect\\_Lepskiy\\_Bronevich.pdf](http://window.edu.ru/resource/files/lect_Lepskiy_Bronevich.pdf) (дата обращения: 05.05.2021).
3. Видеоаналитика и распознавание лиц [Электронный ресурс]. // Wikipedia. – Режим доступа: [https://video-praktik.ru/st\\_videoanalitika.html/](https://video-praktik.ru/st_videoanalitika.html/) (дата обращения: 18.05.2021).
4. Ворона В.А., Костенко В.О. Биометрические технологии идентификации в системах контроля и управления доступом // Computational nanotechnology. 2016. №3. URL: <https://cyberleninka.ru/article/n/biometricheskietehnologii-identifikatsii-v-sistemah-kontrolya-i-upravleniya-dostupom> (дата обращения: 20.05.2021).
5. Гонсалес Р., Вудс Р. Цифровая обработка изображений. М.: Техносфера, 2005. 1072 с. (дата обращения: 18.05.2021).
6. Красильников Н. Н. Цифровая обработка 2D- и 3D- изображений. СПб: БХВ-Петербург, 2011. 608 с. (дата обращения: 18.05.2021).
7. Лукьяница А. А., Шишкин А. Г. Цифровая обработка видеоизображений. М.: Ай-Эс-Эс Пресс, 2009. 518 с.
8. Метод Виолы — Джонса [Электронный ресурс]. // Wikipedia. – Режим доступа: [https://ru.wikipedia.org/wiki/Метод\\_Виолы\\_—\\_Джонса](https://ru.wikipedia.org/wiki/Метод_Виолы_—_Джонса) (дата обращения: 18.05.2021).
9. Метод главных компонент [Электронный ресурс]. // Wikipedia. – Режим доступа: [https://ru.wikipedia.org/wiki/Метод\\_главных\\_компонент](https://ru.wikipedia.org/wiki/Метод_главных_компонент) (дата обращения: 18.05.2021).

10. Метод опорных векторов [Электронный ресурс]. // Wikipedia. – Режим доступа: [https://ru.wikipedia.org/wiki/Метод\\_опорных\\_векторов](https://ru.wikipedia.org/wiki/Метод_опорных_векторов) (дата обращения: 18.05.2021).
11. Обнаружение лиц на цветном растровом изображении с применением модифицированного метода Виолы-Джонса / М. Н. Вязовых, К. И. Зайцев, М.В. Мухортов, А. Н. Перов // URL: <http://sntbul.bmstu.ru/doc/45828>
12. Приоров А. Л., Апальков И. В., Хрящев В. В. Цифровая обработка изображений. Ярославль: ЯрГУ, 2007. 235 с.
13. Скрытая марковская модель [Электронный ресурс]. // Wikipedia. – Режим доступа: [https://ru.wikipedia.org/wiki/Скрытая\\_марковская\\_модель](https://ru.wikipedia.org/wiki/Скрытая_марковская_модель) (дата обращения: 19.05.2021).
14. Татаренков Д. А. Анализ методов обнаружения лиц на изображении [Электронный ресурс]. // Молодой ученый. — 2015. — №4. — С. 270-276. — Режим доступа: <https://moluch.ru/archive/84/15524/> (дата обращения: 11.05.2021).
15. Что такое машинное зрение и чем оно отличается от человеческого? [Электронный ресурс] Режим доступа: <https://meduza.io/feature/2019/03/30/chto-takoe-mashinnoe-zrenie-i-chem-ono-otlichaetsya-ot-chelovecheskogo-seychas-obyasnim-ponyatno> (дата обращения: 05.05.2021).
16. Berthold K.P. Horn. Robot Vision. — MIT Press, 1986. — ISBN 0-262-08159-8 (Б.К.П. Хорн, Зрение роботов: перевод с англ. — М.: Мир, 1989).
17. Freund Y., Schapire R. E. A Short Introduction to Boosting // J. of Japanese Society for Artificial Intelligence. 1999. Vol. 14, iss. 5. P. 771 – 780.
18. Kotropoulos C., Pitas I. Ruled based face detection in frontal views // IEEE Intern. Conf. on Acoustics, Speech, and Signal Processing, Washington DC, USA Piscataway: IEEE Computer Society, 1997. P. 2537 – 2540.
19. Moghaddam B. and Pentland A. Probabilistic Visual Learning for Object Representation. IEEE Transactions on Pattern Analysis and Machine Intelligence 1997, Vol. 19, P. 696-710.

20. Neurotechnology [Электронный ресурс]. // Режим доступа: <http://www.neurotechnology.com/> (дата обращения: 13.05.2021).

21. NIST Mugshot Identification Database - база изображений лиц, подготовленная Национальным институтом стандартов и технологий (США). [Электронный ресурс]. // Режим доступа – [http://www.machinelearning.ru/wiki/index.php?title=NIST\\_Mugshot\\_Identification\\_Database/](http://www.machinelearning.ru/wiki/index.php?title=NIST_Mugshot_Identification_Database/) (дата обращения: 27.05.2021).

22. Viola P. , Jones M. Rapid object detection using a boosted cascade of simple features // IEEE CVPR01. – 2001. – P. 511–518.

23. Viola P., Jones M.J. Robust real-time face detection. // Intern. J. of Computer Vision. 2004. Vol. 57, iss. 2. P 137 – 154

24. Yang G., Huang T. S. Human face detection in complex background. Pattern Recognition // URL: <http://www.sciencedirect.com/science/article/pii/> (дата обращения: 20.05.2021).