

Федеральное государственное учреждение
«Федеральный исследовательский центр
Институт прикладной математики им. М.В. Келдыша
Российской академии наук»

На правах рукописи

Власов Сергей Олегович

**ИССЛЕДОВАНИЕ И РАЗРАБОТКА ГИБРИДНЫХ АЛГОРИТМОВ
РАСПОЗНАВАНИЯ ОБЪЕКТОВ НА ОСНОВЕ ЗРИТЕЛЬНЫХ ДАННЫХ**

Специальность 2.3.5 –

«Математическое и программное обеспечение вычислительных систем,
комплексов и компьютерных сетей»

Диссертация на соискание учёной степени

Кандидата физико-математических наук

Научный руководитель:
доктор физико-математических наук
Богуславский Андрей Александрович

Москва 2025

Оглавление

Введение	4
1. Обзор и описание, классификация существующих алгоритмов распознавания объектов.....	12
1.1 Задачи, решаемые компьютерным зрением, классификация алгоритмов, понятие гибридного алгоритма	12
1.2 Традиционные алгоритмы распознавания.....	18
1.3 Алгоритмы поиска объектов интереса на основе машинного обучения	20
1.4 Гибридные алгоритмы	24
1.4.1 Гибридные алгоритмы, представляющие из себя комбинацию классических методов компьютерного зрения.....	25
1.4.2 Гибридные алгоритмы, представляющие из себя комбинацию методов машинного обучения или их дополнение	26
1.4.3 Гибридные алгоритмы, представляющие из себя комбинацию традиционного компьютерного зрения и машинного обучения	28
1.5 Результаты	31
2. Используемые методы традиционного компьютерного зрения и нейросетевых технологий.....	33
2.1 Методы поиска контуров.....	33
2.1.1 Описание работы алгоритма Хюккеля.....	34
2.1.2 Описание работы алгоритмов детектора границ Кэнни и поиска прямых Хафа.....	39
2.1.3 Тестирование алгоритмов поиска контуров	41
2.1.4 Применение алгоритмов поиска контуров на реальных изображениях ..	45
2.1.5 Построение эталонного изображения	50
2.1.6 Рекомендации по выбору настраиваемых параметров	55

2.2 Детекторы особых точек.....	58
2.2.1 Выбор детектора.....	58
2.3 Нейросетевые технологии	63
2.3.1 Выбор архитектур	65
2.4 Результаты	69
3. Описание предложенных алгоритмов.....	72
3.1 Гибридный алгоритм, основанный на предварительном обнаружении контуров объектов.....	73
3.2 Предварительная обработка нейросетевым методом с поиском особых точек в фрагменте.....	76
3.3 Предварительный поиск области методом особых точек с проверкой нейросетевым методом	80
3.4 Результаты	82
4. Результаты проведенных численных экспериментов	83
4.1 Влияние предварительного обнаружения контуров	83
4.2 Влияние предварительной обработки нейросетью с поиском особых точек в фрагменте	86
4.3 Влияние предварительного поиска области методом особых точек с проверкой нейросетевым методом.....	91
4.4 Результаты	92
Заключение	94
Список литературы	96

Введение

Основную часть информации о внешнем мире человек получает с помощью органов зрения. В результате последующей сложной, но эффективной обработки полученной информации, используя аппарат анализа и интерпретации, человек получает возможность представления трехмерной картины окружающего мира, а именно может распознавать объекты, их границы и расположение.

Область искусственного интеллекта, связанная с решением задач обработки изображений, обнаружения и распознавания объектов на них, называется компьютерным зрением. Также под компьютерным зрением понимают преобразование данных, полученных от фото- или видеокамеры, результатом которого является некоторое решение или новое представление, предназначенное для достижения определенной цели [1].

Как отмечено в работе [2], следует отличать компьютерное зрение и компьютерную графику. Обе области имеют дело со способами отображения и моделирования объектов; главное различие между ними в том, что компьютерное зрение связано с описанием и распознаванием объектов по изображениям, а компьютерная графика— с генерацией изображений по описаниям объектов.

На сегодняшний день основной целью использования систем компьютерного зрения является дополнение или замена человека в таких областях как медицина, системы безопасности, системы контроля качества, создания беспилотных транспортных средств и др. По этой причине задачи распознавания объекта интереса на изображении весьма разнообразны: от поиска лица на изображении, полученном обычной фотокамерой, до классификации антропогенных объектов на спутниковых снимках.

Несмотря на то, что алгоритмы компьютерного зрения могут казаться узкоспециализированными, широкое применение систем технического зрения в различных отраслях промышленности наглядно показывает степень развития данного направления науки и технологий. Развитие математического аппарата компьютерного зрения и появление новых методов обработки изображений идет

параллельно с усовершенствованием существующей и возникновением новой элементной базы и устройств, обеспечивающих работу систем технического зрения.

Под системой реального времени понимают систему, которая должна реагировать на события во внешней по отношению к системе среде или воздействовать на среду в рамках требуемых временных ограничений. Авторы научных работ задаются различным минимальным числом обрабатываемых изображений в единицу времени, необходимым для возможности работы алгоритмов при их применении в системах реального времени. В работах [3] и [4] скорость нейросетевых алгоритмов распознавания объектов для использования в системах реального времени составляет 30 кадров в секунду. Обнаружение объектов интереса со скоростью 20 и 25 изображений в секунду считается достаточным для решения задач реального времени в [5, 6].

Ограниченность вычислительных и временных ресурсов заставляет искать новые пути к решению задач обнаружения объектов в системах реального времени. Для получения удовлетворительного качества распознавания необходима разработка и применение новых методик и алгоритмов.

Гибридный алгоритм - алгоритм, основанный на применении комбинации различных методов «традиционного» компьютерного зрения и нейросетевых технологий, направленный на упрощение их использования и повышение надежности распознавания.

При решении задач в данной работе подразумевается возможность сильного смещения объекта интереса в видеопоследовательности. Также предполагается возможность использования полученных результатов при решении задач реального времени, что накладывает ограничение на время обработки изображения.

В данной работе предложена классификация существующих методов распознавания объектов на изображениях на основе используемого аппарата при решении поставленных задач, проведен обзор предлагаемых решений на основе данной классификации, сформулировано понятие гибридного алгоритма, предложены структуры гибридных алгоритмов, проведены вычислительные

эксперименты на изображениях из открытых источниках, на основе полученных результатов выполнена оценка точностной и временной эффективности предлагаемых гибридных алгоритмов.

Цель работы

Целью работы является анализ существующих и разработка новых гибридных алгоритмов распознавания объектов на изображениях с последующей их реализацией.

Объект исследования

Объектом исследования являются алгоритмы обнаружения и классификации объектов интереса на изображениях в реальном масштабе времени.

Предмет исследования

Предметом исследования являются алгоритмы распознавания, содержащие элементы нейросетевых технологий и классических методов компьютерного зрения.

Задачи

В рамках проведенной работы были поставлены и решены следующие задачи:

1. Провести анализ существующих методов обнаружения объектов;
2. Выполнить исследование способов проектирования и реализации гибридных алгоритмов распознавания объектов;
3. Предложить структуры гибридных алгоритмов, учитывающих достоинства и недостатки используемых методов по отдельности, и реализовать их
4. Апробировать разработанные алгоритмы применительно к решению задач обнаружения и распознавания объектов в системах компьютерного зрения реального времени, в которых требуется распознавать образы объектов интереса в естественной среде.

На защиту выносятся следующие положения:

1. На основе изучения работ из открытых источников предложена классификация алгоритмов обнаружения и классификации объектов на изображениях, основанная на применяемых методах компьютерного зрения.

Сформулировано понятие гибридного алгоритма, рассмотрены существующие реализованные подходы, относящиеся к данному классу алгоритмов.

2. Предложены различные гибридные алгоритмы распознавания объектов интереса на видеопоследовательностях, основанные на совместном применении алгоритмов поиска и классификации объекта, доступных в открытых источниках. Данные алгоритмы основаны на использовании таких методов компьютерного зрения, как поиск контуров объектов на изображениях, применение детекторов особых точек, нейросетевые технологии. Приведено обоснование выбранных архитектур сверточных нейронных сетей, методов традиционного компьютерного зрения.

3. Реализованы три варианта гибридных алгоритмов распознавания объектов на изображениях, выполнены численные эксперименты, проведены временная и точностная оценки полученных результатов работы алгоритмов.

Научная новизна

Проведен анализ предложенных алгоритмов распознавания объектов интереса на изображениях последовательным применением свёрточных нейронных сетей и «традиционных» методов компьютерного зрения, таких как поиск особых точек, их сопоставление, а также обнаружение контуров объектов.

Реализован алгоритм, основанный на предварительном переходе к контурному представлению изображения с последующим поиском объекта интереса нейросетевым методом. В ходе проведенных компьютерных экспериментов показано, что, за время и точность обработки отвечает тип сети, а не наличие предварительной обработки изображений.

Предложен и реализован алгоритм предварительного определения возможных областей расположения малоразмерного объекта на изображении нейросетевым методом с последующей проверкой правильности распознавания путем сопоставления особых точек в данной области с эталонным представлением объекта интереса. Данный подход позволил заметно уменьшить вероятность ложного срабатывания алгоритма поиска и классификации.

Рассмотрен алгоритм предварительного определения возможных областей расположения малоразмерного объекта на изображении детектором особых точек с последующей проверкой правильности распознавания нейросетевым методом. Данный подход позволяет уменьшить время работы с массивами изображений при отсутствии признаков объекта интереса на первом этапе обработки.

Теоретическая и практическая значимость

В рамках данной диссертационной работы предложена новая классификация алгоритмов обнаружения и классификации объектов, предложены и реализованы три варианта гибридных метода распознавания объектов, получены и проанализированы результаты их применения при решении реальных задач компьютерного зрения, проведена оценка возможности из применения в задачах реального времени.

Апробация работы

Основные результаты диссертационной работы докладывались и обсуждались на 2-х конференциях. Среди них международные конференции «35-я Международная научно-техническая конференция «Экстремальная робототехника» (г. Санкт-Петербург, 2024) и «34-я Международная конференция по компьютерной графике и машинному зрению «GraphiCon 2024»» (г. Омск, 2024).

Публикации

По результатам проведенных исследований имеются пять печатных работ [7-11]. Из них три в рецензируемых журналах перечня ВАК. Также принят в печать препринт «Алгоритм Хьюкеля выделения линейных признаков на изображениях» в издании «Препринты ИПМ им.М.В.Келдыша».

Диссертационная работа соответствует паспорту специальности 2.3.5 «Математическое и программное обеспечение вычислительных систем, комплексов и компьютерных сетей» по направлению 7 модели, методы, архитектуры, алгоритмы, форматы, протоколы и программные средства человеко-машинных интерфейсов, компьютерной графики, визуализации, обработки

изображений и видеоданных, систем виртуальной реальности, многомодального взаимодействия в социокиберфизических системах.

Структура и объем диссертации

Диссертация состоит из введения, четырех глав, заключения, списка сокращений и списка литературы. Общий объем работы составляет 107 страниц, содержащих 39 рисунков и 11 таблиц. Список использованной литературы включает 95 источников.

В первой главе представлены результаты проведенного обзора опубликованных работ, посвященных задаче обнаружения и классификации объектов. Приведено описание существующих методов распознавания объектов на изображениях. Сформулированы направления развития систем технического зрения. Рассматривается определение «гибридного алгоритма» и его использование в других работах. Предложена классификация алгоритмов распознавания объектов на основе типов используемых методов обработки визуальной информации. Проведенный обзор показал, что разработка новых алгоритмов обнаружения и классификации объектов на изображениях, а также применение существующих методов к решению новых задач является активно развивающимся направлением научных исследований. Подход, основанный на распознавании объектов интереса с использованием гибридных алгоритмов широко применяется для решения задач реального времени при ограниченных вычислительных ресурсах.

Во второй главе представлено описание традиционных и нейросетевых методов компьютерного зрения, используемых в предложенных в данной работе гибридных алгоритмах распознавания объектов на изображениях. Приведено обоснование выбранных архитектур нейронных сетей, методов поиска контуров и детекторов особых точек. Представлены результаты решения демонстрационных задач, показывающие возможность использования принятых решений в системах реального времени.

В третьей главе описана структура предлагаемых гибридных алгоритмов, принцип работы которых основан на применении классических и нейросетевых методов компьютерного зрения. Представлен алгоритм, основанный на методах поиска контуров, определения прямых линий и применении сверточных нейронных сетей. Описана работа гибридного алгоритма, в котором реализовано предварительное определение возможных областей расположения объекта интереса на изображении с последующей оценкой правильности распознавания путем сравнения с набором эталонных фрагментов традиционным методом компьютерного зрения сопоставления особых точек, целью работы которого является уменьшение числа ложных распознаваний объектов на изображении. Представлена структура алгоритма, основанного на поиске областей возможного расположения объекта интереса на изображении с последующей проверкой правильности распознавания нейросетевым методом.

В четвертой главе представлены результаты работы реализованных гибридных алгоритмов, приведены результаты точностной и временной оценок их эффективности и возможности их применения при решении задач реального времени. Выполнены численные эксперименты на универсальном компьютере, целью которых являлось определение возможности использования данных алгоритмов в системах реального времени. На основе полученных результатов показано, что за время обучения и обработки отвечает тип сети, применение искусственных нейронных сетей ко всему контурному изображению не дает выигрыша в точности или времени обнаружения и классификации объектов. Для алгоритма, основанного на предварительном поиске объекта интереса нейросетевым методом с последующей проверкой правильности распознавания традиционным методом компьютерного зрения определено, что использование детектора особых точек не оказывает значительного влияния на время обработки изображений, в то же время последующая обработка ранее обнаруженных нейросетевым методом объектов снижает число ложных распознаваний. Показана работоспособность алгоритма, основанного на предварительном поиске объекта интереса детектором особых точек с последующей проверкой правильности

распознавания сверточной нейронной сетью. Вследствие малого времени поиска особых точек на изображении, возможно уменьшение времени обработки массива изображений за счет экономии на применении нейросетевого метода. Определено, что значения настраиваемых параметров данных методов оказывают значительное влияние на получаемый результат.

В заключении сформулированы основные результаты работы. На основе проведенного обзора опубликованных исследований и разработок определены основные направления повышения эффективности алгоритмов распознавания объектов на изображениях, предложена классификация алгоритмов обнаружения и распознавания объектов, основанная на используемых методах компьютерного зрения. Сформулировано понятие гибридного алгоритма распознавания объектов на изображениях. Предложены и реализованы три структуры данных алгоритмов, включающих в себя традиционные методы компьютерного зрения, а также использование сверточных нейронных сетей. Полученные результаты оценки эффективности показывают, что разработка алгоритмов данного типа позволяет достигнуть повышения скорости и точности обработки изображений в части распознавания на них антропогенных объектов интереса.

1. Обзор и описание, классификация существующих алгоритмов распознавания объектов

В настоящей главе представлены результаты обзора текущего развития алгоритмов распознавания объектов на изображениях. Предложена классификация алгоритмов на основе используемых методов, рассмотрены предлагаемые подходы к решению конкретных задач.

1.1 Задачи, решаемые компьютерным зрением, классификация алгоритмов, понятие гибридного алгоритма

Современные технологии компьютерного зрения позволяют использовать разработанные решения во многих областях науки и техники, чему посвящено большое количество работ:

- промышленности (контроль качества изготовленной продукции и производственных процессов) [12, 13];
- медицине (анализ полученных на медицинском оборудовании изображений для выявления признаков наличия заболеваний) [14, 15];
- транспортных средствах (анализ ситуации перед автомобилем, контроль трафика на дорогах, определение границ проезжей части перед транспортным средством) [16, 17, 18, 19, 20, 21, 22];
- робототехнике (возможность ориентации объекта в пространстве, формирование рабочих зон, решение навигационных задач) [23, 24];
- распознавании лиц и биометрических данных [25, 26, 27];
- распознавание объектов на снимках со спутников [28, 29].

Увеличение числа решаемых компьютерным зрением задач в различных направлениях науки и техники приводит к разработке новых методов и алгоритмов распознавания объектов на изображениях. Немалую роль играют доступные для каждого отдельно рассматриваемого случая вычислительные мощности,

требования в части точности и надежности работы системы, временные ограничения на обработку данных.

Появление новых алгоритмов и методов, нередко представляющих из себя модификацию ранее разработанных подходов под конкретную задачу, приводит к сложности выбора наиболее подходящего решения на раннем этапе работы с новой задачей.

Авторы обзорных статей, а также монографий, посвященных системам и алгоритмам компьютерного зрения, как правило, либо концентрируются на одном отдельном направлении развития методов распознавания объектов, либо описывают весь путь истории их развития с акцентом на описание конкретных решений. Ниже представлены примеры данных работ.

В [30] Рой Дэвис и Мэтью Тёрк проводят глубокий обзор существующих методов и алгоритмов компьютерного зрения с последующим описанием их работы. В работе делается акцент на использовании машинного обучения в задачах компьютерного зрения. Особое внимание уделяется тому, что последние успехи в области компьютерного зрения в значительной степени связаны с появлением огромных массивов тщательно размеченных данных. Поднимается вопрос вычислительных проблем глубоких нейронных сетей, которые затрудняют их развертывание на оборудовании с ограниченными ресурсами, затрагивается вопрос о перенесении обучения из одной области данных в другую.

Большое количество примеров из практических задач наряду с теоретическими аспектами обработки зрительных данных представлено в работе [2]. Помимо основных разделов компьютерного зрения рассматриваются базы данных изображений и систем виртуальной и дополненной реальности. Приводится большое число примеров использования описываемых методов в промышленности, медицине, землепользовании, мультимедиа и компьютерной графике.

В работе Питера Корка [31] рассмотрены геометрические принципы формирования изображений, связь трехмерного и двумерного изображений.

Акцент выполнен на применении методов компьютерного зрения при решении прикладных задач.

Работа Стивена Л. Брантона и Дж. Натана Куц «Анализ данных в науке и технике» [32] охватывает не только методы, используемые в компьютерном зрении, но и в других направлениях. Также излагается современный взгляд на динамические системы, описываются текущие цели и открытые проблемы.

Обширное описание методов компьютерного зрения и существующих подходов к решению поставленных задач представлено в работе [33].

История развития методов машинного обучения в части использования искусственных нейронных сетей и глубокого обучения, а также возникшие на всех этапах разработки и применения проблемы и пути их решения представлены в [34]. В данной работе, посвященной, в основном, задачам оптического распознавания и машинного зрения при помощи нейронных сетей, отмечено, что сегодняшние приложения, как правило, используют сочетание машинного обучения и классических вычислений [34, с. 18]

Авторы работы [35] считают, что «глубокие модели оказались ключом, который подходит ко всем замкам сразу: новые архитектуры и алгоритмы обучения, а также увеличившиеся вычислительные мощности и появившиеся огромные наборы данных привели к революционным прорывам в компьютерном зрении, распознавании речи, обработке естественного языка и многих других типично «человеческих» задачах машинного обучения».

Отдельно стоит выделить работы, описывающие доступные возможности для решения задач компьютерного зрения с использованием конкретных языков программирования, сред или библиотек с примерами кода для последующей реализации программ [1, 31, 36], а также монографии, целиком посвященные одному из направлений применения компьютерного зрения [37].

Несмотря на успехи, достигнутые за счет применения машинного обучения в компьютерном зрении, для решения многих задач в настоящее время достаточно только традиционных методов, что будет рассматриваться в последующих разделах данной главы.

Помимо работ, в которых приводится подробное описание математического аппарата и используемых в компьютерном зрении подходов при решении прикладных задач, публикуются и работы, посвященные обзору методов компьютерного зрения с целью их классификации.

Обзор и описание классических методов распознавания объектов и свёрточных искусственных нейронных сетей проведен в работе [38, 39].

Описание основных подходов к решению задач компьютерного зрения, с акцентом на глубинное машинное обучение и искусственные нейронные сети, представлено в [40]. Также в данной работе приведен обзор рыночных показателей в данной области и показано, что рынок компьютерного зрения быстро развивается и имеет серьезные перспективы.

Описание и сравнение различных методов компьютерного зрения при решении задачи распознавания лиц на изображении представлено в [41]. В статье [42] показаны результаты анализа сравнительной эффективности архитектур искусственных нейронных сетей открытой библиотеки TensorFlow при решении задачи обнаружения области глаз на фотографии лица человека. На основе проведенных экспериментов установлена наиболее подходящая модель, обученная с нуля.

Краткая история возникновения методов глубокого обучения в части распознавания объектов, описание основных архитектур свёрточных искусственных нейронных сетей и мультимодальных методов машинного обучения, характеристики наиболее популярных массивов изображений для обучения представлены в [43].

В некоторых работах на основе анализа методов распознавания объектов на изображениях в качестве направления дальнейшего развития предлагаются комплексные методики организации систем машинного анализа изображений. Проведенный в [44] анализ современных подходов, применяемых при построении алгоритмов выделения и классификации визуальных объектов в массивах графических данных, предлагает схему построения комплексной методики для организации систем машинного анализа изображений: использование алгоритмов

предварительной обработки входных данных через восстановление исходного изображения, настройку параметров и сегментации, а также алгоритмы выделения визуального объекта, определения его признаков и дальнейшей классификации.

Также стоит отметить, что различные задачи, ставящиеся перед разработчиками систем компьютерного зрения, требуют различного подхода к реализации алгоритмов распознавания объектов. Присутствуют различия в доступном времени обработки поступающих изображений (в задачах реального времени скорость обнаружения искомого положения объекта играет первостепенную роль, в то время как на обработку данных в других проектах могут отводиться часы расчетного времени). Отличаются требования к точности определения границ объекта и степени правильности их классификации на изображении: допустимость ошибок первого и второго рода. Большую роль играет оборудование, на котором должен использоваться разрабатываемый алгоритм, возможности по его оптимизации, распараллеливанию, степень качества входных и выходных изображений.

Вышеописанные ограничения и требования приводят к необходимости комплексного подхода при разработке методов и алгоритмов компьютерного зрения. В то время, как «классическим» подходом ускорения процесса вычислений является увеличение числа расчетных модулей и распараллеливание выполняемых операций, появляются новые, альтернативные, методы повышения скорости обработки вычислений с требованием к ним в части отсутствия снижения точности получаемого результата. Одним из направлений повышения эффективности работы систем распознавания объектов является разработка новых алгоритмов обработки изображений.

На основе проведенного анализа существующих работ выделены три направления повышения скорости и точности обработки изображений в части применяемых алгоритмов:

- повышение точности работы «традиционных» алгоритмов компьютерного зрения;

- модификация методов машинного обучения, в частности, архитектур сверточных нейронных сетей;
- создание алгоритмов и методов, включающих в себя комбинацию различных подходов к решению поставленной задачи.

Можно сформировать следующую классификацию методов распознавания объектов на изображениях:

- классические алгоритмы распознавания;
- алгоритмы поиска объектов интереса на основе машинного обучения;
- гибридные алгоритмы:
 - гибридные алгоритмы, представляющие из себя комбинацию классических методов компьютерного зрения;
 - гибридные алгоритмы, представляющие из себя комбинацию методов машинного обучения или их дополнение;
 - гибридные алгоритмы, представляющие из себя комбинацию традиционного компьютерного зрения и методов машинного обучения.

Под «традиционными» (или «классическими») алгоритмами будут пониматься подходы, не основанные на глубоком обучении, такие как поиск по шаблону, метод Виолы-Джонса, гистограмма направленных градиентов и др.

Под алгоритмами на основе машинного обучения будет пониматься распознавание объектов на изображении с использованием сверточных нейронных сетей и глубокого обучения.

К «гибридными» относятся алгоритмы, основанные на последовательном или параллельном использовании алгоритмов и методов компьютерного зрения, ранее применявшихся по отдельности. Целью данного подхода являются компенсация недостатков применения конкретных методов по отдельности и повышение эффективности системы при их совместном использовании.

Гибридный алгоритм распознавания объектов - алгоритм, основанный на совместном применении различных методов «традиционного» компьютерного зрения и нейросетевых технологий, направленный на упрощение их использования

и повышение надежности обнаружения и классификации объектов на изображении.

Понятие гибридного алгоритма уже использовалось в других работах и не является новым. Гибридный подход, заключающийся в исследовании эффективности применения различных конфигураций совместного поиска соответствий точек и линий для определения относительной ориентации двух изображений, описывается в [45]. В работе [46] предлагается гибридный метод распознавания контуров на изображении. Гибридные методы и алгоритмы для интеллектуальных систем классификации сложноструктурируемых изображений рассматриваются в [47]. Работа [48] посвящена решению задачи снижения потребных вычислительных ресурсов при распознавании объектов на изображениях с использованием гибридных алгоритмов в части совместного применения нейронных сетей и других методов компьютерного зрения.

В данной главе будет проводиться обзор алгоритмов распознавания и их отнесение к той или иной группе с позиции предложенной классификации.

1.2 Традиционные алгоритмы распознавания

Традиционные алгоритмы распознавания включают в себя определение контуров и градиентов присутствующих на сцене объектов, поиск и сравнение особых точек, сопоставление с заранее заданным шаблоном. Для обработки сцены применяются вычисление оптического потока, удаление шумов и помех, присутствующих на изображении, преобразование изображения в полутоновое, пороговая сегментация и т.д. Сложность данного подхода заключается в том, что для каждого изображения необходимо выбрать, какие признаки стоит учитывать [49]. Данный недостаток заметно проявляется при необходимости обнаружения и классификации множества различных объектов на одном изображении, однако при решении задачи распознавания конкретного типа объекта данные методы, как

правило, показывают достаточную точность при малых вычислительных затратах и не требуют наличия большой обучающей выборки.

Несмотря на то, что данные методы используются давно, в настоящее время продолжается их модификация под решение конкретных задач.

В работе [12] предложен алгоритм определения положения камеры и трехмерной реконструкции сцены на основе вычисления оптического потока и фундаментальной матрицы с одновременным определением структуры сцены путем минимизации одного общего функционала для всех неизвестных.

Обнаружение транспортного средства в видеопотоке путем сопоставления объектов на изображении с шаблоном предложено в работе [50], посвященной задаче оценки ситуации на перекрестке беспилотным транспортным средством. В данной работе традиционные методы компьютерного зрения используются на этапах распознавания транспортных средств и их сопровождения на изображениях, поступающих с нескольких камер, установленных на транспортное средство.

Новый подход, основанный на методах традиционного компьютерного зрения, продемонстрирован на тестовых задачах в [51]. В данной работе проведены обработка и анализ графической информации в части разработки двухмасштабной многокомпонентной математической модели цифрового изображения, описание свойств изображения для различных областей анализа, введен показатель размеров объектов, предложены алгоритмы декомпозиции изображений. На основе модели и метода декомпозиции предложен новый подход к анализу и обработке изображений, заключающийся в том, что для последующей работы следует выбирать лишь ту компоненту изображения, в которой должна содержаться искомая информация.

В работе [52] предлагается выполнять локализацию изображения относительно трехмерной модели сцены с минимизацией требований по необходимому объему памяти и времени обработки за счет нового алгоритма сжатия информации: вместо одинаковой обработки всех трехмерных точек предлагается хранение отдельного набора точек с полной информацией об окружающем пространстве, что возможно использовать в беспилотных

транспортных средствах и устройствах дополненной реальности, проведено сравнение с существующими методами.

Применение метода Speeded Up Robust Features, заключающегося в поиске особых точек изображения в задаче поиска объектов в реальном масштабе времени подробно описывается в [53]. Также в данной работе рассмотрены различные алгоритмы предварительной обработки изображений, позволяющие повысить эффективность данного метода, проведены исследования возможности распознавания объекта в условиях различной освещенности, при деформации масштабированием и поворотом, изменении уровня яркости и размытости кадры, изменении точки обзора.

1.3 Алгоритмы поиска объектов интереса на основе машинного обучения

Часто под «машинным обучением» понимают искусственные нейронные сети, однако машинное обучение является более широким термином, охватывающим разнообразные методы, включая деревья решений, случайные леса, метод опорных векторов и многое другое. Нейронные сети представляют собой подмножество машинного обучения. Искусственные нейронные сети, как и все методы, основанные на машинном обучении, работают на основе обучения на заранее собранных и обработанных массивов данных – обучающей выборке, с последующим использованием статистических методов для извлечения закономерностей в данных.

Нейронные сети являются одним из наиболее мощных методов машинного обучения, особенно при работе с данными высокой размерности и сложными задачами. Глубокое обучение является подмножеством нейронных сетей, которые имеют множество слоев. Данный подход обычно используется для решения сложных задач, таких как распознавание образов и обработка естественного языка.

Одной из самых популярных и перспективных задач нейросетей является технология распознавания образов. Для этого собирают большие базы размеченных данных, из которых формируют обучающий и тестовый наборы. Выделив признаки и их комбинации для идентификации похожих объектов, можно натренировать модель машинного обучения распознавать нужные типы закономерностей. Если модели могут неверно распознать некоторые объекты, сеть «дообучают».

Искусственные нейронные сети используются в случае, когда необходимо решать задачи прогнозирования, классификации или автоматизации, для которых нет аналитических методов решения либо они весьма трудоемки в реализации, но есть большое количество статистических данных о результатах решения тем или иным способом.

Для работы с изображениями применяются сверточные нейронные сети, представляющие из себя набор последовательно и параллельно расположенных связанных между собой слоев, каждый из которых выполняет некоторые преобразования полученного изображения и передает их в следующий слой.

Глубокие нейронные сети обнаруживают базовые шаблоны в классах изображений и автоматически вырабатывают наиболее наглядные и характерные особенности в отношении каждого конкретного класса для каждого объекта. Отсюда получается более высокая производительность и точность, по сравнению с традиционными алгоритмами машинного зрения.

Благодаря развитию аппаратного обеспечения и увеличению предельной вычислительной мощности, разработка сверточных нейронных сетей оказала огромное влияние на решение задач распознавания объектов на изображениях.

Потребность в большом наборе данных при решении многих прикладных задач, высокая вариативность параметров архитектуры сети, таких как количество слоев их внутренних настраиваемых параметров, которые оказывают влияние на получаемый результат, являются одними из недостатков сверточных нейронных сетей. Многие параметры нейронных сетей выбираются эмпирически, на основе полученных ранее результатов и опыта разработчика. Глубокое обучение

используется для задач классификации, в которых широкий диапазон потенциальных сигналов должен быть отображен на ограниченное количество категорий, учитывая, что имеется достаточный набор данных [49].

Появление различных архитектур сверточных нейронных сетей и открытых библиотек с возможностью работы с ними позволяет применять достоинства данного метода при решении многочисленных задач по распознаванию объектов на изображениях.

В работе [29] показано решение задачи поиска самолетов на изображениях с использованием TensorFlow Object Detection API. В [14] проведен анализ рентгеновских снимков с использованием сверточной искусственной нейронной сети CheXnet на основе сверточной сети Densenet, с активационной функцией ReLU. В [54] решается задача повышения точности работы с видеопотоком, в котором присутствуют большие смещения объекта интереса, за счет нового алгоритма сопоставления дескрипторов, «глубокого сопоставления», что должно повысить точности работы в системах реального времени, исследуется эффективность построения дескрипторов в структуру, подобную глубоким нейронным сетям. С целью автоматизации и повышения безопасности гражданского судовождения в работе [55] нейросетевой алгоритм применяется для осуществления автоматического обнаружения и распознавания надводных объектов и препятствий заданных классов по всей области изображения. При этом ставится задача его применимости к выполнению в реальном или условно реальном времени для задач оптоэлектронной системы технического зрения. Работа [4] посвящена обнаружению и распознаванию надводных объектов в реальном времени с использованием сверточной искусственной нейронной сети CenterNet, преимущество которой, как отмечено в материалах статьи, заключается в высокой скорости работы даже по сравнению с другими одноэтапными детекторами. Сверточная нейронная сеть LeNet-5 с оптимизатором Adam применяется для распознавания дорожных знаков в работе [56]. Использование нейросетевого подхода для эффективной классификации и распознавания наземных объектов в видеопотоке радиолокационных кадров предложено в [57]. Для решения

поставленной задачи используется сверточная нейронная сеть Faster R-CNN. Целью работы [58] является разработка детектора на основе архитектуры сверточной нейронной сети EfficientDet для обнаружения патологий желудка на видеоизображениях эндоскопических исследований. Также в рамках исследования проведено сравнение полученных результатов с архитектурой SSD300.

Использование нейросетевых технологий для построения контуров объекта и сравнение предложенного метода с традиционным компьютерным зрением представлены в [59]. Решаемая задача в данной работе – определение положения береговой линии поверхностных водных объектов по материалам аэрокосмических съемок. Показано преимущество метода глубокого обучения по сравнению с традиционными методами, предложено несколько способов повышения получаемой точности.

Многие работы по расширению области применения нейронных сетей заключаются не в создании новой структуры или использовании уже существующей архитектуры, а в их модификации. Так, в [28] решается задача обнаружения и распознавания наземных объектов на местности системой искусственного интеллекта БПЛА с помощью модели на основе сверточной нейронной сети с архитектурой YOLO-V3. Также модификация архитектуры tiny-YOLO предлагается в работе [60] для решения задачи детектирования объектов на изображениях мобильными системами компьютерного зрения с целью достижения баланса между точностью и скоростью. Использование архитектуры YOLO системой технического зрения беспилотного автомобиля предложено в [21] в части обнаружения препятствий перед транспортным средством, представленных другими автомобилями. Другие методы компьютерного зрения используются в данном исследовании для определения границ проезжей части. Аналогичная задача подобными методами решается в [61]. В работе [3] искусственная нейронная сеть с архитектурой CenterNet предлагается как основа детектора обнаружения объектов на видео в задачах реального времени. Обеспечение высокой скорости (34 и 37 кадров в секунду) работы метода выполняется за счет применения одноступенчатой нейронной сети, что дает выигрыш во времени по сравнению с

двуступенчатыми подходами. Предложено использование тепловой карты, получаемой CenterNet для улучшения предсказания положения объекта на следующих кадрах. Применение улучшенного метода CenterNet для распознавания посторонних объектов на конвейерных лентах в угольных шахтах рассматривается в [5]. Цель данной работы – обнаружение посторонних предметов в режиме реального времени (20 кадров в секунду). Архитектура CenterNet дополнена фильтром глубины, метод нормализации изображения оптимизирован для снижения памяти компьютера.

Искусственные нейронные сети используются и для решения задачи локализации положения объекта с расположенной на нем камеры наблюдения. В работе [23] данная задача решается для локомотива, а также описывается структура построения систем с алгоритмом визуального распознавания местности и принцип работы нейронной сети NetVLAD.

Большой объем исследований заключается в сравнении эффективности существующих нейросетевых методов при решении определенной задачи. В [62] приведено описание основных типов алгоритмов и методов для поиска объектов, основанных на использовании глубоких нейронных сетей. Работа [63] посвящена анализу решения задачи обнаружения объектов на изображениях с использованием глубоких сверточных нейронных сетей, оценка архитектур глубокой нейронной сети Tiny-YOLO в задаче обнаружения объектов в реальном времени. Сравнение сети MobileNet с другими архитектурами, оценка возможности её использования в задаче распознавания объектов беспилотными летательными аппаратами проведены в [64].

1.4 Гибридные алгоритмы

Под гибридные алгоритмы попадают как комбинации различных реализаций алгоритмов машинного обучения или классических методов компьютерного

зрения, так и алгоритмы, основанные на их совместном применении, с целью упрощения их использования и повышения надежности применения.

Известные структурно-функциональные решения для искусственных нейронных сетей не могут быть эффективно применены при анализе и классификации сложнотруктурируемых изображений ввиду ограниченности объемов обучающих выборок и высокой степени пересечения классов в признаковом пространстве [65, с 4].

В данной работе рассматриваются методы обработки изображений, однако, как отмечено в [66], идея создания гибридных алгоритмов относится не только к задачам распознавания объектов на изображениях, но и, например, в задачах автоматической классификации текста.

1.4.1 Гибридные алгоритмы, представляющие из себя комбинацию классических методов компьютерного зрения

Как было отмечено в предыдущих разделах данной главы, задача обнаружения лица человека на изображении достаточно распространена. В работе [25] предложенный алгоритм объединяет метод сравнения с шаблоном лица и модифицированный алгоритм обнаружения кожи лица: определение цвета кожи с использованием цветовых моделей и блочного метода, исключение областей, которые не относятся к изображению лица, обнаружение границ лица, сравнение с шаблоном, вывод обнаруженной области.

Задача поиска профиля лица человека на изображении решается в [26] – на первом этапе выполняется поиск лица с помощью алгоритма Виолы-Джонса, а затем проводится преобразование изображения в бинарную маску с последующим обнаружением связных областей кожи и отбрасыванием областей, где уже обнаружено лицо. В качестве итоговой обработки выполняется поиск областей, где присутствуют признаки наличия кожи и их сопоставление с результатами, полученными на первом этапе.

В работе [67] предложено использовать модель семейства сверточных искусственных нейронных сетей EfficientDet в комбинированном методе распознавания объектов камерой БПЛА. На первом этапе метода предлагается разбиение исходного изображения на отдельные области и отбрасывание полученных изображений, не содержащих в себе объектов. На следующем этапе фрагменты, содержащие не только фон, обрабатываются нейронной сетью.

Задача обнаружения пешехода на видеофрагменте с камеры видеонаблюдения решается в [16]. Применяются пирамидальное представление кадра для сокращения временных затрат, повышение вероятности правильного обнаружения объектов за счет уточнения предварительной карты векторов движения на основе полученных векторов движения на предыдущем шаге. Особенностью предложенного алгоритма является использование мультипликативной минимаксной функции схожести при анализе векторов движения для повышения точности локализации объектов. Для сопровождения объектов при использовании метода оптического потока предлагается модифицированный алгоритм трассировки.

Объединение методов визуального внимания и сегментации полученной карты внимания для обнаружения наземных объектов системой технического зрения беспилотного летательного аппарата предложено в работе [68]. Рассматриваемый алгоритм не требует наличия шаблона, знаний характеристик объекта или обучения модели на массиве данных.

1.4.2 Гибридные алгоритмы, представляющие из себя комбинацию методов машинного обучения или их дополнение

В работе [17] решается задача поиска и распознавания текста автомобильных номеров - алгоритм детектирования области расположения символов на изображениях со сложной фоновой структурой, отличающийся от аналогов

применением алгоритмической композиции из двух сверточных нейронных сетей, работающих по принципу последовательной классификации.

Сегментация гистологических структур проводится в [69] – на первом этапе осуществляется сегментация изображения посредством самоорганизующейся нейронной сети, на втором осуществляется анализ и классификация патологических изменений в сегментах, выделенных на первом этапе, на основе морфологического анализа и нейронной сети прямого распространения.

В работе [70] за счет построений гибридной нейро-нечеткой модели, состоящей из модуля сверточной нейронной сети и модуля нейро-нечеткого классификатора решается задача распознавания кошек и собак на изображениях.

Задачей работы [71] является разработка трехэтапного гибридного детектора с использованием искусственной нейронной сети EfficientDet. Работа предложенного метода состоит из трех основных этапов: (1) использование модели, основанной на архитектуре EfficientDet для предварительного обнаружения людей на изображении, (2) применение детектора MTCNN для распознавания лиц на полученной ранее исследуемой области, (3) модель FaceNet для классификации конкретного лица, обученная на специальном наборе данных изображений людей. В рамках данного исследования рассматриваются различные модели семейства сверточных сетей EfficientDet.

Стоит отметить, что многие из широко применяемых архитектур нейросетевых методов изначально образовались именно из задачи создания гибридных алгоритмов и в настоящее время уже сами используются как объект дальнейшей модификации. В качестве примера может служить модель Faster R-CNN: в работе [72] описано её образование при параллельном использовании двух методов предварительного определения возможных областей расположения объектов интереса Fast R-CNN и RPN путем модернизации части искусственной нейронной сети с последующим объединением обнаруженных областей для получения более точных контуров объекта.

Использование комбинации искусственной нейронной сети CenterNet и преобразователя изображения Vision Transformer, предназначенного для ускорения

выполнения вычислений путем разбиения исходного изображения на фрагменты с последующим их представлением в векторном виде, в задаче распознавания колосьев пшеницы на изображениях предложено в работе [73]. Рассматриваемый метод на первом этапе применяет CenterNet к входному изображению, на выход поступают карты объектов. На втором этапе Vision Transformer выполняет дополнительную обработку данных и готовит их к классификации на третьем этапе. В данной работе проведено сравнение предложенной комбинации методов с стандартными архитектурами сверточных искусственных нейронных сетей.

1.4.3 Гибридные алгоритмы, представляющие из себя комбинацию традиционного компьютерного зрения и машинного обучения

Данный класс гибридных алгоритмов представляет наибольший интерес при решении задач реального времени так как позволяет искать методы повышения эффективности систем компьютерного зрения за счет точности нейронных сетей и малых трудозатрат при использовании методов традиционного компьютерного зрения.

В задаче поиска лиц на изображении гибридность представлена объединением вейвлет-признаков со скрытой Марковской моделью и с вейвлет-нейронными сетями для распознавания лиц [27].

Предложено [15] решение задачи поиска опухолей на медицинских снимках легких и мозга – расчет порогового значения сегментации объектов, распознавание патологии на основе методов искусственных нейронных сетей, обнаружение её контуров (при наличии).

Объединение методов компьютерного зрения и нейронных сетей для решения задачи поиска и распознавания объектов дорожной инфраструктуры предложено в [74]. Изображение предварительно проходит цветовую сегментацию, проводится контурный анализ с последующим поиском наибольшего по площади контура и поиск области, в которой он располагается методами традиционного

компьютерного зрения. На полученной области изображения выполняется распознавание наличия дорожного знака с использованием предварительно обученной сверточной нейронной сети. Авторами отмечается, что полученный алгоритм чувствителен к входным данным: частичное перекрытие объекта интереса и низкое качество видео приводят к снижению эффективности предложенного метода.

Однако наибольший интерес из-за своей перспективности в части уменьшения требования к вычислительным мощностям при малом времени обработки данные гибридные алгоритмы представляют в направлении робототехники и беспилотных транспортных средств.

Задача распознавания пешеходов решается в [18] и [19]. В [18] используется последовательное применение каскадного классификатора и сверточной искусственной нейронной сети. В [19] реализован комбинированный алгоритм обнаружения пешеходов, представляющий собой комбинацию детектора пешеходов, алгоритма выделения движения и функции слежения. На первом этапе используется алгоритм выделения движения: выбирается опорный кадр и начальная оценка фона, выделение движения методом нахождения разности между текущим кадром и оценкой фона. На втором этапе применяется детектор пешеходов с функцией агрегирования каналов, для которого требуется предварительное обучение. С целью решения проблемы недостаточности обучающей выборки на последнем этапе предложенного метода применяется алгоритм слежения за объектом: выполняется оценка скорости передвижения объекта по межкадровому сдвигу и прогнозирование местоположения объекта на следующих кадрах на основе полученной информации. проводится комбинирование детектора пешеходов и метода выделения движения.

Рассматривается повышение точности работы роботов за счет дополнения методов машинного обучения логическими рассуждениями в части представлений типичного размера объектов интереса [75]. Предварительно выполнялась кластеризация объектов естественной среды по площади поверхности,

соотношению сторон и глубине. Данная информация интегрируется в полученные результаты нейронной сети на этапе постобработки изображения.

В [76] предполагается уменьшение требуемых объемов памяти оборудования при локализации в пространстве за счет алгоритма сжатия сцены. Используются методы сжатия, поиск ближайшего соседа, отбор геометрических совпадений, машинное обучение.

Работа [48] посвящена снижению вычислительной нагрузки и потребного объема памяти при решении задач распознавания объектов на изображениях путем оценки временных и мощностных затрат различных комбинаций алгоритмов совместного применения нейронных сетей и других методов: алгоритма каскада Хаара, локального двоичного шаблона (LBP) и гистограммно-ориентированных градиентов (HOG), для классификации используются алгоритмы сверточной нейронной сети (CNN) и глубокой нейронной сети (DNN). В результате разработаны шесть гибридных структур, таких как Haar Cascade+CNN, LBP+CNN, HOG+CNN и Haar Cascade+DNN, LBP+DNN, HOG+DNN.

В некоторых работах оценивается работа гибридных алгоритмов данного класса при необходимости распознавания объектов на неоднородном фоне, в условиях помех и деформирующих искажений.

В [65] предложены математическое и алгоритмическое обеспечение для гибридных нейросетевых структур, позволяющих осуществлять анализ и классификацию изображений на основе их яркостно-геометрического описания.

В работе [77] решается задача обнаружения техники на смоделированных и реальных изображениях. Предполагается, что данный алгоритм может быть использован для обнаружения объектов на неоднородном фоне в режиме реального времени опико-электронной системой. В данной статье предлагается использование локального усиления контраста для обнаружения инфракрасной малоразмерной цели и подавления фона. После приближения объекта на достаточное расстояние, выполняется обнаружение зон вероятного расположения с применением вейвлет-преобразования. На следующем этапе для создания ограничивающей контур объекта рамки и определения его координат применяется

алгоритм кластеризации данных DBSCAN. В локализованном окне выполняется классификация объекта обученной сверточной нейронной сетью.

Предварительное обнаружение отдельных частей объекта с последующим обнаружением всего объекта интереса предлагается в [78]: предварительная фиксация структурных и внешних характеристик объекта интереса, их поиск, последующий поиск всего объекта интереса в заданной области. Целью данного подхода является повышение качества распознавания объекта при его различных положениях относительно камеры. Поставленная задача решается на наборе данных, содержащем различные типы объектов интереса.

Задача обнаружения объектов на изображении на примере синтезированных объектов решается в работе [79]. На первом этапе с использованием статистически оптимального или нейросетевого алгоритма анализируются локальные участки изображения с целью принятия решения о наличии или отсутствии перепадов яркости. На втором этапе путем анализа результатов первичной обработки локальных участков с использованием алгоритма максимального правдоподобия принимается окончательное решение о наличии объекта на изображении и о локализации его целостной границы.

1.5 Результаты

В результате обзора научных работ по разработке и использованию различных методов обнаружения и классификации объектов на изображениях была предложена следующая классификация алгоритмов:

- классические алгоритмы распознавания;
- алгоритмы поиска объектов интереса на основе машинного обучения;
- гибридные алгоритмы:
 - гибридные алгоритмы, представляющие из себя комбинацию классических методов компьютерного зрения;

- гибридные алгоритмы, представляющие из себя комбинацию методов машинного обучения или их дополнение;

- гибридные алгоритмы, представляющие из себя комбинацию традиционного компьютерного зрения и методов машинного обучения.

Предложено определение понятия гибридного алгоритма: гибридный алгоритм - алгоритм, основанный на применении комбинации различных методов «традиционного» компьютерного зрения и нейросетевых технологий, направленный на упрощение их использования и повышение надежности распознавания.

Проведенный в настоящей главе обзор показал востребованность и актуальность разработки алгоритмов, основанных на совместном использовании традиционного компьютерного зрения и нейросетевых методов. Данное направление активно развивается, что выражается в публикации статей и научных работ, посвященных разработке гибридных алгоритмов.

2. Используемые методы традиционного компьютерного зрения и нейросетевых технологий

В данной главе представлено обоснование выбора используемых архитектур сверточных нейронных сетей и методов классического компьютерного зрения в части обнаружения и классификации объектов на изображении. Приведено сравнение алгоритмов поиска контуров, детекторов особых точек и архитектур нейронных сетей, доступных в открытых библиотеках.

2.1 Методы поиска контуров

В некоторых прикладных задачах распознавания объектов на изображении требуется определить описывающие их контуры. Стоит отметить, что проводятся исследования, в которых алгоритмы поиска контуров объектов на изображениях самостоятельно применяются для решения задачи выделения и построения трехмерных моделей антропогенных объектов за счет сопоставления отрезков на снимках [80]. Сложность решения данной задачи, а также разработки алгоритмов обнаружения границ объектов заключается в зашумленности поступающих на вход изображений, искажении поверхностей объектов и малом значении перепадов яркости на границах различных объектов и фона. При этом стоит отметить, что границы многих антропогенных объектов интереса хорошо описываются набором прямолинейных отрезков.

Алгоритмы обнаружения граничных линий и прямых, описывающих контур объектов на изображении, можно разделить следующим образом:

- по способу обработки входного изображения: целиком или отдельными фрагментами;
- по способу обнаружения прямых: после перехода к контурному представлению изображения или в процессе перехода.

В процессе исследований были рассмотрены два алгоритма поиска контуров объектов с выделением прямых: совместное применение детектора границ Кэнни с последующим использованием алгоритма поиска прямых Хафа и алгоритм Хюккеля, заключающийся в одновременном распознавании контуров и, при их обнаружении, определении их параметров в малой области изображения. Все описанные алгоритмы работают с изображениями, предварительно преобразованными в оттенки серого для уменьшения вычислительных затрат. Был реализован алгоритм Хюккеля, а также проведены численные эксперименты и решены демонстрационные примеры для обоих описанных подходов.

2.1.1 Описание работы алгоритма Хюккеля

Подробно описание принципа работы алгоритма Хюккеля представлено в работах [81, 82, 83]. В данном разделе показаны основные этапы применения данного алгоритма.

Алгоритм Хюккеля работает с отдельными малыми областями изображения по принципу «скользящего окна»: исходное пиксельное изображение разбивается на множество прямоугольных областей, которые могут пересекаться, в каждой из которых формируется входное «окно», наилучшим образом аппроксимирующее окружность. Рекомендации по выбору размеров «окна» и диаметра окружности в нем представлены в [81]. В настоящем исследовании на вход алгоритма поступали квадратные области 10 на 10 пикселей, из которых вырезалось окно, включающее в себя 69 пикселей (рис. 2.1.1). Далее в данной главе под термином «окно» будет пониматься фрагмент изображения, к которому применяется алгоритм Хюккеля и его идеализированное представление.

	1	2	3	4	5	6	7	8	9	10
1			1	2	3	4	5			
2		12	11	10	9	8	7	6		
3	13	14	15	16	17	18	19	20	21	
4	30	29	28	27	26	25	24	23	22	
5	31	32	33	34	35	36	37	38	39	
6	48	47	46	45	44	43	42	41	40	
7	49	50	51	52	53	54	55	56	57	
8		64	63	62	61	60	59	58		
9			65	66	67	68	69			
10										

Рис. 2.1.1 – Круглое «окно» из 69 пикселей, вырезаемое из квадратной области 10 на 10 пикселей

Целью работы алгоритма Хюккеля является представление области изображения в виде идеального элемента, контур на котором будет описан малым количеством параметров (рис. 2.1.2). Идеализированный элемент контура представляется в виде окружности.

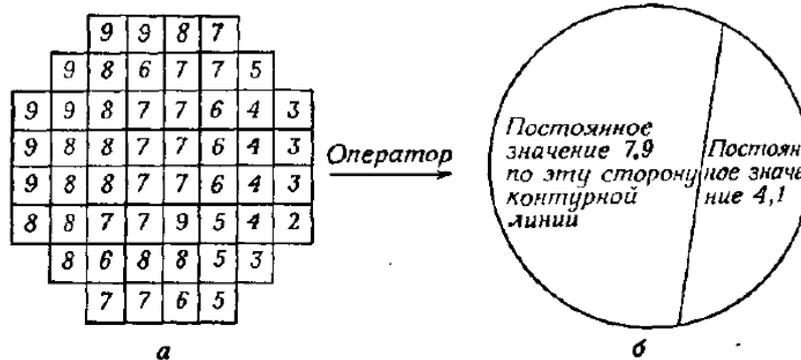


Рис. 2.1.2 – а – эмпирически полученный элемент, б – идеализированный элемент [81, с. 227]

Для каждого пикселя в обрабатываемом окне определяются его координаты x и y . Центр системы координат лежит в центре обрабатываемой области, радиус которой принимается равным единице.

Пусть через $E_p, p \in D$ задаются значения яркости изображения в точке p , при этом D – множество клеток области, наилучшим образом аппроксимирующее круглое окно, принадлежащее изображению. Входом алгоритма является любая функция $\{E_p | p \in D\}$. Первым этапом вычислений, проводимых алгоритмом Хюккеля, является подсчет девяти промежуточных чисел α_i по следующей формуле:

$$\alpha_i = \sum_{p \in D} H_{ip} E_p \quad (i = 0, \dots, 8) \quad (2.1.1)$$

$$\alpha_i = \int_D H_i(x, y) E(x, y) dx dy \quad (i = 0, \dots, 8) \quad (2.1.2)$$

Совокупность чисел H_{ip} образует матрицу из констант, получаемых из аналогичных непрерывных функций $H_i(x, y) (i = 0, \dots, 8)$, совокупность которых является решением системы функциональных уравнений. Данные уравнения выведены и рассчитаны Хюккелем в его работах и имеют единственное решение.

Граница идеального окна определяется следующим условием:

$$D = \{(x, y) | x^2 + y^2 \leq 1\} \quad (2.1.3)$$

При наложении эмпирического элемента на идеализированный, углы некоторых пикселей выходят за пределы окружности, описывающей расчетную область, а углы некоторых внешних пикселей попадают внутрь круга. Для минимизации возникающих ошибок применяются следующие действия:

- выбор рекомендуемых размеров окна;
- переход от (2.1.2) к (2.1.1) осуществляется преобразованием (2.1.4);

$$H_{ip} = \int_p H_i(x, y) dx dy \quad (2.1.4)$$

- ошибка, возникающая из-за выхода некоторых пикселей за пределы окружности, сводится к нулю путем задания условия (2.1.5)

$$H_i(x, y) = 0 \text{ при } x^2 + y^2 > 1 \quad (2.1.5)$$

Набор функций $H_i(x, y) (i = 0, \dots, 8)$ выбирается так, чтобы точкам окна приписывался различный вес. Для области, ограниченной окружностью $x^2 + y^2 = 1$, выбирается весовая функция $[1 - (x^2 + y^2)]^2$.

Базисные функции раскладываются на произведения радиальных и угловых периодических функций. Все функции $H_i (i = 0, \dots, 8)$ обращаются в нуль в периферии окна и на всех дополнительных окружностях и линиях, показанных на рис. 2.1.3. При определении наличия краевых линий и их параметров на результат работы алгоритма оказывают существенное влияние только амплитуды низких частот [81].

Алгоритм Хюккеля построен так, чтобы можно было распознать на входном окне одновременно два элемента контура и описать каждый из них четырьмя параметрами. Затем по величине отношения сигнал/шум решается, какому краю отдать предпочтение, и этот контур считается результатом применения алгоритма.

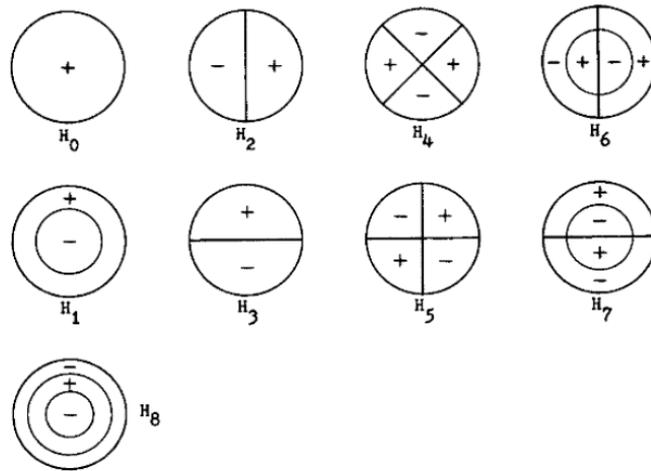


Рис. 2.1.3. Графическое представление базисных функций [82]

Линия контура описывается уравнением $cx + sy = \rho$, при этом для нормализации данного уравнения принимается, что $c^2 + s^2 = 1$. Обозначим b меньшую, а $b + d$ большую из яркостей идеального представления контура, тогда функция, описывающая идеальный контур в рассматриваемой области, принимает вид:

$$F(x, y, c, s, \rho, d, b) = \begin{cases} b, & \text{при } cx + sy \leq \rho, \\ b + d, & \text{при } cx + sy > \rho \end{cases} \quad (2.1.6)$$

Между идеальными краевыми элементами $F(x, y)$ и пятерками чисел (c, s, ρ, d, b) существует взаимно-однозначное соответствие.

В качестве меры оценки соответствия полученного контура перепада яркости реальному Хюккелем был выбран квадрат гильбертова расстояния.

$$\Lambda(c, s, \rho, d, b) = \int_D [E(x, y) - F(x, y, c, s, \rho, d, b)]^2 dx dy \quad (2.1.7)$$

Целью алгоритма является поиск значений $F(x, y, c, s, \rho, d, b)$, минимизирующих $\Lambda(c, s, \rho, d, b)$. Для уменьшения времени обработки критерий Λ был заменен критерием $\bar{\Lambda}$, который может быть минимизирован в соответствии с теоремой о решении, сформулированной в статье [80]. Данный критерий $\bar{\Lambda}$ является взвешенной суммой слагаемых критерия Λ .

$$\bar{\Lambda}(c, s, \rho, d, b) = \sum_{i=0}^{\infty} (w_i (a_i - f_i(c, s, \rho, d, b)))^2, \quad (2.1.8)$$

где w_i – весовые коэффициенты.

Точное определение алгоритма: для данного входа $E(x, y)$ нужно найти такую пятерку чисел (c, s, ρ, d, b) , чтобы $\bar{\Lambda}(c, s, \rho, d, b)$ было минимальным, и выдать эти числа на выход. На практике из полученных коэффициентов определяют угол наклона градиента перепада яркости и расстояние от центра области до края.

В расширенной версии алгоритма Хюккеля, подробно рассматриваемой в работе [82], добавляются коэффициенты, отвечающие за тип краевой линии (край, линия), её ширину и яркость.

Для оценки зашумленности изображения вводится критерий приемлемости шаблона k . В случае идеального краевого элемента его значение равно единице, при $k < 0,9$ считается, что элемент контура слишком зашумлен.

На практике элемент контура рассматривается в зависимости от перепада яркости в исследуемой области; для этого задаются пороговыми значениями $conf$ и $diff$. Обнаруженный перепад считается элементом контура при выполнении условия (2.1.9).

$$1 - k < (1 - conf)d^2 / (diff^2 + d^2) \quad (2.1.9)$$

При начале работы с алгоритмом задаются некоторыми значениями настраиваемых параметров $conf$ и $diff$, или порогами $CONF$ и $DIFF$, по своему смыслу соотносящимися как (2.1.10).

$$conf^{-2} - 1 = CONF \text{ и } diff^2 = DIFF \quad (2.1.10)$$

От выбора настраиваемых параметров зависит насколько сильный яркостной перепад должен быть в расчетной области, чтобы отнести её к контуру с присутствующим краем или линией.

Что будет являться задачей алгоритма – поиск линий или краев на контуре, регулируется соответствующими переменными, от значений которых зависит логика работы алгоритма Хюккеля. Сравнительный анализ полученных решений при реализации алгоритма с результатами, представленными в работах Хюккеля, показан в [83].

2.1.2 Описание работы алгоритмов детектора границ Кэнни и поиска прямых Хафа

Подробно принцип работы алгоритмов детектора границ Кэнни и поиска прямых Хафа рассматриваются в [1, 30]. Также в данных работах представлены примеры их использования.

Детектор границ Кэнни является алгоритмом обнаружения контуров объектов и обрабатывает входное изображение целиком. В качестве настраиваемых параметров данного алгоритма используются верхний и нижний пороги значений градиента яркости и размер ядра Собеля, используемого при вычислениях.

Этапы работы алгоритма:

1. Сглаживание изображения для подавления шума;
2. Вычисление первых производных по x и y , которые затем объединяются в четыре производные по направлению;

3. Подавление немаксимумов: точки, в которых производные достигают локального максимума, считаются кандидатами на включение в границы;
4. Двойная пороговая фильтрация (объединение пикселей в контуры):
 - пиксели со значениями яркости больше или равными верхнему порогу могут считаться достоверными граничными пикселями;
 - пиксели со значениями яркости меньше или равными нижнему порогу, будут подавлены;
 - пиксели значения яркостей которых попали в область между верхним и нижним порогом, принимают фиксированное среднее значение. Пиксель относится к границе, если соприкасается с ней по одному из восьми направлений.

Алгоритм Хафа позволяет обнаруживать различные простые фигуры на изображениях (прямые, окружности и др.). В исследовании используется данный алгоритм применительно к предварительно обработанным преобразованием Кэнни изображениям применительно к задаче поиска прямых.

Преобразование Хафа для поиска прямых заключается в формировании аккумуляторной плоскости, каждая точка которой указывает на вероятность наличия на изображении прямой, проходящей через определенную точку исходного изображения.

При численных расчетах прямая представляется точкой (ρ, θ) в полярной системе координат, так что представленная некоторой точкой прямая проходит через нее перпендикулярно ее радиусу-вектору. Такая прямая описывается уравнением $\rho = x \cos \theta + y \sin \theta$ (рис. 2.1.4).

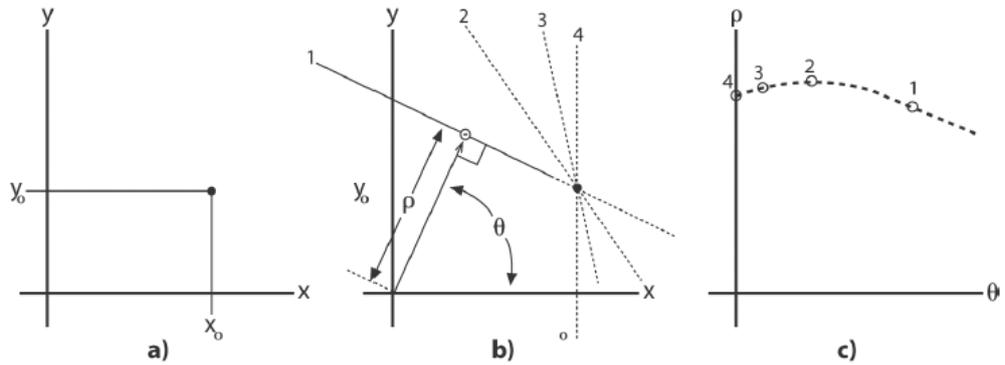


Рис. 2.1.4. Схематичное изображение поиска прямых через точки при использовании алгоритма Хафа [1, с. 302]

Точка (x_0, y_0) на плоскости изображения (а) принадлежит множеству прямых, каждая из которых параметризуется различными значениями ρ и θ (b); на плоскости (ρ, θ) совокупность таких прямых образует кривую характерной формы (с). Построение кривых выполняется для всех точек обрабатываемого изображения. Пересечение кривых в плоскости показывает, что точки принадлежат одной линии. По числу пересечений между кривыми можно сделать заключение о прохождении линии, а по числу пересечений кривых о её длине.

Варьируемым параметром алгоритма поиска прямых Хафа в данном исследовании являлся порог аккумулятора, обозначенный в данной работе как «k». Этот ненормированный аргумент, по существу, говорит сколько точек должно принадлежать прямой, чтобы алгоритм вернул эту прямую [1, с. 304].

2.1.3 Тестирование алгоритмов поиска контуров

Первым этапом проверки точности реализованных при выполнении исследований алгоритмов было их тестирование на сформированных искусственных областях размером 10 на 10 пикселей.

Варьируемыми параметрами являлись угол наклона линии края, величина перепада яркости областей, значение меньшей яркости пикселей для оценки чувствительности реализованного алгоритма. Выполнялось смещение границы перепада по осям X и Y для оценки правильности определения длины

перпендикуляра от центра расчетной области до искомого края, а также влияния расположения области перепада на точность работы алгоритма. Полученные при вычислениях значения направления градиента яркости сравнивались с заданными.

В процессе выполнения теста принимались следующие значения углов наклона: 0, 20, 30, 45, 60, 80, 90, 120, 135, 150, 160, 180 градусов.

Величина перепада яркости варьировалась от 2 до 14, в зависимости от минимальной яркости пикселей на изображении.

Примеры расчетных тестовых областей приведены на рисунке 2.1.5, а получаемых алгоритмом Хюккеля результатов представлены в таблице 2.1.1.

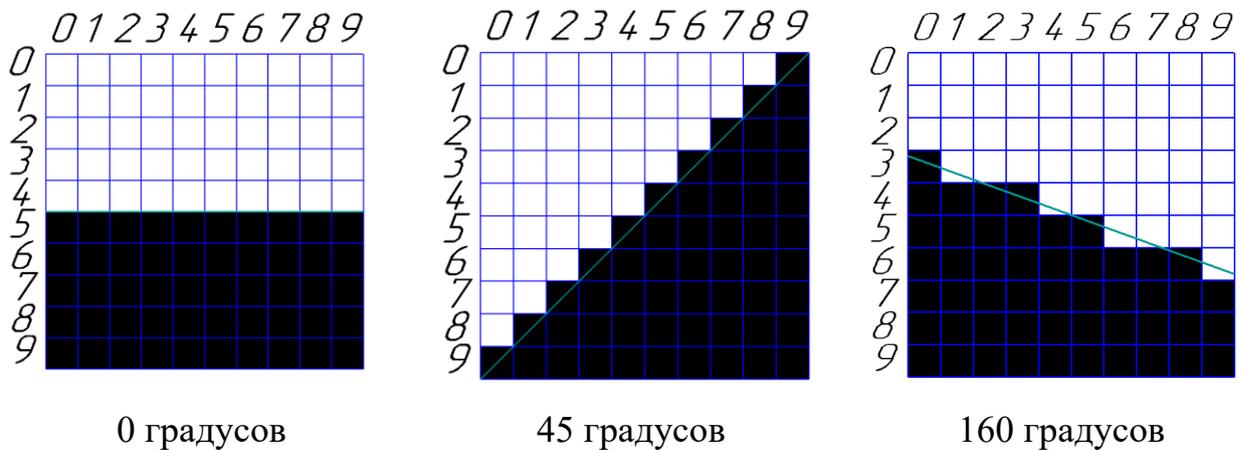


Рис. 2.1.5. Примеры расчетных тестовых областей

Таблица 2.1.1 – Пример полученных результатов использования алгоритма Хюккеля

Угол наклона контура 0°						
Входная яркость		Направление градиента		Полученные значения яркости		
min	max	cos	sin	min	перепад	max
3	5	0	-1	3	2	5
3	7	0	-1	3	4	7
3	9	0	-1	3	6	9
3	11	0	-1	3	8	11
3	13	0	-1	3	10	13
3	15	0	-1	3	12	15
9	11	0	-1	9	2	11
9	13	0	-1	9	4	13
9	15	0	-1	9	6	15

Продолжение таблицы 2.1.1

Угол наклона контура 45°						
Входная яркость		Направление градиента		Полученные значения яркости		
min	max	cos	sin	min	перепад	max
3	5	0,707107	-0,707107	3	2	5
3	7	0,707107	-0,707107	3	4	7
3	9	0,707107	-0,707107	3	6	9
3	11	0,707107	-0,707107	3	8	11
3	13	0,707107	-0,707107	3	10	13
3	15	0,707107	-0,707107	3	12	15
9	11	0,707107	-0,707107	9	2	11
9	13	0,707107	-0,707107	9	4	13
9	15	0,707107	-0,707107	9	6	15
Угол наклона контура 160°						
Входная яркость		Направление градиента		Полученные значения яркости		
min	max	cos	sin	min	перепад	max
3	5	-0.362927	-0.931817	3	2	5
3	7	-0.362907	-0.931825	3	4	7
3	9	-0.362900	-0.931828	3	6	9
3	11	-0.362897	-0.931829	3	8	11
3	13	-0.362895	-0.931830	3	10	13
3	15	-0.362894	-0.931831	3	12	15
9	11	-0.363005	-0.931787	9	2	11
9	13	-0.362947	-0.931810	9	4	13
9	15	-0.362927	-0.931817	9	6	15

Полученные при тестировании результаты показали возможность использования рассматриваемых алгоритмов при решении задач поиска линии контура в малой области.

Для оценки точности обнаружения линий на всём изображении алгоритмами Хюккеля и Хафа решалась следующая тестовая задача: создано искусственное изображение (рис. 2.1.6), на котором проведены две белые параллельные прямые на черном фоне с заданным углом наклона к оси X равным 150°.

К рассматриваемому изображению был применен алгоритм Хюккеля со следующими настраиваемыми параметрами: $n=10$, $conf=0.75$, $diff=20$. Полученное изображение представлено на рисунке 2.1.7.

Данное тестовое изображение также было обработано алгоритмом Хафа поиска прямых. Пример полученного изображения представлен на рис. 2.1.8.

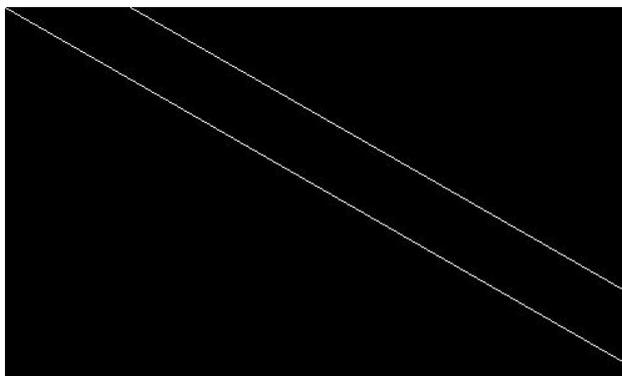


Рис. 2.1.6. Тестовое изображение

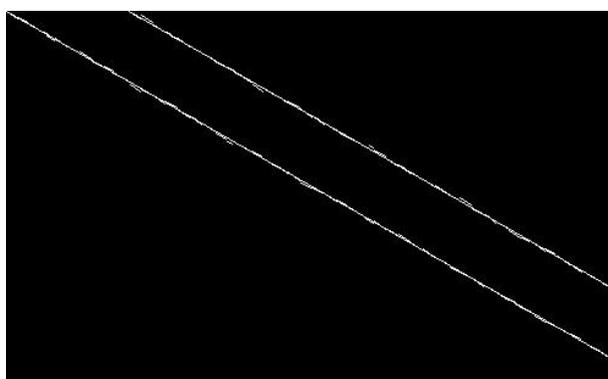


Рис. 2.1.7. Пример изображения, полученного алгоритмом Хьюккеля

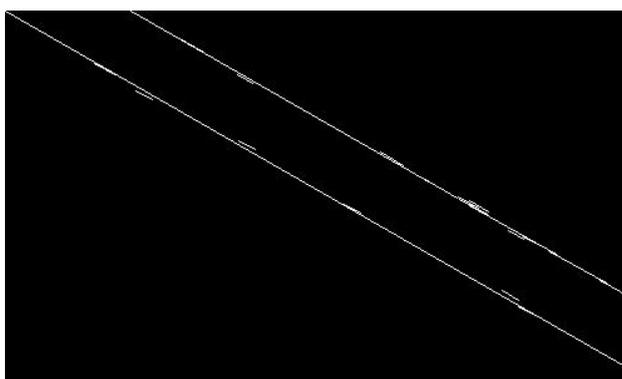


Рис. 2.1.8 Пример изображения, полученного преобразованием Хафа

Полученные на расчетных тестовых областях результаты показали возможность использования данных алгоритмов для решения задач поиска линии контура в малой области. Для изображения, полученного алгоритмом Хьюккеля, было определено, что ошибка менее 2° достигается в 48% полученных результатов, а превышает 5° в 12%.

Алгоритм Хафа дал ошибку менее 2° для всех обнаруженных прямых.

Полученные результаты демонстрируют возможность применения реализованного алгоритма Хюккеля для решения задачи поиска градиента яркости. Проведенные расчеты показали, что точность обнаружения границы и определения её параметров зависит от расположения края относительно центра обрабатываемой области. В последующем при обработке изображений при решении прикладных задач шаг смещения окна обработки будет принят равным не более 9 пикселей.

Решение задачи обнаружения контуров на искусственно созданном изображении показало большую точность определения угла наклона прямой при использовании алгоритма Хафа. Зависимость качества получаемых результатов в случае применения обоих вариантов алгоритмов от настраиваемых параметров показывает, что для решения конкретной прикладной или демонстрационной задачи перед проведением оценки затраченного времени на обработку изображения или оценки точности нахождения контура реального объекта необходимо предварительно уточнить значения данных параметров с целью нахождения наибольшего числа линий контура объекта интереса при минимальной зашумленности получаемого изображения.

2.1.4 Применение алгоритмов поиска контуров на реальных изображениях

Для сравнения результатов работы алгоритма Хюккеля и совместного применения алгоритмов Кэнни и Хафа при работе с реальным изображением проведен серия вычислительных экспериментов на серии спутниковых снимков, полученных из открытых источников [84]. Оценка возможности обработки спутниковых изображений алгоритмом Хюккеля проведена в работе [85], в данной работе рассматривалось влияние размера обрабатываемой области на количество обнаруженных точек граничных линий и время обработки.

В приведенном примере объектом интереса является комплекс сельскохозяйственных построек, границы которых хорошо различимы на исходном изображении, что в дальнейшем позволит построить точный контур эталонного представления объекта интереса.

Размер исходного изображения составляет 1156 на 721 пиксель. К изображению был применен алгоритм Хюккеля при различных значениях настраиваемых параметров. Также проводилась обработка алгоритмом Кэнни с последующим использованием алгоритма Хафа поиска прямых.

Для оценки влияния размера исходного изображения на скорость работы рассматриваемых алгоритмов выполнено масштабирование исходного изображения в 2, 4 и 8 раз.

Рассмотрены следующие варианты обработки изображений:

- 1) Применен алгоритм Кэнни для всего изображения, затем для всего изображения применен алгоритм Хафа;
- 2) Для областей размером 10x10 пикселей с шагом n применен алгоритм Хюккеля.

Значение шага сдвига обрабатываемой алгоритмом Хюккеля области варьировались $n = 3, 5, 7, 9$ пикселей при обработке исходного изображения, а также изображений, полученных уменьшением масштаба в 2 и 4 раза. Для изображений с разрешением, уменьшенным в 4 и 8 раз, также проводилась оценка при сдвиге области обработки на $n = 2, 3, 4, 5$ пикселей.

На первом этапе каждое изображение было обработано при широком диапазоне варьирования настраиваемых параметров для определения их оптимального значения в зависимости от размера изображения и перепада яркости искомого контура. При дальнейшей обработке и оценке результатов применения алгоритма использовались конкретные предпочтительные значения настраиваемых параметров алгоритмов, позволяющие обнаружить наибольшее число контуров объектов.

Исходное изображение и изображение с отмеченным объектом интереса представлены на рисунке 2.1.9.



А)



Б)

Рис. 2.1.9. Исходное изображение (А) и изображение с выделенным объектом интереса (Б)

Примеры полученных изображений при использовании алгоритма Хюккеля, а также совместного применения алгоритмов Кэнни и Хафа к исходному, а также отмасштабированному в 2 и 4 раза изображениям показаны на рисунке 2.1.10.

Из полученных результатов видно, что обнаружение объекта интереса возможно как алгоритмом Хюккеля, так и при совместном применении алгоритмов Кэнни и Хафа как на исходном изображении, так и при уменьшенном в 2 и 4 раза масштабе.

При уменьшении разрешения изображения в 8 раз только совместное применение алгоритмов Кэнни и Хафа позволяют определить наличие объекта интереса подобного класса.

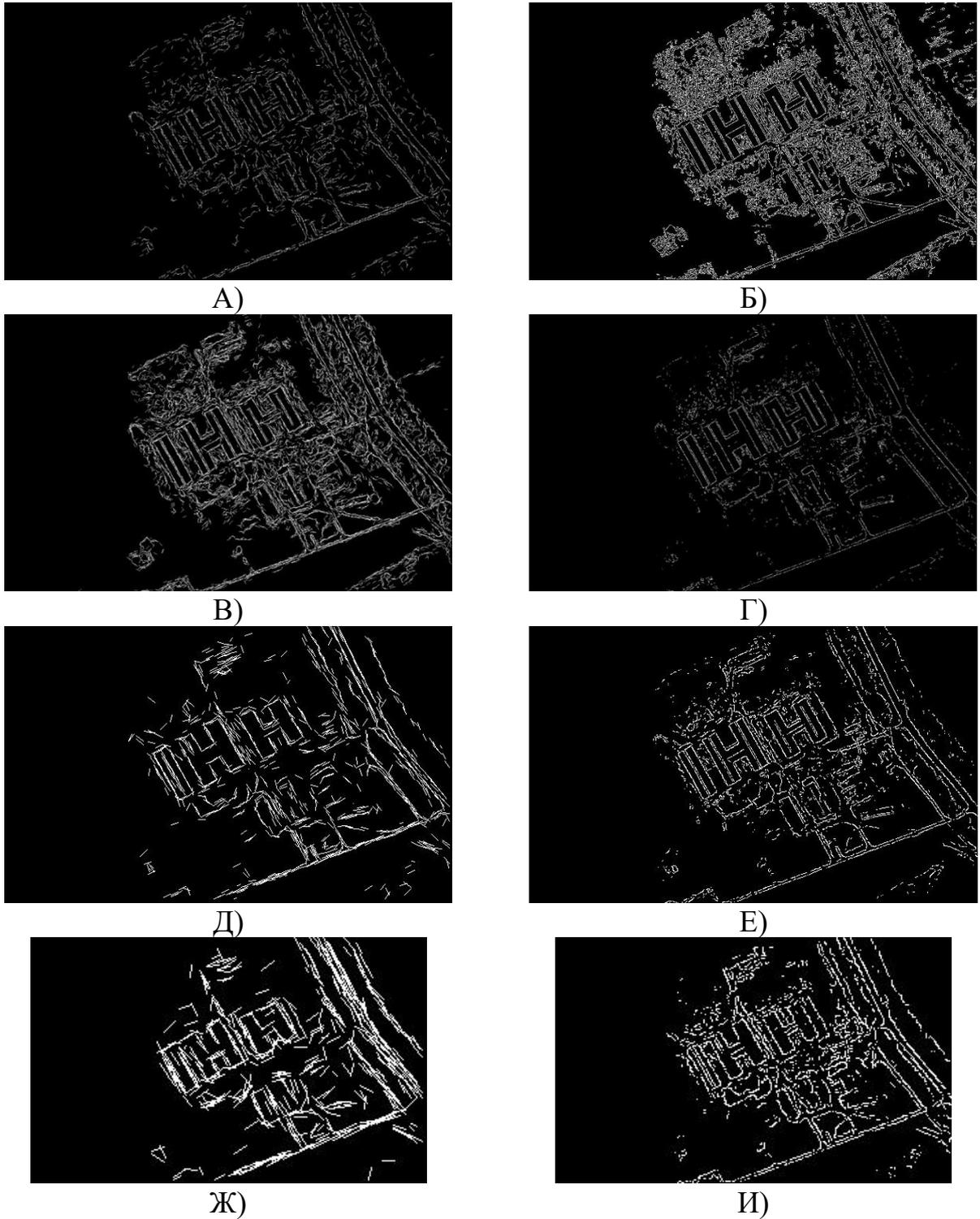


Рис. 2.1.10. Изображения, полученные алгоритмом Хюккеля (А, В, Д, Ж) и совместным использованием алгоритмов Кэнни и Хафа (Б, Г, Е, И) при различных масштабах

Была проведена оценка затрачиваемого времени на обработку изображений. Полученные результаты представлены в таблицах 2.1.2 и 2.1.3. Для каждого заданного набора параметров выполнялось 100 вычислений и бралось среднее время обработки изображения.

Таблица 2.1.2 – Результаты временной оценки работы алгоритмов

	Настраиваемые параметры, Хюккель	Настраиваемые параметры, Кэнни+Хаф	Среднее время Хюккель, с	Среднее время Хаф, с
Исходное изображение (1156x721)	n=9, conf=0.15, diff=20	k=8, Canny(60;140)	0.155893	0.562933
	n=7, conf=0.15, diff=20	k=6, Canny(60;140)	0.224592	0.570257
	n=5, conf=0.15, diff=20	k=4, Canny(60;140)	0.322393	0.559069
Разрешение уменьшено в 2 раза (578x360)	n=5, conf=0.15, diff=20	k=6, Canny(60;140)	0.079626	0.172817
	n=3, conf=0.15, diff=20	k=4, Canny(60;140)	0.194198	0.173226
Разрешение уменьшено в 4 раза (289x180)	n=3, conf=0.15, diff=20	k=4, Canny(60;140)	0.044446	0.056029

Таблица 2.1.3 – Результаты оценки числа обрабатываемых изображений в секунду

	Настраиваемые параметры, Хюккель	Настраиваемые параметры, Кэнни+Хаф	Среднее кадр/с Хюккель	Среднее кадр/с Хаф
Исходное изображение (1156x721)	n=9, conf=0.15, diff=20	k=8, Canny(60;140)	6.414656207	1.77641033
	n=7, conf=0.15, diff=20	k=6, Canny(60;140)	4.452518344	1.75359530
	n=5, conf=0.15, diff=20	k=4, Canny(60;140)	3.10180432	1.78868797
Разрешение уменьшено в 2 раза (578x360)	n=5, conf=0.15, diff=20	k=6, Canny(60;140)	12.55871198	5.78646776
	n=3, conf=0.15, diff=20	k=4, Canny(60;140)	5.149383619	5.77280546
Разрешение уменьшено в 4 раза (289x180)	n=3, conf=0.15, diff=20	k=4, Canny(60;140)	22.49921253	17.8479001

Алгоритм Хюккеля выполняет обнаружение контуров объекта интереса за заметно меньшее время при обработке исходного изображения. Уменьшение разрешения изображения приводит к уменьшению различий во времени обработки. Большое значение на время обработки изображения алгоритмом Хюккеля оказывает шаг обработки, влияние других настраиваемых параметров не оказывает значительного влияния и для случая алгоритма Хюккеля, ни при совместном использовании алгоритмов Кэнни и Хафа.

2.1.5 Построение эталонного изображения

Для оценки точности работы алгоритмов на исходном изображении определён эталонный контур объекта интереса, заданный набором прямых, расположенных под углами 29° и 119° . Для областей, в которых алгоритм обнаружил линию контура, из выходных данных алгоритма определен угол наклона линии контура. На основе сравнения с углом наклона эталонного контура объекта интереса формируется распределение ошибки определения угла наклона краевой линии.

Результаты применения алгоритмов Хюккеля с значениями настраиваемых параметров $\text{conf} = 0.15$ $\text{diff} = 20$ и Хафа с заданным пороговым значением $k=8$ и минимальной длиной обнаруженных прямых равной 4 пикселям, ко всему изображению, представлены на рисунках 2.1.11 и 2.1.12. Обнаруженные линии контура наложены на исходное изображение.



Рис. 2.1.11. Результат применения алгоритма Хюккеля к исходному изображению



Рис. 2.1.12. Результат применения алгоритма Хюккеля к исходному изображению

Эталонный контур объекта интереса представлен на рисунке 2.1.13 (область расположения объекта интереса увеличена).



Рис. 2.1.13. Эталонный контур объекта интереса

Для оценки получаемой точности был применен алгоритм Хюккеля со значениями настраиваемых параметров $conf = 0.25$ $diff = 100$ к областям, описанным уравнениями (2.1.11), (2.1.12), (2.1.13), (2.1.14). Перекрытие областями поиска контуров алгоритмом Хюккеля эталонной границы объекта интереса показан на рисунке 2.1.14 (область расположения объекта интереса увеличена).



Рис. 2.1.14. Перекрытие областями поиска контуров алгоритмом Хюккеля эталонной границы объекта интереса.

Данные уравнения составлены так, чтобы в начальной точке элементов эталонного контура – прямых находился центр области обработки алгоритма Хюккеля. Переменные x_0 и y_0 – начальные точки; i – номер области (для первого или второго набора прямых соответственно); k – шаг сдвига области обработки в пикселах; n – сдвиг центра начальной области обработки по вертикали.

$$x_{i1} = x_{01} - 5 + i_1 \cdot k; \quad (2.1.11)$$

$$y_{i1} = y_{01} - 5 - n - (i_1 \cdot k \cdot \tan(29^\circ)); \quad (2.1.12)$$

$$x_{i2} = x_{02} - 5 + i \cdot k; \quad (2.1.13)$$

$$y_{i2} = y_{02} - 5 - n - (i_2 \cdot k \cdot \tan(119^\circ)); \quad (2.1.14)$$

$n = -1; 0; 1$ – начальный сдвиг области обработки от эталонного контура по координате y относительно y_{01} и y_{02} .

$k = 5$ – шаг области обработки по координате x .

На рисунке 2.1.15 показаны найденные алгоритмом Хюккеля краевые линии в обрабатываемых областях, наложенные на исходное изображение (область расположения объекта интереса увеличена).



Рис. 2.1.15. Найденные алгоритмом Хюккеля контуры в обрабатываемых областях

Для обрабатываемых областей, в которых алгоритм Хюккеля обнаружил линию контура, из выходных данных алгоритма определено направление градиента перепада яркости. Определялось значение угла наклона краевой линии к оси X исходного изображения и проводилось сравнение с эталонным контуром. Признак обнаружения контура присутствовал во всех обрабатываемых областях. Полученное распределение ошибки представлено на рисунке 2.1.16.

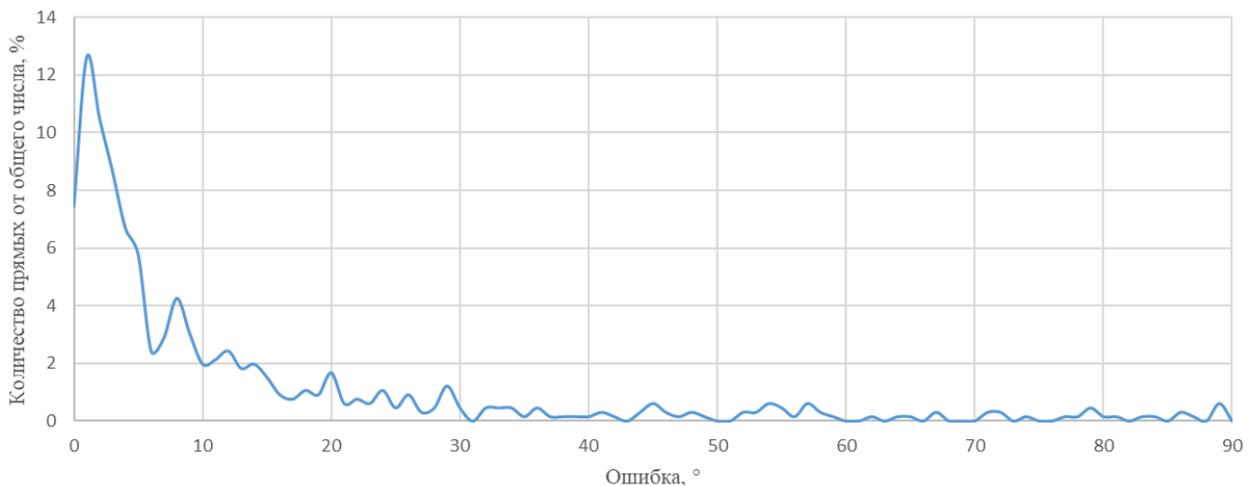


Рис. 2.1.16. Распределение ошибки определенных углов алгоритмом Хюккеля

Проведена обработка исходного изображения алгоритмом Кэнни, после чего, к областям изображения, перекрывающим эталонный контур объекта интереса

(рис. 2.1.14), применен алгоритм поиска прямых Хафа. Расчет проводился с настраиваемыми параметрами алгоритмов Кэнни (60; 140) и Хафа (1, $\pi/180$, 8).

По аналогии с описанной выше оценкой алгоритма Хюккеля, выполнено сравнение результатов, полученных алгоритмом Хафа с эталонным изображением. Выполнялась обработка областей 10 на 10 пикселей, найденные в обрабатываемых областях контуры представлены на рисунке 2.1.17 (область расположения объекта интереса увеличена). Полученное распределение ошибки представлено на рисунке 2.1.18



Рис. 2.1.17. Найденные алгоритмом Хафа контуры в обрабатываемых областях

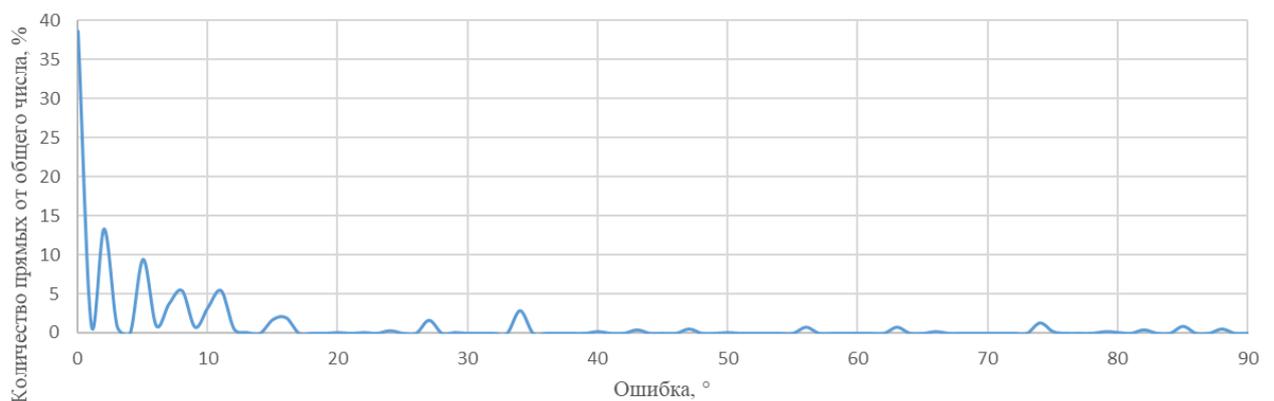


Рис. 2.1.18. Распределение ошибки определенных углов алгоритмом Хафа

Для областей, в которых алгоритм Хафа обнаружил линию контура, из выходных точек алгоритма определялся угол наклона краевой линии к оси X исходного изображения и проводилось сравнение с заданными значениями эталонного контура.

Значения ошибок определения угла наклона линии контура, не превышающих 5° и 10° , по сравнению с эталоном, алгоритмами Хюккеля и Хафа, представлены в таблице 2.1.4.

Таблица 2.1.4 – Значения ошибок определения угла линий контура

Значение ошибки	Алгоритм Хюккеля	Алгоритм Хафа
$\leq 5^\circ$	52%	64%
$\leq 10^\circ$	66%	78%

На основе полученных результатов видно, что совместное применение алгоритмов Кэнни и Хафа как к изображению целиком, так и к отдельным областям позволяет обнаруживать перепады яркости на изображении с большей точностью. Стоит учитывать, что значения ошибок могут быть завышены, так как в расчетные области также попадали контуры объектов, не относящихся к объекту интереса.

2.1.6 Рекомендации по выбору настраиваемых параметров

В зависимости от значений настраиваемых параметров рассмотренных алгоритмов получаемые изображения могут быть как зашумленными, так попадающими в диапазон четкого различия контуров, также возможно отсутствие распознавания краевых линий из-за завышенных требований к величине перепада яркости, что приводит к потере контуров объекта интереса.

На основе решения демонстрационной задачи путем варьирования настраиваемых параметров алгоритмов определено, что при малых значениях заданных переменных алгоритма Хюккеля или низких пороговых значениях для алгоритмов Кэнни и Хафа алгоритмы реагируют даже на незначительные перепады яркости. В результате контурное изображение получается зашумленным. При

дальнейшем увеличении значений настраиваемых параметров начинается диапазон, подходящий для четкого различия краевых линий отдельных объектов, после которого для обнаружения контура объекта становится необходимым высокий перепад яркости пикселей.

На рисунке 2.1.19 показано графическое представление распределения полученных результатов в зависимости от заданных настраиваемых параметров алгоритмов Хюккеля и совместного применения алгоритмов Кэнни и Хафа (в качестве примера представлено распределение для фиксированных параметров алгоритма Хафа).

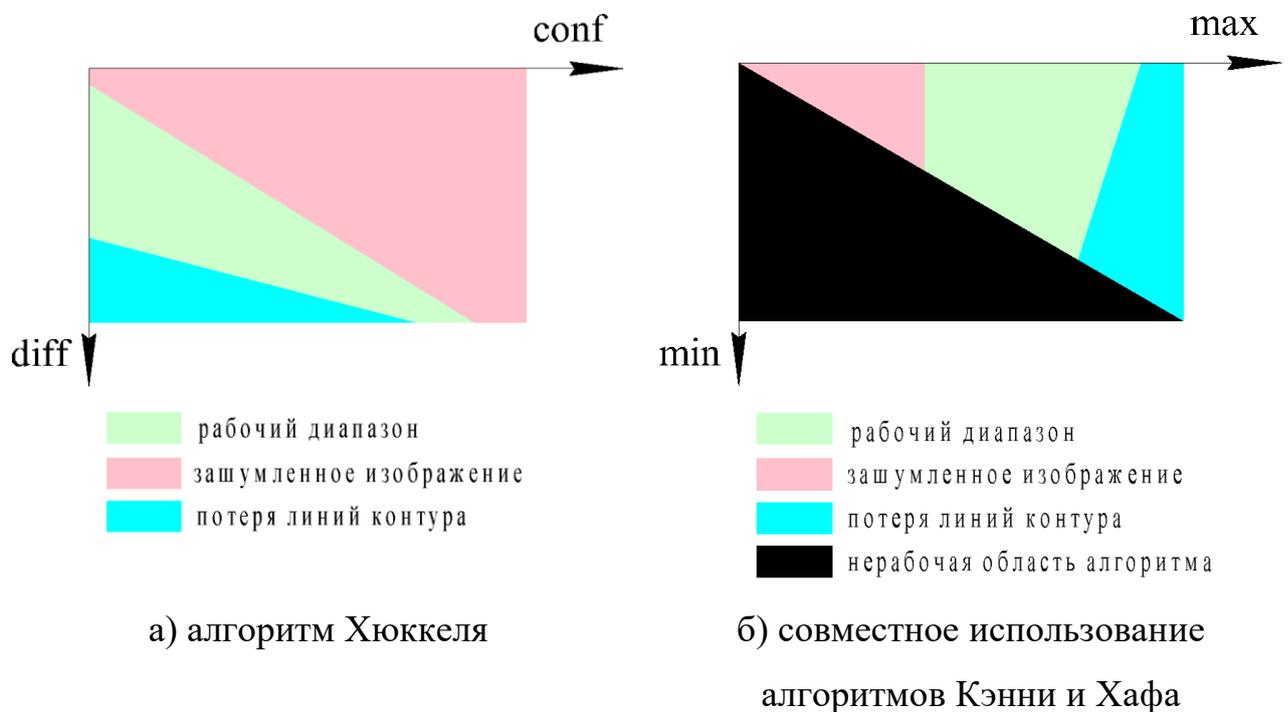


Рис. 2.1.19 Графическое представление полученных результатов от настраиваемых параметров для алгоритмов: а) Хюккеля и б) Кэнни+Хаф

Конкретные предпочтительные значения настраиваемых параметров зависят от типа объекта интереса и сцены, качества изображения, но на основе полученных при решении демонстрационной задачи результатов возможно сформировать общие рекомендации по начальному выбору настраиваемых параметров, с которых следует начать поиск оптимальных значений для конкретно выбранного объекта интереса на предоставленной сцене.

Для алгоритма Хюккеля – шаг сдвига окна обработки 5 – 7 пикселей, начальные значения настраиваемых параметров рекомендуется брать из диапазонов, указанных в таблице 2.1.5.

Таблица 2.1.5 – Рекомендуемые начальные диапазоны варьирования настраиваемых параметров алгоритма Хюккеля

conf	diff
0.15	80 – 140
0.25	100 – 160
0.35	120 – 180
0.45	160 – 200

Для совместного использования алгоритмов Кэнни и Хафа значение порогового коэффициента алгоритма Хафа предпочтительно брать 6 пикселей, а также в зависимости от размера объекта интереса задаться минимальной длиной искомым прямым. Значения минимального и максимального пороговых значений алгоритма Кэнни для объектов интереса и сцен, подобных рассмотренной в данной демонстрационной задаче, рекомендуется брать из диапазона, представленного в таблице 2.1.6.

Таблица 2.1.6 – Рекомендуемые начальные диапазоны варьирования пороговых значений алгоритма Кэнни

min	max
100	200
160	220

2.2 Детекторы особых точек

Помимо антропогенных объектов, контур которых возможно описать набором прямых, выделяются объекты интереса, образ которых не содержит характерных прямолинейных признаков, но в достаточной степени описывается точечными признаками. Для распознавания объектов данного класса одним из подходов традиционного компьютерного зрения является применение детекторов особых точек. Данный метод позволяет обнаружить на различных изображениях особые точки расположенных объектов, а затем по количеству совпадений сделать заключение о наличии объекта в той или иной области обрабатываемого изображения.

Особые точки представляют собой небольшие фрагменты изображения с высокой плотностью информации о том, что на нем изображено. Они позволяют перейти от пиксельного представления изображения к более компактному.

Дескрипторы представляют из себя некую математическую конструкцию, описывающую особую точку и позволяющую выполнить в дальнейшем сопоставление двух особых точек.

Задача сопоставления двух изображений заключается в обнаружении особых точек на каждом из них, создании дескриптора для каждой точки, проведении сравнения полученных дескрипторов.

В настоящее время существует множество детекторов обнаружения особых точек, различающихся скоростью работы, мощностью, инвариантностью относительно вращений, масштабирования, а иногда и малых аффинных преобразований, и методов их сопоставления [1].

2.2.1 Выбор детектора

При выполнении данной работы предварительно выполнялась оценка скорости и точности работы таких детекторов особых точек, как FAST, SIFT, SURF

и ORB [1]. Рассматривались детекторы, реализованные в библиотеке OpenCV, а в качестве объекта интереса использовался объект типа «БЛА».

Алгоритм обнаружения признаков FAST основан на прямом сравнении точки со множеством точек внутри небольшой описанной вокруг нее окружности. По результатам сравнения выполняется классификация соседних точек на основе пороговой величины. В результате оценки количества подходящих пикселей в заданной области делается заключение о применимости данной точки для дальнейшего использования в работе. Достоинством детектора FAST является быстрота работы, но при этом отсутствуют признаки ориентации. Пример обнаружения особых точек на изображении детектором FAST представлен на рисунке 2.2.1.



Рис. 2.2.1 Пример применения детектора особых точек FAST

Алгоритм SIFT используется при решении задач распознавания и сопровождения. Достоинствами данного метода являются инвариантность

относительно масштаба и направления. К недостаткам данного алгоритма относят дороговизну проводимых вычислений. Пример применения детектора особых точек SIFT представлен на рисунке 2.2.2.



Рис. 2.2.2. Пример обнаружения особых точек на изображении детектором SIFT

Детектор признаков SURF является более эффективным с вычислительной точки зрения продолжением алгоритма SIFT. Также разработчикам данного алгоритма удалось повысить устойчивость к изменениям ориентации и освещения.

Особенностью вычислителя дескрипторов BRIEF является вычисление дескрипторов особых точек, найденных другим методом. Главной идеей дескриптора BRIEF является то, что признак описывается в виде последовательности тестов, в каждом из которых сравнивается один пиксель в окрестности признака с каким-то другим пикселем, в результате чего получается бинарный ответ в зависимости от яркости пикселей. Данный дескриптор представляет собой совокупность результатов n таких тестов. Для подавления излишней чувствительности дескриптора к шуму, алгоритм сначала сглаживает изображение, свернув его с ядром Гаусса. Одним из способов выбрать сравниваемые пары, образующие дескриптор BRIEF, является разбиение на случайные пары, при этом сначала выбирается точка из нормального распределения с центром в особой точке, а затем – вторая точка из нормального распределения с центром в первой выбранной точке. Число генерируемых тестов может варьироваться.

В качестве более быстрой альтернативой методам SIFT и SURF для решения задач обнаружения особых точек и их дескрипторов был разработан детектор

признаков ORB. Данный детектор похож на FAST, но используемый дескриптор основан на BRIEF и дополнен вычислением ориентации, которая придает признакам ORB инвариантность относительно вращения. На первом этапе работы алгоритма потенциальные признаки обнаруживаются с помощью FAST, но с дополнением поиска в масштабируемом пространстве. Также была добавлена к методу FAST ориентация найденных особых точек. Для достижения инвариантности относительно вращения выполняется вычисление вектора признака относительно его направления. Также важным отличием ORB от BRIEF является однократное проведение анализа тестовых пар с определенными свойствами. Пример обнаружения особых точек на изображении методом ORB представлен на рисунке 2.2.3.

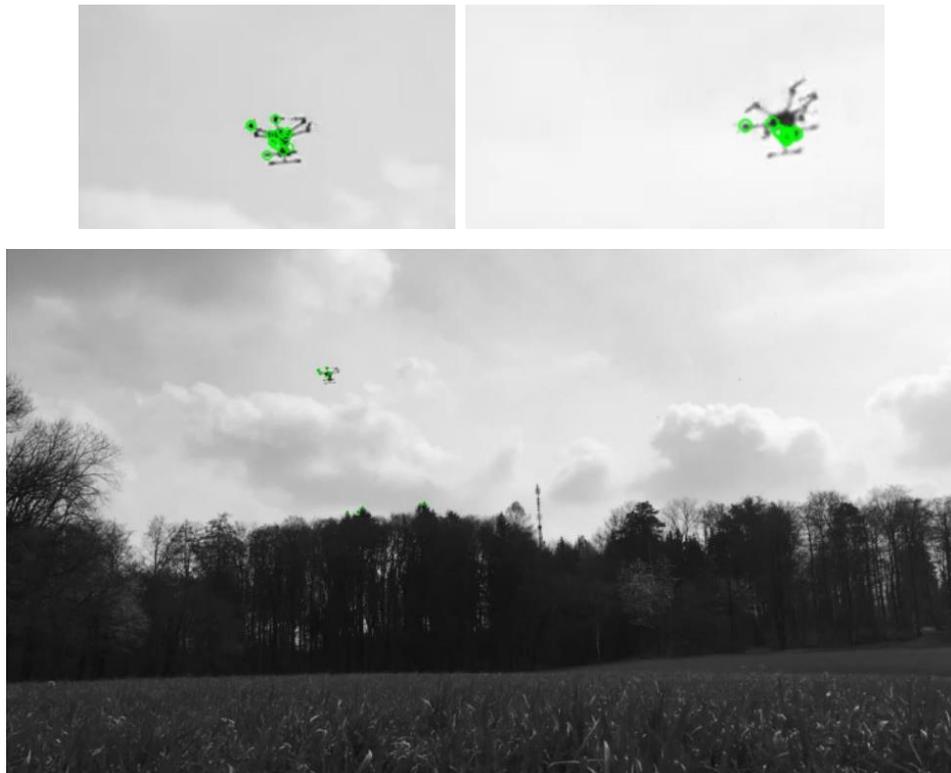


Рис. 2.2.3. Пример обнаружения особых точек на изображении методом ORB

Для сопоставления особых точек используются различные методы, но среди них есть принципиально различные способы: полный перебор и поиск ближайших соседей.

При решении задачи полным перебором сопоставитель берет каждый дескриптор в опрашиваемом наборе и сравнивает его с каждым дескриптором в обучающем наборе. При работе с данным методом задается метрика вычисления расстояний для сравнения.

Быстрый приближенный поиск ближайших соседей для сопоставления дескрипторов предлагает использовать различные варианты индексирования, такие как индексирование с помощью kd-деревьев, иерархического метода К средних, локально-чувствительное хэширование и др. Подробно методы поиска ближайших соседей описаны в [1]. Пример сопоставления особых точек различных изображений с присутствующим на них объектом интереса показан на рисунке 2.2.4.

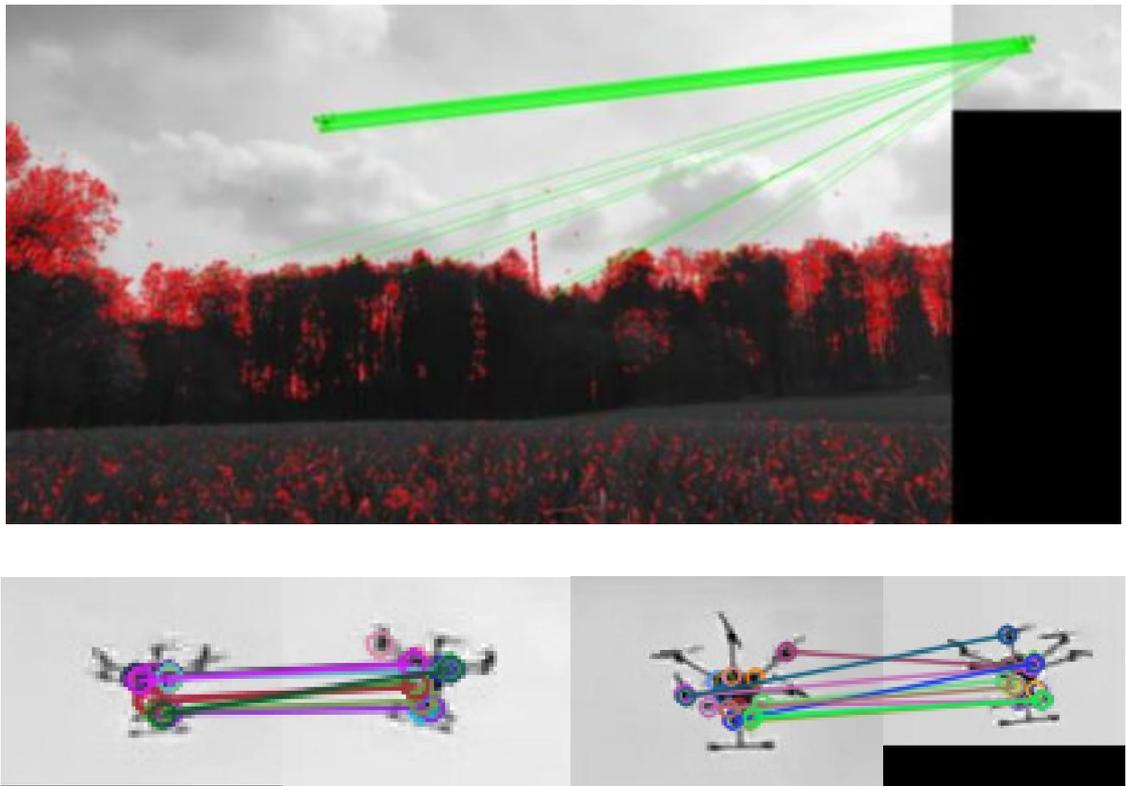


Рис. 2.2.4 Пример сопоставления особых точек

2.3 Нейросетевые технологии

Для работы с изображениями в части решения задач классификации и обнаружения широко применяются сверточные нейронные сети.

Различие между алгоритмами обнаружения объекта и алгоритмами классификации заключается в том, что в алгоритмах обнаружения выполняется построение ограничивающего прямоугольника вокруг интересующего объекта, выделяющего его на изображении. Кроме того, количество объектов заранее неизвестно. В задачах классификации необходимо только сделать заключение о наличии объекта на изображении или отнесении изображения к какому-либо классу.

Две основные концепции, используемые в свёрточных нейронных сетях:

- свёртка;
- операция подвыборки (pooling, max pooling).

Принцип работы свертки заключается в следующем [1, 36]:

- 1) Входное изображение интерпретируется как двумерный массив пикселей;
- 2) Для работы с нейронной сетью выполняется нормализация исходного изображения – все значения пикселей приводятся к интервалу от 0 до 1;
- 3) Суть свертки заключается в создании другого набора значений, который называется ядром (фильтром) меньшего размера (как правило, используется область 3x3 пиксела). Задаются значения всех элементов сверточного ядра и выполняется «сканирование» исходного изображения данным ядром. Сверточный слой применяется к каждому пикселу входного изображения: пиксель и сверточное ядро центрируются, каждый пиксель изображения умножается на соответствующий пиксель ядра и складываются все значения произведений, а результирующее значение пикселя присваивается новому изображению;
- 4) Подобная операция проводится со всеми пикселями исходного изображения, для граничных пикселей изображение достраивается по одному из

множества методов (как правило, выполняется «обнуление» пикселей вне исходного изображения).

Свертка изображения является полноценным слоем сверточной искусственной нейронной сети.

Операция субдискретизации (подвыборки, pooling, max pooling) — это операция уменьшения пространственных размеров изображения путем обработки значений блоков пикселей [1]:

- 1) Определяется размер выборки (размер прямоугольной сетки) и величина шага по изображению. Пример: размер выборки 3x3 и шаг в 3 пикселя;
- 2) Находится максимальное значение пикселя, попадающее в выделенную сетку. Это значение «переносится» в регенерируемое новое изображение. Сетка сдвигается на один шаг и процесс выборки максимального значения и его переноса на новое изображение повторяется;
- 3) По результатам проделанной операции получено изображение меньшего размера по сравнению с оригинальным. Размер итогового изображения варьируется в зависимости от выбора размера сетки и шага.

Процесс обучения сверточной ИНС заключается в оптимизации значений в сверточных слоях.

Главным достоинством нейросетевого метода является возможность распознавания различных объектов на сложном фоне. В то же время при использовании данного метода обнаружения и классификации имеется вероятность возникновения ошибок распознавания первого и второго рода. За скорость вычислений, обучения и точности работы нейронной сети отвечают не только характеристики вычислительной системы и входных данных, но и её архитектура. Как правило, архитектуры нейронных сетей, позволяющие достичь высокой точности, требуют больших вычислительных мощностей технической системы. В то же время модели, позволяющие выполнять решение поставленной задачи с малой затратой ресурсов и за меньшее время обычно обладают меньшей точностью работы и могут обнаружить не все объекты на обрабатываемом изображении. Важным аспектом при обучении сверточной нейронной сети

являются размер и качество обучающего и тестового набора изображений. Как правило, при решении задач нейросетевым методом, используются большие массивы данных. В данной работе предполагается отсутствие наличия такой выборки.

2.3.1 Выбор архитектур

Для решения задачи распознавания объектов на изображении рассматривались архитектуры искусственных нейронных сетей, представленных в «зоопарке моделей» открытой библиотеки Tensorflow2 [86].

В библиотеке представлены одноэтапные и двухэтапные модели сверточных нейронных сетей.

Одноэтапные сверточные нейронные сети одновременно прогнозируют классы объектов и смещения ограничивающих рамок. К достоинствам данного класса нейронных сетей относят быструю скорость обработки изображений, однако, как правило, более низкую точность распознавания по сравнению с двухэтапными моделями.

Двухэтапные сети состоят из двух модулей: первый модуль представляет собой глубокую сверточную ИНС, которая определяет регионы предполагаемых объектов, а второй модуль является детектором, использующим ранее определенные регионы для классификации объектов внутри данных областей.

Стоит отметить, что в «зоопарке моделей» представлены архитектуры нейронных сетей, позволяющие выполнять сегментацию изображений или поиск по ключевым точкам. Также имеется возможность использовать модель различного масштаба. При проведении последующего анализа рассматривались только модели, результатом работы которых является вывод координат ограничивающего обнаруженный объект прямоугольника. Определение ключевых точек и сегментация требуют увеличения объема разметки обучающих данных.

Перечень основных архитектур представлен в таблице 2.3.1. Полный список доступных нейронных сетей открытых библиотек Tensorflow1 и Tensorflow2 представлен в [86, 87].

Таблица 2.3.1 – Основные архитектуры нейронных сетей, представленные в «зоопарке моделей» Tensorflow2

Название модели	Скорость, мс	COCO mAP
CenterNet HourGlass104	70	41.9
CenterNet Resnet50 V1 FPN	27	31.2
CenterNet Resnet50 V2	27	29.5
CenterNet MobileNetV2 FPN	6	23.4
EfficientDet D0... D7	39	33.6
SSD MobileNet v2	19	20.2
SSD MobileNet V1 FPN	48	29.1
SSD MobileNet V2 FPNLite	22	22.2
SSD ResNet50 V1 FPN (RetinaNet50)	87	38.3
SSD ResNet101 V1 FPN (RetinaNet101)	57	35.6
SSD ResNet152 V1 FPN (RetinaNet152)	80	35.4
Faster R-CNN ResNet50 V1	53	29.3
Faster R-CNN ResNet101 V1	55	31.8
Faster R-CNN ResNet152 V1	85	37.6
Faster R-CNN Inception ResNet V2	206	37.7

Для каждой архитектуры в таблице выше указаны следующие параметры:

- Название;
- Скорость модели - время работы в мс с изображением размером 600x600 (включая всю предварительную и постобработку). Стоит отметить, что указанное время сильно зависит от конкретной конфигурации оборудования и во многих случаях его следует рассматривать скорее, как относительное.
- Точность работы детектора на подмножестве набора проверки COCO, измеренном с помощью показателя mAP, зависящего от набора данных и

используемого оборудования. Указывается только mAP ограничивающего прямоугольника, округленного до ближайшего целого числа.

Распределение представленных в «зоопарке моделей» архитектур в зависимости от скорости и точности с учетом и без учета масштабируемости моделей показано на рисунках 2.3.1 и 2.3.2.

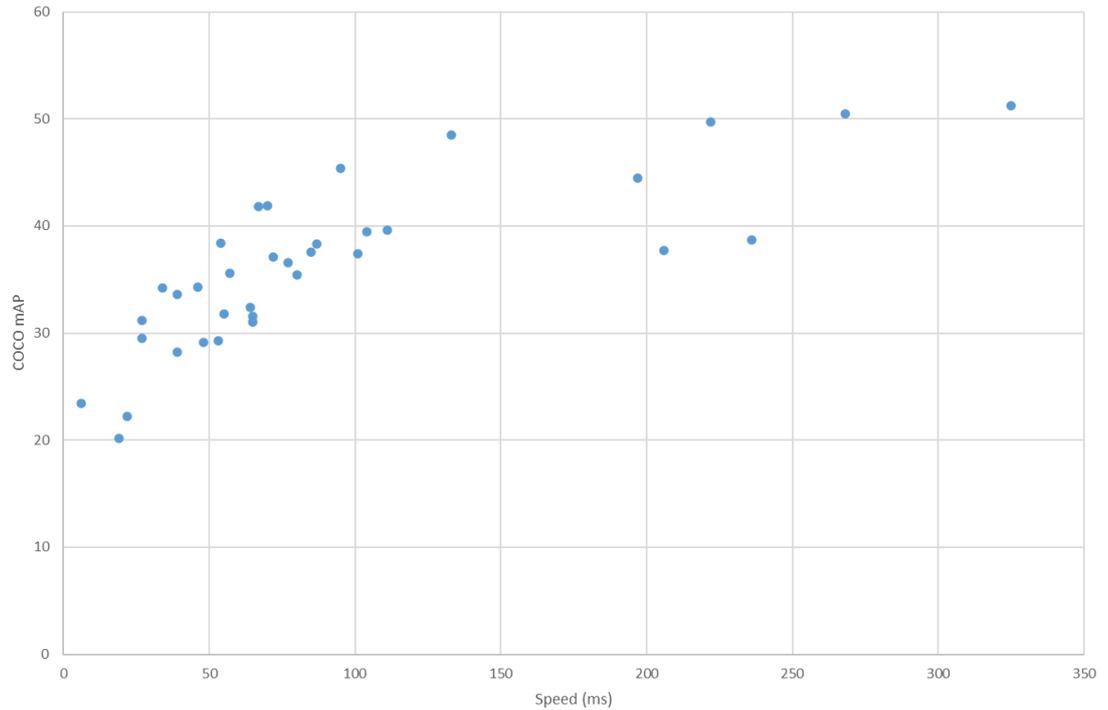


Рисунок 2.3.1 – Распределение архитектур по скорости и точности

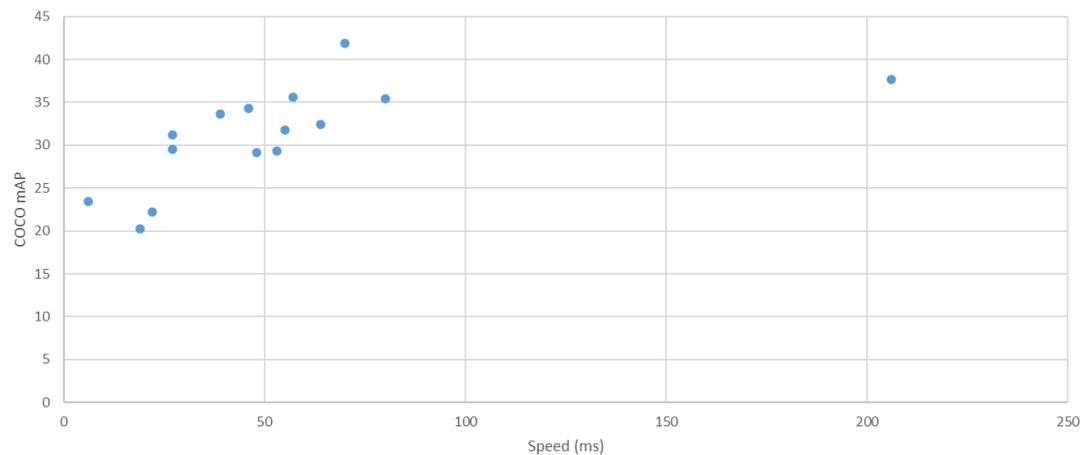


Рисунок 2.3.2 – Распределение архитектур по скорости и точности без учета масштабируемости моделей

Из распределения видно, что часть архитектур выделяется низкой точностью работы и высокой скоростью работы. Также отдельные архитектуры отличаются высокой точностью работы при крайне низкой скорости обработки.

Стоит отметить, что поставленную задачу предполагается решать в реальном масштабе времени, что накладывает временные ограничения на скорость работы нейронной сети. На основе соотношения точности и скорости работы, а также для проведения сравнения эффективности применения одноэтапных и двухэтапных сверточных нейронных сетей, в последующих компьютерных экспериментах использовались следующие модели:

- EfficientDet: семейство одноэтапных моделей, использующих метод сложного масштабирования для оптимизации размера модели и вычислительных затрат при сохранении высокой производительности. Подробное описание принципа работы и строения семейства сверточных нейронных сетей EfficientDet, а также сравнение результатов работы предложенных архитектур с другими архитектурами проведено в работе [88]. Данная архитектура уже активно используется при разработке алгоритмов для решения различных прикладных задач [58, 89], рассматривается возможность её применения в комбинации с другими методами [67], в том числе в задачах реального времени [71].

- Centernet_resnet50_fpn: нейронная сеть, подход работы которой заключается в определении центральной точки объектов и их размеров для повышения эффективности проводимых вычислений. Подробное описание принципа работы нейронной сети CenterNet, её отличия от других архитектур, используемых для распознавания объектов на изображениях представлено в [90], а пример её комбинирования с другими методами в [73]. Применение данной модели для решения задач с ограниченным временем обработки показано в работах [3, 5]. Для улучшения возможности обнаружения объектов различного масштаба используется архитектура FPN, предназначенная для улучшения обнаружения объектов путем решения проблем изменения масштаба изображений. Он создает пирамиду функций с семантикой высокого уровня, позволяющую обнаруживать объекты в разных масштабах

- Faster_rcnn_inception_v2: широко используемая, универсальная модель обнаружения объектов. Она сочетает в себе сеть поиска областей для формирования потенциальных предложений по ограничивающим рамкам объектов

и сверточную нейронную сеть для классификации и уточнения этих предложений. Faster R-CNN обеспечивает высокую точность, но может потребовать больше вычислительных ресурсов по сравнению с некоторыми другими моделями. Inception - архитектура, известная использованием начальных модулей, которые эффективно фиксируют многомасштабные функции. Она способна достигать высокой точности при решении задач классификации изображений.

– SSD_mobilenet: данный детектор использует одну нейронную сеть для одновременной классификации объектов и смещений нескольких ограничивающих рамок различных размеров и соотношений сторон, выделяется своей скоростью и эффективностью, сохраняя при этом конкурентоспособную точность. Используется в том числе для решения задач, в которых скорость имеет решающее значение для эффективной обработки данных на устройствах с малой вычислительной мощностью. Решение различных задач с применением архитектуры SSD и сравнение с другими алгоритмами проводится в работах [42, 91]. MobileNet предназначен для эффективной обработки данных на устройстве, что делает его пригодным для мобильных и периферийных приложений. Он использует отделимые по глубине свертки для сокращения вычислений при сохранении точности.

2.4 Результаты

В данной главе было проведено сравнение алгоритма Хюккеля и совместного применения алгоритмов Кэнни и Хафа, предназначенных для обнаружения контуров объектов на изображении и выделения из них прямых линий.

На основе проведенных численных экспериментов определено, что алгоритм Хюккеля применим для задач, требующих получения общего описания контуров объекта интереса на изображении большого разрешения за малый промежуток времени в условиях, когда не требуется большая точность их описания. Также

искомый контур должен обладать четкими линейными признаками, размер которых сопоставим с размером области обработки алгоритма.

При обработке изображений малого разрешения или обнаружения объекта сложной формы, а также при необходимости точного описания контура объекта предпочтительно совместное использование алгоритмов Кэнни и Хафа.

Полученные результаты по числу обрабатываемых кадров в секунду (17...22 изображений в секунду при размере входных изображений 289 на 180 пикселей) для рассматриваемых алгоритмов показывают возможность их использования при решении задач реального времени в качестве элементов алгоритмов, сочетающих в себе применение различных подходов компьютерного зрения в части распознавания объектов на изображениях.

Настраиваемые параметры обработки для обоих алгоритмов получают исходя из зашумленности и перепадов яркости, присутствующих на изображении.

В результате обзора доступных в библиотеке компьютерного зрения OpenCV детекторов особых точек на изображении для дальнейшей работы был выбран алгоритм ORB, разработанный для решения задач обнаружения особых точек и их дескрипторов в задачах реального времени. Инвариантность относительно вращения, пониженная чувствительность к зашумленности изображения позволяют использовать его для решения широкого круга задач.

По результатам проведенного анализа архитектур сверточных нейронных сетей из библиотеки открытого доступа «зоопарк моделей» Tensorflow для дальнейшей работы предпочтение было отдано следующим моделям:

- Efficientdet_d0;
- Centernet_resnet50_v1_fpn;
- SSD_mobilenet_v1, SSD_mobilenet_v2;
- Faster_rcnn_inception_v2.

Практически все рассматриваемые архитектуры нейронных сетей – одноэтапные. Однако, для проведения сравнения эффективности применения моделей при решении конкретной задачи, были проведены эксперименты и с двухэтапной моделью.

Реализуемые в последующих главах гибридные алгоритмы обнаружения и классификации объектов на изображениях будут основаны на рассмотренных в данной главе традиционных и нейросетевых методах компьютерного зрения.

3. Описание предложенных алгоритмов

В данной главе описана структура исследуемых гибридных алгоритмов, представляющих из себя совместное использование методов традиционного компьютерного зрения и сверточных нейронных сетей. В рамках проведенной работы были предложены и реализованы три структуры алгоритмов, представленных на рисунке 3.1.

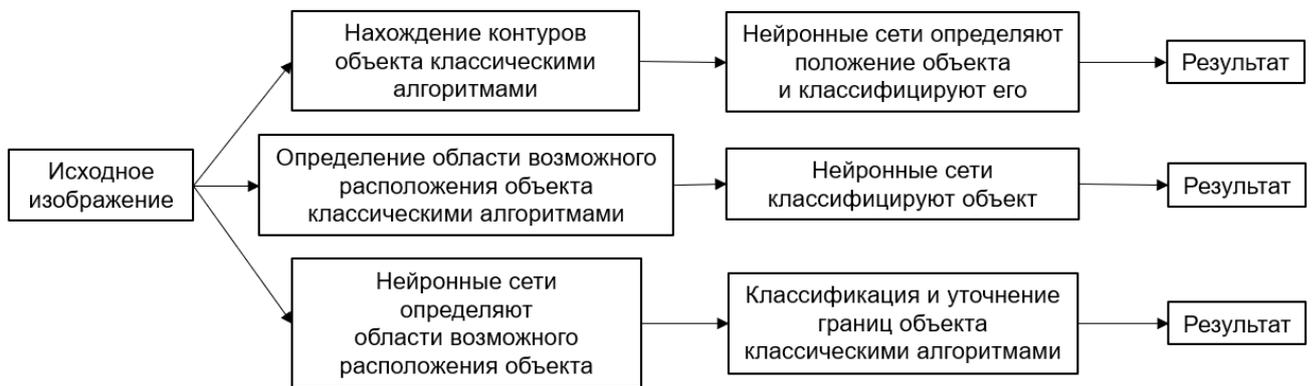


Рисунок 3.1 - Предложенные структуры гибридных алгоритмов

Как было отмечено в предыдущих главах, у каждого метода есть свои достоинства: возможность обнаружения объектов под разными ракурсами нейросетевыми методами, в том числе на сложном фоне; высокая скорость обработки изображений детекторами особых точек; предварительное обнаружение контуров объектов на нескольких изображениях в секунду и т.д. Предполагается, что в описанных далее алгоритмах, при правильном их использовании в части задания настраиваемых параметров и качестве обучения при малом наборе исходных данных, удастся достичь лучших результатов в части распознавания объектов на изображении по сравнению с применением каждого метода по отдельности.

3.1 Гибридный алгоритм, основанный на предварительном обнаружении контуров объектов

В данном алгоритме методы поиска контуров и определения прямых линий использовались для оценки влияния предварительного перехода к контурному представлению объектов на точность работы нейросетевых методов.

На вход алгоритма поступает изображение, содержащее в себе не менее одного объекта интереса. Формально решаемая данным методом задача состоит из следующих шагов:

- Формирование наборов изображений для обучения и тестирования сверточных нейронных сетей (используемый в работе набор изображений собран из открытых источников и размечен авторами самостоятельно с использованием программы LabelImg [92]);

- Обработка набора изображений алгоритмами обнаружения контуров, подбор настраиваемых параметров;

- Обучение сверточных нейронных сетей различных архитектур на необработанных изображениях;

- Обучение сверточных нейронных сетей различных архитектур на предварительно обработанных алгоритмами поиска контуров изображениях;

- Распознавание объектов на тестовых обработанных и необработанных изображениях, которые нейронная сеть видит впервые;

- Сравнение полученных результатов.

В качестве характерного примера различий получаемых результатов при использовании различных моделей нейронных сетей, в данной главе показаны результаты использования двух архитектур: `Faster_rcnn_inception` и `SSD_mobilenet`.

На рисунке 3.1.1 показан результат применения двух сверточных сетей различной архитектуры к одному и тому же входному изображению. Обучение проводилось на одинаковых наборах данных и оборудовании до начала переобучения моделей. Двухэтапная нейронная сеть архитектуры `Faster_rcnn_inception_v2`, сочетающая в себе сеть предложений регионов для

генерации потенциальных предложений по ограничивающим рамкам объектов и сверточную нейронную сеть для классификации и уточнения этих предложений, показывает более высокую точность распознавания объекта интереса, чем одноэтапная нейронная сеть SSD_mobilenet, использующая одну нейронную сеть для одновременного прогнозирования классов объектов и смещений для нескольких ограничивающих рамок по умолчанию различных размеров и соотношений сторон. Однако применение нейронной сети второго типа архитектуры позволило получить результат за приблизительно в восемь раз меньший промежуток времени, так как она предназначена для обработки данных на устройствах с малой вычислительной мощностью. Стоит отметить, что в настоящее время ведутся работы по разработке систем обучения одноступенчатых моделей с подходом имитации двухэтапных функций, целью чего является преодоление разрыва в точности между данными типами моделей [93].

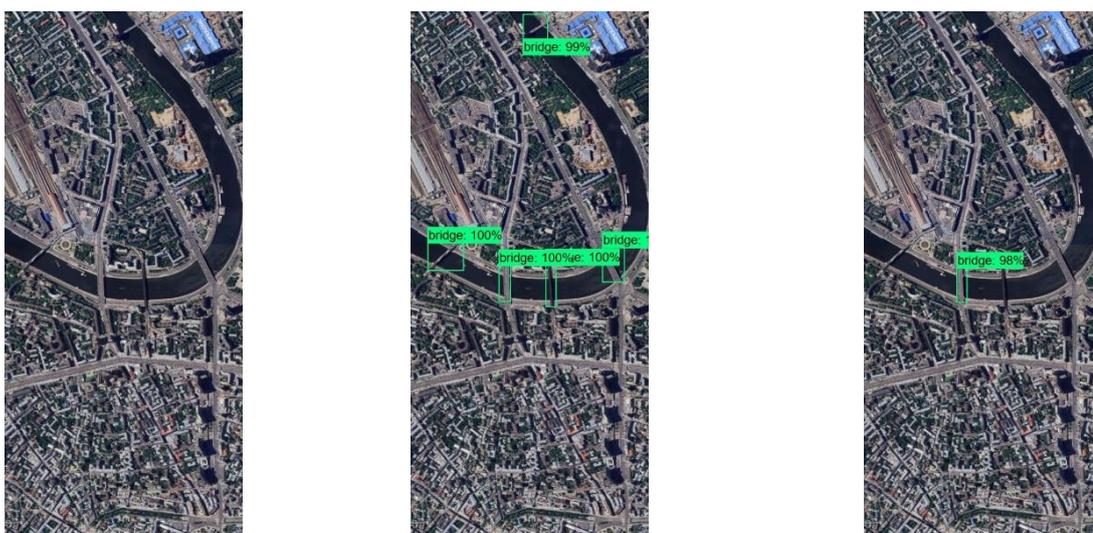


Рис. 3.1.1 Результаты обработки исходного изображения (слева) нейронными сетями Faster_rcnn_inception (по центру) и SSD_mobilenet (справа)

В рамках проведенных исследований предполагается, что для определенного набора объектов интереса предварительная обработка изображения алгоритмом поиска контуров может увеличить вероятность их дальнейшего успешного распознавания. Как показано на рисунке 3.1.2, объект интереса типа «мост» четко выделяется на фоне при переходе к контурному представлению изображения.

Помимо существующих алгоритмов распознавания прямых линий на изображении, ведутся исследования по разработке алгоритмов, позволяющих обнаруживать прямые линии контура объектов в реальном времени [94].

Рассматриваемые алгоритмы определения контуров объектов и выделения прямых линий из них на изображениях:

- алгоритм Хафа с предварительной обработкой алгоритмом Кэнни;
- алгоритм Хюккеля.

Достоинства и недостатки данных алгоритмов определения границ объектов на изображениях, а также перечень настраиваемых параметров представлены в главе 2. Значения настраиваемых параметров обработки получены исходя из зашумленности и перепадов яркости, присутствующих на изображении.



Рис. 3.1.2 Результаты обработки исходного изображения с объектом интереса типа «мост» (сверху) алгоритмами поиска контуров Хюккеля (по центру) и Хафа (снизу)

В ходе решения тестовых задач на отдельных изображениях определено, что переход к контурному изображению позволяет выполнять успешное распознавание объекта интереса, как показано на рисунке 3.1.3.

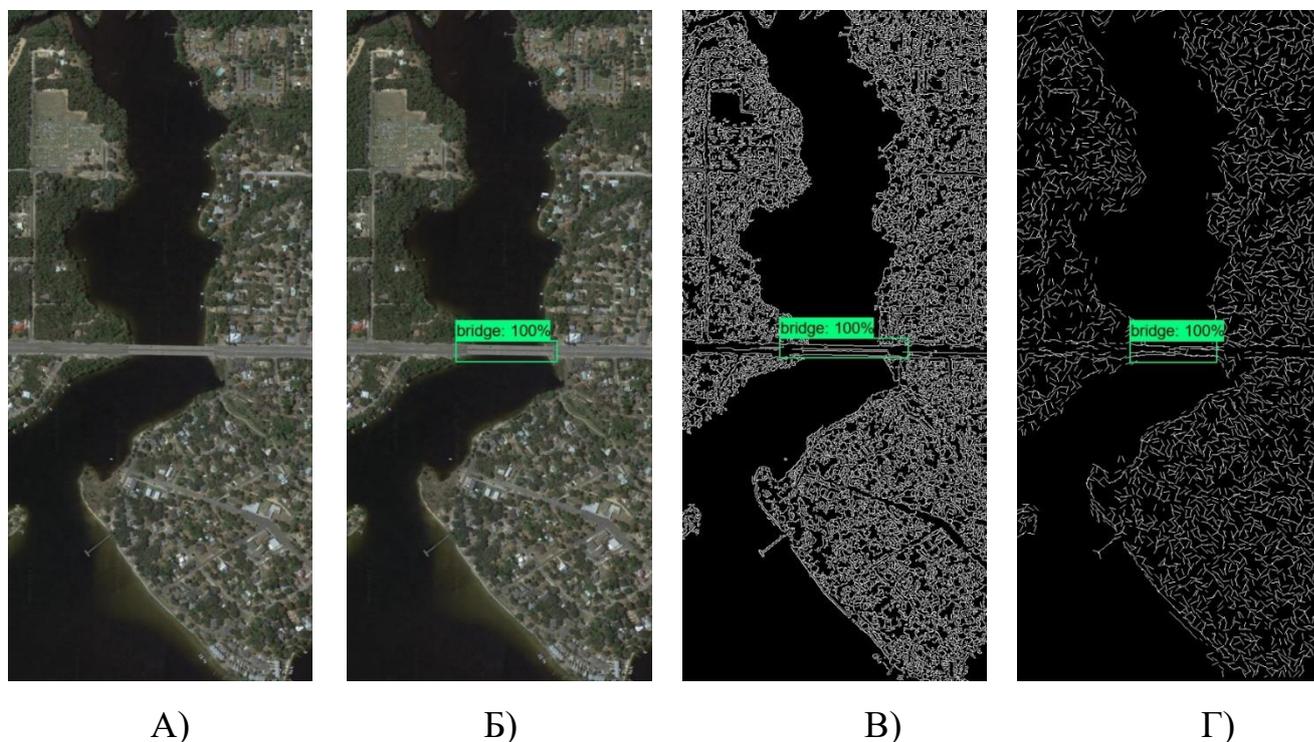


Рис. 3.1.3 – Распознавание объекта на изображении: А) – исходное изображение; Б) – результат распознавания без обработки; В) – результат распознавания на изображении, обработанном алгоритмом Хафа; Г) - результат распознавания на изображении, обработанном алгоритмом Хюккеля

Результаты численных экспериментов на массиве данных представлены в главе 4.

3.2 Предварительная обработка нейросетевым методом с поиском особых точек в фрагменте

Работа второго из рассматриваемых гибридных алгоритмов заключается в предварительном определении возможных областей расположения объекта интереса на изображении с последующей оценкой правильности распознавания путем сравнения с набором эталонных фрагментов традиционным методом компьютерного зрения сопоставления особых точек.

Общая схема предложенного алгоритма обнаружения объектов представлена на рисунке 3.2.1 и состоит в следующем:

- на вход алгоритма поступает изображение, на котором может присутствовать как объект интереса, так и объекты, не являющиеся целью поиска;
- выполняется поиск объектов интереса с использованием технологии сверточных нейронных сетей;
- выделяются области изображения (фрагменты), предположительно содержащие интересующий объект;
- на каждом обнаруженном фрагменте проводится поиск особых точек;
- выполняется сопоставление найденных особых точек с примерами «эталонных» наборов таких точек для объекта интереса, отобранных предварительно;
- на основе проведенного сравнения подтверждается или отвергается предположение о наличии объекта интереса в данной области;
- вывод полученных результатов.



Рис. 3.2.1. Схема работы предложенного гибридного алгоритма

Как было отмечено ранее, для решения задачи распознавания объектов на изображении рассматривались архитектуры искусственных нейронных сетей, представленных в «зоопарке моделей» открытой библиотеки Tensorflow2 [86].

При условии, что поставленную задачу предполагается решать в масштабе реального времени, накладываются временные ограничения на скорость работы сверточной нейронной сети и реализованного алгоритма в целом. При реализации данного варианта гибридного алгоритма применялись модели следующих архитектур:

- EfficientDet.
- Centernet_resnet50_fpn.
- SSD_mobilenet.

В качестве «традиционной» составляющей гибридного алгоритма при его реализации использовался детектор особых точек ORB, особенности работы которого были описаны в главе 2.

Для тестирования рассматриваемого гибридного алгоритма набор данных был взят из открытого источника [95] и размечен с использованием инструментария [92]. Этот набор данных содержит три видеофрагмента, содержащих малый антропогенный движущийся объект типа «БЛА» на естественном фоне. Часть обрабатываемых изображений содержат также объекты, изменяющие свое положение, но не представляющие интереса.

Для всех трех рассматриваемых в данном варианте гибридного алгоритма архитектур сверточных нейронных сетей предварительно выполнялось обучение и оценка точности работы. Пример обнаружения объекта интереса обученной нейронной сетью на сложном фоне представлен на рисунке 3.2.2. Обучение проводилось на одинаковом наборе данных: все обучающие и тестовые изображения брались из одного видеофрагмента, число обучающих изображений – 2063, тестовых – 365. Обучение проводилось с использованием видеоадаптера NVIDIA GeForce GTX 1650 SUPER.



Рис. 3.2.2 – Пример обнаружения объекта нейронной сетью на сложном фоне

Недостатком нейросетевого метода является возможное возникновение ложных распознаваний, что показано на рисунке 3.2.3. Для снижения числа ошибок

в данном исследовании выполнялась последующая проверка правильности классификации в областях возможного расположения объектов методом поиска и сопоставления особых точек.



Рис. 3.2.3 – Примеры корректного (справа) и некорректного (слева сверху) обнаружения объекта интереса нейронной сетью

Предварительно для оценки возможности обнаружения и распознавания объекта интереса на исходных изображениях к ним был применен детектор ORB. Пример полученных результатов, показывающий, что обнаруженные точки присутствуют на объекте интереса, представлен на рисунке 3.2.4. Найденные особые точки выделены круговыми маркерами.



Рис. 3.2.4 – Пример работы детектора особых точек ORB на изображении, содержащем объект интереса

Из первого видеофрагмента было отобрано пять изображений, содержащих объект интереса в разных ракурсах. Данные изображения были использованы как эталонные образцы для сравнения особых точек с фрагментами возможного расположения объекта интереса, предварительно обнаруженными нейронными сетями. Был определен диапазон количества особых точек, достаточного для сопоставления двух объектов на фрагментах. Данные изображения с отмеченными зеленым особыми точками представлены на рисунке 3.2.5.

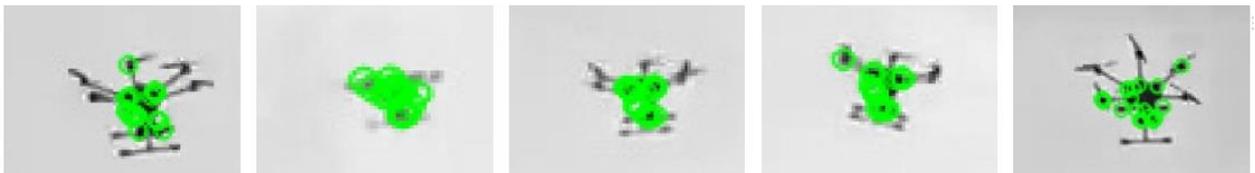


Рис. 3.2.5 – Эталонные изображения с выделенными особыми точками

3.3 Предварительный поиск области методом особых точек с проверкой нейросетевым методом

Третий, последний, из реализуемых в данной работе гибридных алгоритмов, основывается на поиске областей возможного расположения объекта интереса на изображении с последующей проверкой правильности распознавания нейросетевым методом. Схема работы данного гибридного алгоритма представлена на рисунке 3.3.1.



Рис. 3.3.1 – Схема работы предложенного гибридного алгоритма

Было рассмотрено два варианта реализации данного алгоритма. Общим в них является использование 5 фрагментов из первого видео, содержащих объект интереса в качестве «эталонов» (рис. 3.2.5). Использовались те же архитектуры нейронных сетей и детектор особых точек, что и в главе 3.2.

Формальное описание первой реализации алгоритма:

1. Исходное изображение делится на большое количество областей размера, сопоставимого с предполагаемыми размерами объекта интереса;
2. В каждом фрагменте выполняется поиск особых точек. При нахождении соответствий больше заданного порога обрабатываемый фрагмент передается на вход сверточной нейронной сети. Число искомых точек для сопоставления было принято равным 40, необходимое количество совпадений варьировалось от 15 до 25;
3. Нейронная сеть проводит проверку наличия объекта интереса в каждом переданном ей фрагменте. Пороговое значение вероятности наличия объекта на фрагменте при тестовом решении задачи принимался равным 0.3;
4. При распознавании объекта детектором особых точек и подтверждении наличия объекта интереса сверточной нейронной сетью обрабатываемый фрагмент выделяется зеленым цветом.

Решение тестовых задач при реализации данного варианта алгоритма показали время обработки менее 3 кадров/сек при низкой точности работы алгоритма, по этой причине дальнейшие вычислительные эксперименты не проводились. Примеры полученных результатов показаны на рисунке 3.3.2.



Рис. 3.3.2 – Пример полученных результатов первого варианта применения описанного алгоритма – большое число ложных срабатываний, низкая точность выделения размеров объекта интереса

Работа второго рассматриваемого гибридного алгоритма данного типа:

1. На изображении выполняется поиск особых точек, строится область, ограничивающая возможное расположение объекта интереса. При этом учитываются только точки, имеющие не менее заданного числа «соседей» на

определенном расстоянии (по условию задачи объект занимает небольшую часть исходного изображения, на основе этого выбирается пороговое значение максимального расстояния между найденными особыми точками, тем самым уменьшив возможную область расположения объекта);

2. Полученный фрагмент сравнивается с набором эталонов (рис. 3.2.5), при числе совпадений больше заданного делается вывод о необходимости проверки области сверточной нейронной сетью;

3. Выполняется проверка наличия объекта интереса в заданной области нейросетевым методом.

Варьируемые параметры при данной реализации алгоритма:

- количество искомым точек на исходном изображении;
- минимальное количество особых точек, расположенных на заданном расстоянии друг от друга;
- количество минимального числа сопоставленных с эталоном точек;
- пороговое значение вероятности наличия объекта интереса при использовании нейросетевого метода;
- настраиваемые параметры нейронной сети.

3.4 Результаты

В данной главе представлено описание структуры трех реализованных гибридных алгоритмов, принцип работы которых основан на применении классических и нейросетевых методов компьютерного зрения, описанных в главе 2. Выполнены тестовые обработки используемых в дальнейшем изображений, показывающие работоспособность реализованных алгоритмов.

4. Результаты проведенных численных экспериментов

В данной главе представлены результаты применения описанных в главе 3 гибридных алгоритмов обнаружения и классификации объектов интереса на изображениях при решении конкретных задач. По результатам проведенных вычислительных экспериментов проведена оценка точностной и временной эффективности реализованных алгоритмов. Особенностью решаемых задач является требование к возможности выполнения вычислений в реальном времени, ограниченность вычислительных ресурсов, относительно малые обучающая и тестовая выборки при обучении нейронных сетей.

4.1 Влияние предварительного обнаружения контуров

В качестве объекта интереса при оценке эффективности предварительной обработки исходного изображения алгоритмом выделения прямолинейных контуров с последующим поиском объекта интереса на всем изображении рассматривался объект типа «мост», образ которого содержит прямолинейные признаки, что отличает его от большого количества объектов, описываемых точечными признаками, но при этом характерно для объектов искусственного происхождения с простой геометрической формой.

Была проведена серия расчетов для двух типов искусственных сверточных нейронных сетей: на первом этапе выполнялась разметка данных и классификация объектов на изображениях без дополнительной обработки, на втором этапе процесс повторялся для заранее обработанных изображений алгоритмами Хафа и Хюккеля.

Поиск объектов сверточной искусственной нейронной сетью проводился на всем полученном изображении.

При обработке большого набора изображений выявлены недостатки предложенного метода: возможно возникновение дополнительных ложных распознаваний при работе с предварительно обработанными изображениями.

Однако, при правильном подборе настраиваемых параметров и отсутствии на изображении объектов, похожих на объект интереса (в рассматриваемой задаче дорога может быть подобна мосту), задача решается корректно. Также стоит отметить, что ложное распознавание дороги как объекта интереса присутствует и при отсутствии предварительной обработки изображений алгоритмами выделения контуров.

Пример ложного распознавания при поиске объекта на необработанном изображении, а также примеры корректных и ложных распознаваний представлены на рисунке 4.1.1.

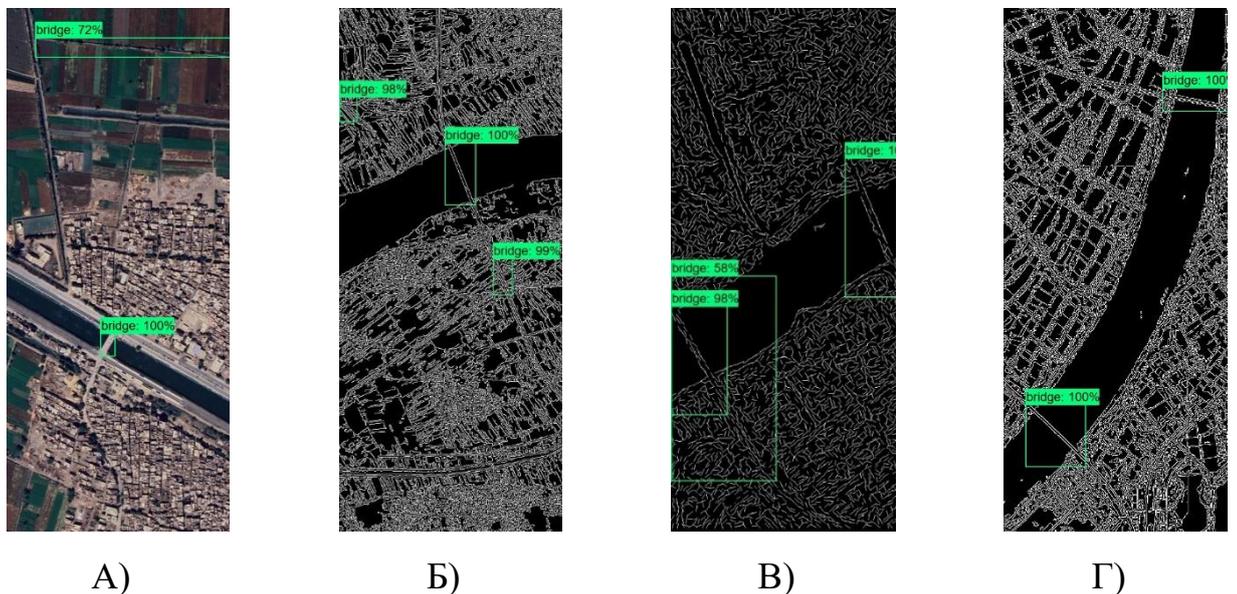


Рис. 4.1.1 – Примеры ложного распознавания на исходном изображении (А), на обработанном изображении (Б), успешного распознавания на предварительно обработанных изображениях (В и Г)

Полученные временная и точностная оценки предварительной обработки изображений алгоритмами поиска контуров представлены в таблице 4.1.1. Вычисления проводились на процессоре Intel(R) Core(TM) i5-10400F CPU @ 2.90GHz, размер поступающих на вход изображений 450 на 1000 пикселей. Обучающая выборка сверточных нейронных сетей содержала 200 размеченных изображений, тестовая – 50. Оценка времени и точности обработки проводилась на ранее используемых, а также 35 новых изображениях. Каждое изображение содержало не менее одного объекта интереса. Пороговое значение вероятности,

при котором выполнялся вывод области распознавания объекта, было принято равным 50%. Разброс по времени и точности работы с предварительно обработанными изображениями связан с влиянием настраиваемых параметров.

Таблица 4.1.1 – Оценка влияния предварительной обработки на время и точность распознавания

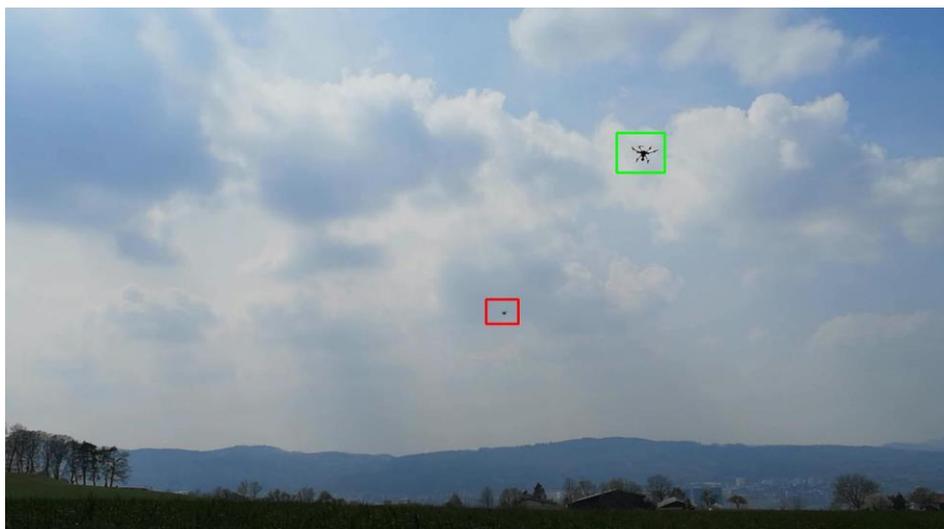
	faster_rcnn_inception		ssd_mobilenet	
	Исходные изображения	Обработанные изображения	Исходные изображения	Обработанные изображения
Время предварительной обработки, с	-	0,14...0,56	-	0,14...0,56
Время обработки одного изображения сетью, с	0,36	0,35...0,37	0,045	0,042...0,044
Время обнаружения объектов интереса на изображении, с	0,36	0,51...1,05	0,045	0,144...0,602
% успешного обнаружения	92	75...91	79	61...74
% ложного обнаружения от всех найденных	15	15...18	1,5	5...7

Полученные результаты показывают, что использование одноэтапных сверточных нейронных сетей позволяет достичь времени обработки изображений, удовлетворяющего требованиям систем реального времени. Точность сетей данного класса при обработке всего входного изображения уступает модели двухэтапной архитектуры. Результаты, представленные в таблице 4.1.1, показывают, что за время обучения и обработки отвечает тип сети, а не наличие предварительной обработки изображений. Применение искусственных нейронных сетей ко всему контурному изображению не дает выигрыша в точности или времени обнаружения и классификации объектов.

4.2 Влияние предварительной обработки нейросетью с поиском особых точек в фрагменте

В рамках проведенной работы была выполнена серия вычислительных экспериментов на тестовых видеопоследовательностях с размером изображений 1920x1080 пикселей применительно к распознаванию объекта интереса размерами от 0,02% до 0,4% от площади кадра. Образ объекта интереса типа «беспилотный летательный аппарат» на видеопоследовательности изменяется в результате его смещения и поворота. На части кадров обрабатываемых видеопоследовательностей объект интереса отсутствовал, перекрывался фоном или присутствовали другие объекты.

Каждый участок изображения, на котором с вероятностью более 30% нейронная сеть обнаруживала объект интереса, после обработки детектором особых точек ORB, подавался на сопоставление с пятью изображениями – «эталонами» (рисунок 3.2.5). Количество искомых точек на обоих изображениях было принято равным 40. При наличии совпадений с одним из фрагментов менее заданного числа особых точек (варьировались значения от 15 до 25) принимается решение о ложном распознавании и область выделяется красным прямоугольником, при большем количестве подтверждается наличие объекта интереса и фрагмент выделяется зеленым. Пример изображений с успешным применением предложенного гибридного алгоритма представлен на рисунке 4.2.1.



а)



б)

Рис. 4.2.1. – Результаты работы гибридного алгоритма: а) определено одно ложное распознавание, б) определены два ложных распознавания

В таблице 4.2.1 представлены полученные результаты по влиянию порогового значения числа соответствующих эталонным фрагментам особых точек для трех видеофрагментов при обработке изображений различными архитектурами сверточных нейронных сетей. Вычисления проводились на процессоре Intel(R) Core(TM) i5-10400F CPU @ 2.90GHz. Из полученных результатов видно, что увеличение порогового значения приводит к уменьшению числа ложных распознаваний, и увеличению количества областей изображений, верно обработанных нейросетевым методом, но ошибочно отброшенных детектором

особых точек. В зависимости от поставленной задачи, типа объекта и сложности фона необходимо предварительно определять необходимое значение порога.

В данной таблице приняты следующие обозначения:

* – процентное соотношение относительно количества кадров с объектом интереса;

*1 – процентное уменьшение количества ложных срабатываний детектора;

Кадров без объекта – количество кадров видеофрагмента, не содержащих объект интереса.

Таблица 4.2.1 – Влияние числа соответствующих точек на результат обработки изображений

Нейросетевая архитектура	№ видео, порог по особым точкам	Ложные обнаружения		Не обнаружено объектов		Кадров без объекта
		Было	Стало	Было	Стало	
EfficientDet	№1, 15 точек	20	1; 95% ^{*1}	181; 7%*	185; 7,2%*	1825
	№2, 25 точек	100	14; 86% ^{*1}	125; 3,6%*	133; 3,8%*	1252
	№3, 25 точек	65	6; 90,7% ^{*1}	174; 4,1%*	1080; 25,4%*	75
	№3, 15 точек	65	33; 49,2% ^{*1}	174; 4,1%*	302; 7,1%*	75
	№3, 20 точек	65	20; 69,2% ^{*1}	174; 4,1%*	543; 12,8%*	75
Centernet_resnet50_fpn	№1, 15 точек	19	0; 100% ^{*1}	5; 0,2%*	19; 0,7%*	1825
	№2, 15 точек	92	37; 40,2% ^{*1}	12; 0,3%*	14; 0,4%*	1252
	№2, 20 точек	92	21; 59,8% ^{*1}	12; 0,3%*	14; 0,4%*	1252
	№3, 18 точек	20	8; 60% ^{*1}	1004; 23,6%*	1044; 24,5%*	75
SSD_mobilenet	№1, 18 точек	25	3; 88% ^{*1}	43; 1,7%*	62; 2,4%*	1825
	№2, 18 точек	116	30; 74,1% ^{*1}	15; 0,4%*	31; 0,9%*	1252
	№3, 18 точек	1466	706; 51,8% ^{*1}	248; 5,8%*	451; 10,6%*	75

С помощью предложенного алгоритма число ложных обнаружений объекта было снижено на 40 – 100% по сравнению с чисто нейросетевым подходом для различных видеопоследовательностей и значений настраиваемых параметров

Для каждого видеофрагмента для различных значений требуемого сопоставления количества особых точек была проведена оценка временных затрат на выполнение отдельных частей программы. При выполнении данной оценки не учитывались вспомогательные операции: преобразование форматов изображений, обработка признаков наличия объектов, настройка вывода изображений, вычисление параметров изображения. Полученные результаты представлены в таблице 4.2.2.

Таблица 4.2.2 – Временная оценка предложенного алгоритма

Нейросетевая архитектура	№ видео	Время, с		% от общего	На кадр	Кадров/с
EfficientDet	1, 4377 кадров	Общее	1565.3	100	0.36	2.8
		Поиск особых точек	8.3	0.53	0.0019	527
		Нейро	502	32	0.115	8.72
		Сохранение файла	879.4	56.18	0.2	4.98
	2, 4749 кадров	Общее	1772.8	100	0.37	2.68
		Поиск особых точек	12.7	0.72	0.0027	373.94
		Нейро	547.4	30.88	0.12	8.68
		Сохранение файла	992.7	56	0.21	4.78
	3, 4332 кадров	Общее	1613.9	100	0.37	2.68
		Поиск особых точек	12.9	0.80	0.0030	335.81
		Нейро	497.2	30.81	0.11	8.71
		Сохранение файла	898.8	55.69	0.21	4.82
Centernet_resnet50_fpn	1, 4377 кадров	Общее	1456.9	100	0.33	3.00
		Поиск особых точек	8.9	0.61	0.0020	491.80
		Нейро	381.8	26.21	0.09	11.46
		Сохранение файла	886.9	60.88	0.20	4.94
	2, 4749 кадров	Общее	1581.5	100	0.33	3.00
		Поиск особых точек	13.1	0.83	0.0028	362.52
		Нейро	414.5	26.21	0.09	11.46
		Сохранение файла	960.7	60.75	0.20	4.94
	3, 4332 кадров	Общее	1480	100	0.34	2.93
		Поиск особых точек	10.1	0.68	0.0023	428.91
		Нейро	374.4	25.30	0.09	11.57
		Сохранение файла	899.9	60.80	0.21	4.81

Продолжение таблицы 4.2.2

Нейросетевая архитектура	№ видео	Время, с		% от общего	На кадр	Кадров/с
SSD_mobilenet	1, 4377 кадров	Общее	1505	100	0.34	2.91
		Поиск особых точек	8.8	0.58	0.0020	497.39
		Нейро	426.7	28.35	0.10	10.26
		Сохранение файла	890.6	59.18	0.20	4.91
	2, 4749 кадров	Общее	1646.5	100	0.35	2.88
		Поиск особых точек	12.7	0.77	0.0027	373.94
		Нейро	403.2	24.49	0.08	11.78
		Сохранение файла	1008.4	61.25	0.21	4.71
	3, 4332 кадров	Общее	1521.7	100	0.35	2.85
		Поиск особых точек	16.2	1.06	0.0037	267.41
		Нейро	383	25.17	0.09	11.31
		Сохранение файла	921.5	60.56	0.21	4.70

Полученные результаты показывают, что по сравнению с нейросетевым методом введение в гибридный алгоритм работы детектора особых точек не оказывает значительного влияния на время обработки изображений, а сохранение файла и дополнительная обработка занимают значительный промежуток общего затраченного времени.

Скорость обработки с учетом загрузки кадра, обработки видео, сохранения полученного изображения и преобразований в одноканальный тип составляет приблизительно 2,6...3,0 кадра/с.

Скорость проводимых нейронной сетью вычислений составляет приблизительно 8,68...11,78 кадров/с.

Скорость обработки детектором особых точек и сопоставления с пятью фрагментами составляет приблизительно 267,41...527 кадров/с.

Выполнение обнаружения и классификации объекта интереса на изображении, а также проверки верности распознавания без сохранения файла (непосредственно расчетов) возможно со скоростью до 7,7 кадров/с.

4.3 Влияние предварительного поиска области методом особых точек с проверкой нейросетевым методом

При реализации гибридного алгоритма, основанного на предварительном поиске области возможного расположения объекта интереса с последующей проверкой правильности обнаружения нейросетевым методом, использовались свёрточная нейронная сеть архитектуры Efficientdet и детектор особых точек ORB.

Были обработаны изображения из двух видеофрагментов, используемых при оценке работоспособности алгоритма в главе 4.2. Примеры полученных изображений представлены на рисунке 4.3.1. В таблице 4.3.1 приведена временная оценка реализованного алгоритма применительно к одному изображению. На изображениях, полученных при обработке, желтым выделена область возможного расположения объекта интереса, определенная детектором особых точек, а синим - результат последующего применения сверточной нейронной сети.

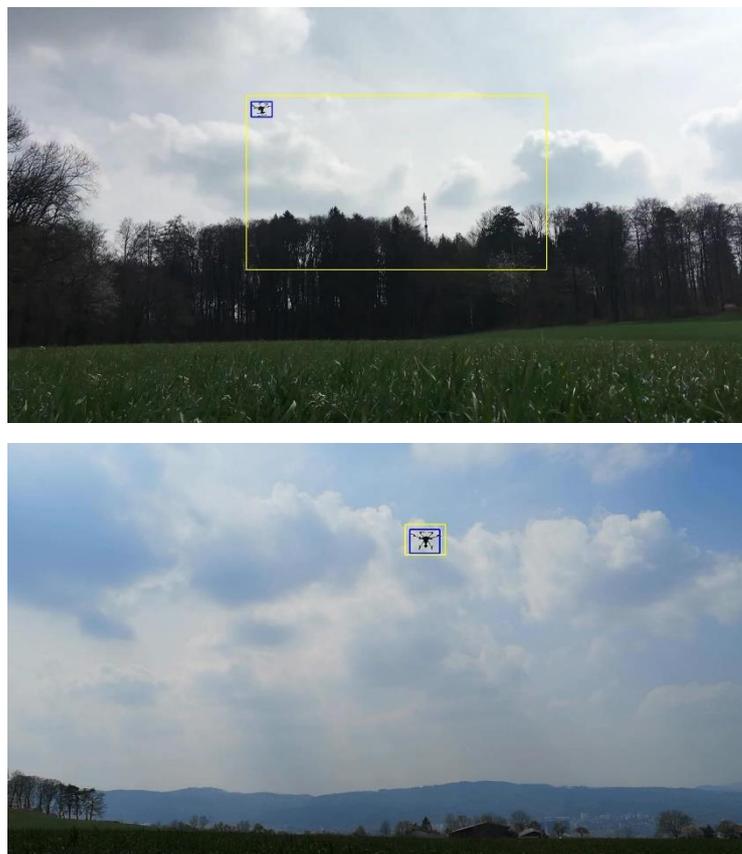


Рис. 4.3.1 - Примеры полученных результатов: желтым выделена область обнаружения детектором особых точек, синим – нейросетевым методом

Таблица 4.3.1 - Временная оценка реализованного алгоритма применительно к одному изображению

Параметр	Кадров/с
Сравнение фрагмента с эталоном	до 126,39
Обработка изображения детектором особых точек	до 16,72
Обработки фрагмента свёрточной нейронной сетью	до 9,29

Результаты, представленные в таблице 4.3.1, показывают, что полученная скорость работы детектора особых точек в 1,8 раза выше, чем использование свёрточной нейронной сети при размере обрабатываемых изображений 1920x1080 пикселей. На основе проведенных расчетов можно сделать вывод, что предварительное применение детектора особых точек, при отсутствии обнаружения признаков наличия объекта интереса на изображении, может уменьшить среднее время обработки массива изображений за счет отсутствия необходимости применения нейросетевого подхода .

4.4 Результаты

По гибриднему алгоритму, основанному на предварительном обнаружении контуров объектов с последующим применением свёрточных нейронных сетей: выполнен численный эксперимент на универсальном компьютере, целью которого являлось определение возможности использования данных алгоритмов в системах реального времени. На основе проведенных оценок времени и точности работы описанного алгоритма в сравнении с применением только сверточных нейронных сетей сделан вывод, что за время обучения и обработки отвечает тип сети, а не наличие предварительной обработки изображений.

По гибриднему алгоритму, основанному на предварительной обработке изображения нейросетевым методом с последующим поиском особых точек в фрагменте: было проведено сравнение получаемых результатов по скорости и точности работы комбинаций различных типов архитектур сверточных нейронных сетей и детектора особых точек. Отмечается возможность использования предложенного алгоритма в системах реального времени на базе компьютеров

универсальной архитектуры, намечаются пути его дальнейшего совершенствования. Использование детектора особых точек не оказывает значительного влияния на время обработки изображений (в проведенных расчетах не превышало 1% от общего времени обработки изображений с учетом сохранения файлов и преобразований изображений, 4% от общего времени расчетов), в то же время, последующая обработка ранее обнаруженных нейросетевым методом объектов снижает число ложных распознаваний на 40...100%. Использование детектора особых точек ORB не позволяет уменьшить число необнаруженных объектов интереса. Более того, требуется точный подбор числа совпадений количества особых точек с эталонными образами объектов. Полученная скорость выполнения обнаружения и классификации объекта интереса на изображении, а также проверки верности распознавания (непосредственно расчетов) возможно со скоростью до 7,7 кадров в секунду при размере обрабатываемых изображений 1920x1080 пикселей. Данная скорость обработки одного изображения может считаться недостаточной для решения задач реального времени, однако, последующая доработка реализации предложенного алгоритма, а также его применение к изображениям другого разрешения или на других вычислительных мощностях позволят применять его в системах реального времени.

По гибриднему алгоритму, основанному на предварительном поиске области возможного расположения объекта интереса методом особых точек с последующей проверкой нейросетевым методом: реализованный алгоритм работоспособен. Вследствие малого времени поиска особых точек, возможно уменьшение времени обработки массива изображений за счет сокращений применении нейросетевого метода. Значение заданных настраиваемых параметров оказывает значительное влияние на получаемый результат.

Заключение

В процессе выполнения работы были получены следующие результаты:

1) Определены основные направления повышения эффективности алгоритмов распознавания объектов на изображениях:

- повышение точности работы «традиционных» алгоритмов компьютерного зрения;

- модификация архитектур сверточных нейронных сетей;

- создание алгоритмов и методов, включающих в себя комбинацию различных подходов к решению поставленной задачи (гибридные алгоритмы)

2) Предложена классификация алгоритмов обнаружения и распознавания объектов, основанная на используемых методах компьютерного зрения, сформулированы типы гибридных алгоритмов на основе опубликованных работ;

3) Предложены и реализованы три структуры гибридных алгоритмов, включающих в себя выделение прямолинейных контуров на изображениях, детектор поиска особых точек, а также использование сверточных нейронных сетей;

4) Показано, что алгоритм Хюккеля применим для задач, требующих получения общего описания контуров объекта на изображении большого разрешения без высокой точности за малый промежуток времени. При обработке изображений малого разрешения или обнаружения объекта сложной формы, а также при необходимости точного описания контура объекта предпочтительно совместное использование алгоритмов Кэнни и Хафа. Скорость работы алгоритмов (17...22 изображений в секунду при размере входных изображений 289 на 180 пикселей) показывает возможность их использования при решении задач реального времени;

5) Выполнено сравнение доступных в открытом доступе архитектур сверточных нейронных сетей библиотеки Tensorflow. Выбраны модели для реализуемых гибридных алгоритмов. Показано, что использование одноэтапных

нейронных сетей позволяет достичь времени обработки изображений, удовлетворяющего требованиям систем реального времени;

б) На основе проведенных оценок времени и точности работы гибридного алгоритма, основанного на предварительном переходе к контурному представлению изображения, в сравнении с применением только сверточных нейронных сетей, сделан вывод, что за время и точность обработки отвечает тип сети, а не наличие предварительной обработки изображений;

7) Разработан гибридный алгоритм поиска объектов интереса на отдельных фрагментах изображения, объединяющий обнаружение объектов интереса нейросетевым методом, с последующей проверкой правильности распознаванием детектором особых точек. Использование детектора особых точек не оказывает значительного влияния на время обработки изображений, в то же время разработанный алгоритм позволил уменьшить число ложных распознаваний на 40...100%. Получена скорость выполнения расчетов на универсальном компьютере до 7,7 кадров в секунду при размере обрабатываемых изображений 1920x1080 пикселей.

8) На основе исследования гибридного алгоритма, сочетающего предварительное обнаружение областей возможного расположения объекта интереса на изображении детектором особых точек с последующей обработкой фрагмента свёрточной нейронной сетью показана возможность гибридных алгоритмов сократить время обработки массива изображений при использовании тех же вычислительных ресурсов (полученная скорость работы детектора особых точек в 1,8 раза выше, чем использование свёрточной нейронной сети при размере обрабатываемых изображений 1920x1080 пикселей).

9) Полученные результаты оценки эффективности реализованных гибридных алгоритмов показывают, что разработка алгоритмов данного типа позволяет достигнуть повышения скорости и точности обработки изображений в части распознавания на них антропогенных объектов интереса.

Список литературы

1. Кэлер А., Брэдски Г. Изучаем OpenCV 3 / пер. с англ. А. А. Слинкина. – М.: ДМК Пресс, 2017. – 826 с.: ил.
2. Деривативное электронное издание на основе печатного аналога: Компьютерное зрение / Л. Шапиро, Дж. Стокман ; пер. с англ.—М. : БИНОМ. Лаборатория знаний, 2006.—752 с. : ил., [8] с. цв. вкл.—(Лучший зарубежный учебник).
3. Xu Z., Hrustic E., Vivet D. CenterNet heatmap propagation for real-time video object detection // Proceedings of the computer vision (ECCV). – Springer, Cham, 2020. – С. 220-234. – DOI:10.1007/978-3-030-58595-2_14.
4. Алгоритм нейросетевого распознавания надводных объектов в реальном времени / В. А. Бондаренко, В. А. Павлова, В. А. Тупиков, Н. Г. Холод // Известия Тульского государственного университета. Технические науки. – 2021. – № 1. – С. 19-33.
5. Ma G. et al. Intelligent detection of foreign matter in coal mine transportation belt based on convolution neural network // Scientific Programming. – 2022. – Т. 3. – DOI: 10.1155/2022/9740622.
6. Черников А.А. Алгоритм обнаружения и классификации объектов на неодно-родном фоне для оптико-электронных систем // Труды МАИ. 2023. № 129. DOI: 10.34759/trd-2023-129-26.
7. Решение задачи обнаружения объекта с помощью нейросетевых технологий / С.О. Власов [и др.] // Препринты ИПМ им. М.В.Келдыша.2023. № 16. 27 с. <https://doi.org/10.20948/prepr-2023-16>
8. Агеев А.В., Богуславский А.А., Власов С.О, Соколов С.М. Компьютерное зрение. Лабораторный практикум. М.: ИПМ им.М.В.Келдыша, 2024. 60 с. <https://keldysh.ru/e-biblio/ageev/>
9. Власов, С. О. Гибридные алгоритмы обнаружения объектов в системах компьютерного зрения реального времени / С. О. Власов, А. А. Богуславский // GraphiCon 2024 : Материалы 34-й Международной конференции по компьютерной

графике и машинному зрению, Омск, 17–19 сентября 2024 года. – Омск: Омский государственный технический университет, 2024. – С. 476-483.

10. Власов, С. О. Гибридный алгоритм распознавания объекта в системе компьютерного зрения / С. О. Власов, А. А. Богуславский, С. М. Соколов // Экстремальная робототехника. – 2024. – № 1(35). – С. 300-307.

11. Власов С.О. Гибридный алгоритм распознавания объекта в системе компьютерного зрения / С.О. Власов, А.А. Богуславский, С.М. Соколов // Робототехника и техническая кибернетика. – Т. 13. - № 2. – Санкт-Петербург : ЦНИИ РТК. – 2025. – С. 121-128.

12. Беляков, П. В. Алгоритм построения пространственной структуры сцены / П. В. Беляков // Современные информационные технологии и ИТ-образование. – 2019. – Т. 15, № 2. – С. 331-339. – DOI 10.25559/SITITO.15.201902.331-339.

13. Михалёв О.Н. Искусственный интеллект в автоматизации технологических процессов / О.Н. Михалёв, А.С. Янюшкин // Актуальные проблемы в машиностроении. – 2021. – Т. 8. – № 1-2. – С. 7-13. – Текст: непосредственный.

14. Васильченко, В. А. Повышение качества и оперативности идентификации специальных состояний мониторируемых объектов на основе разработки математического и программного обеспечения обработки компьютерных изображений с использованием больших баз данных / В. А. Васильченко, В. Л. Бурковский // Моделирование, оптимизация и информационные технологии. – 2020. – Т. 8, № 2(29). – DOI 10.26102/2310-6018/2020.29.2.030.

15. Хамад, Ю. А. Разработка и исследование алгоритмов сегментации и распознавания объектов на медицинских изображениях на основе шварц-преобразования и нейронных сетей : специальность 05.13.17 "Теоретические основы информатики" : диссертация на соискание ученой степени кандидата технических наук / Хамад ЮСИФ АХМЕД, 2020. – 171 с.

16. Богуш, Р. П. Комбинирование блочных алгоритмов вычисления оптического потока для обнаружения и сопровождения движущихся объектов на видеопоследовательностях / Р. П. Богуш, В. Ю. Лысенко, Г. А. Самощенко //

Вестник Полоцкого государственного университета. Серия С. Фундаментальные науки. – 2011. – № 4. – С. 2-7.

17. Друки, А. А. Алгоритмы нейросетевого детектирования и распознавания символов на сложном фоне : специальность 05.13.01 "Системный анализ, управление и обработка информации (по отраслям)" : диссертация на соискание ученой степени кандидата технических наук / Друки Алексей Алексеевич, 2016. – 216 с.

18. Косатый, Д.Н. Методы создания гибридных классификаторов на основе каскадов и глубоких нейронных сетей: магистерская дис.: защищена 27.05.2016 / Косатый Дмитрий Николаевич. - СПб., 2016. - 28 с.

19. Стротов, В. В. Комбинирование нескольких алгоритмов для увеличения точности обнаружения и локализации пешеходов / В. В. Стротов, П. Е. Жгутов // Труды Международной конференции по компьютерной графике и зрению "Графикон". – 2022. – № 32. – С. 628-635. – DOI 10.20948/graphicon-2022-628-635.

20. Бернс, Лоуренс. Автономия. Как появился автомобиль без водителя и что это значит для нашего будущего / Лоуренс Бернс, Кристофер Шулган; перевод с английского К. Вантуха. — Москва : Бомбора™, 2021. — 399 с. : 22 см — (Книги Политеха. Идеи и технологии).

21. Goenka U. et al. Threat detection in self-driving vehicles using computer vision // arXiv arXiv: 2209.02438. 2022. DOI:10.48550/arXiv.2209.02438 – URL: <https://arxiv.org/abs/2209.02438>.

22. Holder C. J., Breckon T. Learning to Drive: Using Visual Odometry to Bootstrap Deep Learning for Off-Road Path Prediction // IEEE Intelligent Vehicles Symposium (IV), 2018. P. 2104-2110.

23. Машенко, П. Е. Метод визуального распознавания местности NetVLAD для локализации локомотива / П. Е. Машенко, П. П. Ширяев // Автоматика, связь, информатика. – 2020. – № 10. – С. 14-17. – DOI 10.34649/AT.2020.10.10.003.

24. Носков В.П. Комплексированные СТЗ автономных роботов / В.П. Носков, И.В. Рубцов, В.С. Лапшов // Робототехника и техническая кибернетика. – Т. 8. - №3. – Санкт-Петербург : ЦНИИ РТК. – 2020. – С. 233-240.

25. Мохаммед, М. М. Г. Алгоритмы обнаружения лица на основе анализа и обработки изображений : специальность 05.13.01 "Системный анализ, управление и обработка информации (по отраслям)" : диссертация на соискание ученой степени кандидата технических наук / Мохаммед Мамдух Мохаммед Гомаа, 2017. – 118 с.

26. Алиев, М. В. Идентификация лица человека по заданному примитиву / М. В. Алиев, Д. А. Шовгенов // Вестник Адыгейского государственного университета. Серия 4: Естественно-математические и технические науки. – 2014. – № 3(142). – С. 185-188.

27. Ван, Л. Программные средства с использованием методов и алгоритмов кратномасштабного вейвлет анализа для обработки и поиска изображений : специальность 05.13.11 "Математическое и программное обеспечение вычислительных машин, комплексов и компьютерных сетей" : диссертация на соискание ученой степени кандидата технических наук / Ван Лянпэн, 2020. – 138 с.

28. Модель системы распознавания наземных объектов на основе нейронной сети / С. В. Войцеховский, Я. В. Бугаева, М. Ю. Касавцев, А. Б. Кириенко // Известия Тульского государственного университета. Технические науки. – 2022. – № 7. – С. 115-123. – DOI 10.24412/2071-6168-2022-7-115-123.

29. Митрофанова, Е. Ю. Поиск объектов на изображениях с использованием TensorFlow Object Detection / Е. Ю. Митрофанова // Информатика: проблемы, методы, технологии : Материалы XXII Международной научно-практической конференции им. Э.К. Алгазинова, Воронеж, 10–12 февраля 2022 года / Под редакцией Д.Н. Борисова. – Воронеж: Общество с ограниченной ответственностью "Вэлборн", 2022. – С. 680-687.

30. Дэвис Р., Терк М. Компьютерное зрение. Современные методы и перспективы развития / пер. с англ. В.С. Яценкова. — М.: ДМК Пресс, 2022. — 690 с.

31. Корк П. Машинное зрение. Основы и алгоритмы с примерами на Matlab / пер. с англ. В. С. Яценкова. – М.: ДМК Пресс, 2023. – 584 с.: ил.

32. Брантон С. Л., Куц Дж. Н. Анализ данных в науке и технике / пер. с англ. А. А. Слинкина. – М.: ДМК Пресс, 2021. – 574 с.: ил.
33. Szeliski, R.: Computer Vision: Algorithms and Applications (2nd edn.). Springer, Berlin (2022).
34. Лекун Я. Как учится машина: Революция в области нейронных сетей и глубокого обучения / Ян Лекун. — Пер. с фр. — М.: Альпина ПРО, 2021.
35. Николенко С., Кадурын А., Архангельская Е. Глубокое обучение. — СПб.: Питер, 2018. — 480 с.: ил. — (Серия «Библиотека программиста»).
36. Шолле Ф. Глубокое обучение на Python — СПб.: Питер, 2018. — 400 с.
37. Шаошань Лю, Лиюнь Ли, Цзе Тан, Шуаш Ву, Жан-Люк Годье. Разработка беспилотных транспортных средств / науч. ред. В. С. Яценков; пер. с англ. П. М. Бомбаковой. – М.: ДМК Пресс, 2022. – 246 с.: ил.
38. Андриянов, Н.А. Обнаружение объектов на изображении: от критериев Байеса и Неймана–Пирсона к детекторам на базе нейронных сетей EfficientDet / Н.А. Андриянов, В.Е. Дементьев, А.Г. Ташлинский // Компьютерная оптика. – 2022. – Т. 46, № 1. – С. 139-159.
39. Прокопеня. А. С. Современные методы распознавания изображений / А. С. Прокопеня, И. С. Азаров // BIG DATA and Advanced Analytics = BIG DATA и анализ высокого уровня : сборник материалов V Международной научно-практической конференции, Минск, 13–14 марта 2019 г. В 2 ч. Ч. 1 / Белорусский государственный университет информатики и радиоэлектроники; редкол. : В. А. Богуш [и др.]. – Минск, 2019. – С. 351 – 359.
40. Горячкин Б.С. Компьютерное зрение / Б.С. Горячкин, М.А. Китов // E-Scio. - 2020. - №9 (48). - С. 317-345.
41. Ульянов, Р. Д. Компьютерное зрение. Методы распознавания лиц / Р. Д. Ульянов // Дни науки студентов Владимирского государственного университета имени Александра Григорьевича и Николая Григорьевича Столетовых : Сборник материалов заочных научно-практических конференций, Владимир, 15–30 апреля 2020 года. – Владимир: Владимирский государственный университет имени Александра Григорьевича и Николая Григорьевича Столетовых, 2020. – С. 333-348.

42. Бурдуковский С. О. Эффективность моделей Tensorflow в приложении к задаче обнаружения глаз на фотографии / С. О. Бурдуковский // Вестник НГУЭУ. – 2022. – № 2. – С. 228-238. – DOI: 10.34020/2073-6495-2022-2-228-238.

43. Hazarika P., Kumari Ms. Evolution of Modern Deep Learning Methods of Object Recognition // Research in Computing Science, 2019 No. 148. P. 71-79. DOI: 10.13053/rcs-148-3-6.

44. Мостовщиков, Д. Н. Комплексные алгоритмы классификации визуальных объектов на базе методов машинного обучения / Д. Н. Мостовщиков, Р. К. Могилатов, О. В. Бабкин // Проблемы современной науки и образования. – 2021. – № 6(163). – С. 15-22.

45. Hraby P. et al. Handbook on Leveraging Lines for Two-View Relative Pose Estimation // IEEE International Conference on 3D Vision (3DV), 2024, pp. 376–386.

46. Клестов, Р. А. Гибридный метод распознавания контуров на изображении на основе технологий компьютерного зрения / Р. А. Клестов, В. Ю. Столбов // GraphiCon 2017 : Труды 27-й Международной конференции по компьютерной графике и машинному зрению, Пермь, 24–28 сентября 2017 года / Пермский государственный национальный исследовательский университет. – Пермь: Издательский центр Пермского государственного национального исследовательского университета, 2017. – С. 208-211.

47. Томакова, Р. А. Гибридные методы и алгоритмы для интеллектуальных систем классификации сложноструктурируемых изображений : специальность 05.13.17 "Теоретические основы информатики" : диссертация на соискание ученой степени доктора технических наук / Р. А. Томакова. – Белгород, 2013. – 296 с.

48. Taspinar Y., Murat S. Object recognition with hybrid deep learning methods and testing on embedded systems // International Journal of Intelligent Systems and Applications in Engineering. – 2020. – Т. 8. – С. 71-77. – DOI: 10.18201/ijisae.2020261587.

49. Сеничев, А. В. Сравнение глубокого обучения с традиционными методами компьютерного зрения в задачах идентификации дефектов / А. В.

Сеничев, А. И. Новикова, П. В. Васильев // Молодой исследователь Дона. – 2020. – № 4(25). – С. 64-67.

50. Javed A. et al. (2007). A Vision-Based System for a UGV to Handle a Road Intersection // Proceedings of the Twenty-Second AAAI Conference on Artificial Intelligence, July 22-26, Vancouver, British Columbia, Canada, 2007. P. 1077-1082.

51. Чочиа, П. А. Теория и методы обработки видеoinформации на основе двухмасштабной модели изображения: специальность 05.13.18 "Математическое моделирование, численные методы и комплексы программ" : диссертация на соискание ученой степени доктора технических наук / Чочиа Павел Антонович, 2016. – 302 с.

52. Camposeco F. et al. Hybrid Scene Compression for Visual Localization // IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), Long Beach, CA, USA, 2019, pp. 7645-7654. DOI: 10.1109/CVPR.2019.00784.

53. Пименова, М. Б. Распознавание характерных объектов на местности с использованием метода SURF / М. Б. Пименова // Политехнический молодежный журнал. – 2019. – № 10(39). – С. 3. – DOI 10.18698/2541-8009-2019-10-540.

54. Weinzaepfel P. et al. DeepFlow: Large Displacement Optical Flow with Deep Matching // 2013 IEEE International Conference on Computer Vision, Sydney, NSW, Australia, 2013, pp. 1385-1392. DOI: 10.1109/ICCV.2013.175.

55. Нейросетевой алгоритм полнокадрового распознавания надводных объектов в реальном времени / В. А. Бондаренко, В. А. Павлова, В. А. Тупиков, Н. Г. Холод // Известия ЮФУ. Технические науки. – 2020. – № 1(211). – С. 188-199. – DOI 10.18522/2311-3103-2020-1-188-199.

56. Kanagaraj N. et al. Deep learning using computer vision in self driving cars for lane and traffic sign detection // International Journal of System Assurance Engineering and Management, Vol. 12, No. 6, 2021, pp. 1011-1025. DOI:10.1007/s13198-021-01127-6.

57. Ненашев, В. А. Классификация и распознавание наземных объектов в потоке радиолокационных кадров на основе нейросетевого подхода в передней зоне обзора бортовых РЛС многопозиционной системы / В. А. Ненашев, С. А. Ненашев // Труды Международной конференции по компьютерной графике и

зрению "Графикон". – 2023. – № 33. – С. 572-580. – DOI 10.20948/graphicon-2023-572-580.

58. Хрящев, В. В. Использование нейронной сети EfficientDet в задаче обнаружения патологий желудка на видеоизображениях эндоскопического исследования / В. В. Хрящев, А. Л. Приоров // Модели, системы, сети в экономике, технике, природе и обществе. – 2023. – № 2(46). – С. 185-192. – DOI 10.21685/2227-8486-2023-2-12.

59. Нгуен, Т. Д. Разработка методики определения положения береговой линии поверхностных водных объектов по материалам аэрокосмических съемок с использованием технологии глубокого обучения / Т. Д. Нгуен, В. А. Малинников // Мониторинг. Наука и технологии. – 2021. – № 2(48). – С. 32-40. – DOI 10.25714/MNT.2021.48.003.

60. Береснев, А. П. Исследование свёрточных нейронных сетей класса YOLO для мобильных систем детектирования объектов на изображениях / А. П. Береснев, И. В. Зоев, Н. Г. Марков // Труды Международной конференции по компьютерной графике и зрению "Графикон". – 2018. – № 28. – С. 196-199.

61. Eadhunath V. et al. Self-Driving Car using Convolutional Neural Network and Road Lane Detector // International Journal of Engineering Research & Technology (IJERT). Vol.8, No. 05, 2019, pp. 951-954.

62. Сирота, А. А. Анализ алгоритмов поиска объектов на изображениях с использованием различных модификаций сверточных нейронных сетей / А. А. Сирота, Е. Ю. Митрофанова, А. И. Милованова // Вестник Воронежского государственного университета. Серия: Системный анализ и информационные технологии. – 2019. – № 3. – С. 123-137.

63. Денисенко, А. А. Модификация архитектуры Tiny-YOLO для задачи обнаружения объектов в реальном времени / А. А. Денисенко // Международный журнал прикладных и фундаментальных исследований. – 2021. – № 4. – С. 53-57. – DOI 10.17513/mjpf.13203.

64. Авраменко, В. С. Распознавание объектов беспилотными летательными аппаратами на основе нейронной сети MOBILENET / В. С. Авраменко, Е. С. Чичков // Экстремальная робототехника. – 2024. – № 1(34). – С. 261-266.

65. Томакова, Р. А. Гибридные методы и алгоритмы для интеллектуальных систем классификации сложноструктурируемых изображений : специальность 05.13.17 "Теоретические основы информатики" : автореферат диссертации на соискание ученой степени доктора технических наук / Томакова Римма Александровна. – Белгород, 2013. – 42 с.

66. Мешкова, Е. В. Методика построения классификатора текста на основе гибридной нейросетевой модели / Е. В. Мешкова // Известия ЮФУ. Технические науки. – 2008. – № 4(81). – С. 212-215.

67. Varga L. A., Zell A. Tackling the background bias in sparse object detection via cropped windows // Proceedings of the IEEE/CVF International Conference on Computer Vision Workshops (ICCVW) –2021 – С. 2768-2777.

68. Федосов, В. П. Обнаружение и отслеживание объектов на основе карты внимания для БПЛА / В. П. Федосов, Р. Р. Ибадов, С. Р. Ибадов // Журнал радиоэлектроники. – 2022. – № 11. – DOI 10.30898/1684-1719.2022.11.7.

69. Анализ гистологических изображений посредством морфологических операторов, синтезированных на основе преобразования Фурье и нейросетевого моделирования / Р. А. Томакова, С. А. Филист, С. А. Горбатенко, Н. А. Швецова // Биотехносфера. – 2010. – № 3(9). – С. 54-60.

70. Korablyov M. et al. Hybrid Neuro-Fuzzy Model with Immune Training for Recognition of Objects in an Image // International Conference on Information Control Systems & Technologies, 2020. pp. 267-281.

71. Reddy K. S. S. et al. A real-time automated system for object detection and facial recognition // E3S Web of Conferences. – 2023. – №. 430. – DOI: 10.1051/e3sconf/202343001076.

72. Ren S. et al. Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks // IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 39, no. 6, pp. 1137-1149, 1 June 2017. DOI: 10.1109/TPAMI.2016.2577031.

73. Suma K. G. et al. CETR: CenterNet-Vision transformer model for wheat head detection // *Journal of Autonomous Intelligence*. – 2024. – Т. 7. – №. 3. – С. 1-12. – DOI: 10.32629/jai.v7i3.1189.

74. Смольянинов, В. А. Проектирование программного комплекса обнаружения и распознавания дорожных знаков в потоковом видео / В. А. Смольянинов, Ю. С. Белов // *Научное обозрение. Технические науки*. – 2021. – № 4. – С. 16-21.

75. Chiatti A. Fit to Measure: Reasoning about Sizes for Robust Object Recognition // *arXiv arXiv: 2010.14296*. 2020. DOI:10.48550/arXiv.2010.14296 – URL: <https://arxiv.org/abs/2010.14296>.

76. Lynen S. et al. Large-scale, real-time visual-inertial localization revisited // *arXiv arXiv: 1907.00338*. 2019. DOI: 10.48550/arXiv.1907.00338 – URL: <https://arxiv.org/abs/1907.00338>.

77. Черников А.А. Алгоритм обнаружения и классификации объектов на неоднородном фоне для оптико-электронных систем // *Труды МАИ*. 2023. № 129. DOI: 10.34759/trd-2023-129-26.

78. Zhang D. Q., Chang S.-F. A Generative-Discriminative Hybrid Method for Multi-View Object Detection // *2006 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'06)*, New York, NY, USA, 2006, pp. 2017-2024. DOI: 10.1109/CVPR.2006.27.\

79. Сирота, А. А. Двухэтапный алгоритм обнаружения и оценивания границы объектов на изображениях в условиях аддитивных помех и деформирующих искажений / А. А. Сирота, А. И. Соломатин, Е. В. Воронова // *Компьютерная оптика*. – 2010. – Т. 34, № 1. – С. 109-117.

80. Новоторцев, Л. В. Исследование и разработка алгоритмов поиска и восстановления населенных пунктов на наборе аэрофотоснимков: специальность 05.13.11 "Математическое и программное обеспечение вычислительных машин, комплексов и компьютерных сетей" : диссертация на соискание ученой степени кандидата физико-математических наук / Новоторцев Леонид Владимирович, 2021. – 101 с.

81. Хюккель, М. Оператор нахождения контуров на кодированных изображениях / М. Хюккель // Интегральные роботы. Вып.1. – М.: Мир, -1973. - С.225-240.

82. Manfred H. Hueckel. A Local Visual Operator Which Recognizes Edges and Lines. Journal of the Association for Computing Machinery, vol. 20, No. 4, October 1973.

83. Солодовщиков А.Ю., Платонов А.К. Исследование метода Карунена-Лоэва // Препринты ИПМ им. М.В.Келдыша. 2006. № 19. 29 с.

84. <https://yandex.ru/maps> (дата обращения: 07.12.2023).

85. Abduljabbar H.M. Analysing of Hueckel Edge Detector Performance Using Satellite Image // Energy Procedia 157, 2019, pp. 533–540.

86. TensorFlow 2 Detection Model Zoo // Google. URL: https://github.com/tensorflow/models/blob/master/research/object_detection/g3doc/tf2_detection_zoo.md (дата обращения: 20.05.2024).

87. TensorFlow 1 Detection Model Zoo // Google. — https://github.com/tensorflow/models/blob/master/research/object_detection/g3doc/tf1_detection_zoo.md. (дата обращения: 20.05.2024).

88. Tan M. X., Pang R. M., Le Q. V. EfficientDet: Scalable and efficient object detection // Proceedings of IEEE/CVF Conference on Computer Vision and Pattern Recognition. – 2020. С.10778–10787. – DOI: 10.1109/CVPR42600.2020.01079

89. Song S. et al. EfficientDet for fabric defect detection based on edge computing // Journal of engineered fibers and fabrics. – 2021. – Т. 16 – №. 1. – С. 155892502110083 – DOI: 10.1177/1558925021100834.

90. Duan K. et al. Centernet++ for object detection // arXiv arXiv:2204.08394 – 2022. – DOI: 10.48550/arXiv.2204.08394 – URL: <https://arxiv.org/abs/2204.08394>.

91. Lim J., Astrid M. Small object detection using context and attention // arXiv arXiv:1912.06319 – 2019. DOI: 10.48550/arXiv.1912.06319 – URL: <https://arxiv.org/abs/1912.06319>.

92. LabelImg URL: <https://github.com/heartexlabs/labelImg> (дата обращения: 21.05.2024).

93. Lu X. et al. MimicDet: Bridging the Gap Between One-Stage and Two-Stage Object Detection // arXiv arXiv:2009.11528 – 2020. DOI: 10.48550/arXiv.2009.11528 URL: <https://arxiv.org/abs/2009.11528>.

94. Huang S. et al. Tp-lsd: Tri-points based line segment detector // arXiv arXiv: 2009.05505 – 2020. DOI: 10.48550/arXiv.2009.05505 URL: <https://arxiv.org/abs/2009.05505>.

95. Drone-tracking-datasets URL: <https://github.com/CenekAlbl/drone-tracking-datasets> (дата обращения: 27.05.2024).