

	Содержание	
Перв. примен.	Введение.....	4
	Глава 1.Обзор существующих подходов к обнаружению объектов на изображении	7
	1.1 Структура алгоритмов для обнаружения объектов.....	7
	1.2 Методы построения признакового описания.....	8
	1.2.1 Моменты изображения.....	9
	1.2.2 Гистограммы ориентированных градиентов.....	10
	1.2.3 Совместная гистограмма ориентированных градиентов.....	11
	1.2.4 Алгоритм GIST.....	12
	1.3 Методы локализации объектов на изображении	13
	1.3.1 Локализация на основе цвета.....	13
	1.3.2 Локализация на основе форм.....	13
	1.3.3 Локализация с использованием машинного обучения.....	14
	1.4 Выводы по первой главе.....	19
	Глава 2. Информационная модель распознавания дорожных знаков на изображении	20
	2.1 Схема системы распознавания дорожных знаков на изображении.....	20
	2.1.1 Предварительная обработка изображения.....	21
	2.1.2 Локализация дорожных знаков на изображении	22
	2.1.3 Построение глобального признакового описания по ТАБ.....	25
	2.2 Выводы по второй главе.....	34
	Глава 3. Вычислительный эксперимент	35
	3.1 Описание эксперимента	35
	3.2 Описание программной системы.....	36
	3.2.1 Описание программного обеспечения для клиента.....	40
	3.2.2 Описание программного обеспечения для сервера	40
	3.3 Анализ результатов эксперимента	43
	3.4 Выводы по третьей главе.....	47
	Заключение	48
	Библиографический список	49
Справ. №		
Подпись и дата		
Инв. № дубл.		
Взам. инв. №		
Подпись и дата		
Ине. № подл.		

Изм

Лист

№ докум.

Подпись

Дат

Разраб.

Густякова А.П

02.07.20

Провер.

Гай В.Е.

02.07.20

Реценз.

Н. Контр.

Гай В.Е.

02.07.20

Утверд.

Жевнерчук Д.В.

02.07

БКР-НГУ-09.04.01-(М18-ИБТ-3)-004-2020 (ПЗ)

Лит.

Лист

Листов

3

Модель и алгоритмы
распознавания дорожных знаков

НГУим. Р.Е. Алексеева

[illegible]

Перв. примен.	<h2 style="text-align: center;">Введение</h2> <h3 style="text-align: center;">Актуальность темы</h3> <p>На сегодняшний день обработка изображений имеет множество областей применения, таких как медицина, искусственный интеллект, робототехника и т. д. Системы распознавания дорожных знаков являются одними из наиболее актуальных областей, использующих компьютерное зрение. По мере того, как мир движется к беспилотным транспортным средствам, все большие требования возникают к безопасности дорожного движения. Увеличение вычислительной мощности машин сделало возможным обнаружение и распознавание дорожных знаков в режиме реального времени. Темой данной работы стало распознавание дорожных знаков для парковки. Система, распознающая такие знаки, полезна водителям для толкования значения знаков. Также, подобные технологии используются в беспилотных автомобилях. Такая система поможет избегать ошибок при парковке.</p> <p>Однако при обработке изображений, содержащих дорожные знаки, возникают некоторые проблемы. Ими являются низкое разрешение камеры, несовершенное состояние знака, освещенность, условия погоды. Все это может повлиять на цветовой анализ и извлечение формы знаков. Чтобы правильно распознать признаки, вышеупомянутые проблемы должны быть должным образом решены. В большинстве опубликованных работ система разработана с использованием трехэтапного последовательного подхода, состоящего из предварительной обработки, детектирования и принятия решения.</p> <p>Данная работа посвящена применению инструментария теории активного восприятия для решения задачи распознавания дорожных знаков на изображениях.</p> <h3 style="text-align: center;">Цель работы и задачи исследования</h3> <p>Целью выпускной квалификационной работы является разработка и исследование модели и алгоритмов распознавания дорожных знаков на основе теории активного восприятия.</p> <p>Для достижения данной цели необходимо решить следующие задачи:</p> <ol style="list-style-type: none"> 1. обзор и анализ известных алгоритмов решения задачи распознавания дорожных знаков; 2. обзор и анализ известных признаковых описаний дорожных знаков; 3. разработка модели и алгоритмов извлечения и хранения признакового описания, формирования входных данных, а также классификации; 4. выбор средств разработки; 5. программная реализация системы; 6. тестирование реализованной системы. 					
						Справ. №
Подпись и дата						
	Инв. № дубл.					
Взам. инв. №						
	Подпись и дата					
Инв. № подл.						
	Изм.	Лист	№ докум.	Подпись	Дата	ВКР-НГТУ-09.04.01-(М18-ИВТ-3)-004-2020(ПЗ)

Перв. примен.	<p>Объект исследования Объектом исследования являются цифровые изображения.</p> <p>Предмет исследования Предметом исследования являются модель и алгоритмы распознавания дорожных знаков.</p> <p>Методы исследования Для решения поставленной задачи в работе были использованы методы детектирования дорожных знаков, теория активного восприятия. Метод вычислительного эксперимента был использован для практической апробации разработанных алгоритмов и реализован на языках программирования <i>R, Python и Kotlin</i>.</p> <p>Положения, выносимые на защиту На защиту выносятся следующие положения:</p> <ul style="list-style-type: none"> Информационная модель глобального признакового описания дорожных знаков, построенная на основе положений теории активного восприятия об алгебре групп. Результаты вычислительного эксперимента. <p>Научная новизна Научная новизна данного метода решения задачи распознавания дорожных знаков на изображении - это использование нового подхода к формированию глобального признакового описания изображения. Для формирования глобального признакового описания в работе используются положения алгебры групп теории активного восприятия.</p> <p>Теоретическая и практическая ценность Теоретическая ценность работы представляется в разработанной информационной модели глобального признакового описания объекта. Практическая ценность работы представляется в спроектированном, реализованном и проверенном программном обеспечении, использующем новый алгоритм составления признакового описания и в результатах проведенного вычислительного эксперимента.</p> <p>Обоснованность и достоверность Обоснованность и достоверность данного исследования, полученных результатов и выводов доказывается проведением вычислительного эксперимента, сравнением полученных результатов с известными и апробацией основных положений выпускной квалификационной работы в докладе на международной научно-технической конференции.</p>														
	Справ. №														
Подпись и дата															
Инв. № дубл.															
Взам. инв. №															
Подпись и дата															
Инв. № подл.															
<table border="1"> <tr> <td>Изм.</td> <td>Лист</td> <td>№ докум.</td> <td>Подпись</td> <td>Дата</td> </tr> <tr> <td></td> <td></td> <td></td> <td></td> <td></td> </tr> </table>					Изм.	Лист	№ докум.	Подпись	Дата						<p>ВКР-НГТУ-09.04.01-(М18-ИВТ-3)-004-2020(ПЗ)</p> <p>Лист 5</p>
Изм.	Лист	№ докум.	Подпись	Дата											

Перв. примен.	<p>Публикации</p> <p>А.П. Густякова, В.Е. Гай - Модель и алгоритмы распознавания дорожных знаков // Материалы XXVI международной научно-технической конференции «Информационные системы и технологии - 2020», ИСТ-2020, Россия, Н. Новгород, 2020г.</p> <p>Объём и структура</p> <p>Выпускная квалификационная работа состоит из введения, трёх глав основной части, заключения и списка литературы. Общий объём работы – 50 страниц. Диссертация содержит 25 рисунков, 16 таблиц, 3 формулы. Библиографический список состоит из 23 наименований.</p>						
	Справ. №						
						Подпись и дата	
						Инв. № дубл.	
						Взам. инв. №	
						Подпись и дата	
Инв. № подл.							
Изм.	Лист	№ докум.	Подпись	Дата	ВКР-НГТУ-09.04.01-(М18-ИВТ-3)-004-2020(ПЗ)	Лист 6	

Перв. примен.	<h1>Глава 1. Обзор существующих подходов к обнаружению объектов на изображении</h1>														
	<h2>1.1 Структура алгоритмов для обнаружения объектов</h2>														
Справ. №	<p>В большинстве случаев, обнаружение объектов на изображении производят с помощью последовательных этапов:</p> <ul style="list-style-type: none"> • Предварительная обработка изображения • Составление признакового описания искомого объекта • Принятие решения на основе этапа признакового описания 														
	<p>Предварительная обработка - это общее название для операций с изображениями на самом низком уровне абстракции. Целью предварительной обработки является улучшение данных изображения, которое подавляет нежелательные искажения или улучшает некоторые особенности изображения, важные для дальнейшей обработки.</p> <p>Составление признакового описания - способ уменьшения размерности, который эффективно представляет интересные части изображения в виде компактного вектора признаков. Этот подход полезен, когда размеры изображений велики, и для быстрого выполнения таких задач, как сопоставление и поиск изображений, требуется уменьшенное представление объектов. Обнаружение признаков, извлечение признаков и сопоставление часто сочетаются для решения общих проблем компьютерного зрения, таких как обнаружение и распознавание объектов, поиск изображений на основе содержимого, обнаружение и распознавание лиц и классификация текстов.</p> <p>Принятие решения производится посредством сравнения признакового описания объекта с эталонным. Для этого этапа предварительно выполняются перечисленные выше шаги для базы эталонных изображений.</p> <p>Схема перечисленных этапов представлена на рисунке 1.1.</p>														
Подпись и дата															
Инв. № дубл.															
Взам. инв. №															
Подпись и дата															
Инв. № подл.															
<table border="1"> <tr> <td>Изм.</td> <td>Лист</td> <td>№ докум.</td> <td>Подпись</td> <td>Дата</td> </tr> <tr> <td></td> <td></td> <td></td> <td></td> <td></td> </tr> </table>					Изм.	Лист	№ докум.	Подпись	Дата						<div>ВКР-НГТУ-09.04.01-(М18-ИВТ-3)-004-2020(ПЗ)</div> <div>Лист 7</div>
Изм.	Лист	№ докум.	Подпись	Дата											

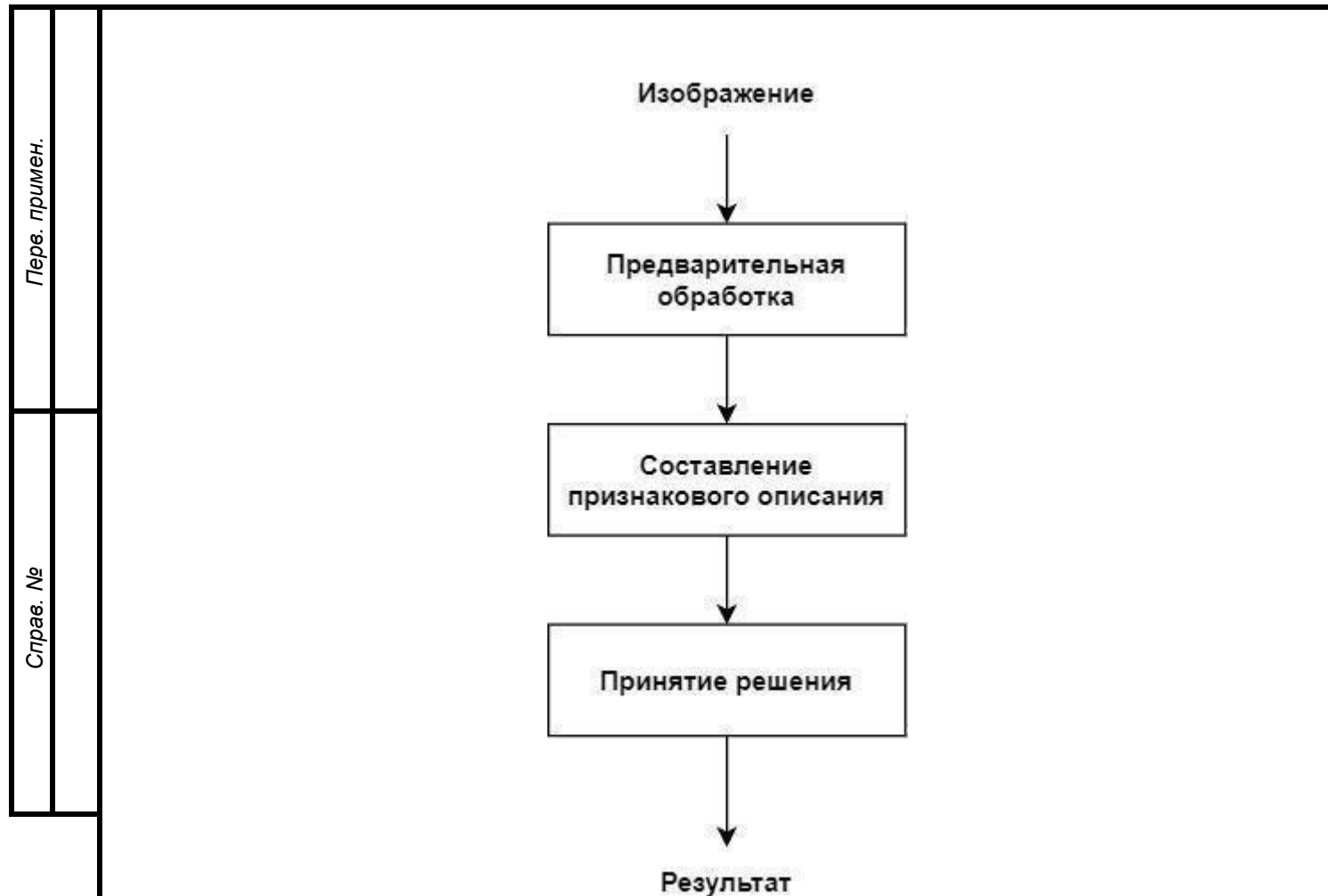


Рисунок 1.1. - Схема обнаружения объекта на изображении

1.2 Методы построения признакового описания

Признаковое описание - это информация, извлеченная из изображений в виде числовых значений. Как правило, признаки, извлеченные из изображения, имеют гораздо меньший размер, чем исходное изображение. Уменьшение размерности уменьшает накладные расходы на обработку пакета изображений.

В основном есть два типа признаковых описаний, получаемых на основе изображений. Это локальные и глобальные признаковые описания. Признаки также называют дескрипторами. Глобальные дескрипторы обычно используются для поиска изображений, обнаружения и классификации объектов, в то время как локальные дескрипторы используются для распознавания и идентификации объектов. Существует большая разница между обнаружением и идентификацией. Обнаружение - это определение существования объекта на изображении (видео), а распознавание - это определение самого объекта.

Глобальные характеристики описывают изображение в целом для обобщения всего объекта, а локальные элементы описывают участки изображения (ключевые точки на изображении) объекта.

Перв. примен.	<p>Глобальные функции включают в себя представления контуров, дескрипторы форм и объекты текстуры, а локальные элементы представляют ключевые точки изображения. Матрицы формы, моменты, гистограммно-ориентированные градиенты (HOG) и Co-HOG являются некоторыми примерами глобальных дескрипторов. SIFT, SURF, LBP, BRISK, MSER и FREAK - некоторые примеры локальных дескрипторов.</p> <p>Рассмотрим наиболее известные способы вычисления признаков описаний.</p>								
	<p align="center">1.2.1 Моменты изображения</p>								
Справ. №	<p>Моменты изображения используются для описания объектов на изображении. Используя моменты изображения [1], можно рассчитать такие значения, как площадь объекта, центр тяжести (центр объекта в терминах координат x, y) и информацию о том, как объект вращается. Обычно моменты изображения рассчитываются на основе контура изображения. Библиотека OpenCV предоставляет функцию <code>HuMoments</code> [2], которую можно использовать для характеристики структуры и формы объекта.</p> <p>Однако в пакете <code>mahotas</code> можно найти более мощные дескрипторы форм - <code>zernike_moments</code>. Подобно моментам H_u, моменты Цернике используются для описания формы объекта; однако, поскольку полиномы Цернике ортогональны друг другу, между моментами нет избыточности информации.</p> <p>При использовании моментов Цернике для описания формы стоит обратить внимание на масштабирование и перемещение объекта на изображении. В зависимости от того, где изображение переведено в изображение, моменты Цернике будут существенно отличаться. Точно так же, в зависимости от того, насколько велик или мал объект на изображении, моменты Цернике не будут идентичны. Однако величины моментов Цернике не зависят от вращения объекта, что является чрезвычайно хорошим свойством при работе с дескрипторами формы.</p> <p>Чтобы избежать дескрипторов с различными значениями, основанными на переводе и масштабировании изображения, обычно сначала выполняется сегментация. Сегментируется передний план (интересующий объект на изображении) от фона. Получив сегментацию, можно сформировать плотную ограничивающую рамку вокруг объекта и обрезать его, получая неизменность перевода.</p> <p>Наконец, можно изменить размер объекта до постоянной $N \times M$ пикселей, получая масштабную инвариантность. Отсюда легко применить моменты Цернике, чтобы охарактеризовать форму объекта.</p>								
	<p align="center">1.2.2 Гистограммы ориентированных градиентов</p>								
Ине. № подл.	<p>Гистограмма ориентированных градиентов, или сокращенно HOG, - это дескрипторы, в основном используемые в компьютерном зрении и машинном обучении для обнаружения объектов. Однако их также можно использовать для количественного определения и представления, как форм, так и текстур. Функции HOG впервые были представлены Далалом и Триггсом в их статье [3].</p>								
	<table border="1"> <tr> <td>Изм.</td> <td>Лист</td> <td>№ докум.</td> <td>Подпись</td> <td>Дата</td> </tr> </table>					Изм.	Лист	№ докум.	Подпись
Изм.	Лист	№ докум.	Подпись	Дата					
Ине. № дубл.	<p align="center">ВКР-НГТУ-09.04.01-(М18-ИВТ-3)-004-2020(ПЗ)</p>				<p align="center">Лист</p> <p align="center">9</p>				
Взам. ине. №									
Подпись и дата									

В своей работе Далал и Триггс предложили HOG и пятиступенчатый дескриптор для классификации людей на изображениях. Данные пять этапов заключаются в следующих шагах:

1. Нормализация изображения
2. Вычисление градиентов в обоих направлениях x и y
3. Получение взвешенных голосов в пространственных и ориентационных ячейках
4. Используя контраст, нормализуется перекрывающиеся пространственные ячейки
5. Собираются все гистограммы ориентированных градиентов, чтобы сформировать окончательный вектор признаков.

В большинстве реальных приложений HOG используется в сочетании с методом Linear SVM для обнаружения объектов. Причина, по которой HOG используется так интенсивно, заключается в том, что внешний вид и форму локального объекта можно охарактеризовать с помощью распределения локальных градиентов интенсивности.

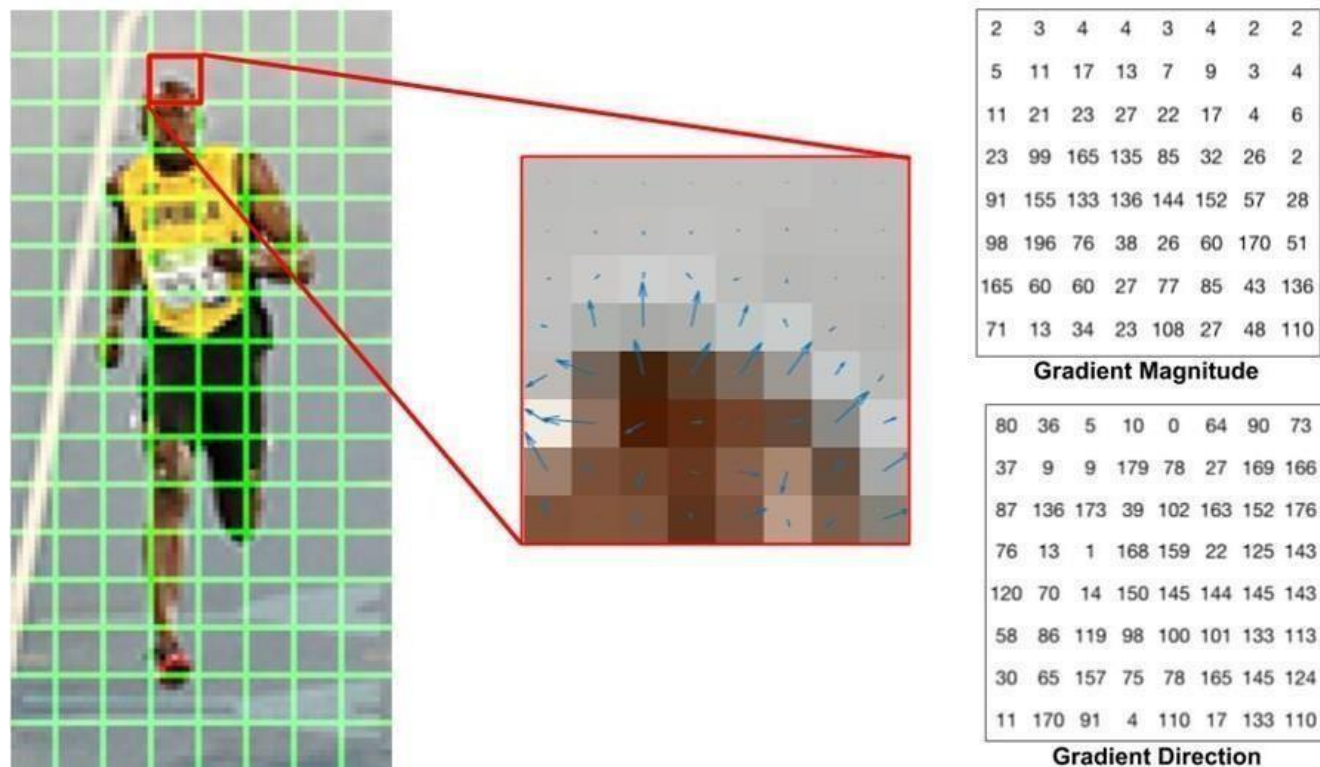
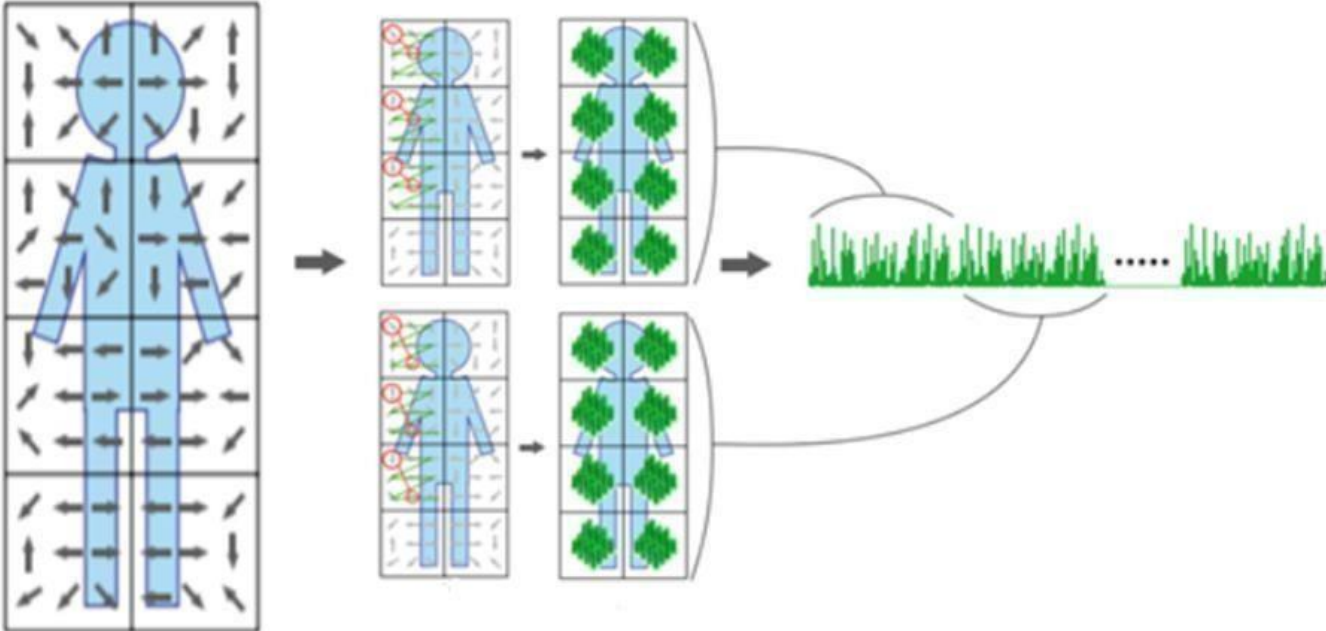


Рисунок 1.2 - Визуализация алгоритма HOG

Перв. примен.	<h3>1.2.3 Совместная гистограмма ориентированных градиентов</h3> <p>Авторы работы [4] предлагают многомерную функцию совместного использования гистограмм ориентированных градиентов (CoHOG). Данный алгоритм использует пары градиентных ориентаций в качестве единиц, из которого он строит гистограммы. В дальнейшем гистограмма называется матрицей смежности. Матрица смежности выражает распределение ориентаций градиента при заданном смещении по изображению. Комбинации соседних градиентных ориентаций могут подробно выражать фигуры.</p> <p>Алгоритм CoHOG обладает устойчивостью к деформации и дисперсии. Также преимуществом алгоритма CoHOG относительно HOG является его реализация. Рассчитать CoHOG, можно только итерируя компоненты матриц совместного использования, в то время как расчет HOG включает в себя дополнительные процедуры, такие как взвешивание по ориентации, нормализация гистограммы, перекрытие регионов и т. д. Таким образом, алгоритм CoHOG может достичь высокой производительности без сложных процедур.</p>  <p>Рисунок 1.3 - Этапы алгоритма CoHOG</p> <p>На рисунке 1.3 изображены основные этапы алгоритма CoHOG: вычисление направленных градиентов, составление комбинаций, матрицы смежности и векторизация.</p>				
	Справ. №				
Подпись и дата					
	Инв. № дубл.				
Взам. инв. №					
Подпись и дата					
Инв. № подл.					

					ВКР-НГТУ-09.04.01-(М18-ИВТ-3)-004-2020(ПЗ)	Лист 11
Изм	Лист	№ докум.	Подпись	Дата		

Перв. примен.	<h2 style="text-align: center;">1.2.4 Алгоритм GIST</h2> <p>Алгоритм GIST был представлен в статье [5]. В последнее время дескриптору GIST уделяется все больше внимания в контексте распознавания сцен, алгоритм показывает хорошие результаты при распознавании изображений. Перед извлечением дескрипторов GIST, изображение разделяется сеткой фиксированного размера, для которой вычисляются ориентированные гистограммы. Операция свертки производится на разных частях изображения, после чего результаты собираются для получения глобального описания. Таким образом, алгоритм представляет собой низко размерное представление сцены, которое не требует какой-либо формы сегментации.</p> <p>В сравнении с алгоритмом SIFT, который представлен в работе [6], алгоритм SIFT работает локально, описывая ключевые точки изображения, число которых может измеряться тысячами, в то время как алгоритм GIST на выходе формирует один вектор, описывающий изображение глобально. SIFT был разработан для масштабной и аффинной инвариантности в задачах широкого сопоставления базовых изображений. Впоследствии он был использован для классификации изображений и показал хорошие результаты. GIST был разработан для формирования целостного дескриптора, который обеспечивает более простое представление, сначала использовался для визуальной классификации сцен, которая поддается GIST. Выбор между этими дескрипторами зависит от цели применения.</p>				
Подпись и дата	Име. № дубл.	Взам. инв. №	Подпись и дата	Име. № подл.	<div style="display: flex; justify-content: space-between;"> <div> <div>Изм</div> <div>Лист</div> <div>№ докум.</div> <div>Подпись</div> <div>Дата</div> </div> <div> ВКР-НГТУ-09.04.01-(М18-ИВТ-3)-004-2020(ПЗ) </div> <div> <div>Лист</div> <div>12</div> </div> </div>

Перв. примен.	<h3>1.3 Методы локализации объектов на изображении</h3> <p>Локализация объекта на изображении – это оценка положения объекта на изображении и его границ. Далее разобраны различные подходы к решению задачи локализации.</p> <h4>1.3.1 Локализация на основе цвета</h4> <p>Методы обнаружения на основе цвета направлены на поиск области интереса с помощью основе методов пороговых значений или сегментирования. В середине 1990-х годов несколько исследователей использовали модель цветового пространства HIS. Авторы в работе [7] использовали пороговые значения цветов для сегментирования изображения. Они реализовали пороговое значение с использованием 16-разрядной справочной таблицы.</p> <p>С.Н. Lai и С.С.Yu в [8] выполнили определение цвета в пространстве HSV. Они пришли к выводу, что насыщенность цвета показывает значительные изменения, когда изображения получены с разных устройств. Авторы работы [9] изучали влияние света на цвет дорожных знаков в течение дня и ночи и пришли к выводу, что цвет придорожного изображения может быть искажен из-за света, и это может повлиять на качество изображений.</p> <p>Тем не менее, они подчеркнули, что освещение не может повлиять на различия компонентов RGB для дорожных знаков.</p> <p>К сожалению, все алгоритмы сегментации, основанные на пороговых значениях цветов, приходится постоянно корректировать пороговые значения. Поэтому цветовая сегментация всегда требует фиксации порога.</p> <h4>1.3.2 Локализация на основе форм</h4> <p>Для локализации объекта на основе его формы применяется преобразование Хафа. Преобразование Хафа - это метод, который можно использовать для выделения элементов определенной формы на изображении. Поскольку требуется, чтобы искомые характеристики были заданы в некоторой параметрической форме, классическое преобразование Хафа чаще всего используется для обнаружения регулярных кривых, таких как линии, круги, эллипсы и т. д.</p> <p>Обобщенное преобразование Хафа может использоваться в приложениях, где простое аналитическое описание объекта невозможно. Несмотря на ограничения, классическое преобразование Хафа имеет множество применений, поскольку большинство объектов содержат границы признаков, которые можно описать с помощью регулярных кривых.</p>											
	Справ. №											
Подпись и дата		Инв. № дубл.	Взам. инв. №	Подпись и дата	Инв. № подл.							
<table border="1"> <tr> <td>Изм.</td> <td>Лист</td> <td>№ докум.</td> <td>Подпись</td> <td>Дата</td> <td> ВКР-НГТУ-09.04.01-(М18-ИВТ-3)-004-2020(ПЗ) </td> <td> Лист 13 </td> </tr> </table>						Изм.	Лист	№ докум.	Подпись	Дата	ВКР-НГТУ-09.04.01-(М18-ИВТ-3)-004-2020(ПЗ)	Лист 13
Изм.	Лист	№ докум.	Подпись	Дата	ВКР-НГТУ-09.04.01-(М18-ИВТ-3)-004-2020(ПЗ)	Лист 13						

Перв. примен.	<p>Основное преимущество метода преобразования Хафа состоит в том, что он допускает пропуски в описании границ объектов и относительно не подвержен влиянию шума изображения. Техника Хафа особенно полезна для вычисления глобального описания функции, где число классов решений не должно быть известно априори, с учетом локальных измерений.</p> <p>Идея, лежащая в основе метода Хафа для обнаружения линий, состоит в том, что каждое входное измерение, например, координатная точка, указывает его вклад в глобально согласованное решение, например, физическую линию, которая дала начало этой точке изображения. Однако, данный алгоритм имеет ряд недостатков, таких как:</p> <ul style="list-style-type: none"> Алгоритм ищет только один тип объекта. Вычислительно сложен для объектов с множеством параметров 														
	Справ. №	<p>1.3.3 Локализация с использованием машинного обучения</p> <p>В машинном обучении существует множество подходов к локализации искомого объекта. Не все из них универсальны и показывают стабильные результаты. Так, известный алгоритм Виолы-Джонса [10] сильно чувствителен к поворотам объекта и условиям освещения. Методы машинного обучения, основанные на SVM, не имеют высокой точности обнаружения. Методы глубокого обучения, основанные на нейронной сети, направлены на повышение точности обнаружения целей, поэтому они были рассмотрены подробно в данной главе.</p>													
Подпись и дата		<p>Сверточные нейронные сети (R-CNN), основанные на регионах</p> <p>По мере развития области глубокого обучения, стояла задача заменить классификаторы на основе HOG более точными классификаторами, основанными на сверточных нейронных сетях (CNN). Проблема заключалась в том, что CNN были слишком медленными и вычислительно очень дорогими. Было невозможно запустить CNN на стольких регионах, генерируемых детектором скользящего окна. R-CNN [11] решает эту проблему с помощью алгоритма предложения объектов (выборочный поиск), который уменьшает количество ограничивающих блоков, которые подаются в классификатор, чтобы приблизиться к числу предложенных регионов, приблизительно равному двум тысячам. Выборочный поиск работает на локальных подсказках, такие как текстура, интенсивность, цвета и т. д., чтобы генерировать все возможные местоположения объекта.</p>													
	Име. № подл.	<table border="1"> <tr> <td>Изм.</td> <td>Лист</td> <td>№ докум.</td> <td>Подпись</td> <td>Дата</td> </tr> <tr> <td></td> <td></td> <td></td> <td></td> <td></td> </tr> </table>					Изм.	Лист	№ докум.	Подпись	Дата				
Изм.		Лист	№ докум.	Подпись	Дата										
<p align="center">ВКР-НГТУ-09.04.01-(М18-ИВТ-3)-004-2020(ПЗ)</p>					<p align="center">Лист</p> <p align="center">14</p>										

Пространственные пирамидальные сети (SPP-Net)

Сети RCNN показывает не лучшие результаты по скорости вычислений, поэтому для решения данной проблемы был предложен алгоритм SPP-Net [12]. Запуск CNN на 2000 предложенных регионах, сгенерированных при выборочном поиске, занимает много времени. SPP-Net пытался это исправить. В SPP-Net рассчитывается представление CNN по изображению всего один раз и можно использовать его для вычисления представления CNN для каждой карты признаков, полученной в результате выборочного поиска.

Фиксированный размер ввода для слоев CNN, вычисляется с помощью пространственного объединения после последнего сверточного слоя, в отличие от традиционно используемого максимального объединения. Слой SPP разделяет область произвольного размера на постоянное количество элементов разрешения, и максимальный пул выполняется для каждого элемента разрешения. Поскольку количество бинов остается неизменным, в результате создается вектор постоянного размера.

Тем не менее, у сети SPP есть один большой недостаток: выполнение обратного распространения через слой пространственного пула является нетривиальной задачей. Следовательно, сеть SPP настраивает только полностью подключенную часть сети.

Fast R-CNN

Fast RCNN [13] использует идеи из SPP-net и RCNN и устраняет ключевую проблему в SPP-Net, выполняя обучение из конца в конец. Чтобы распространять градиенты через пространственное объединение, он использует простое вычисление обратного распространения, которое схоже с операцией вычисления градиента максимального пула, за исключением того, что области объединения перекрываются, и поэтому ячейка может иметь градиенты, взятые из нескольких областей. Еще одна вещь, которую выполнил алгоритм Fast RCNN - это добавление регрессии к ограничивающей рамке в само обучение нейронной сети, выполняя одновременно задачи локализации и классификации. Эти все позволило сократить общее время обучения и увеличить точность по сравнению с сетью SPP.

Faster R-CNN

Преимущество алгоритма Faster R-CNN [14] перед вышесказанными заключается в его скорости. Самая медленная часть алгоритма Fast-RCNN – это селективный поиск.

Перв. примен.														
Справ. №														
Подпись и дата														
Инв. № дубл.														
Взам. инв. №														
Подпись и дата														
Инв. № подл.														
<table border="1"> <tr> <td>Изм.</td> <td>Лист</td> <td>№ докум.</td> <td>Подпись</td> <td>Дата</td> </tr> <tr> <td></td> <td></td> <td></td> <td></td> <td></td> </tr> </table>					Изм.	Лист	№ докум.	Подпись	Дата					
Изм.	Лист	№ докум.	Подпись	Дата										
ВКР-НГТУ-09.04.01-(М18-ИВТ-3)-004-2020(ПЗ)														
<div style="text-align: right;">Лист 15</div>														

Faster R-CNN заменяет выборочный поиск очень маленькой сверточной сетью, называемой Region Proposal Network, для генерирования областей интересов.

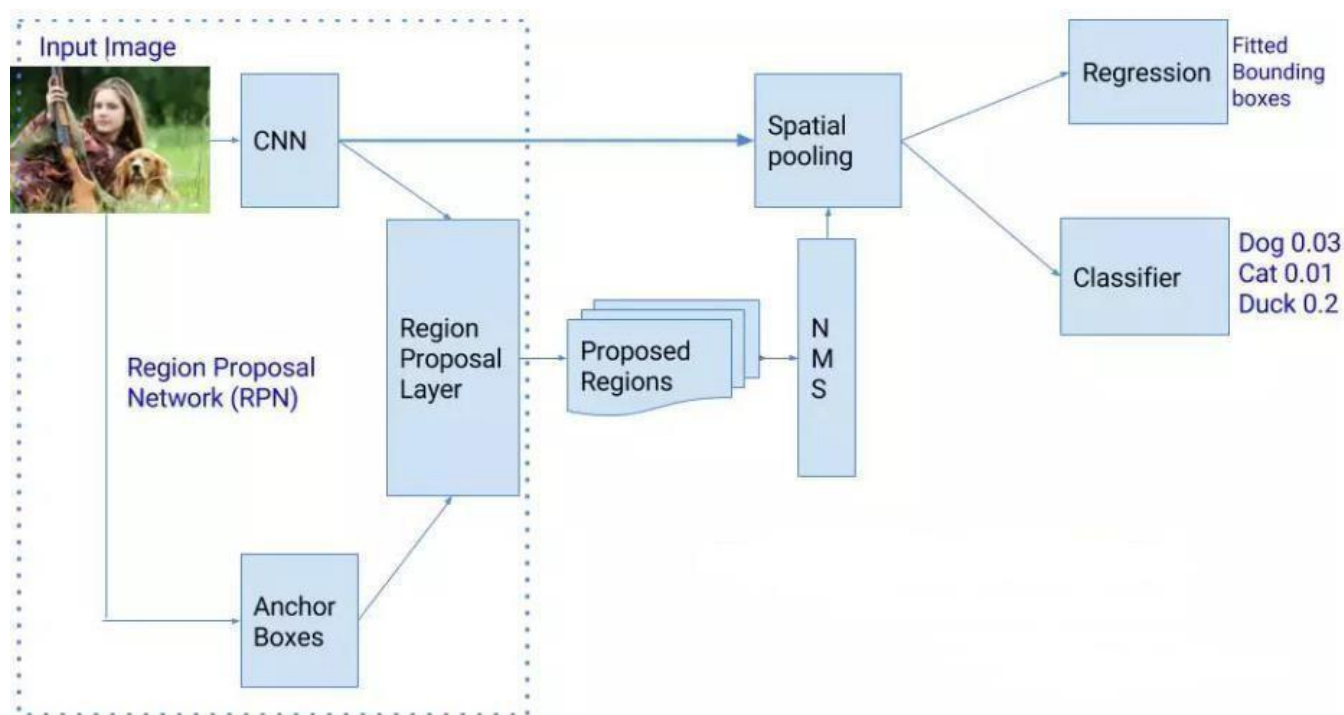


Рисунок 1.4 - Последовательность этапов Faster-RCNN

Faster R-CNN заменяет выборочный поиск очень маленькой сверточной сетью, называемой Region Proposal Network, для генерирования областей интересов.

Faster R-CNN вводит идею «якорей» в рамках сети RPN. Якорь - это коробка. В стандартной конфигурации Faster R-CNN в позиции изображения имеется 9 якорей. На следующем графике показано 9 якорей в положении (320, 320) изображения с размером (600, 800).

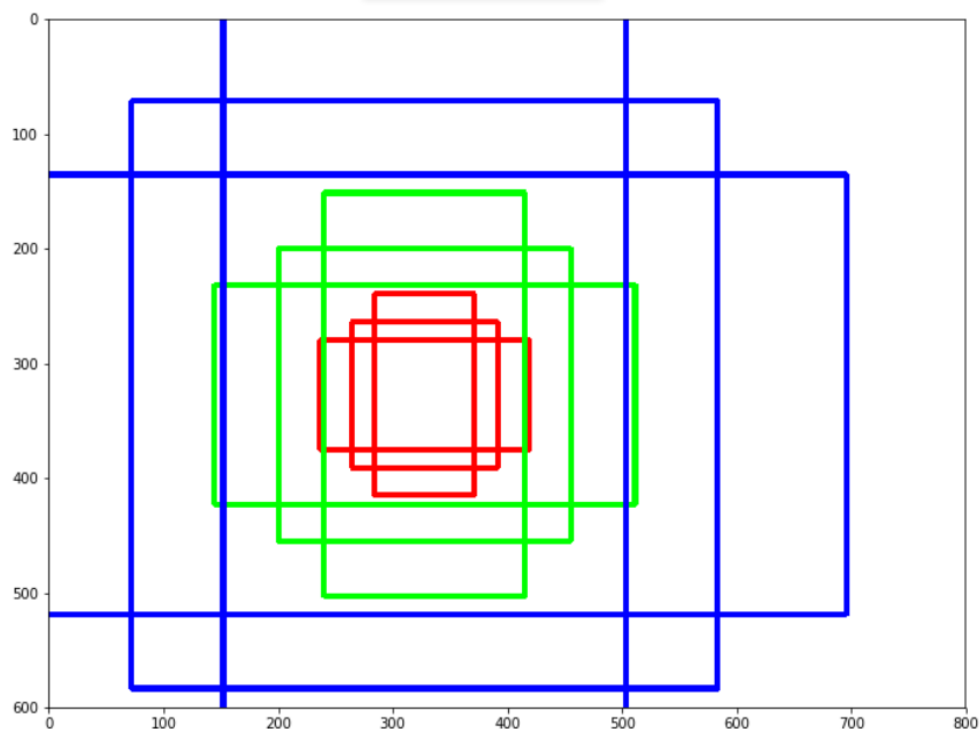


Рисунок 1.5 - Девять якорей на изображении 600x800

На основе якорей, сеть RPN предсказывает вероятности классов для полученных рамок. Оставшаяся сеть похожа на работу Fast-RCNN. Faster-RCNN работает в 10 раз быстрее, чем Fast-RCNN с такой же точностью наборов данных, как VOC-2007. Поэтому Faster-RCNN является одним из самых точных алгоритмов обнаружения объектов.

Сравним работу алгоритмов RCNN, Fast-RCNN и Faster-RCNN.

Таблица 1.1 Сравнение скорости работы алгоритмов RCNN, Fast-RCNN и Faster-RCNN

	RCNN	Fast-RCNN	Faster-RCNN
Тестовое время на изображение	50 секунд	2 секунды	0.2 секунды
Ускорение	1x	25x	250x

YOLO (You only Look Once)

Существуют методы, которые ставят проблему определения объекта как задачу регрессии. Два самых популярных из них - YOLO и SSD. YOLO

[15] делит каждое изображение на сетку $S \times S$, и каждая сетка предсказывает N ограничивающих рамок и достоверность.

Достоверность отражает точность ограничивающего прямоугольника и то, содержит ли ограничивающий прямоугольник объект (независимо от класса). YOLO также предсказывает оценку классификации для каждой рамки для каждого класса в обучении.

Таким образом, прогнозируется общее количество блоков $S \times S \times N$. Тем не менее, большинство из этих блоков имеют низкие показатели достоверности, и если установить порог, скажем, 30% достоверности, можно удалить большинство из них, как показано в примере ниже:

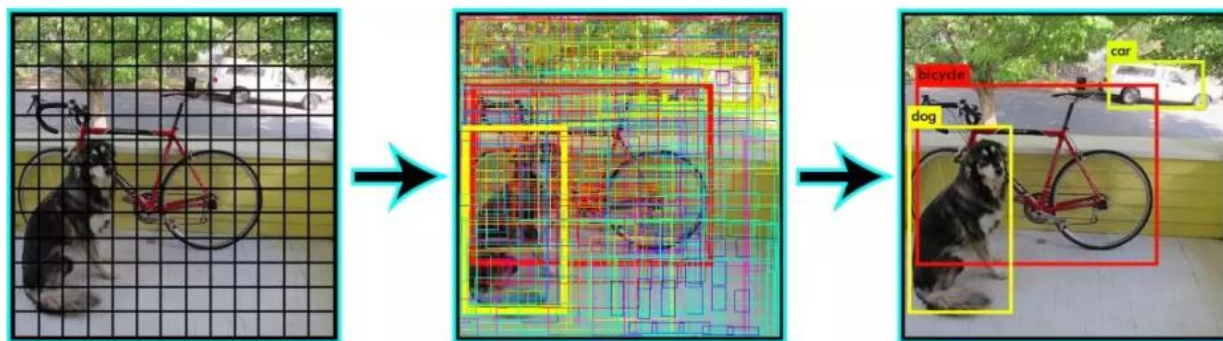


Рисунок 1.6 – Визуализация алгоритма YOLO

Во время выполнения, изображение проходит через операцию CNN только один раз. Следовательно, YOLO работает очень быстро и может работать в режиме реального времени. Еще одно ключевое отличие состоит в том, что YOLO сразу видит полное изображение, а не только предложенных регионов. Таким образом, эта контекстная информация помогает избежать ложных срабатываний. Однако одно ограничение для YOLO состоит в том, что он предсказывает только один тип класса в одной сетке, следовательно, с очень маленькими объектами работает не лучшим образом.

SSD (Single Shot Detector)

Single Shot Detector является компромиссом в плане баланса скорости и точности. SSD запускает сверточную сеть на входном изображении только один раз и рассчитывает карту признаков. Теперь запускается небольшое сверточное ядро размером 3×3 на этой карте характеристик, чтобы предсказать ограничивающие рамки и вероятность классификации. SSD также использует якорные блоки с различным соотношением сторон, аналогичные Faster-RCNN, и изучает смещение, а не поле. Чтобы справиться с масштабом, SSD предсказывает ограничивающие рамки после нескольких сверточных слоев. Поскольку каждый сверточный слой работает в разном масштабе, он способен обнаруживать объекты различного масштаба.

1.3 Выводы по первой главе

В данной главе были рассмотрены подходы к построению признаковых описаний объектов, а также их локализации на изображении. Приведены преимущества и недостатки этих методов. Количество и качество существующих решений для детектирования объектов подтверждает актуальность исследований в этой области, и подтверждает важность формирования новых решений. В ходе обзора, был выявлен ряд проблем, возникающих при детектировании объектов на изображениях:

Проблема формирования глобального признакового описания. Система признаков должна обеспечивать точность, простоту и скорость вычислений. Признаковое описание является ключевым этапом распознавания образов. Задача нахождения оптимального по всем критериям подхода является актуальной до сих пор.

Проблема принятия решения о локализации объекта. Локализация объекта на изображении представляет собой сложную задачу, которая требует вычислительных ресурсов. На данный момент лучше всего с задачей справляются решения с использованием нейронных сетей.

Таким образом, данная работа посвящена разработке нового подхода, решающего задачу детектирования объектов на изображении. Теоретические основы предлагаемого подхода описаны во второй главе.

Перв. примен.

Справ. №

Подпись и дата

Инв. № дубл.

Взам. инв. №

Подпись и дата

Инв. № подл.

ВКР-НГТУ-09.04.01-(М18-ИВТ-3)-004-2020(ПЗ)

Лист

19

Изм Лист № докум. Подпись Дата

Глава 2. Информационная модель распознавания дорожных знаков на изображениях

Предлагаемый в данной работе подход к решению задачи детектирования объекта на изображении заключается в применении метода локализации объекта с помощью нейронной сети, и с применением положений теории активного восприятия об алгебре групп [20] на этапе формирования глобального признакового описания выделенного объекта. На этапе формирования признакового описания для изображений применяется операция вычисления замкнутых групп, согласно теории ТАВ. Полученные группы являются признаками, которые впоследствии используются для классификации изображений.

На этапе принятия решения о распознанном на изображении объекте применяется градиентный бустинг для классификации.

2.1 Схема системы распознавания дорожных знаков на изображениях

Общий алгоритм работы системы следует разделить на два этапа:

1. Начальная подготовка системы, обучение и построение моделей
2. Детектирование дорожных знаков на основе построенных моделей

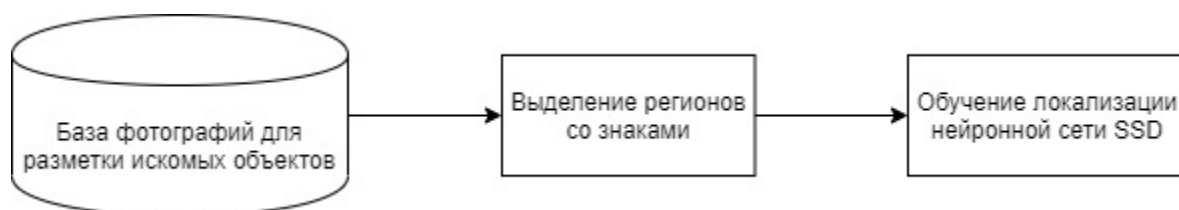


Рисунок 2.1 – Инициализация системы для локализации

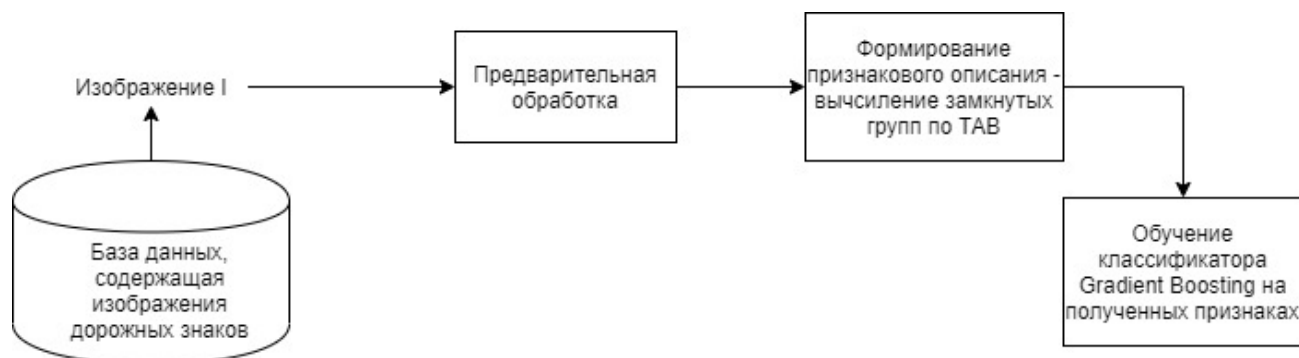
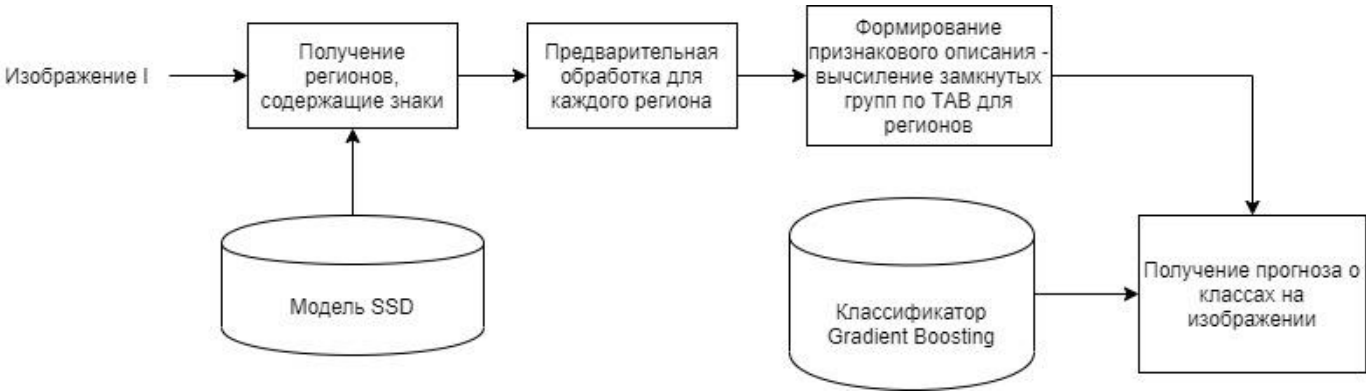


Рисунок 2.2 – Инициализация системы для классификации

Перв. примен.	<div></div>														
Справ. №	<p>Рисунок 2.3 – Схема распознавания знаков на изображении</p> <p>Первичная инициализация предполагает обучение локализации сети SSD, также вычисление замкнутых групп (признаков) по базе эталонных изображений и составление модели, классифицирующей знаки на изображениях на основе полученных признаков – замкнутых групп.</p> <p>Второй этап ожидает на вход изображение, содержащее объекты, которые нужно распознать (дорожные знаки). Это изображение проходит все описанные выше этапы предварительной обработки, формирования признакового описания, которое впоследствии используется для принятия решения. Результатом является название распознанных классов.</p> <p>Данная работа посвящена распознаванию дорожных знаков. Таким образом, база эталонных объектов содержит изображения распознаваемых знаков. База анализируемых изображений содержит фотографии, на которых присутствуют данные объекты.</p> <p>Далее приведено более детальное описание этапов.</p>														
Подпись и дата															
Инв. № дубл.	<p>2.1.1 Предварительная обработка изображения</p>														
Взам. инв. №	<p>Изображение, подаваемое на вход системы, находится в цветовом пространстве RGB. В таком цветовом пространстве каждый пиксель изображения является вектором из трех элементов – яркостей красного, зеленого и синего цвета. На основе этого, функция яркости выглядит следующим образом:</p>														
Подпись и дата	$I(x,y)=\frac{(R_{xy}+G_{xy}+B_{xy})}{3},x\in X,y\in Y\quad (2,1)$														
Инв. № подл.	<table><tr><td></td><td></td><td></td><td></td><td></td></tr><tr><td>Изм</td><td>Лист</td><td>№ докум.</td><td>Подпись</td><td>Дата</td></tr></table> <div>ВКР-НГТУ-09.04.01-(М18-ИВТ-3)-004-2020(ПЗ)</div> <div>Лист 21</div>										Изм	Лист	№ докум.	Подпись	Дата
Изм	Лист	№ докум.	Подпись	Дата											

Значения R_{xy}, G_{xy}, B_{xy} представляют собой яркости цветов на пикселе с координатами (x,y);

X,Y – это множества координат данного изображения.

Отсюда, функции яркости изображения для конкретного пикселя вычисляется как среднее значение яркостей компонент RGB этого пикселя. Матрица яркостей изображения представляется набором значений функции $I(x,y)$ по всему изображению на X и Y.

2.1.2 Локализация дорожных знаков на изображении

Учитывая перечисленные преимущества и недостатки, для локализации была использована нейронная сеть SSD. Нейронная сеть SSD хорошо работает на графических процессорах среднего уровня и имеет оптимальное соотношение скорости и точности для данной задачи.

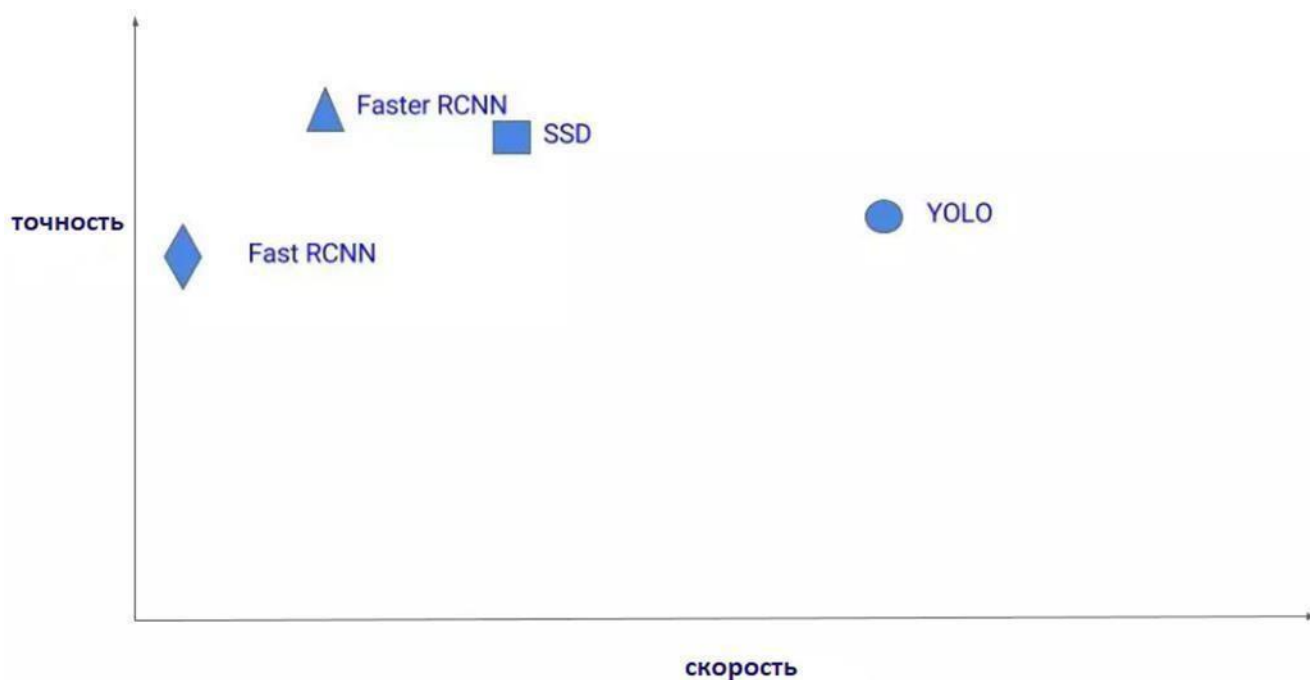


Рисунок 2.4 – Соотношение скорости и точности вычислений нейронных сетей

Обнаружение объекта в SSD происходит в два этапа:

1. Извлекаются карты признаков
2. Применяются фильтры свертки для детектирования объектов

SSD использует сеть VGG16 для извлечения карт признаков. Затем он обнаруживает объекты, используя слой Conv4_3.

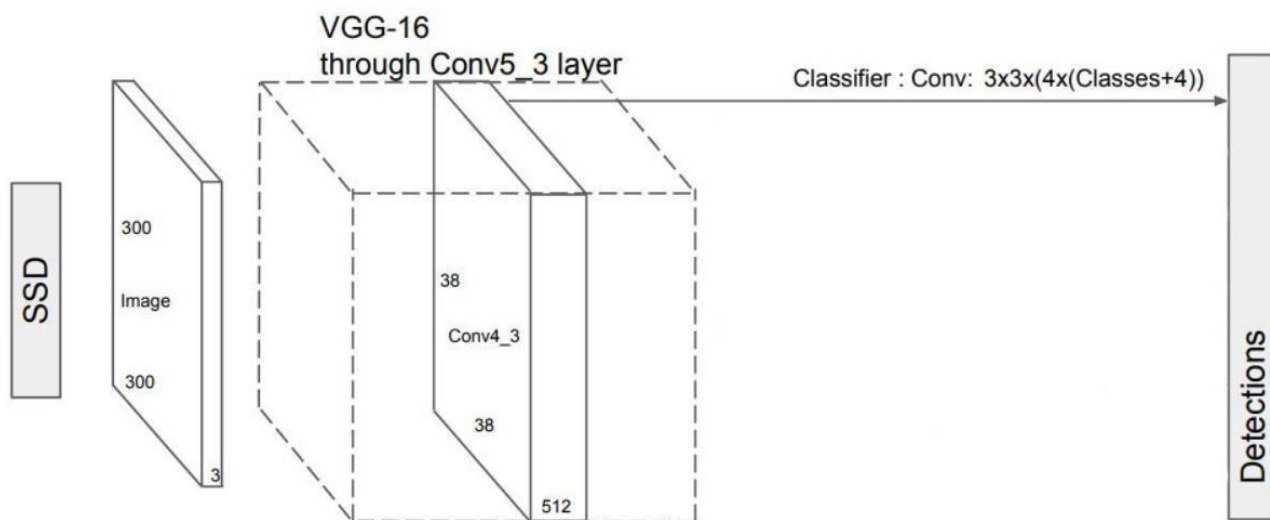
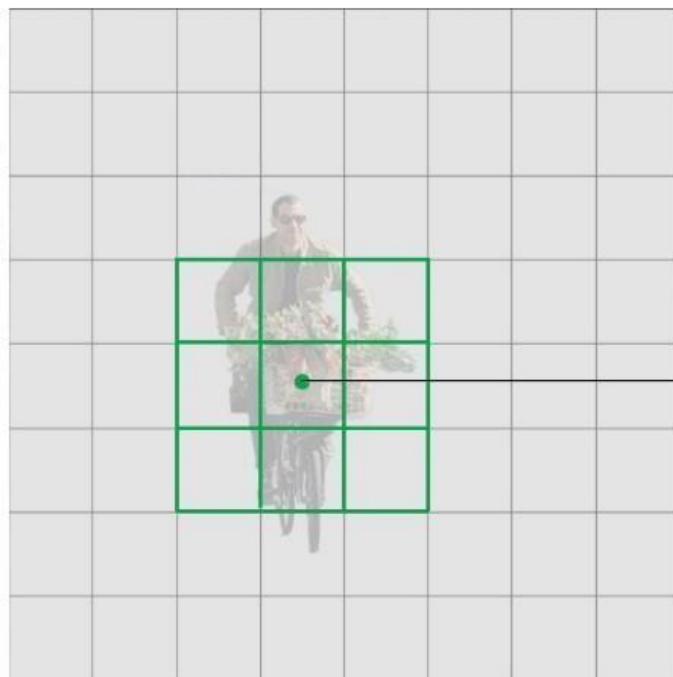


Рисунок 2.5 – SSD сеть

Каждое предсказание состоит из рамки и оценки для каждого класса, выбирается наивысшая оценка класса для объекта. Conv4_3 делает в общей сложности $38 \times 38 \times 4$ прогнозов: четыре прогноза на ячейку независимо от глубины карт признаков. SSD резервирует класс «0», чтобы указать, что у него нет объектов. В данной работе дополнительно используется класс «1» чтобы указать, что объект – дорожный знак. SSD вычисляет местоположение и оценки класса, используя маленькие фильтры свертки. После извлечения карт признаков, SSD применяет фильтры свертки 3×3 для каждой ячейки, чтобы делать прогнозы. (Эти фильтры вычисляют результаты точно так же, как обычные фильтры CNN.)


 $(\Delta cx, \Delta cy, w, h)$

+

person: 0.86
bike: 0.75
...
car: 0.5

Рисунок 2.6 – Использование фильтра 3x3 для детектирования объекта

SSD использует несколько слоев (многомасштабные карты признаков) для независимого обнаружения объектов. По мере того, как CNN постепенно уменьшает пространственное измерение, разрешение карт объектов также уменьшается. SSD использует слои с более низким разрешением для обнаружения объектов более крупного масштаба.

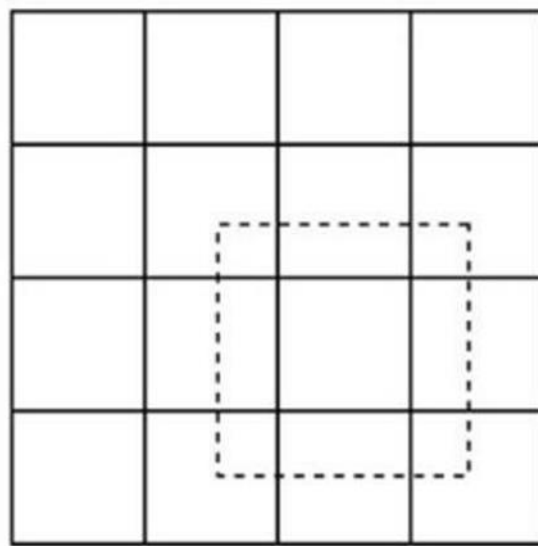
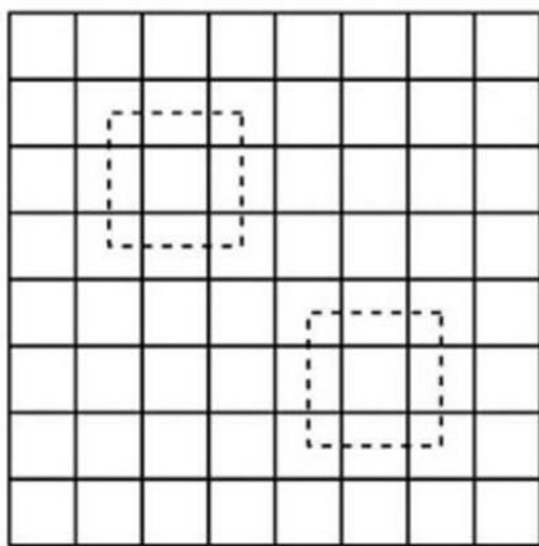


Рисунок 2.7 – Использование фильтра с более низким расширением для обнаружения крупных объектов

SSD добавляет еще 6 вспомогательных слоев свертки после VGG16. Пять из них будут добавлены для обнаружения объектов.

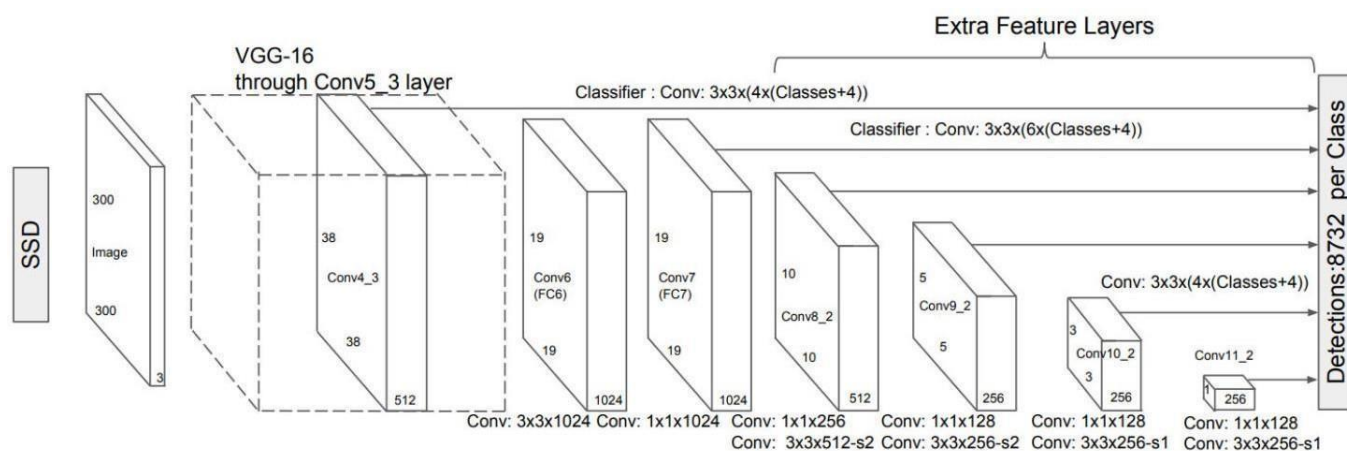


Рисунок 2.8 – SSD сеть со вспомогательными слоями

2.1.3 Построение глобального признакового описания по ТАВ

Обзор теории активного восприятия

Теория активного восприятия [20] была разработана Владимиром Александровичем Утробиним, профессором НГТУ им. Р.Е. Алексеева. ТАВ применяется при решении задач составления признакового описания, обработки и анализа изображений. Одним из основных моментов теории активного восприятия является операция U-преобразования, которая применяется по изображениям U-преобразование - это последовательное применение операций интегрирования и дифференцирования по изображению, на выходе такого преобразования составляется набор спектральных коэффициентов, или, другими словами, признаковое описание изображения. В теории активного восприятия существует понятие «Алгебра групп». Раздел алгебры групп служит для нахождения зависимостей между спектральными коэффициентами, полученными на выходе вычисления U-преобразования. В данной работе, зависимости, полученные из алгебры групп, используются при формировании глобального признакового описания изображения.

U-преобразование и Q-преобразование

Подаваемое на вход система изображение сформировано как функция яркости $I(x, y)$,

$x \in X, y \in Y, X$ и Y - конечные множества. После выполнения U-преобразования по изображению, находится спектральное представление изображения:

$$D = \{d_i\}, i \in \overline{1, M} \quad (2,2)$$

M – число сегментов, на которые было разбито исследуемое изображение;
 d_i – спектральное представление сегмента i , состоящее из L спектральных коэффициентов (L – число фильтров, используемых на этапе дифференцирования);
 $d_i\{k\}$ – k -й спектральный коэффициент представления сегмента I ,
 $k \in \overline{1, L}$

Формирование исходного описания с позиции теории активного восприятия изображения является операцией интегрирования, и включает в себя два таких шага:

- Нормирование изображения
- Вычисление матрицы «визуальных масс»

Нормирование изображения делается с помощью путем вычитания из каждого элемента матрицы минимального элемента матрицы, а затем делением каждого элемента матрицы на максимальный элемент матрицы. Для получения матрицы «визуальных масс», согласно ТАВ нужно совершить Q-преобразование. Для совершения Q-преобразования нужно произвести разбиение матрицы на части размером $N/4 * M/4$. Q-преобразование описывается формулой (2,3):

$$B_{kl} = \sum_{i=1}^2 \sum_{j=1}^2 A_{kl}[i, j] \quad A_{kl}[i, j] \quad k = \overline{1, 4} \quad l = \overline{1, 4} \quad (2,3)$$

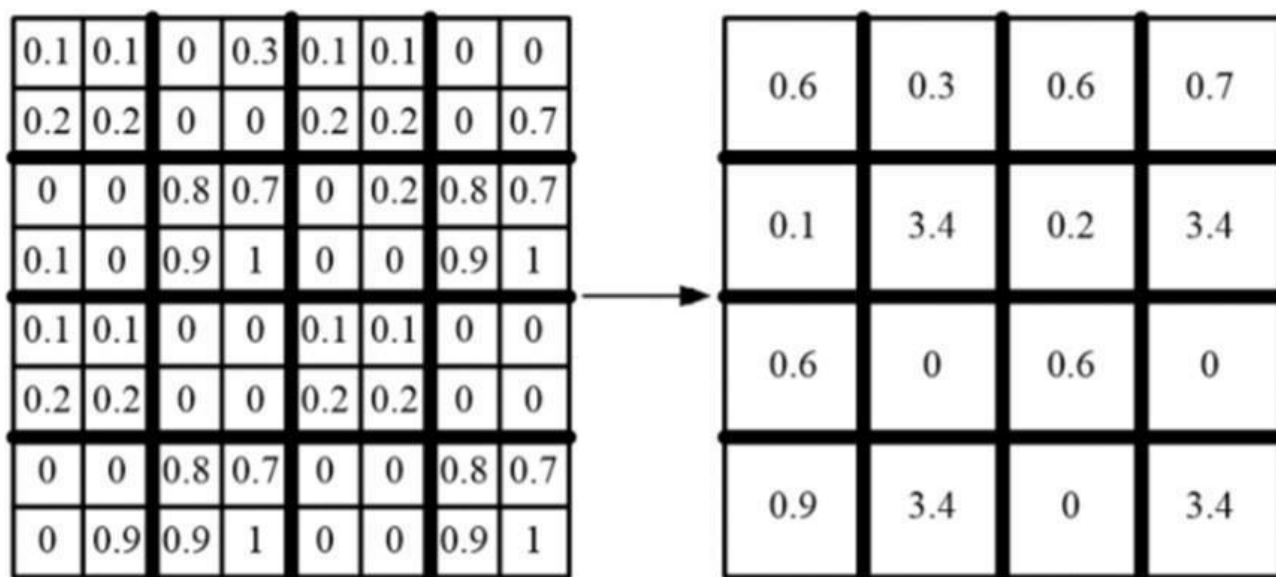


Рисунок 2.9 – Q-преобразование

Создание системы признаков с согласно ТАВ - это операция дифференцирования. Результатом операции дифференцирования является вектор «спектральных коэффициентов» $\mu = (\mu_0, \mu_1, \mu_2, \dots, \mu_{15})$

Дифференцирование выполняется с использованием 16 фильтров (рисунок 2.11), которые применяются к полученной до этого матрице визуальных масс, на рисунке серый компонент значит умножение элемента матрицы визуальных масс на -1, белый компонент – умножение элемента матрицы визуальных масс на +1. Значения, полученные после операций с фильтрами, складываются между собой, и их сумма на выходе является спектральным коэффициентом.

Далее приведен пример нахождения U-преобразования по одному сегменту изображения с применением фильтров для дифференцирования. Сначала производится интегрирование – изображение делится на 16 частей одинакового размера. Значения пикселей в каждой такой части изображения складываются, и получается матрица визуальных масс. Данный этап изображен на рисунке 2.10:



1209	1446	1027	741
1786	2686	1955	1273
426	688	500	317
288	365	281	243

Рисунок 2.10 – Формирование матрицы «визуальных масс»

Как видно на рисунке, чем светлее участок изображения – тем выше значение матрицы визуальных масс, чем темнее – тем значение матрицы наоборот, ниже. После производится дифференцирование. При выполнении этого этапа матрица визуальных масс поэлементно умножается на фильтры, изображенные на рисунке 2.11. Белый элемент фильтра значит «1», серый элемент значит «-1».

Изм.	Лист	№ докум.	Подпись	Дата	ВКР-НГТУ-09.04.01-(М18-ИВТ-3)-004-2020(ПЗ)	Лист 28

		Перв. примен.			
		Справ. №			
				</	

Перв. примен.	Таким образом, после умножения на фильтр под номером восемь, матрица визуальных масс выглядит так, как показан на рисунке 2.12:																																						
	<table><tr><td>1209</td><td>-1446</td><td>-1027</td><td>741</td></tr><tr><td>-1786</td><td>2686</td><td>1955</td><td>-1273</td></tr><tr><td>-426</td><td>688</td><td>500</td><td>-317</td></tr><tr><td>288</td><td>-365</td><td>-281</td><td>243</td></tr></table>					1209	-1446	-1027	741	-1786	2686	1955	-1273	-426	688	500	-317	288	-365	-281	243																		
1209	-1446	-1027	741																																				
-1786	2686	1955	-1273																																				
-426	688	500	-317																																				
288	-365	-281	243																																				
Справ. №	Рисунок 2.12 – Матрица визуальных масс																																						
	Спектральный коэффициент, соответствующий восьмому фильтру, вычисляется как алгебраическая сумма всех элементов матрицы после применения этого фильтра. Также вычисляются все шестнадцать спектральных коэффициентов по изображению. Результаты применения U-преобразования по изображению приведены в таблице 2.1.																																						
Таблица 2.1 – Спектральные коэффициенты																																							
<table><tr><th>Фильтр</th><th>Спектральный коэффициент</th></tr><tr><td>0</td><td>15232</td></tr><tr><td>1</td><td>-2556</td></tr><tr><td>2</td><td>9015</td></tr><tr><td>3</td><td>-1703</td></tr><tr><td>4</td><td>2667</td></tr><tr><td>5</td><td>4030</td></tr><tr><td>6</td><td>1547</td></tr><tr><td>7</td><td>-526</td></tr><tr><td>8</td><td>1388</td></tr><tr><td>9</td><td>-285</td></tr><tr><td>10</td><td>2524</td></tr><tr><td>11</td><td>-50</td></tr><tr><td>12</td><td>-191</td></tr><tr><td>13</td><td>-308</td></tr><tr><td>14</td><td>730</td></tr><tr><td>15</td><td>-228</td></tr></table>						Фильтр	Спектральный коэффициент	0	15232	1	-2556	2	9015	3	-1703	4	2667	5	4030	6	1547	7	-526	8	1388	9	-285	10	2524	11	-50	12	-191	13	-308	14	730	15	-228
Фильтр	Спектральный коэффициент																																						
0	15232																																						
1	-2556																																						
2	9015																																						
3	-1703																																						
4	2667																																						
5	4030																																						
6	1547																																						
7	-526																																						
8	1388																																						
9	-285																																						
10	2524																																						
11	-50																																						
12	-191																																						
13	-308																																						
14	730																																						
15	-228																																						
Ине. № подл.					Лист 30																																		
	Изм.	Лист	№ докум.	Подпись																																			
Ине. № дубл.	ВКР-НГТУ-09.04.01-(М18-ИВТ-3)-004-2020(ПЗ)																																						
Взам. инв. №																																							
Подпись и дата																																							
Подпись и дата																																							

Перв. примен.	<p>Таким образом, были найдены получены 16 спектральных коэффициентов после применения U-преобразования. Полученные коэффициенты составляют признаковое описание.</p> <p style="text-align: center;">Алгебра групп</p> <p>Алгебра групп – это один из разделов теории активного восприятия. Алгебра групп существует для нахождения зависимостей между спектральными коэффициентами. Для выявления таких зависимостей, данный раздел вводит определения операторов, полных и замкнутых групп. Множество бинарных операторов $\{Vi\}$ вычисляется на основе множества фильтров $\{Fi\}$, изображенных на рисунке 2.11. Для получения оператора по фильтру нужно сделать следующее соответствие между значениями элементов фильтра и оператора: $(+1 \rightarrow 1)$, $(-1 \rightarrow 0)$. Пример нахождения бинарного оператора 3 на основе фильтра 3 приведен на рисунке 2.13:</p> <div style="display: flex; align-items: center; justify-content: center; margin: 20px 0;"> <table border="1" style="border-collapse: collapse; text-align: center;"> <tr><td>-1</td><td>-1</td><td>1</td><td>1</td></tr> <tr><td>-1</td><td>-1</td><td>1</td><td>1</td></tr> <tr><td>1</td><td>1</td><td>-1</td><td>-1</td></tr> <tr><td>1</td><td>1</td><td>-1</td><td>-1</td></tr> </table> <div style="margin: 0 20px; font-size: 24px;">→</div> <table border="1" style="border-collapse: collapse; text-align: center;"> <tr><td>0</td><td>0</td><td>1</td><td>1</td></tr> <tr><td>0</td><td>0</td><td>1</td><td>1</td></tr> <tr><td>1</td><td>1</td><td>0</td><td>0</td></tr> <tr><td>1</td><td>1</td><td>0</td><td>0</td></tr> </table> </div> <p>Рисунок 2.13 – Вычисление бинарного оператора 3 на основе фильтра 3</p> <p>На основании полученных операторов формируются множества полных и замкнутых групп:</p> <ul style="list-style-type: none"> •Множество полных групп $Pn = \{Pni\}$, $Pni = \{ , Vj , Vk \}$ мощностью 35. •Множество замкнутых групп $Ps = \{Psi\}$, $Psi = \{ , Vj , Vk, Vr \}$ мощностью 105. <p>Количество существующих полных и замкнутых групп можно увеличить с помощью использования в их составе инверсных операторов. В итоге получается общее количество полных групп – 140, замкнутых – 840.</p> <p>С помощью таких групп можно выполнять спектрально-корреляционный анализ изображения. Полные группы находят корреляционные связи между операторами, замкнутые группы – связи между полными группами. Отношения операторов, полных и замкнутых групп изображены на рисунке 2.14:</p>					-1	-1	1	1	-1	-1	1	1	1	1	-1	-1	1	1	-1	-1	0	0	1	1	0	0	1	1	1	1	0	0	1	1	0	0
	-1	-1	1	1																																	
-1	-1	1	1																																		
1	1	-1	-1																																		
1	1	-1	-1																																		
0	0	1	1																																		
0	0	1	1																																		
1	1	0	0																																		
1	1	0	0																																		
Справ. №																																					

Изн.	Лист	№ докум.	Подпись	Дата	Изн.	Лист	№ докум.	Подпись	Дата	ВКР-НГТУ-09.04.01-(М18-ИВТ-3)-004-2020(ПЗ)	Лист 31

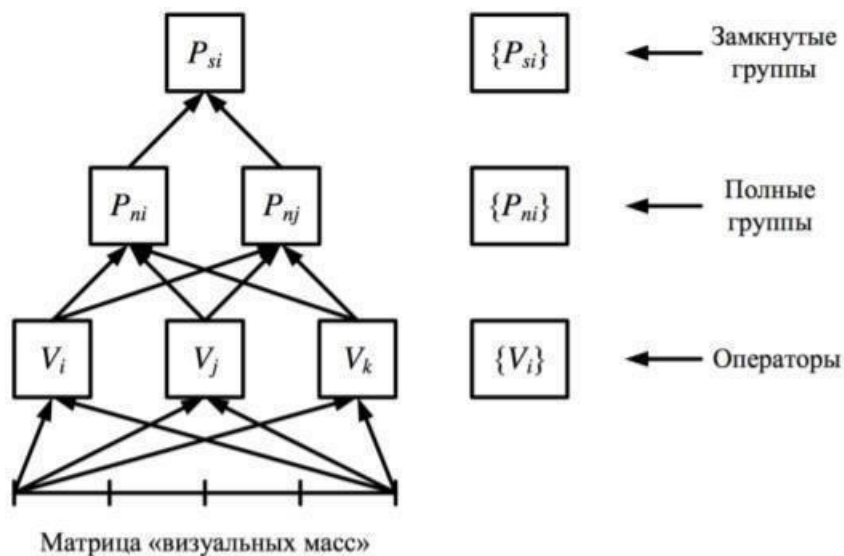


Рисунок 2.14 – Иерархическая структура операторов, полных и замкнутых групп

Для оператора или группы ставится в соответствие понятие образа оператора или группы. Образ оператора 3 изображен на рисунке 2.13. На основании таких образов производится вычисление масс операторов, полных и замкнутых групп. Пример нахождения массы замкнутой группы по сегменту изображения (рисунок 2.10) изображен на рисунке 2.15:

$$\Sigma \left(\begin{array}{c} \text{Матрица визуальных масс} \\ \begin{array}{|c|c|c|c|} \hline 1209 & 1446 & 1027 & 741 \\ \hline 1786 & 2686 & 1955 & 1273 \\ \hline 426 & 688 & 500 & 317 \\ \hline 288 & 365 & 281 & 243 \\ \hline \end{array} \end{array} \right) \times \left(\begin{array}{c} \text{Образ группы } V_1 V_3 V_7 \overline{V_{12}} \\ \begin{array}{|c|c|c|c|} \hline 0 & 0 & 1 & 1 \\ \hline 0 & 0 & 1 & 1 \\ \hline 0 & 0 & 1 & 1 \\ \hline 1 & 1 & 0 & 0 \\ \hline \end{array} \end{array} \right) = 6466$$

Рисунок 2.15 – Вычисление массы замкнутой группы по сегменту изображения

Так находятся массы операторов, полных и замкнутых групп, вычисленных на основе спектрального представления исследуемого изображения. Информация о наличии тех или иных полных и замкнутых групп в спектре изображения или его сегмента, а также массы данных групп - это базис, на котором формируется новый вид признакового описания изображения, предлагаемый ТАВ.

Име. № подл.	Подпись и дата	Взам. инв. №	Инв. № дубл.	Подпись и дата	Справ. №	Перв. примен.	<p>Такой способ составления признакового описания является более сложным и массивным по сравнению с U-преобразованием. При этом, такое признаковое описание, полученное с помощью алгебры групп включает больше информации, данные о корреляционных связях между спектральными коэффициентами. Таким образом получается более подробное признаковое описание.</p>				
Изм.	Лист	№ докум.	Подпись	Дата	ВКР-НГТУ-09.04.01-(М18-ИВТ-3)-004-2020(ПЗ)						
					Лист 33						

2.2 Выводы по второй главе

Во второй главе был сделан теоретический обзор подхода к обнаружения дорожных знаков на изображении согласно ТАВ. В данной главе были представлены схемы этапов по решению задач локализации и классификации. Этапы решения задач можно описать так:

1. Предварительная обработка изображения – нормализация, нахождение функции яркости.
2. Составление глобального признакового описания на основе алгебры групп по теории активного восприятия
3. Принятие решения о локализации дорожного знака – нейронная сеть SSD.
4. Классификация знака - выполняется с использованием градиентного бустинга.

Применение таких этапов для решения задачи обнаружения дорожных знаков на изображении доказывает научную новизну предлагаемого подхода. В третьей главе подробно рассмотрены эксперименты по реализации этапов решения задачи и оценены итоговые результаты работы системы.

Перв. примен.

Справ. №

Подпись и дата

Инв. № дубл.

Взам. инв. №

Подпись и дата

Инв. № подл.

ВКР-НГТУ-09.04.01-(М18-ИВТ-3)-004-2020(ПЗ)

Лист

34

Изм. Лист № докум. Подпись Дата

Глава 3. Вычислительный эксперимент

В данном разделе описаны эксперименты по комбинации методов теории активного восприятия на этапе предварительной обработки и формирования признакового описания, применению нейронной сети для локализации дорожных знаков на фотографии и использование различных моделей машинного обучения для классификации.

Вычислительный эксперимент служит доказательством работоспособности и применимости выполненного исследования по решению задачи обнаружения дорожных знаков на фотографии.

Анализ результатов позволяет сделать выводы о точности и конкурентоспособности реализованного подхода.

3.1 Описание эксперимента

Для проведения эксперимента был собран набор данных, состоящий из изображений искомых знаков. В качестве объектов выбраны конкретно знаки, связанные с парковкой автомобиля. В базе содержатся изображения 10 различных дорожных знаков, в качестве которых взяты наиболее основные и часто встречающиеся знаки, имеющих отношение к парковке автомобиля. Для каждого из классов в наборе данных содержится по 300 изображений для обучения. Также было собрано и размечено 200 изображений для обучения локализации знаков сети SSD.



Рисунок 3.1 – Пример изображений в наборе данных для расчета признаков

Перв. примен.	<div></div> <p>Рисунок 3.2 – Разметка изображения для обучения сети SSD локализации</p>														
Справ. №	<h3>3.2 Описание программной системы</h3> <p>Для проведения описанного выше вычислительного эксперимента была разработана система клиент-сервер. Клиент представляет собой Android-приложение, разработанное на языке Kotlin. Серверное приложение разработано с помощью языков R и Python. Работа система изображена на рисунке 3.3.</p> <div></div> <p>Рисунок 3.3 – Схема сообщения компонентов системы</p> <h3>Набор исходных данных</h3> <p>Как было сказано прежде, для обучения системы необходим набор данных, для которых будут вычисляться признаки. Размер набора был выявлен в ходе вычислительного эксперимента, в результате сравнения зависимости количества изображений на класс от точности классификатора на таком наборе. Зависимость показана в таблице 3.1. Для таблицы взяты данные классификатора Gradient Boosting.</p>														
Подпись и дата															
Инв. № дубл.															
Взам. инв. №															
Подпись и дата															
Инв. № подл.	<div><table><tr><td>Изм.</td><td>Лист</td><td>№ докум.</td><td>Подпись</td><td>Дата</td></tr><tr><td></td><td></td><td></td><td></td><td></td></tr></table><div>ВКР-НГТУ-09.04.01-(М18-ИВТ-3)-004-2020(ПЗ)</div><div>Лист 36</div></div>					Изм.	Лист	№ докум.	Подпись	Дата					
Изм.	Лист	№ докум.	Подпись	Дата											

Таблица 3.1 – Зависимость классификатора от данных

Количество изображений на класс	Точность классификатора на тестовой выборке
50	0.50
100	0.68
200	0.77
300	0.99

Классификатор

Для того, чтобы система могла определять класс знака, необходимо подобрать и обучить классификатор. Для наиболее лучших результатов, были выбраны и протестированы наиболее используемые классификаторы, такие как SVM, k-NN, Decision Tree и Gradient Boosting.

Результаты эксперимента с каждым из перечисленных классификаторов на стандартном наборе параметров представлены в таблице, с учетом того, что на каждый класс приходится 300 эталонных изображений.

Таблица 3.2 – Зависимость точности от классификатора

Классификатор	Точность классификатора на тестовой выборке	Точность классификатора на тренировочной выборке
KNeighborsClassifier	0.78	0.84
SVC	0.64	0.99
DecisionTreeClassifier	0.85	0.99
GradientBoostingClassifier	0.95	0.99

В соответствии с полученными результатами, для работы был выбран классификатор GradientBoostingClassifier. Далее рассмотрим его работу и настройку параметров.

Настройка параметров для классификатора

Рассмотрим настройку основных параметров классификатора для более точных результатов.

Перв. примен.

Справ. №

Подпись и дата

Инв. № дубл.

Взам. инв. №

Подпись и дата

Инв. № подл.

Скорость обучения (learning rate) - в машинном обучении и статистике скорость обучения является параметром настройки в алгоритме оптимизации, который определяет размер шага на каждой итерации при движении к минимуму функции потерь.

Рассмотрим зависимость точного выбранного классификатора от следующих значений параметра learning rate:

Таблица 3.3 – Точность классификации на разных значениях параметра learning rate

Learning rate	Точность классификатора на тестовой выборке	Точность классификатора на тренировочной выборке
0.05	0.87	0.99
0.075	0.95	0.99
0.1	0.97	0.99
0.25	0.97	1.99
0.5	0.97	1.00
0.75	0.95	0.99
1	0.22	0.24

Из таблицы видно, что лучшие значения классификатор показывает при параметре learning rate=0.5.

Далее рассмотрим настройку параметра n_estimators.

n_estimators - Количество последовательных деревьев для моделирования. Хотя выбранный классификатор довольно устойчив к большому количеству деревьев, он все же может переобучаться в определенный момент. Следовательно, параметр должен быть настроен для определенной скорости обучения, в данном случае этот параметр = 0.5.

Перв. примен.	Таблица 3.4 – Точность классификации на разных значениях параметра n-estimators																											
	<table border="1"> <thead> <tr> <th>n-estimators</th> <th>Точность классификатора на тестовой выборке</th> <th>Точность классификатора на тренировочной выборке</th> </tr> </thead> <tbody> <tr> <td>10</td> <td>0.63</td> <td>0.70</td> </tr> <tr> <td>50</td> <td>0.92</td> <td>0.97</td> </tr> <tr> <td>100</td> <td>0.97</td> <td>1.00</td> </tr> <tr> <td>500</td> <td>0.97</td> <td>1.00</td> </tr> <tr> <td>1000</td> <td>0.97</td> <td>1.00</td> </tr> <tr> <td>1500</td> <td>0.98</td> <td>1.00</td> </tr> <tr> <td>2000</td> <td>0.98</td> <td>0.99</td> </tr> </tbody> </table>					n-estimators	Точность классификатора на тестовой выборке	Точность классификатора на тренировочной выборке	10	0.63	0.70	50	0.92	0.97	100	0.97	1.00	500	0.97	1.00	1000	0.97	1.00	1500	0.98	1.00	2000	0.98
n-estimators	Точность классификатора на тестовой выборке	Точность классификатора на тренировочной выборке																										
10	0.63	0.70																										
50	0.92	0.97																										
100	0.97	1.00																										
500	0.97	1.00																										
1000	0.97	1.00																										
1500	0.98	1.00																										
2000	0.98	0.99																										
Справ. №	<p>По результатам эксперимента, параметр n-estimators выбран равным 1500. Настройка параметра max_depth приведена в следующей таблице.</p> <p>max_depth - максимальная глубина, ограничивает количество узлов в дереве. Настройка этого параметра нужна для лучшей производительности; лучшее значение зависит от взаимодействия входных переменных.</p>																											
Подпись и дата																												
Инв. № дубл.																												
Взам. инв. №																												
Подпись и дата																												
Инв. № подл.																												
<table border="1"> <tr> <td>Изм.</td> <td>Лист</td> <td>№ докум.</td> <td>Подпись</td> <td>Дата</td> </tr> <tr> <td></td> <td></td> <td></td> <td></td> <td></td> </tr> </table>					Изм.	Лист	№ докум.	Подпись	Дата						ВКР-НГТУ-09.04.01-(М18-ИВТ-3)-004-2020(ПЗ)													
Изм.	Лист	№ докум.	Подпись	Дата																								
					Лист 39																							

Таблица 3.5 – Точность классификации на разных значениях параметра max_depth

max_depth	Точность классификатора на тестовой выборке	Точность классификатора на тренировочной выборке
1	0.63	0.70
2	0.92	0.97
3	0.98	1.00
4	0.98	1.00
5	0.98	1.00
6	0.99	1.00
7	0.98	0.99

По результатам эксперимента, параметр max_depth = 6.

Таким образом, настройка параметров позволила улучшить точность на тестовой выборке с 0.95 до 0.99, и на тренировочной выборке с 0.99 до 1.00.

3.2.3 Описание программного обеспечения для клиента

Как было сказано выше, программа для клиента написана на языке Kotlin. Kotlin - это кроссплатформенный, статически типизированный язык программирования общего назначения.

7 мая 2019 года Google объявил, что язык программирования Kotlin теперь является предпочтительным языком для разработчиков приложений для Android.

Приложение-клиент предоставляет изображение для анализа серверу. С помощью приложения на телефона можно сделать фотографию, и получить расшифровку знаков на ней. Основные этапы по распознаванию объектов реализует серверная часть системы.

3.2.4 Описание программного обеспечения для сервера

Данный продукт написан на языках R и Python. Далее будут рассмотрены его основные функции.

Метод *runRecognition* – вызывается при получении сервером изображения. Изображение читается из директории, входных параметров в методе нет.

Таблица 3.6 – Описание метода *runRecognition*

№	Имя параметра	Тип параметра	Комментарий
1	Result	Выходной	JSON-строка для отправки клиенту

Метод *getBoundingBoxes* – обращается к обученной модели SSD, для получения предсказаний с регионами интереса.

Таблица 3.7 – Описание метода *getBoundingBoxes*

№	Имя параметра	Тип параметра	Комментарий
1	<i>ImagePath</i>	Входной	Путь до изображения

Метод *getClosedGroup* – вычисление замкнутых групп по изображению, применяется к изображениям, полученным из регионов интереса

Таблица 3.8 – Описание метода *getClosedGroup*

№	Имя параметра	Тип параметра	Комментарий
1	<i>ImagePath</i>	Входной	Путь до изображения
2	<i>sdata</i>	Входной	Отклики фильтров, полученные по изображению
3	<i>iMass</i>	Входной	Матрица визуальных масс
4	<i>clsGrp</i>	Входной	Список замкнутых групп
5	<i>oper</i>	Входной	Операторы

Метод *getPredictions* – по замкнутым группам, для каждого региона интереса определяет класс объекта, который на нем изображен.

Таблица 3.9 – Описание метода *getPredictions*

№	Имя параметра	Тип параметра	Комментарий
1	<i>Classes</i>	Выходной	Массив, содержащий предсказанные классы для изображения

Перв. примен.	<p>Метод <i>getSignMeaning</i> – составляет строку о значении знака, либо знаков, на основании номеров классов</p> <p>Таблица 3.10 – Описание метода <i>getSignMeaning</i></p> <table border="1"> <tr> <th>№</th> <th>Имя параметра</th> <th>Тип параметра</th> <th>Комментарий</th> </tr> <tr> <td>1</td> <td>Classes</td> <td>Входной</td> <td>Массив, содержащий предсказанные классы для изображения</td> </tr> <tr> <td>2</td> <td>signsMeaning</td> <td>Выходной</td> <td>Массив строк, содержащий значения классов</td> </tr> </table> <p>Пример работы программы показан на рисунке 3.14:</p>					№	Имя параметра	Тип параметра	Комментарий	1	Classes	Входной	Массив, содержащий предсказанные классы для изображения	2	signsMeaning	Выходной	Массив строк, содержащий значения классов
	№	Имя параметра	Тип параметра	Комментарий													
1	Classes	Входной	Массив, содержащий предсказанные классы для изображения														
2	signsMeaning	Выходной	Массив строк, содержащий значения классов														
Справ. №																	

Подпись и дата	Име. № дубл.	Взам. име. №	Подпись и дата	Име. № подл.

Изм.	Лист	№ докум.	Подпись	Дата	ВКР-НГТУ-09.04.01-(М18-ИВТ-3)-004-2020(ПЗ)	<div>Лист</div> <div>42</div>

Ив. № подл.	Подпись и дата	Взам. инв. №	Инв. № дубл.	Подпись и дата	Справа. №	Перв. примен.

Рисунок 3.4 – Пример работы программы

Изм.	Лист	№ докум.	Подпись	Дата

ВКР-НГТУ-09.04.01-(М18-ИВТ-3)-004-2020(ПЗ)

Лист

43

3.3 Анализ результатов эксперимента

Для проведения вычислительного эксперимента было сделано три выборки изображений, в каждой по 12 фотографий. Выборки разделены между собой условиями съемки:

- Фотографии сделаны при дневном свете, хорошей погоде, без наклона
- Фотографии сделаны при дневном свете, хорошей погоде, с разным углами поворота
- Фотографии сделаны в дождь/снег, в темное время суток, при искусственном освещении

Выборка 1:

Таблица 3.11 – Точность на 1 выборке

№	Точность локализации	Точность классификации	Время
1	100%	99%	0.12 с
2	100%	100%	0.15 с
3	87%	98%	0.22 с
4	100%	99%	0.10 с
5	82%	89%	0.20 с
6	98%	99%	0.13 с
7	97%	99%	0.11 с
8	95%	99%	0.18 с
9	100%	100%	0.14 с
10	100%	99%	0.15 с
11	99%	100%	0.12 с
12	92%	100%	0.18 с

Перв. примен.

Справ. №

Подпись и дата

Инв. № дубл.

Взам. инв. №

Подпись и дата

Инв. № подл.

Перв. примен.		<div style="margin-bottom: 10px;">Выборка 2:</div> <div style="text-align: center; margin-bottom: 5px;">Таблица 3.12 – Точность на 2 выборке</div> <table border="1" style="width: 100%; border-collapse: collapse;"> <tr> <th style="width: 10%;">№</th> <th style="width: 20%;">Точность локализации</th> <th style="width: 20%;">Точность классификации</th> <th style="width: 50%;">Время</th> </tr> <tr><td>1</td><td>94%</td><td>99%</td><td>0.18 с</td></tr> <tr><td>2</td><td>89%</td><td>97%</td><td>0.24 с</td></tr> <tr><td>3</td><td>97%</td><td>92%</td><td>0.19 с</td></tr> <tr><td>4</td><td>92%</td><td>89%</td><td>0.14 с</td></tr> <tr><td>5</td><td>99%</td><td>100%</td><td>0.15 с</td></tr> <tr><td>6</td><td>98%</td><td>99%</td><td>0.18 с</td></tr> <tr><td>7</td><td>92%</td><td>98%</td><td>0.11 с</td></tr> <tr><td>8</td><td>95%</td><td>91%</td><td>0.11 с</td></tr> <tr><td>9</td><td>99%</td><td>98%</td><td>0.17 с</td></tr> <tr><td>10</td><td>91%</td><td>97%</td><td>0.19 с</td></tr> <tr><td>11</td><td>99%</td><td>99%</td><td>0.12 с</td></tr> <tr><td>12</td><td>99%</td><td>100%</td><td>0.15 с</td></tr> </table>					№	Точность локализации	Точность классификации	Время	1	94%	99%	0.18 с	2	89%	97%	0.24 с	3	97%	92%	0.19 с	4	92%	89%	0.14 с	5	99%	100%	0.15 с	6	98%	99%	0.18 с	7	92%	98%	0.11 с	8	95%	91%	0.11 с	9	99%	98%	0.17 с	10	91%	97%	0.19 с	11	99%	99%	0.12 с	12	99%	100%	0.15 с
		№	Точность локализации	Точность классификации	Время																																																					
1	94%	99%	0.18 с																																																							
2	89%	97%	0.24 с																																																							
3	97%	92%	0.19 с																																																							
4	92%	89%	0.14 с																																																							
5	99%	100%	0.15 с																																																							
6	98%	99%	0.18 с																																																							
7	92%	98%	0.11 с																																																							
8	95%	91%	0.11 с																																																							
9	99%	98%	0.17 с																																																							
10	91%	97%	0.19 с																																																							
11	99%	99%	0.12 с																																																							
12	99%	100%	0.15 с																																																							
Справ. №																																																										

Подпись и дата		Инв. № дубл.		Взам. инв. №		Подпись и дата		Инв. № подл.		<div style="margin-bottom: 10px;">Выборка 3:</div> <div style="text-align: center; margin-bottom: 5px;">Таблица 3.13 – Точность на 3 выборке</div> <table border="1" style="width: 100%; border-collapse: collapse;"> <tr> <th style="width: 10%;">№</th> <th style="width: 20%;">Точность локализации</th> <th style="width: 20%;">Точность классификации</th> <th style="width: 50%;">Время</th> </tr> <tr><td>1</td><td>94%</td><td>89%</td><td>0.21 с</td></tr> <tr><td>2</td><td>89%</td><td>99%</td><td>0.24 с</td></tr> <tr><td>3</td><td>91%</td><td>89%</td><td>0.19 с</td></tr> <tr><td>4</td><td>97%</td><td>92%</td><td>0.17 с</td></tr> <tr><td>5</td><td>89%</td><td>99%</td><td>0.20 с</td></tr> <tr><td>6</td><td>98%</td><td>97%</td><td>0.16 с</td></tr> <tr><td>7</td><td>95%</td><td>97%</td><td>0.19 с</td></tr> <tr><td>8</td><td>95%</td><td>85%</td><td>0.18 с</td></tr> <tr><td>9</td><td>92%</td><td>88%</td><td>0.19 с</td></tr> <tr><td>10</td><td>99%</td><td>99%</td><td>0.22 с</td></tr> <tr><td>11</td><td>85%</td><td>99%</td><td>0.23 с</td></tr> <tr><td>12</td><td>92%</td><td>100%</td><td>0.19 с</td></tr> </table>					№	Точность локализации	Точность классификации	Время	1	94%	89%	0.21 с	2	89%	99%	0.24 с	3	91%	89%	0.19 с	4	97%	92%	0.17 с	5	89%	99%	0.20 с	6	98%	97%	0.16 с	7	95%	97%	0.19 с	8	95%	85%	0.18 с	9	92%	88%	0.19 с	10	99%	99%	0.22 с	11	85%	99%	0.23 с	12	92%	100%	0.19 с
										№	Точность локализации	Точность классификации	Время																																																					
1	94%	89%	0.21 с																																																															
2	89%	99%	0.24 с																																																															
3	91%	89%	0.19 с																																																															
4	97%	92%	0.17 с																																																															
5	89%	99%	0.20 с																																																															
6	98%	97%	0.16 с																																																															
7	95%	97%	0.19 с																																																															
8	95%	85%	0.18 с																																																															
9	92%	88%	0.19 с																																																															
10	99%	99%	0.22 с																																																															
11	85%	99%	0.23 с																																																															
12	92%	100%	0.19 с																																																															
Изм.	Лист	№ докум.	Подпись	Дата																																																														

ВКР-НГТУ-09.04.01-(М18-ИВТ-3)-004-2020(ПЗ)

Лист
45

Перв. примен.	Сравним результаты с уже известными подходами. Данные для сравнения взяты из статьи [19].																	
	<p>Таблица 3.14 – Сравнение алгоритмов</p> <table border="1"> <thead> <tr> <th>Алгоритм</th> <th>Точность</th> </tr> </thead> <tbody> <tr> <td>HOG</td> <td>70%</td> </tr> <tr> <td>Viola-Jones</td> <td>91%</td> </tr> <tr> <td>Modified GHT with preprocessing</td> <td>97%</td> </tr> <tr> <td>Modified GHT without preprocessing</td> <td>89%</td> </tr> <tr> <td>ConvNet</td> <td>99%</td> </tr> <tr> <td>Предлагаемый метод</td> <td>96%</td> </tr> </tbody> </table>					Алгоритм	Точность	HOG	70%	Viola-Jones	91%	Modified GHT with preprocessing	97%	Modified GHT without preprocessing	89%	ConvNet	99%	Предлагаемый метод
Алгоритм	Точность																	
HOG	70%																	
Viola-Jones	91%																	
Modified GHT with preprocessing	97%																	
Modified GHT without preprocessing	89%																	
ConvNet	99%																	
Предлагаемый метод	96%																	
Справ. №	Анализируя результаты в таблице 3.14, можно сделать вывод о том, что алгоритм устойчив к условиям съемки и показывает точность, сравнимую с существующими методами.																	
Подпись и дата																		
Инв. № дубл.																		
Взам. инв. №																		
Подпись и дата																		
Инв. № подл.																		
					Лист													
ВКР-НГТУ-09.04.01-(М18-ИВТ-3)-004-2020(ПЗ)					46													
Изм.	Лист	№ докум.	Подпись	Дата														

Перв. примен.		<div>1.4 Выводы по третьей главе</div> <div>В третьей главе было выполнено описание разработанного продукта для обнаружения дорожных знаков и оценка его работы. Была описана настройка параметров для компонентов системы и их влияние на работу системы в целом.</div> <div>Полученные результаты были сравнены с результатами работы наиболее известных подходов к детектированию дорожных знаков. В ходе сравнения было выявлено, что предлагаемый метод имеет сопоставимую точность работы. Это подтверждает конкурентную способность алгоритма.</div>					
Справ. №							
Подпись и дата							
Инв. № дубл.							
Взам. инв. №							
Подпись и дата							
Инв. № подл.							
Изм.	Лист	№ докум.	Подпись	Дата	ВКР-НГТУ-09.04.01-(М18-ИВТ-3)-004-2020(ПЗ)		Лист
							47

Перв. примен.	<h2 style="text-align: center;">Заключение</h2> <p>В результате выполнения выпускной квалификационной работы был выполнен обзор уже существующих подходов к решению задачи обнаружения дорожных знаков на фотографии. На основе обзора были сделаны выводы об общей структуре подходов, определены их сильные и слабые стороны. С учетом полученных знаний был разработан собственный подход решения данной задачи.</p> <p>Разработанный метод применяет совокупность подходов решения задачи. Используется алгебра групп на этапе формирования признакового описания согласно теории ТАВ. Используются нейронные сети и модели машинного обучения для решения задач локализации и классификации.</p> <p>Для реализации предложенного метода был разработан программный продукт на языках R, Python и Kotlin. Программный продукт создан как для мобильного устройства, так и для сервера. Программное обеспечение было проверено путем ряда вычислительных экспериментов.</p> <p>Полученные результаты эксперимента свидетельствуют о корректной работе разработанной системы, а также о конкурентоспособности нового алгоритма. При сравнении точности работы предлагаемый подход оказался в некоторых случаях даже лучше существующих и используемых решений.</p> <p>Таким образом, была выполнена поставленная задача по решению проблемы обнаружения дорожных знаков на фотографии в полном объеме, и цель работы была достигнута.</p>				
Подпись и дата	Инв. № дубл.	Взам. инв. №	Подпись и дата	Инв. № подл.	
Изм.					Лист

ВКР-НГТУ-09.04.01-(М18-ИВТ-3)-004-2020(ПЗ)
Лист

48

Перв. примен.	Библиографический список																
	<div>1. Khotanzad, Y.H. Hong Invariant image recognition by Zernike moments // IEEE Transactions on Pattern Analysis and Machine Intelligence – 1990. – С. 489 – 497.</div> <div>2. Буэно, Суарес, Эспиноса. Обработка изображений с помощью OpenCV - ДМК-Пресс, 2016. -210 с.</div> <div>3. N. Dalal, B. Triggs Histograms of oriented gradients for human detection // IEEE Computer Society Conference on Computer Vision and Pattern Recognition. – 2005. – С.387.</div> <div>4. Tomoki Watanabe, Satoshi Ito, and Kentaro Yokoi Co-occurrence Histograms of Oriented Gradients for Pedestrian Detection // Lecture Notes in Computer Science book series (LNCS, volume 5414) – 2009. - С. 37-45.</div> <div>5. Aude Oliva, Antonio Torralba Building the Gist of a Scene: The Role of Global Image Features in Recognition // International Journal of Computer Vision 42(3) – 2001. – С.145-175</div> <div>6. Lowe, David G. Object recognition from local scale-invariant features // Proceedings of the International Conference on Computer Vision – 1999. -С. 1150–1157.</div> <div>7. Arturo Escalera, Lius Moreno, Miguel Salichs and Jose Armingol, Road traffic Sign Detection and Classification // IEEE Transactions on Industrial Electronics, Vol.44 – 1997.-С.848 859</div> <div>8. Ching-Hao Lai and Chia-Chen Yu, An efficient real-time traffic sign recognition system for intelligent vehicles with smart phones, In Technologies and Applications of Artificial Intelligence (TAAI), // IEEE International Conference – 2010 -С.195-202.</div> <div>9. M. Benallal and J. Meunier, Real-time color segmentation of road signs // IEEE Canadian Conference on Electrical and Computer Engineering, Vol.3 - 2003. –С.1823-1826.</div> <div>10. Viola, Jones: Robust Real-time Object Detection, IJCV – 2001 –С. 1,3.</div> <div>11. Ross Girshick, Jeff Donahue, Trevor Darrell, Jitendra Malik: Rich feature hierarchies for accurate object detection and semantic segmentation – 2013.</div> <div>12. Kaiming He, Xiangyu Zhang, Shaoqing Ren, Jian Sun: Spatial Pyramid Pooling in Deep Convolutional Networks for Visual Recognition – 2015.-С. 1904-1916.</div> <div>13. Ross Girshick: Fast R-CNN – 2015. –С. 1-10</div> <div>14. Shaoqing Ren, Kaiming He, Ross Girshick, Jian Sun: Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks // IEEE Transactions on Pattern Analysis and Machine Intelligence – 2017 – С. 1137-1149.</div> <div>15. Joseph Redmon, Santosh Divvala, Ross Girshick, Ali Farhadi: You Only Look Once: Unified, Real-Time Object Detection – 2016</div>																
Справ. №																	
Подпись и дата																	
Инв. № дубл.																	
Взам. инв. №																	
Подпись и дата																	
Инв. № подл.																	
<table><tr><td></td><td></td><td></td><td></td><td></td><td rowspan="2">ВКР-НГТУ-09.04.01-(М18-ИВТ-3)-004-2020(ПЗ)</td><td rowspan="2">Лист 49</td></tr><tr><td>Изм</td><td>Лист</td><td>№ докум.</td><td>Подпись</td><td>Дата</td></tr></table>											ВКР-НГТУ-09.04.01-(М18-ИВТ-3)-004-2020(ПЗ)	Лист 49	Изм	Лист	№ докум.	Подпись	Дата
					ВКР-НГТУ-09.04.01-(М18-ИВТ-3)-004-2020(ПЗ)	Лист 49											
Изм	Лист	№ докум.	Подпись	Дата													

Перв. примен.		Справ. №		<p>16.Chengcheng Ning ; Huajun Zhou ; Yan Song ; Jinhui Tang: Inception Single Shot MultiBox Detector for object detection // 2017 IEEE International Conference on Multimedia & Expo Workshops – 2017</p> <p>17.J.Greenhalgh and M. Mehdi: Real-Time Detection and recognition of road traffic signs // IEEE Transactions on intelligent transportation systems – 2012 – C.1498-1506</p> <p>18.Karunalithika,R.P.Jayasundra,M.A.Rasamjan ,D.N. Senayanke, V.N.Vithana: Road sign identification application using image processing and augmented reality, International journal of advanced computer technology – 2015-C. 79-93</p> <p>19.Alexander Shustanov , Pavel Yakimov: CNN Design for Real-Time Traffic Sign Recognition, ScienceDirect - Procedia Engineering 201 - 2017</p> <p>20.Утробин В.А. Элементы теории активного восприятия изображений // Труды Нижегородского государственного технического университета им. Р.Е. Алексеева. - 2010. - Т. 81, №2. - С. 61-69.</p> <p>21.Утробин В.А. Компьютерная обработка изображений. Анализ и синтез. – Нижний Новгород: НГТУ им. Р.Е. Алексеева, 2003. – С. 228.</p> <p>22.Утробин В.А. Компьютерная обработка изображений. Принятие решений в пространстве эталонов. – Нижний Новгород: НГТУ им. Р.Е. Алексеева, 2004. – С. 221.</p> <p>23.Утробин В.А. Компьютерная обработка изображений. Информационные модели этапа понимания. – Нижний Новгород: НГТУ им. Р.Е. Алексеева, 2006. – С. 247.</p>					
Подпись и дата		Инв. № дубл.		Взам. инв. №		Подпись и дата		Инв. № подл.	
<div> <div>Изм</div> <div>Лист</div> <div>№ докум.</div> <div>Подпись</div> <div>Дата</div> </div> <div> <div>ВКР-НГТУ-09.04.01-(М18-ИВТ-3)-004-2020(ПЗ)</div> <div>Лист</div> <div>50</div> </div>									