

ЮГО-ЗАПАДНЫЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ

На правах рукописи

Хомяков Олег Олегович

**МОДЕЛЬ, МЕТОД, АЛГОРИТМ И ВЫЧИСЛИТЕЛЬНЫЙ МОДУЛЬ
ОБРАБОТКИ ИЗОБРАЖЕНИЙ СИМВОЛЬНОЙ МАРКИРОВКИ**

Специальность 2.3.2. Вычислительные системы и их элементы

Диссертация на соискание ученой степени

кандидата технических наук

Научный руководитель: кандидат
технических наук, доцент
Панищев Владимир Славиевич

Курск–2024

ОГЛАВЛЕНИЕ

ВВЕДЕНИЕ	4
ГЛАВА 1. АНАЛИТИЧЕСКИЙ ОБЗОР МОДЕЛЕЙ, МЕТОДОВ, АЛГОРИТМОВ И ВЫЧИСЛИТЕЛЬНЫХ МОДУЛЕЙ ОБРАБОТКИ ИЗОБРАЖЕНИЙ СИМВОЛЬНОЙ МАРКИРОВКИ	10
1.1. Специализированные системы компьютерного зрения обработки символьной информации	10
1.2. Методы и алгоритмы, применяемые в специализированных системах компьютерного зрения	15
1.2.1. Методы поиска и выделения объекта	19
1.2.2. Методы поиска текста	20
1.2.3. Методы распознавания текста	22
1.2.4. Методы классификации характеристик объекта	23
1.3. Архитектуры вычислительных модулей, используемых для обработки символьной информации на изображениях	25
1.4. Альтернативные методы распознавания характеристик объектов продукции	31
1.5. Противоречие между точностью и скоростью методов обработки символьной информации	34
Вывод по первой главе	36
ГЛАВА 2. МАТЕМАТИЧЕСКАЯ МОДЕЛЬ И МЕТОД ОБРАБОТКИ ИЗОБРАЖЕНИЙ, СОДЕРЖАЩИХ СИМВОЛЬНУЮ ИНФОРМАЦИЮ.....	38
2.1. Математическая модель обработки изображений	38
2.2. Метод обработки изображений обработки изображений, содержащих символьную информацию	45
Вывод по второй главе	51
ГЛАВА 3. КОНВЕЙЕРНО-ПАРАЛЛЕЛЬНЫЙ АЛГОРИТМ ОБРАБОТКИ ИЗОБРАЖЕНИЙ, СОДЕРЖАЩИХ СИМВОЛЬНУЮ ИНФОРМАЦИЮ.....	53

3.1.	Общая структура алгоритма	53
3.2.	Поиск объекта для распознавания	59
3.3.	Отслеживание объекта для распознавания	61
3.4.	Выделение объекта для распознавания	63
3.5.	Поиск областей текста и выделение символов	71
3.6.	Распознавание текста	74
3.7.	Классификация характеристик.....	77
3.8.	Допущения и ограничения.....	78
3.9.	Исследование точности выделения объекта и распознавания символов	80
	Вывод по третьей главе	82
	ГЛАВА 4. КОНВЕЙЕРНО-ПАРАЛЛЕЛЬНЫЙ ВЫЧИСЛИТЕЛЬНЫЙ МОДУЛЬ ОБРАБОТКИ СИМВОЛЬНОЙ ИНФОРМАЦИИ.....	83
4.1.	Структурная организация вычислительного модуля	83
4.2.	Тестирование комплексной характеристики «точность × пропускная способность» вычислительного модуля	93
	Вывод по четвертой главе	98
	Список использованной литературы.....	101

ВВЕДЕНИЕ

Актуальность темы. Научно-технический прогресс определил применение специализированных вычислительных систем (ВС) и их элементов для решения различных расчетно-инженерных, научно-исследовательских, проблемно-ориентированных задач, задач управления промышленными объектами, имеющих в своем составе модули-обработчики по различным видам информации. Символьная информация и вычислительные элементы (средства) ее получения и обработки служат основой для работы высокопроизводительных ВС, способных расширять их функциональные возможности за счет формализации интеллектуальных действий, свойственных человеку в части распознавания и интерпретации текстов и текстовых образов. В связи с этим в архитектуру высокопроизводительных ВС при решении данного класса задач целесообразно вводить вычислительные модули (ВМ) для подготовки видеоданных, предобработки изображений в части устранения искажений, шумов, избыточности и распознавания символьных данных по их изображениям.

Специализированные системы компьютерного зрения (ССКЗ), отвечающие за распознавание символьной информации (СИ) на этапе ее подготовки и ввода, являются частью специализированных ВС промышленного применения в гибких производственных системах (ГПС), ведущих потоковую (на конвейере) обработку символьной маркировки продукции. При этом вычислительные процессы распознавания и обработки символьной информации (ОСИ) в условиях априорных искажений символов, смешанных алфавитов, неопределенности стилей и характеристик записи символов являются недостаточно проработанными как с алгоритмической, так и с аппаратной точек зрения.

Используемые ССКЗ создаются преимущественно с использованием программных решений или на базе стандартных архитектур ВС последовательного типа, что ограничивает их скоростные характеристики. Известные аппаратные решения для распознавания и классификации символьных данных по их изображениям отличаются повышенной аппаратной сложностью, зависимостью

качества распознавания от исходных (тестовых) наборов. Как следствие, ВМ распознавания и классификации символьной маркировки продукции (характеристик объекта) по своей производительности становятся узким местом для параллельных архитектур, специализированных ВС. Исходя из этого, существует **объективное противоречие** между точностью и скоростью распознавания символьной маркировки продукции на **вычислительных модулях** (ВМ) для специализированных ВС.

Научно-техническая задача разработки методов и аппаратных средств распознавания и классификации маркировки символьной продукции в условиях неопределённости (несколько алфавитов) и искажений/зашумлений изображений символов является актуальной.

Степень разработанности темы. В настоящее время по вопросам разработки ССКЗ и обработки СИ опубликовано большое количество научно-исследовательских работ. В частности, существует широкий выбор методов для распознавания текста и классификации символов. В научной литературе известно большое количество работ, посвященных этой теме, как зарубежных, так и отечественных авторов Ю.И. Журавлев, И.Б. Гуревич, В.А. Сойфер, А.А. Орлов, В.В. Корнеев, В.Г. Хорошевский, И.А. Каляев, W. Pratt, R. Gonzalez, R. Woods, R.O.Duda, P.E. Hart. Тем не менее в известных трудах недостаточно проработаны вопросы и технические решения, обеспечивающие **в комплексе** потоковое распознавание и классификацию символьной маркировки продукции по их изображениям. Большинство реализованных подходов к построению ССКЗ ограничены использованием моноалфавитов (латиница) и не учитывают различные виды искажений, что снижает точность распознавания. Таким образом, в данной области существует необходимость в разработке новых моделей, методов, алгоритмов и ВМ для вычислительных систем.

Соответствие паспорту специальности. Согласно паспорту специальности 2.3.2. Вычислительные системы и их элементы, проблематика, рассмотренная в диссертации, соответствует пункту 5 и 6 паспорта специальности (5. Разработка научных методов и алгоритмов организации арифметической,

логической, символьной и специальной обработки данных, хранения и ввода-вывода информации, в части разработки метода и алгоритма обработки изображений, содержащих СИ. 6. Разработка научных подходов и методов, архитектурных и структурных решений, обеспечивающих эффективную техническую реализацию аппаратно-программных систем и комплексов за счет оптимизации применяемой электронной компонентной базы, элементов вычислительных систем и встраиваемого программного обеспечения, в части разработки архитектурного решения для вычислительного модуля обработки изображений, обеспечивающего повышение комплексной характеристики «точность × пропускная способность»).

Объектом исследования являются элементы вычислительной системы для обработки изображений и распознавания символьных данных (маркировки) на них.

Предметом исследования являются модели, методы, алгоритмы и вычислительные модули обработки символьной информации на изображениях маркировки продукции.

Цель диссертационной работы. Повышение комплексной характеристики «точность × пропускная способность» обработки символьной информации на маркировке продукции.

Задачи.

1. Провести сравнительный анализ существующих моделей, методов, алгоритмов и вычислительных модулей обработки СИ, постановка задачи обработки изображений с маркировкой продукции.

2. Разработать математическую модель и метод распознавания символьной маркировки с возможностью настройки на конкретные виды маркировок.

3. На основе математической модели и метода создать конвейерно-параллельный алгоритм обработки изображений, содержащих СИ, реализующий распознавание символьной маркировки продукции.

4. Разработать структурную организацию вычислительного модуля обработки символьной информации.

5. Провести экспериментальные исследования разработанного алгоритма и вычислительного модуля.

Научная новизна и положения, выносимые на защиту.

1. Математическая модель обработки изображения, отличающаяся использованием последовательности разработанных частных математических моделей для выделения объекта, предобработки изображения объекта, распознавания текста и классификации характеристик объекта, в совокупности позволяющая получить характеристики объекта, содержащего символьную маркировку.

2. Метод обработки изображений, содержащих символьную информацию, отличающийся применением двумерных дескрипторов и выделения объекта на основе комбинации методов Канни и Рамера-Дугласа-Пекера, а также применением рекуррентной сверточной нейронной сети, позволяющий вести комплексную обработку областей текста в условиях геометрической, информационной, цветовой неопределенностей для последующей классификации объектов.

3. Конвейерно-параллельный алгоритм обработки изображений, содержащих СИ, осуществляющий распознавание символьной маркировки продукции, основанный на разработанной модели и методе обработки изображений, отличающийся распараллеливанием процесса обработки, настройкой к различным видам характеристик и устойчивостью к искажениям, что обеспечивает повышенную точность распознавания.

4. Структурная организация вычислительного модуля обработки символьной информации, которая учитывает особенности созданного алгоритма обработки изображений, содержащих СИ, отличающаяся применением конвейерно-параллельной обработки и позволяющая повысить комплексный показатель «точность × пропускная способность».

Практическая значимость. Разработанные модель, метод, алгоритм и вычислительный модуль обработки символьной информации на маркировке продукции может быть применен в составе ССКЗ для решения широкого спектра

задач, таких как обнаружение и распознавание символьной информации в условиях конвейеризированных производств, а также автоматизации учета и инвентаризации складских помещений. Разработанное решение обеспечивает повышение комплексного показателя «точность × пропускная способность», в сравнении с аналогами до 39%-57% в зависимости от уровня шума на изображении.

Результаты диссертационной работы внедрены в ООО «Ценозавр» (г. Курск), а также используются в образовательном процессе кафедры вычислительной техники Юго-Западного государственного университета (г. Курск) при проведении занятий по дисциплинам «Цифровая обработка и анализ изображений» и «Архитектура систем обработки, анализа и интерпретации данных» направления подготовки 09.04.01 Информатика и вычислительная техника.

Методология и методы исследования. Для решения поставленных задач используются: методы проектирования вычислительных систем и их элементов, методы цифровой обработки изображений, метод компьютерного моделирования, теории алгоритмов, проектирования ЭВМ, математического моделирования

Обоснованность и достоверность научных положений и выводов основывается на использовании общепринятых методов компьютерного зрения, машинного обучения и математического моделирования с экспериментальной проверкой отдельных гипотез на тестовой выборке, содержащей различные виды исследуемых изображений, и сопоставлением полученных данных с результатами теоретических исследований.

Публикации.

По теме диссертационного исследования опубликовано 11 работ [1,2,3,4,5,6,7,8,9,10,11], из которых 3 публикации входит в перечень изданий, утвержденных ВАК Минобрнауки России, получено положительное решение о выдаче патента [12].

Апробация

Результаты диссертационного исследования докладывались, обсуждались и получили положительную оценку на 8 Международных, Всероссийских и Региональных научно-технических конференциях: Труды Международной научно-

технической конференции «Информационные технологии и математическое моделирование систем 2019» (Одинцово, 2019), Международный научно-исследовательский конкурс «Лучший исследовательский проект 2020» (Петрозаводск, 2020), национальный суперкомпьютерный форум «НСКФ-2021» (Переславль-Залесский, 2021), XIV Всероссийская межвузовская научная конференция «Наука и образование в развитии промышленной, социальной и экономической сфер регионов России» (Муром, 2022), Международная научная конференция молодых ученых «Исторические, философские, методологические проблемы современной науки» (Курск, 2022), VI Всероссийская научно-практическая конференция «Актуальные вопросы современной науки: теория и практика научных исследований» (Пенза, 2022), XVII Международная научно-техническая конференция «Оптико-электронные приборы и устройства в системах распознавания образов и обработки изображений» (Курск, 2023), Международный научный форум «Наука и инновации – современные концепции» (Москва, 2024). Работа представлялась на конкурсе инновационных проектов «Инновация и изобретение года» (номинация «Инновационная идея»), (Курск, 2021).

Личный вклад автора. В работах, опубликованных в соавторстве, лично соискателем предложены: [1, 3, 4] – математическая модель обработки изображений, [2, 5] – метод обработки изображений, содержащих символьную информацию, [3, 6, 7, 10] – конвейерно-параллельный алгоритм обработки изображений, содержащих СИ, [9, 12] – структурная организация вычислительного модуля обработки символьной информации.

Структура и объем работы. Диссертация состоит из введения, четырех разделов, заключения, списка литературы из 117 наименований и приложения. Основная часть работы изложена на 113 страницах машинописного текста, содержит 24 рисунка и 11 таблиц.

ГЛАВА 1. АНАЛИТИЧЕСКИЙ ОБЗОР МОДЕЛЕЙ, МЕТОДОВ, АЛГОРИТМОВ И ВЫЧИСЛИТЕЛЬНЫХ МОДУЛЕЙ ОБРАБОТКИ ИЗОБРАЖЕНИЙ СИМВОЛЬНОЙ МАРКИРОВКИ

1.1. Специализированные системы компьютерного зрения обработки символьной информации

Специализированные системы компьютерного зрения (ССКЗ) предназначены для автоматизации анализа визуальных данных, таких как изображения и видеопотоки, с целью извлечения и интерпретации полезной информации. Эти системы играют критическую роль в различных областях, включая промышленное производство, контроль качества продукции, инвентаризацию и логистику [13, 14, 15, 16, 17, 18, 19]. Основная задача ССКЗ – обеспечение точного и быстрого распознавания и анализа визуальных данных, что позволяет принимать обоснованные решения и выполнять необходимые действия в реальном времени [20].

В промышленном производстве ССКЗ используются для автоматического контроля качества продукции. Они позволяют обнаруживать дефекты, отслеживать соответствие стандартам и предотвращать выпуск бракованной продукции. Например, на производственных линиях ССКЗ могут контролировать размеры, форму, цвет и наличие повреждений на изделиях, обеспечивая тем самым высокое качество конечного продукта [21]. Эти системы могут также следить за процессом сборки, проверять правильность выполнения монтажных операций и выявлять отклонения от технологических норм, что способствует снижению затрат и повышению производительности.

Одним из основных применений ССКЗ в производстве является контроль качества продукции путем распознавания маркировок, серийных номеров, сроков годности и других критически важных данных, нанесенных на изделия путем использования специализированных вычислительных модулей. Например, на конвейерных линиях ССКЗ используются для автоматического считывания штрихкодов и QR-кодов, а также для проверки правильности нанесения текстовой

информации. Это позволяет своевременно выявлять и устранять ошибки в маркировке, предотвращая выпуск бракованной продукции.

ССКЗ также применяются для отслеживания и документирования производственных партий. Благодаря автоматическому распознаванию символьной информации системы могут фиксировать данные о каждой единице продукции, что обеспечивает прозрачность и прослеживаемость на всех этапах производства, в отраслях, где требуется строгий контроль качества и соответствие продукции нормативным требованиям, таких как пищевая промышленность, фармацевтика и автомобилестроение [22].

Еще одно важное применение ССКЗ в производстве связано с обеспечением соответствия продукции установленным стандартам и спецификациям. Системы компьютерного зрения могут автоматически проверять наличие и правильность всех необходимых маркировок, включая инструкции по применению, предупредительные надписи и другую важную информацию.

Важным аспектом использования ССКЗ в производстве является их интеграция с ГПС, включающими вычислительные системы для выполнения различных операций. Это позволяет автоматически передавать данные, полученные в процессе распознавания символьной информации, в централизованные системы управления, что обеспечивает оперативное обновление информации и автоматизирует процесс принятия решений. Например, данные о текущем состоянии продукции могут быть использованы для оптимизации графиков производства, управления запасами и планирования логистических операций [23, 24].

Кроме того, ГПС позволяют автоматизировать процессы упаковки и маркировки продукции. Они могут управлять роботизированными системами, которые наносят текстовую информацию на упаковку, а затем проверяют ее корректность и читаемость. Что ускоряет процесс упаковки и снижает количество ошибок, связанных с ручным нанесением маркировки [25].

В сфере инвентаризации ССКЗ играют важную роль в автоматизации учета и отслеживания товаров. Они могут использоваться для автоматического считывания

СИ, определения количества и состояния товаров на складе. Это позволяет значительно сократить время и трудозатраты на проведение инвентаризаций, а также снизить вероятность ошибок, связанных с ручным учетом [26]. ССКЗ могут также использоваться для автоматической идентификации и сортировки товаров на складах, что ускоряет процессы комплектации и отгрузки заказов.

Вычислительные системы компьютерного зрения могут сканировать метки на продуктах, коробках и паллетах, автоматически идентифицируя каждый элемент. Это позволяет существенно ускорить процесс инвентаризации, сводя к минимуму необходимость ручного ввода данных, что, в свою очередь, снижает вероятность ошибок и повышает точность учета. Например, при поступлении товара на склад ССКЗ могут автоматически зарегистрировать его, сравнив с данными в системе управления складом, что позволяет оперативно обновлять информацию о наличии и расположении товаров [27].

ССКЗ также используются для проведения регулярных инвентаризаций. Они позволяют быстро и точно сканировать большое количество товаров, значительно сокращая время, необходимое для проверки запасов. Системы компьютерного зрения могут быть интегрированы с мобильными устройствами, такими как планшеты и смартфоны, что позволяет сотрудникам склада проводить инвентаризацию с использованием технологий ОСИ. Это применимо для крупных складских комплексов, где точный и своевременный учет товаров играет важную роль [28].

Кроме того, ССКЗ играют важную роль в автоматизации складских операций. Они используются для управления роботизированными системами, такими как автоматические погрузчики и разгрузчики, роботы для перемещения товаров по складу и системы для автоматической укладки товаров на стеллажи. В этом контексте ССКЗ позволяют значительно повысить скорость и точность выполнения складских операций, что способствует оптимизации использования складских площадей и снижению затрат на эксплуатацию складов [29].

ССКЗ могут использоваться для отслеживания движения товаров внутри склада. Например, они могут фиксировать перемещение продукции из одной зоны

хранения в другую, автоматически обновляя данные в системе управления складом. Это позволяет обеспечить высокую точность учета и предотвратить потерю или неправильное размещение товаров. Системы компьютерного зрения также могут автоматически отслеживать состояние товаров, например, путем считывания информации о сроках годности и предупреждая о необходимости ротации запасов [30].

Важным преимуществом использования ССКЗ в инвентаризации является их интеграция с системами управления складом и системами планирования ресурсов предприятия, за счет использования специализированных ВМ. Это обеспечивает централизованное управление и анализ данных, полученных в процессе инвентаризации, что позволяет принимать более обоснованные решения и оптимизировать складские операции. Например, данные о текущем уровне запасов могут быть использованы для автоматического формирования заказов на пополнение или перераспределения товаров между различными складскими зонами [31].

Системы компьютерного зрения также могут использоваться для автоматизации процессов комплектации заказов. Они могут сканировать и проверять товары на этапах комплектации и упаковки, обеспечивая соответствие заказу и предотвращая ошибки. Сферой применения подобных систем является автоматизация электронных коммерческих складов, где точность и скорость выполнения заказов являются ключевыми факторами конкурентоспособности.

В логистике ССКЗ применяются для оптимизации процессов погрузки и разгрузки товаров, контроля за состоянием грузов и мониторинга транспортных средств. Они могут автоматически распознавать и отслеживать контейнеры, паллеты и отдельные товары, обеспечивая точный учет и контроль на всех этапах логистической цепочки. ССКЗ могут также использоваться для анализа изображений транспортных средств и выявления повреждений или нарушений условий транспортировки, что способствует повышению безопасности и сохранности грузов.

ССКЗ также играют важную роль в обеспечении точности и своевременности отслеживания грузов на всех этапах логистического процесса. Системы компьютерного зрения могут считывать информацию с транспортных этикеток и сопроводительных документов, что позволяет отслеживать движение товаров в режиме реального времени. ОСИ служит для управления поставками, где отслеживание грузов позволяет минимизировать задержки и предотвращать потерю товаров. Интеграция ССКЗ с системами управления транспортом обеспечивает централизованное управление и анализ данных о перемещении грузов, что способствует улучшению планирования и координации логистических операций.

Важным аспектом применения ССКЗ в логистике является их способность обеспечивать автоматическую проверку и верификацию информации. Например, системы компьютерного зрения могут считывать и проверять данные на транспортных документах, таких как накладные и счета-фактуры, обеспечивая точность и соответствие информации. Это позволяет избежать ошибок, связанных с ручным вводом данных, и снижает административные затраты. Верификация данных также помогает в обеспечении соответствия нормативным требованиям и стандартам, что особенно важно для международных перевозок и логистических операций.

ССКЗ могут быть интегрированы с роботизированными системами и автоматизированными транспортными средствами, путем использования дополнительных вычислительных модулей, что позволяет автоматизировать процесс перемещения товаров внутри логистических центров. Автоматизация перемещения товаров с использованием ССКЗ позволяет сократить затраты на трудовые ресурсы и повысить пропускную способность логистических центров.

Таким образом, специализированные системы компьютерного зрения являются неотъемлемой частью современных технологических решений, обеспечивая автоматизацию и повышение эффективности процессов в различных отраслях. Их использование позволяет достичь высокой точности и надежности в анализе визуальных данных, что способствует улучшению качества продукции,

повышению уровня безопасности и оптимизации производственных и логистических процессов [32].

1.2. Методы и алгоритмы, применяемые в специализированных системах компьютерного зрения

Основные методы, используемые в ССКЗ, можно разделить на четыре группы: поиск и выделение объекта, поиск текста, распознавание текста и классификация характеристик объекта.

Эти подходы разработаны различными группами исследователей и непрерывно дорабатываются для повышения точности и надежности распознавания визуальных данных [33, 34, 35, 36, 37, 38, 39, 40]. Прогресс в области глубокого обучения и компьютерного зрения позволяет значительно расширять функциональные возможности ССКЗ.

Методы поиска и выделения объекта направлены на идентификацию и изоляцию интересующего объекта на изображении или в видеопотоке. Для этого применяются алгоритмы шаблонного сравнения, машинного обучения, адаптивного порогового выделения, сегментации, фильтрации, извлечения признаков, контурные методы и детекторы объектов. В частности, данные методы часто используются для повышения точности обнаружения объектов. Исследования показывают, что использование таких методов позволяет достигать высокой точности в различных задачах, связанных с распознаванием объектов на изображениях [41, 42, 43, 44, 45]. Основные методы и алгоритмы их реализующие представлены в таблице 1.1.

Таблица 1.1 – Методы и алгоритмы поиска и выделения объекта.

Методы	Шаблонное сравнение Методы машинного обучения Методы адаптивного порогового выделения Методы сегментации Методы выделения признаков Алгоритмы фильтрации Извлечение признаков Контурные методы
Примеры алгоритмов	Normalized Cross-Correlation (NCC) Сверточные нейронные сети (CNN) Алгоритмы основных точек и дескрипторов, такие как SIFT или SURF Алгоритмы, основанные на методе кластеризации, k-means для сегментации изображений. Алгоритм адаптивного порогового выделения, метод Otsu. Использование ключевых точек (Harris corner detection) Операторы Собеля и Кэнни

Методы поиска текста предназначены для обнаружения текстовой информации на изображениях и в видеопотоках. В данной области применяются методы сегментации текста, детекции текста на основе границ, поиска горизонтальных и вертикальных линий и нейронные сети, в совокупности используются для повышения точности и скорости обнаружения текстовых областей [46, 47, 48]. Основные методы и алгоритмы реализующие их представлены в таблице 1.2.

Таблица 1.2 – Методы и алгоритмы поиска текста

Методы	Сегментация текста Детекция текста на основе границ (Stroke Width Transform) Поиск горизонтальных или вертикальных линий Нейронные сети
Примеры алгоритмов	Методы пороговой сегментации, использование адаптивного порога. Сегментация на основе цветовой информации. SWT для поиска текста на изображении. Применение фильтрации по направлению для выделения горизонтальных или вертикальных структур. Faster R-CNN, SSD (Single Shot Multibox Detector), EAST

Методы распознавания текста применяются для преобразования обнаруженных текстовых областей в машинно-читаемый формат. Оптическое распознавание символов является основным подходом в этой области. Современные системы OCR, используют рекуррентные нейронные сети (RNN), сети долгой краткосрочной памяти (LSTM), сверточные нейронные сети (CNN), методы шаблонного сопоставления, языковые модели для повышения точности распознавания в условиях низкого качества изображений и сложных шрифтов [49, 50, 51, 52, 53, 54, 55]. Основные методы и алгоритмы реализующие их представлены в таблице 1.3.

Таблица 1.3 – Методы и алгоритмы распознавания текста

Методы	Рекуррентные нейронные сети (RNN) и долгой краткосрочной памяти (LSTM) Использование сверточных нейронных сетей (CNN) Методы шаблонного сопоставления Использование языковых моделей
Примеры алгоритмов	Использование рекуррентных слоев в нейронных сетях. Применение сверточных слоев для извлечения локальных признаков из изображений с текстом. Использование шаблонных изображений символов для сравнения с обнаруженными символами. Применение языковых моделей для коррекции ошибок OCR и улучшения точности распознавания.

Методы классификации характеристик объекта направлены на анализ распознанной информации и определение ключевых характеристик объекта. Алгоритмы машинного обучения и глубокого обучения для классификации, использование правил и шаблонов, сопоставление с базой данных, графы и связи характеристик и их признаков применяются для определения типа объекта, его состояния, принадлежности к определенной категории и других атрибутов. Эти методы позволяют автоматизировать процесс принятия решений на основе визуальной информации, что подтверждается различными исследованиями [56, 57, 58]. Основные методы и алгоритмы реализующие их представлены в таблице 1.4.

Таблица 1.4 – Методы и алгоритмы классификации характеристик объекта

<p style="text-align: center;">Методы</p>	<p>Алгоритмы машинного обучения для классификации</p> <p>Использование правил и шаблонов</p> <p>Сопоставление с базой данных</p> <p>Использование графов и связей</p>
<p style="text-align: center;">Примеры алгоритмов</p>	<p>Метод опорных векторов (SVM), случайные леса, нейронные сети для задачи классификации текста.</p> <p>Создание словарей ключевых слов для каждой характеристики объекта и проверка их наличия в распознанном тексте.</p> <p>Создание базы данных, содержащей текстовую информацию и соответствующие характеристики.</p> <p>Графовые базы данных для представления и анализа связей между текстовой информацией и характеристиками объекта.</p>

Таким образом, специализированные системы компьютерного зрения включают в себя комплекс методов, направленных на поиск и выделение объектов, поиск и распознавание текста, а также классификацию характеристик объектов.

1.2.1. Методы поиска и выделения объекта

Методы поиска и выделения объекта в специализированных системах компьютерного зрения (ССКЗ) направлены на идентификацию и изоляцию

интересующего объекта на изображении или в видеопотоке. Эти методы необходимы для решения нескольких ключевых задач: определения движения в кадре, выделения фона и объекта, определения присутствия искомого объекта в кадре и отслеживания его движения в видеопотоке [59].

Определение движения в кадре включает анализ последовательности изображений для выявления изменений, указывающих на движение объектов. Для этого используются методы разности кадров. Разность кадров вычисляет разницу между последовательными кадрами, позволяя выявить движущиеся объекты. Методы данной категории оценивают направление и скорость движения каждого пикселя, предоставляя детализированную информацию о движении [60].

Для выделения фона и объекта применяются методы сегментации и адаптивного порогового выделения. Адаптивные пороговые методы, помогают различить фон и объект, даже при изменяющемся освещении. Сегментация изображения разделяет его на области, соответствующие фону и объектам, для идентификации и отслеживания [61, 62].

Для определения, находится ли в кадре искомый объект, применяются методы извлечения признаков, поиска дескрипторов и их сравнения. Дескрипторы, полученные из обрабатываемого кадра и шаблонного изображения искомого объекта позволяют распознать объекты на основе их характерных признаков. Эти методы обеспечивают высокую точность при идентификации объектов в различных условиях, в том числе при изменяющемся освещении и наличии искажений [63, 64, 65].

Отслеживание движения объекта включает непрерывное определение его положения в последовательности кадров. Эти алгоритмы предсказывают положение объекта в следующем кадре, что позволяет отслеживать его движение даже при частичных перекрытиях и изменениях условий съемки [66].

В специализированных системах компьютерного зрения, предназначенных для распознавания текстовых маркировок на продукции, вышеуказанные функции играют поисковую роль и направлены на обеспечение потока кадров, содержащих объект для распознавания на последующие этапы.

В контексте распознавания текстовых маркировок на движущемся конвейере, определение движения помогает системе зафиксировать момент, когда продукт с маркировкой попадает в поле зрения камеры. Это позволяет синхронизировать процесс съемки и анализа изображения, обеспечивая считывание текста в момент, когда объект появляется в кадре [67].

Отделение объекта, содержащего текстовую маркировку от фона и определение присутствия искомого объекта в кадре являются критически важными для обеспечения распознавания только тех кадров, которые содержат искомый объект, что позволяет повысить общую пропускную способность ССКЗ за счет пропуска кадров, не содержащих объект для распознавания.

В случае, когда продукция движется по конвейеру, отслеживание ее движения позволяет системе непрерывно следить за текстовой маркировкой, обеспечивая стабильное считывание текста независимо от скорости движения. Это особенно важно для производственных линий, где продукция движется с высокой скоростью.

1.2.2. Методы поиска текста

Методы поиска текста в специализированных системах компьютерного зрения (ССКЗ) предназначены для обнаружения текстовой информации на изображениях и в видеопотоках. Эти методы решают несколько ключевых задач: определение наличия текста в изображении, локализация текстовых областей слов и символов для последующего распознавания [68, 69].

Определение наличия текста в изображении. Первый шаг в процессе поиска текста заключается в определении, присутствует ли текст на изображении. Для этого используются методы пороговой сегментации машинного обучения и нейронные сети. Нейронные сети, такие как EAST (An Efficient and Accurate Scene Text Detector) [70], обученные на больших наборах данных изображений с текстом и без текста, могут классифицировать изображения на те, которые содержат текст, и те, которые не содержат. Сегментация текста делит изображение на отдельные регионы, среди которых выделяются текстовые области. Адаптивные методы

пороговой сегментации, такие как метод Otsu, и методы на основе нейронных сетей позволяют выделить текстовые области, даже если фон изображения неоднороден.

Локализация областей слов и символов. После определения наличия текста необходимо локализовать области, где находится текст. Для этого применяются методы сегментации текста, такие как детекция текста на основе границ, поиск базовой линии и методы поиска горизонтальных и вертикальных линий [71]. Детекция текста на основе границ (Stroke Width Transform, SWT) анализирует ширину линий на изображении, выделяя области с характерными для текста линиями. Этот метод применим для обнаружения текста в условиях низкого контраста и шумных изображениях. Для выделения текстовых строк часто используется фильтрация по направлению, которая помогает обнаружить горизонтальные или вертикальные структуры, характерные для текста. Такие алгоритмы, как операторы Кэнни и Собеля, позволяют выделять линейные структуры, соответствующие тексту. Поиск базовой линии текста позволяет обеспечить выделение текста на изображениях с вращением текста. Этот метод анализирует геометрические характеристики текста, чтобы определить его основные линии, что помогает корректировать и выделять текстовые области, даже если текст расположен под углом [72].

В контексте распознавания текстовых маркировок, определение наличия текста в кадре позволяет системе фильтровать участки изображения, содержащие текст. Локализация текстовых областей и выделение конкретных символов значительно повышает точность последующего распознавания. Разбиение текста на компоненты позволяет адаптировать работу нейросетей распознавания, так как нейронные сети, осуществляющие распознавание текста, специализируются на распознавании изображений, содержащих один символ либо одну строку текста, что снижает вероятность ошибок и увеличивает общую точность распознавания. Эти методы позволяют ССКЗ работать с высоким уровнем точности, в условиях наличия шумов и искажений на изображениях [73].

1.2.3. Методы распознавания текста

Методы распознавания текста в специализированных системах компьютерного зрения предназначены для преобразования обнаруженных текстовых областей в машинно-читаемый формат. В этой области применяются различные подходы, включая оптическое распознавание символов, использование сверточных нейронных сетей (CNN), рекуррентных нейронных сетей (RNN) и их комбинаций, таких как сверточная рекуррентная нейронная сеть (CRNN) [74, 75].

Сверточные слои (CNN) в CRNN используются для извлечения локальных признаков из изображения, таких как формы, углы и текстуры. Рекуррентные слои (RNN), такие как сети долгой краткосрочной памяти (LSTM), позволяют анализировать последовательность извлеченных признаков. Благодаря комбинации CNN и RNN, CRNN способна извлекать детализированные признаки текста и учитывать их последовательность, что значительно улучшает точность распознавания по сравнению с использованием только CNN или RNN. В том числе данное преимущество проявляется при распознавании поли-алфавитов и декоративных, нестандартных шрифтов [76].

CRNN показывает высокую устойчивость к шумам и искажениям на изображениях. Сверточные слои фильтруют локальные шумы и искажения, выделяя ключевые признаки текста, а рекуррентные слои позволяют анализировать последовательности даже при наличии искажений, сохраняя контекст и взаимосвязь между признаками [77, 78].

CRNN архитектура хорошо справляется с задачами распознавания текста в сложных условиях, таких как плохое освещение, низкое разрешение изображений, наличие различных шрифтов и углов наклона текста.

1.2.4. Методы классификации характеристик объекта

Методы классификации характеристик объекта в специализированных системах компьютерного зрения (ССКЗ) предназначены для анализа распознанной информации и определения ключевых характеристик объекта. Эти методы

применяются для определения типа объекта, принадлежности к определенной категории и других атрибутов (примером могут служить: серия, номер модели, характеристики продукции). Основные подходы включают использование алгоритмов машинного обучения, глубокого обучения, правил и шаблонов, сопоставление с базой данных, а также графовые модели [79, 57, 80, 81].

Алгоритмы машинного обучения, такие как метод опорных векторов (SVM), случайные леса и нейронные сети, широко применяются для классификации характеристик объекта.

SVM эффективно используется для классификации текстовых данных, разделяя объекты в многомерном пространстве на основе заданных характеристик [82]. Случайные леса объединяют множество решений деревьев для улучшения точности классификации, что позволяет эффективно обрабатывать большие объемы данных и учитывать различные признаки объекта. Глубокие нейронные сети (DNN), включая сверточные нейронные сети (CNN), способны извлекать и классифицировать сложные признаки объектов на основе обученных моделей [83].

Для классификации характеристик объектов также применяются методы на основе правил и шаблонов. Этот подход включает создание словарей ключевых слов для каждой характеристики объекта и проверку их наличия в распознанном тексте.

Создание правил и шаблонов для определения характеристик объекта позволяет точно идентифицировать и классифицировать объекты на основе текстовой информации. Например, определенные последовательности символов могут указывать на тип продукта или его состояние. Данный подход применим, когда объектом распознавания является продукция, содержащая набор характеристик со сходными значениями, что позволяет объединить их для создания шаблонов и правил классификации характеристик.

Для повышения точности классификации характеристик объектов применяется метод сопоставления с базой данных [84]. Этот подход включает создание базы данных, содержащей текстовую информацию и соответствующие характеристики объектов. При распознавании текстовых маркировок на продуктах,

информация из базы данных позволяет точно определить характеристики объекта, такие как дата производства, номер партии и другие атрибуты. Сопоставление распознанного текста с данными из базы позволяет обеспечить точную идентификацию и классификацию объектов, в условиях работы с ограниченным набором распознаваемых объектов.

Графовые модели используются для представления и анализа связей между текстовой информацией и характеристиками объекта. В графовых базах данных информация представляется в виде узлов и связей между ними, что позволяет эффективно анализировать сложные зависимости и взаимосвязи между различными характеристиками объекта [85]. Этот подход особенно полезен для анализа больших объемов данных и выявления закономерностей, в условиях, когда распознаваемые объекты содержат характеристики, которые в совокупности могут быть использованы, для однозначного определения характеристик.

Символьные характеристики касаются непосредственно текста и символов, представленных на изображении, и включают в себя формы символов, их взаимное расположение, а также используемые шрифты.

Методы на основе правил и шаблонов обеспечивают дополнительную проверку и коррекцию распознанной информации. Сопоставление с базой данных и использование графовых моделей позволяют обрабатывать большие объемы данных и учитывать сложные зависимости между различными характеристиками объектов. В совокупности данные методы позволяют точно классифицировать объекты, даже если текстовые маркировки содержат шумы и искажения.

На производственных линиях, где продукция движется с высокой скоростью, важно не только распознавать текстовые маркировки, но и правильно классифицировать объекты на основе распознанной информации. Это позволяет автоматизировать процессы идентификации продукции, отслеживания партий товаров и принятия решений на основе визуальной информации.

1.3. Архитектуры вычислительных модулей, используемых для обработки символьной информации на изображениях

Архитектуры вычислительных модулей являются критическим компонентом специализированных систем компьютерного зрения (ССКЗ), так как они определяют эффективность и скорость обработки символьной информации на изображениях. В контексте ССКЗ, используемых для распознавания текстовых маркировок на продукции, выбор архитектуры вычислительных модулей напрямую влияет на пропускную способность, точность и надежность системы. Важными факторами при выборе архитектуры являются типы операций, которые будут выполняться, объемы данных, требующие обработки, и необходимость в реальном времени выполнения задач [86, 87, 88, 89, 90].

Существует несколько основных групп архитектур, каждая из которых имеет свои преимущества и недостатки. В таблице 1.5 приведены основные группы архитектур вычислительных модулей, используемых в ССКЗ, а также их преимущества и недостатки.

Представленные группы архитектур вычислительных модулей являются параллельными позволяют значительно повысить пропускную способность модуля обработки изображений символьной маркировки. Параллельная обработка данных позволяет выполнять множество операций одновременно, что существенно ускоряет процесс анализа изображений и распознавания текста [91, 92].

Таблица 1.5 – Основные группы архитектур вычислительных модулей, используемых для обработки символьной информации на изображениях.

Архитектура	Преимущества и недостатки
Dataflow	<p>Преимущества: обработка данных выполняется по мере их доступности, что позволяет увеличить пропускную способность системы.</p> <p>Недостатки: сложность управления потоком данных и синхронизации операций.</p>
SIMD	<p>Преимущества: использование параллелизма для простых операций над пикселями.</p> <p>Недостатки: ограничения в применимости к сложным операциям обработки изображений, требующим доступа к различным частям изображения одновременно.</p>
MIMD	<p>Преимущества: гибкость и возможность параллельной обработки различных частей изображения.</p> <p>Недостатки: требуется сложная система управления и синхронизации процессоров.</p>
Систолический массив	<p>Преимущества: подходит для задач с постоянной скоростью обработки и постоянным потоком данных.</p> <p>Недостатки: ограничения в применимости к задачам с переменными потоками данных или операциями.</p>

Кроме того, существуют гибридные архитектурные решения, которые предполагают использование центрального процессора (CPU) и нескольких графических процессоров (GPU). В таких системах центральный процессор выполняет роль координационного центра, распределяя задачи между графическими процессорами, которые выполняют параллельную обработку данных.

Гибридные архитектуры CPU-GPU. В гибридных системах центральный процессор (CPU) и графические процессоры (GPU) работают совместно для выполнения различных типов задач. CPU обрабатывает задачи общего назначения и координирует работу всей системы, а GPU выполняет интенсивные вычисления, связанные с предобработкой изображений и обработкой информации с помощью нейросетей. Такая организация позволяет эффективно распределять нагрузку и повышать пропускную способность модуля [93, 94].

Графические процессоры используются для выполнения задач предобработки изображений, таких как фильтрация, сегментация и преобразования. Эти задачи могут быть параллелизованы и выполнены с большей точностью на GPU по сравнению с CPU. Использование GPU для предобработки данных позволяет подготовить изображения для последующего анализа.

Нейросетевые модели, такие как сверточные нейронные сети (CNN) и рекуррентные нейронные сети (RNN), требуют значительных вычислительных ресурсов. GPU, благодаря своей архитектуре, позволяют значительно ускорить применение нейросетей [95]. В гибридных системах GPU используются для выполнения задач распознавания текста, что позволяет повысить точность и скорость обработки данных [96].

Таким образом, гибридные архитектуры, сочетающие CPU и GPU, обеспечивают высокую пропускную способность и производительность специализированных систем компьютерного зрения [97]. Это достигается за счет эффективного распределения задач предобработки и обработки информации между центральным и графическими процессорами. Гибридный подход позволяет использовать преимущества каждой архитектуры для выполнения специфических задач, что делает системы более гибкими и производительными [98].

Модульный подход в архитектуре вычислительных модулей является перспективным направлением в разработке специализированных систем компьютерного зрения [89, 91, 99, 87], в контексте обработки символьной информации на изображениях. Использование модульной архитектуры позволяет

создавать системы, которые не только гибко адаптируются к различным задачам, но и обеспечивают высокую производительность за счет эффективного разделения процессов на вычислительные модули.

Каждый модуль в такой системе выполняет строго определенные задачи, что упрощает их разработку, тестирование и последующую интеграцию в общую систему. Применение модульного подхода особенно эффективно в гибридных архитектурах, где CPU и GPU работают в тандеме. В ССКЗ модули, выполняющие отдельные задачи, такие как распознавание текста, его обработка и дальнейшая передача в ГПС, работают независимо, что увеличивает общую производительность системы и обеспечивает возможность адаптации под различные требования и условия эксплуатации.

Таким образом, вычислительные модули на основе гибридных архитектур CPU-GPU предоставляют мощный инструмент для создания эффективных и масштабируемых ССКЗ в составе ГПС, что делает этот подход перспективным в контексте современных и будущих разработок в области обработки символьной информации.

Существуют патенты [100, 101, 102, 103, 104], в которых предлагаются решения задач обработки символьной информации (ОСИ) с использованием специализированных вычислительных модулей.

Устройство, представленное в патенте [100] осуществляет ОСИ для обеспечения компоновки текста документа на экране произвольного размера, и реализуется с использованием модульной структуры, представленной на рисунке 1.1.

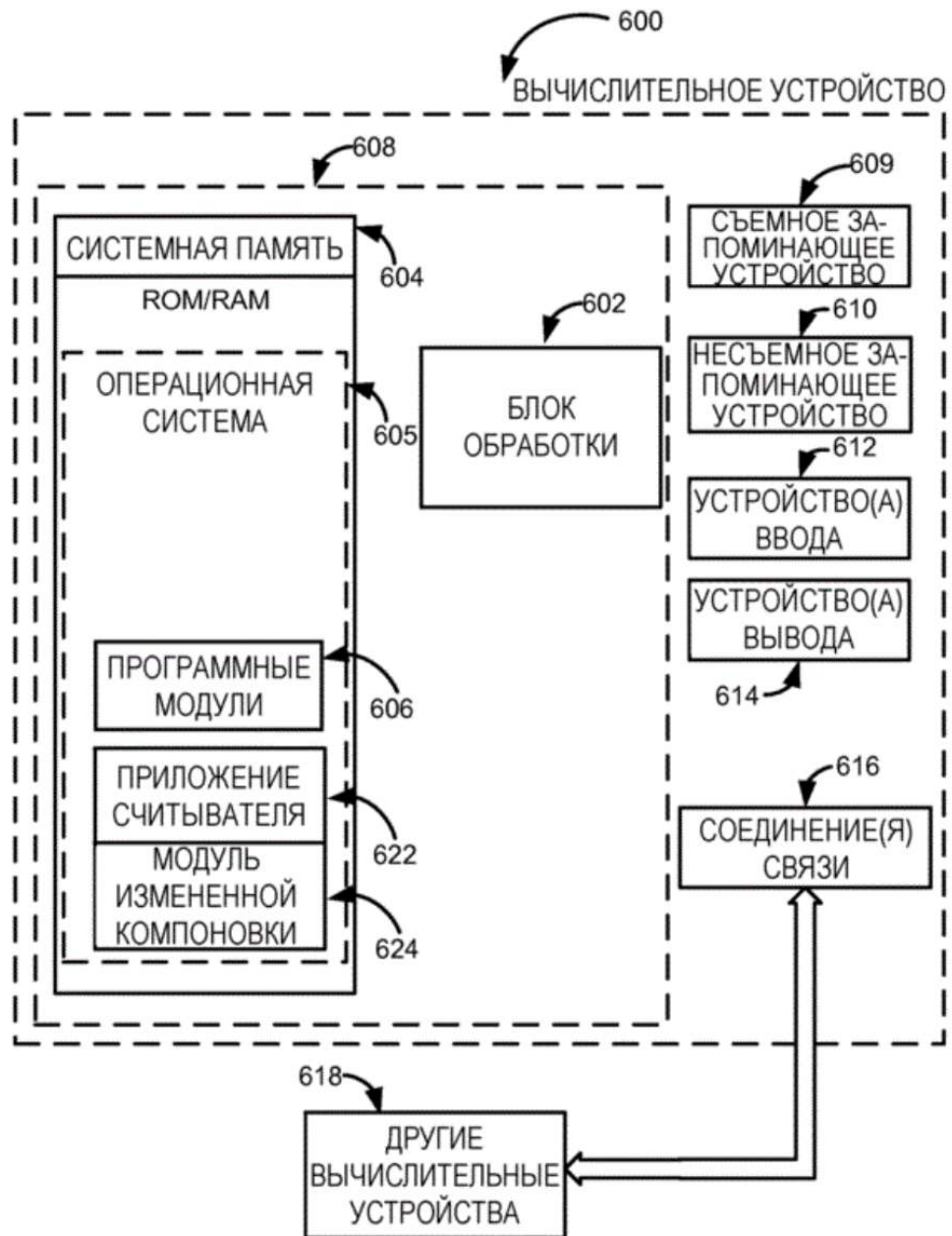


Рисунок 1.1 – Пример модульного устройства обработки СИ

Устройство, представленное в патенте [102] осуществляет верификации идентичности, которая удовлетворяет требованию уровня безопасности для обработки услуги онлайн. Устройство выполнено на модульной основе, обобщенная архитектура и состав модулей представлены на рисунке 1.2.

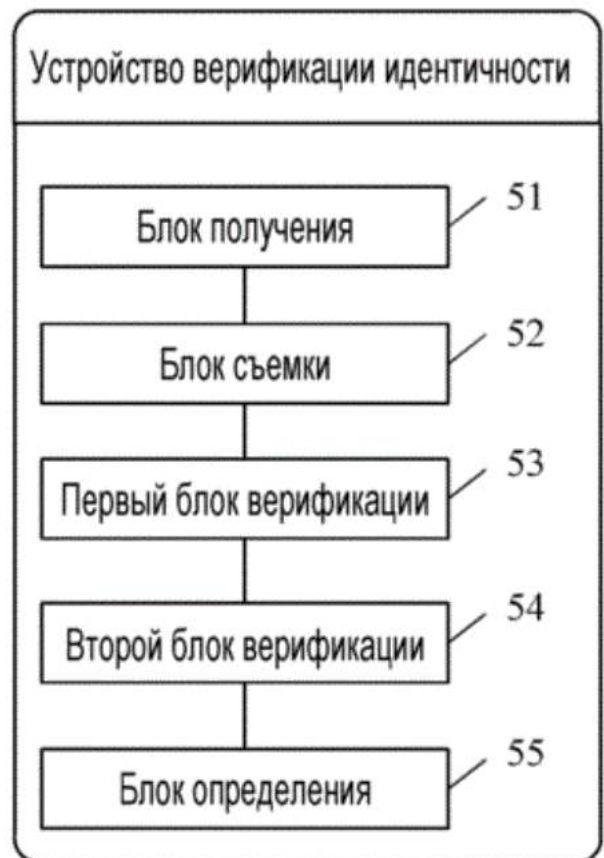
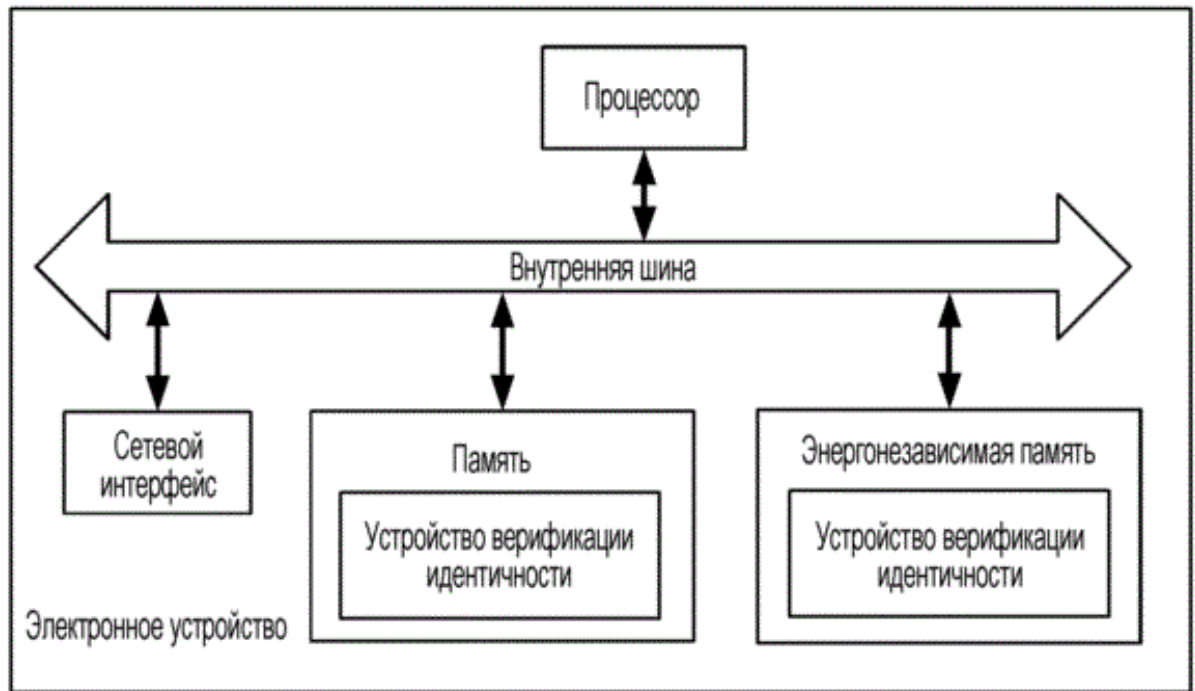


Рисунок 1.2 – Пример устройства осуществления верификации идентичности выполненного на модульной основе

Устройство, представленное в патенте [103] относится к средствам идентификации и инспекции вагонов. Технический результат достигается за

счет обработки СИ, нанесенной на транспортные вагоны, таким образом устройство может быть использовано в составе систем управления железнодорожным транспортом. Состав и взаимодействие модулей, используемых в устройстве представлен на рисунке 1.3.

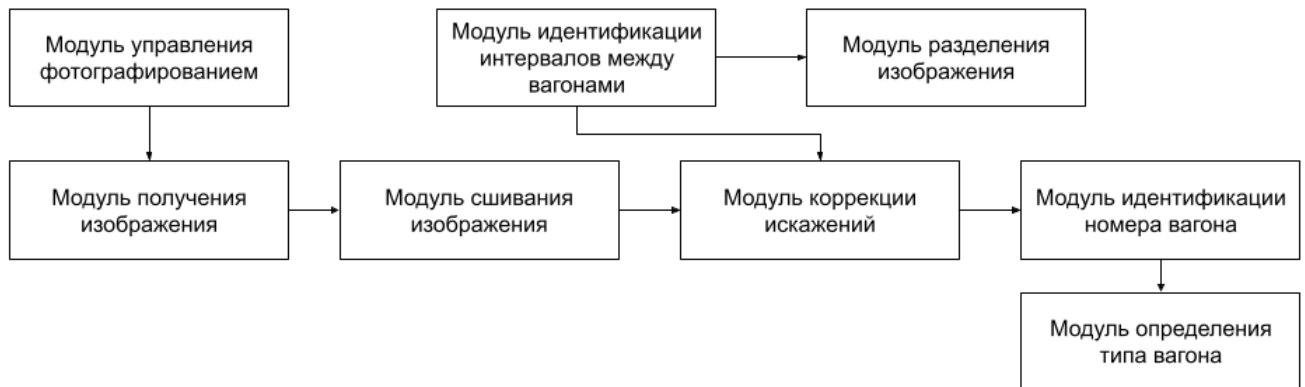


Рисунок 1.3 – Состав и взаимодействие модулей, использующихся в устройстве идентификации вагонов.

Эти патенты описывают различные методы и архитектуры, направленные на повышение эффективности и точности распознавания символов на изображениях. В частности, патенты охватывают широкий спектр технологий, от алгоритмов предобработки и сегментации изображений до применения нейросетевых моделей для классификации и распознавания символов. Такие патентные разработки демонстрируют активное развитие области и подтверждают высокую актуальность исследования архитектур вычислительных модулей для ОСИ.

1.4. Альтернативные методы распознавания характеристик объектов продукции

В качестве альтернатив специализированным системам компьютерного зрения для распознавания характеристик объектов продукции часто используются не только методы обработки и анализа изображений, но и альтернативные методы, такие как штрихкоды, QR-коды и RFID-метки [105, 106, 107, 108]. Эти методы позволяют автоматизировать процесс сбора и интерпретации данных о продукции.

Штрихкоды представляют собой одномерные графические коды, состоящие из параллельных линий разной ширины и промежутков между ними. Они содержат

информацию в цифровом формате, которая может быть считана с помощью лазерных или оптических сканеров.

Простота нанесения и считывания штрихкодов делают их распространенными в различных отраслях промышленности и торговли. Ограниченная емкость данных и чувствительность к повреждениям являются основными недостатками штрихкодов, так как поврежденные или загрязненные штрихкоды могут быть трудночитаемыми или нечитаемыми.

QR-коды представляют собой двумерные матричные штрихкоды, содержащие информацию в виде черно-белых квадратов. QR-коды способны хранить значительное количество данных, включая текстовую информацию и URL-адреса. Высокая емкость данных и устойчивость к повреждениям делают QR-коды эффективными для использования в различных приложениях. Они могут быть легко считаны с помощью обычных камер и специализированных сканеров, что упрощает их интеграцию в производства. Однако, для правильного считывания QR-коды должны иметь достаточный размер, что может ограничивать их применение на объектах маленького размера.

RFID-метки (Radio-Frequency Identification) используют радиочастотную идентификацию для хранения и передачи данных. Для использования RFID маркировки необходимо наличие метки (тега) и считывателя, который использует радиоволны для чтения информации с метки. RFID-метки могут быть считаны на расстоянии без необходимости прямой видимости, что позволяет считывать информацию через упаковку или другие непрозрачные материалы. Высокая емкость данных и возможность перезаписи делают RFID-метки эффективными для управления складскими запасами и автоматизации логистических процессов. Однако, внедрение RFID-систем может быть дорогостоящим из-за высокой стоимости меток и считывателей, а их работа может быть подвержена интерференции от других радиочастотных устройств, что снижает точность считывания.

Использование специализированных систем компьютерного зрения для ОСИ обладает рядом существенных преимуществ по сравнению с применением QR-

кодов, штрихкодов и RFID-меток. Эти преимущества делают ССКЗ более гибкими и универсальными для решения задач автоматизации и анализа продукции.

Преимущество ССКЗ заключается в их способности работать с различными типами информации. В отличие от QR-кодов, штрихкодов и RFID-меток, которые могут хранить ограниченное количество заранее закодированных данных, ССКЗ способны анализировать произвольные тексты, символы и графические элементы, присутствующие на изображениях. Это позволяет распознавать не только стандартные маркировки, но и более сложные символы и надписи, такие как логотипы, графические элементы, текстовые описания и другие визуальные признаки.

Второе преимущество состоит в гибкости ССКЗ в отношении различных условий эксплуатации. ССКЗ могут эффективно работать с изображениями, содержащими шум, искажения, загрязнения и повреждение нанесенных данных, плохое освещение или сложные фоны. Благодаря использованию методов обработки изображений и нейронных сетей, таких как сверточные нейронные сети (CNN) и рекуррентные нейронные сети (RNN), ССКЗ обеспечивают высокую точность распознавания даже в сложных условиях. Это делает их подходящими для применения на производственных линиях, где качество изображений может варьироваться.

ССКЗ обладают возможностью интегрирования в существующие производственные и логистические процессы. ССКЗ могут быть легко интегрированы в системы контроля качества, отслеживания продукции и управления запасами, предоставляя реальную информацию в режиме реального времени. Это позволяет автоматизировать процессы принятия решений, улучшить точность и скорость обработки данных, а также снизить затраты на ручной труд.

ССКЗ способны предоставлять комплексную информацию о характеристиках объектов продукции. Использование алгоритмов классификации, позволяющих получать сведения геометрического, топологического, цветового и символического характера, улучшает общую точность и пропускную способность

системы. Это позволяет более эффективно решать задачи, связанные с контролем качества продукции, управлением запасами и логистикой.

Таким образом, специализированные системы компьютерного зрения для обработки символьной информации обладают рядом значительных преимуществ по сравнению с альтернативными методами, такими как QR-коды, штрихкоды и RFID-метки, включающих способность работать с различными типами информации, гибкость в отношении условий эксплуатации, возможность интеграции в существующие процессы, высокую пропускную способность и предоставление комплексной информации о характеристиках объектов продукции. Это делает ССКЗ более универсальными инструментами для решения задач автоматизации и анализа продукции.

1.5. Противоречие между точностью и скоростью методов обработки символьной информации

Вопрос о противоречии между точностью и скоростью обработки данных является ключевым аспектом при разработке специализированных систем компьютерного зрения (ССКЗ) для обработки символьной информации (ОСИ). Данные, рассмотренные в предыдущих разделах, демонстрируют, что методы поиска и выделения объекта, поиска текста, распознавания текста и классификации характеристик объекта, как правило, либо решают свои задачи с высокой точностью, либо обеспечивают высокую пропускную способность [41].

Методы поиска и выделения объекта, такие как алгоритмы оптического потока и адаптивного порогового выделения, направлены на идентификацию и изоляцию интересующего объекта на изображении или в видеопотоке. Эти методы могут обеспечивать высокую точность при распознавании объектов, но требуют значительных вычислительных ресурсов и времени на обработку, что снижает их пропускную способность [46].

Методы поиска текста, включая детекцию текста на основе границ и использование нейронных сетей, также сталкиваются с проблемой баланса между точностью и скоростью. Алгоритмы, такие как Stroke Width Transform и нейронные

сети, обеспечивают высокую точность обнаружения текстовых областей, но могут быть медленными при обработке больших объемов данных.

Методы распознавания текста, использующие сверточные рекуррентные нейронные сети (CRNN), позволяют достигать высокой точности при распознавании текста на изображениях, содержащих шумы и искажения. Однако такие методы требуют значительных вычислительных ресурсов и времени на обучение и обработку, что снижает их пропускную способность [99, 34, 99, 91].

Методы классификации характеристик объекта, такие как алгоритмы машинного обучения и глубокого обучения, позволяют анализировать и интерпретировать информацию о геометрических, топологических, цветовых и символьных характеристиках объектов. Эти методы обеспечивают высокую точность классификации, но их сложность и требования к вычислительным ресурсам могут ограничивать их пропускную способность.

В качестве решения противоречия между точностью и скоростью обработки данных в специализированных системах компьютерного зрения (ССКЗ) для обработки символьной информации (ОСИ) предлагается использование конвейерно-параллельной архитектуры вычислительного модуля. Эта архитектура представляет собой интеграцию конвейерного и параллельного подходов к обработке данных, что позволяет существенно улучшить общую производительность системы.

Конвейерная обработка данных организует последовательность выполнения задач таким образом, что каждая стадия обработки выполняется одновременно с другими, но на разных этапах потока данных. Это значит, что как только один блок завершает свою операцию, он сразу же передает результат на следующий этап обработки, не дожидаясь завершения всей задачи. Таким образом, конвейерность позволяет существенно увеличить пропускную способность системы, уменьшая время ожидания и максимально используя доступные вычислительные ресурсы. В контексте ОСИ, конвейерный подход может быть применен к этапам обработки, начиная от предобработки изображений и заканчивая классификацией и

распознаванием символов. Это помогает сократить общее время обработки, не снижая точности.

Параллельная архитектура, в свою очередь, предполагает разделение задач на подзадачи, которые могут выполняться одновременно на различных вычислительных устройствах. В ССКЗ это достигается за счет использования современных графических процессоров, которые способны обрабатывать множество потоков данных параллельно. В ОСИ параллельная обработка может быть использована для одновременного распознавания текста на разных частях изображения или для параллельной работы нескольких моделей нейронных сетей, что значительно увеличивает скорость обработки без потери точности.

Интеграция этих подходов в единую конвейерно-параллельную архитектуру позволяет вычислительному модулю эффективно балансировать между точностью и скоростью, преодолевая ограничения каждого из методов в отдельности. Такая архитектура обеспечивает высокую точность распознавания символов и их классификации за счет использования сложных алгоритмов и моделей, при этом достигается высокая пропускная способность системы при обработке больших объемов данных.

Вывод по первой главе

Проведенный анализ показал, что вычислительная сложность средств распознавания и классификации символьных данных в условиях структурной неопределённости и искажений является критичной, что негативно сказывается на производительности работы ССКЗ в целом. Рассмотренные архитектуры реализации параллелизма при использовании НС-решений обеспечивают распределение вычислительных задач между различными вычислительными ресурсами, что позволяет эффективно использовать ресурсы аппаратного обеспечения и ускорить процесс обработки данных [109, 110].

В данной работе изучается возможность компенсации данных недостатков с целью повышения комплексной характеристики «точность × пропускная способность». Используемым подходом к решению данной проблемы

является использование гибридных архитектур вычислительных модулей, включающих центральные и графические процессоры. Такой подход позволяет распределять задачи предобработки изображений с помощью различных алгоритмов на центральный процессор и задачи обработки информации с помощью нейросетей на графические процессоры, что значительно увеличивает пропускную способность модуля.

Кроме того, внедрение конвейерно-параллельной архитектуры вычислительного модуля, позволяет эффективно использовать параллелизм и конвейеризацию для ускорения обработки данных. Это дает возможность обрабатывать большие объемы изображений, не теряя при этом в точности распознавания и классификации. Таким образом, рассматриваемый подход позволяет найти баланс между точностью и скоростью методов ОСИ, повышая общую комплексную характеристику «точность × пропускная способность» специализированных систем компьютерного зрения.

ГЛАВА 2. МАТЕМАТИЧЕСКАЯ МОДЕЛЬ И МЕТОД ОБРАБОТКИ ИЗОБРАЖЕНИЙ, СОДЕРЖАЩИХ СИМВОЛЬНУЮ ИНФОРМАЦИЮ

2.1. Математическая модель обработки изображений

Представлена математическая модель обработки изображений маркировок продукции, включающая все этапы обработки и анализа данных. Данная модель охватывает процессы выделения объектов, поиска и распознавания текста, а также классификации характеристик объектов на основе полученной информации. Модель разрабатывалась с учетом необходимости повышения комплексной характеристики «точность × пропускная способность» распознавания.

Предложены оригинальные частные модели, направленные на решение специфических задач, возникающих при обработке изображений маркировок продукции. Эти частные модели включают в себя методы предобработки изображений, алгоритмы выделения и классификации объектов, а также подходы к улучшению распознавания текста, учитывающие особенности различных типов маркировок и условий их применения [1, 3, 4].

Таким образом, данная модель предоставляет теоретическую основу для построения систем обработки изображений маркировок продукции, а также описывает подходы к реализации систем ОСИ с целью повышения их точности и пропускной способности.

Входными данными модели является кадр видеопотока (RGB изображение), выходными - классифицированная текстовая информация о характеристиках объекта.

Математическую модель обработки изображения M_{result} составляют следующие частные модели:

- математическая модель кадра видеопотока I_{src} ;
- математическая модель M_{roi} получения области изображения, содержащего объект I_{roi} ;
- функция $F_{\text{move}}(I_{\text{src}}, \vec{I}_{\text{src}})$ определения области изображения, содержащего объект;

- функция F_{contours} (F_{move}) поиска контуров объекта;
- функция F_{con} (F_{contours}) выделения объекта;
- функция F_{desc} (F_{con}) формирования дескрипторов объекта;
- функция F_{comp} ($F_{\text{desc}}, E_{\text{desc}}$) сравнения дескрипторов объекта с эталонными дескрипторами;
- математическая модель M_{obj} предобработки изображения объекта:
 - функция F_{hsv} (I_{roi}) трансформации RGB изображения в HSV изображение;
 - функция F_{bin} (F_{hsv}, L_b, L_t) бинаризации по нижней и верхней границам;
 - функция F_{filter} (F_{bin}) удаления шумов;
 - функция F_{dilate} (F_{filter}) дилатации изображения с уменьшенным уровнем шума;
 - функция F_{bord} (F_{dilate}) поиска границ объекта;
 - функция $F_{\text{transform}}$ ($F_{\text{bord}}, I_{\text{roi}}$) исправления аффинных и проекционных искажений;
 - математическая модель M_{text} поиска и распознавания текста:
 - функция F_{find} (I_{obj}) поиска области текста I_{text} ;
 - функция F_{base} (I_{text}) поиска базовой линии;
 - функция F_{split} ($F_{\text{base}}, I_{\text{text}}$) разбиения на символы;
 - функция F_{rec} (F_{split}) распознавания текста;
 - математическая модель M_{class} получения характеристик:
 - функция F_{info} (T, I_{text}) получения дополнительной информации геометрического и цветового характера, описывающей область текста;
 - функция F_e (T, E, F_{info}) сравнения признаков области текста и эталонных признаков;
 - функция F_{out} (F_e, T) формирования результата.

Математическая модель кадра изображения представлена в виде матрицы I_{src}

интенсивности цвета точек изображения в RGB представлении:

$$I_{src} = \begin{pmatrix} i_{0,0} & \cdots & i_{x_{src},0} \\ \vdots & \ddots & \vdots \\ i_{0,y_{src}} & \cdots & i_{x_{src},y_{src}} \end{pmatrix}, \quad (2.1)$$

где $i_{x_{src},y_{src}}$ – вектор, содержащий интенсивность цвета точки (ИЦТ) изображения в RGB представлении, используемый диапазон интенсивности находится в пределах между 0 и 255 (стандартное 8-битное изображение), x_{src} и y_{src} – размеры изображения в пикселях по вертикали и горизонтали.

Математическая модель M_{roi} получения области изображения, содержащего объект формирует двумерную матрицу I_{roi} размерности x_{roi} и y_{roi} , элементы которой представляют собой ИЦТ и в общем виде представляется как:

$$M_{roi} = F_{comp}(F_{desc}(F_{con}(F_{contours}(F_{move}(I_{src}, \vec{I}_{src}(t_{start}, t_{end}))))), E_{desc}(I_e)), \quad (2.2)$$

где \vec{I}_{src} – вектор, содержащий матрицы изображений I_{src} , полученные в моменты времени от t_{start} до t_{end} , E_{desc} – функция получения дескрипторов шаблонного изображения, I_e – матрица шаблонного изображения.

Функция $F_{move}(I_{src}, \vec{I}_{src})$ определения области изображения, содержащего объект формирует вектор размерности n_{move} , элементы которого представляют собой координаты точек, изменивших значения яркости относительно $\vec{I}_{src}(t_{start}, t_{end})$.

Функция $F_{contours}(F_{move})$ поиска контуров объекта формирует вектор C_{roi} , представляющий собой последовательность координат точек, описывающих границы объекта, для которых выполняется одно из следующих условий:

$$\begin{aligned} i_{move}(x_{move}, y_{move}) &= \{x_{move} = \max(x_0, x_1, \dots, x_n)\} \\ \text{или } i_{move}(x_{move}, y_{move}) &= \{x_{move} = \min(x_0, x_1, \dots, x_n)\} \end{aligned} \quad (2.3)$$

при $x_i \in F_{move}(I_{src}, \vec{I}_{src}(t_{start}, t_{end}))$ и $y_{move} = const$,

где i_{move} – элемента вектора C_{roi} , x_{move} , y_{move} – координаты точки по горизонтали и вертикали

Функция F_{con} (F_{contours}) выделения объекта формирует двумерную матрицу, элементы которой представляют собой ИЦТ, описывающие область, в котором находится объект и рассчитывается по выражению:

$$F_{\text{con}}(F_{\text{contours}}) = \begin{pmatrix} i_{\min_x(C_{roi}), \min_y(C_{roi})} & \cdots & i_{\max_x(C_{roi}), \min_y(C_{roi})} \\ \vdots & \ddots & \vdots \\ i_{\min_x(C_{roi}), \max_y(C_{roi})} & \cdots & i_{\max_x(C_{roi}), \max_y(C_{roi})} \end{pmatrix}, \quad (2.4)$$

где C_{roi} – вектор точек, в которых происходило движение, \min и \max – функции поиска точки с минимальным и максимальным значением координат по горизонтали и вертикали.

Функция F_{desc} (F_{con}) формирования дескрипторов объекта создает дескрипторы, представляющие собой двумерную матрицу, содержащую координаты 500 особых точек дескрипторов:

$$F_{\text{desc}}(F_{\text{con}}) = \begin{pmatrix} d_{x_1}, d_{x_2}, \dots, d_{x_{500}} \\ d_{y_1}, d_{y_2}, \dots, d_{y_{500}} \end{pmatrix}, \quad (2.5)$$

где d_x и d_y – координаты особых точек изображения, определяемых распределением по Гауссу относительно центра изображения.

Функция F_{comp} (F_{desc} , E_{desc}) сравнения дескриптора объекта с эталонным дескриптором формирует двумерную матрицу I_{roi} в случае, если удовлетворяется следующее условие:

$$\frac{E_{11} + E_{00}}{E_{00} + E_{01} + E_{10} + E_{11}} > T_{\text{desc}}, \quad (2.6)$$

где T_{desc} – пороговое значение соответствия дескрипторов; E_{01} – количество точек, для которых результат бинарного теста дескриптора области с объектом равен 0, а эталонного 1; E_{10} – количество точек, для которых результат бинарного теста дескриптора области с объектом равен 1, а эталонного 0; E_{11} – количество точек, для которых результат бинарного теста двух дескрипторов равен 1; E_{00} – количество точек, для которых результат бинарного теста двух дескрипторов равен 0; при этом бинарный тест τ определяется как:

$$\tau(d_x, d_x d'_x d'_y) = \begin{cases} 1, & \text{при } i_{d_x d_y} < i_{e_{d'_x d'_y}}; \\ 0, & \text{при } i_{d_x d_y} \geq i_{e_{d'_x d'_y}} \end{cases} \quad (2.7)$$

где d'_x и d'_y – координаты особых точек эталонного изображения, $i_{e_{d'_x d'_y}}$ – ИЦТ эталонного изображения.

Математическая модель M_{obj} предобработки изображения формирует изображение I_{obj} , представляющее собой двумерную матрицу, элементами которой являются ИЦТ и в общем виде может быть записана как

$$M_{obj} = F_{transform} \left(F_{bord} \left(F_{dilate} \left(F_{filter} \left(F_{bin} (F_{hsv}(I_{roi}), L_b, L_t) \right) \right) \right), I_{roi} \right). \quad (2.8)$$

Функция $F_{hsv}(I_{roi})$ трансформации RGB изображения в HSV изображение формирует двумерную матрицу изображения I_{hsv} :

$$I_{hsv} = \begin{pmatrix} i_{hsv_{0,0}} & \cdots & i_{hsv_{x_{hsv},0}} \\ \vdots & \ddots & \vdots \\ i_{hsv_{0,y_{hsv}}} & \cdots & i_{hsv_{x_{hsv},y_{hsv}}} \end{pmatrix}, \quad (2.9)$$

где $i_{hsv_{x_{hsv},y_{hsv}}}$ – вектор содержащий цветовой тон, насыщенность и яркость изображения в цветовой модели HSV, x_{hsv} и y_{hsv} – размерности изображения по горизонтали и вертикали.

Функция $F_{bin}(F_{hsv}, L_b, L_t)$ бинаризации по нижней и верхней границам преобразует I_{hsv} в двумерную матрицу I_{bin} , с элементами i_{bin} формируемыми по следующему условию:

$$i_{bin} = \begin{cases} 1, & \text{при } L_b \leq i_{hsv} \leq L_t \\ 0, & \text{в ином случае} \end{cases}. \quad (2.10)$$

Функция $F_{filter}(F_{bin})$ удаления шумов преобразует элементы I_{hsv} , формируя двумерную матрицу I_{filter} размерности I_{hsv} , элементы которой формируются как $g(x_g, y_g)$

$$g(x_g, y_g) = \sum_{u=-k}^k \sum_{v=-k}^k \frac{1}{2\pi\sigma^2} \exp\left(-\frac{u^2 + v^2}{2\sigma^2}\right) i_{bin_{x_g-u, y_g-v}}. \quad (2.11)$$

Функция $F_{dilate}(F_{filter})$ дилатации изображения с уменьшенным уровнем шума применяет к I_{filter} операцию дилатации с использованием матрицы T_{dilate}

$$T_{dilate} = \begin{pmatrix} 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 1 & 1 & 0 \\ 1 & 1 & 1 & 1 & 1 \\ 0 & 1 & 1 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 \end{pmatrix}. \quad (2.12)$$

Функция F_{bord} (F_{dilate}) поиска границ объекта формирует вектор C_{obj} размерности n_c , содержащий последовательность точек C_{obj} , описывающих границы объекта на изображении, полученном в результате F_{dilate} :

$$C_{obj} = (C_{obj_0}, C_{obj_1}, \dots, C_{obj_{n_c}}). \quad (2.13)$$

Функция $F_{transform}$ (F_{bord} , I_{roi}) исправления аффинных и проекционных искажений формирует двумерную матрицу I_{obj} , элементами которой являются вектора i_{obj} , содержащие ИЦТ таким образом, что:

$$I_{obj}(x_{obj}, y_{obj}) = I_{roi} \left(\frac{P_{00}x_{obj} + P_{01}y_{obj} + P_{02}}{P_{20}x_{obj} + P_{21}y_{obj} + P_{20}}, \frac{P_{10}x_{obj} + P_{11}y_{obj} + P_{12}}{P_{20}x_{obj} + P_{21}y_{obj} + P_{22}} \right), \quad (2.14)$$

где P – матрица трансформации размерности 3×3 , получаемая с использованием вектора границ объекта C_{obj} , x_{obj} и y_{obj} координаты точек по вертикали и горизонтали.

Математическая модель M_{text} поиска и распознавания текста формирует вектор, содержащий распознанный текст T и в общем виде может быть записана как:

$$M_{text} = F_{rec}(F_{split}(F_{base}(F_{find}(I_{obj}))), F_{find}(I_{obj}))). \quad (2.15)$$

Функция F_{find} (I_{obj}) поиска области текста I_{text} формирует вектор, содержащий координаты области текста

$$I_{text} = (T_{bb0}, T_{bb1}, \dots, T_{bbn})$$

$$T_{bbi} = (x_{left}, y_{top}, x_{right}, y_{bottom}), \quad (2.16)$$

где x_{left} и y_{top} , x_{right} и y_{bottom} – координаты, описывающие верхние левые, и правые нижние координаты области текста.

Функция F_{base} (I_{text}) поиска базовой линии формирует координаты линии, по которой будет происходить разбиение на символы.

Функция F_{split} (F_{base} , I_{text}) разбиения на символы использует базовую линию для поиска координат, описывающих область с символом I_{char} .

$$I_{char_n} = \begin{pmatrix} x_{charL} = na_{mx} - m_w/2 \\ y_{charT} = na_{my} - m_h/2 \\ x_{charR} = na_{mx} - m_w/2 \\ y_{charB} = na_{my} - m_w/2 \end{pmatrix}, \quad (2.17)$$

Где x_{charL} , y_{charT} , x_{charR} , y_{charB} – координаты левого верхнего и нижнего правого углов прямоугольника, описывающего символ текста, n – номер символа в строке; a_{mx} , a_{my} – координаты середины символа по горизонтали и вертикали, рассчитываемые как среднее от верхней, нижней и базовой линий текста; m_w , m_h – максимальный размер символов текста по горизонтали и вертикали.

Функция F_{rec} (F_{split}) распознавания текста формирует вектор T , содержащий распознанные символы текста.

Математическая модель M_{class} получения характеристик в общем виде представлена как

$$M_{class} = F_{out}(F_e(T, E, F_{info}(T, I_{text})), T). \quad (2.18)$$

Функция F_{info} (T , I_{text}) получения дополнительной информации геометрического и цветового характера, описывающего область текста формирует вектор:

$$F_{info}(T, I_{text}) = \begin{pmatrix} I_{color} \\ I_{size} \\ I_{relations} \end{pmatrix}, \quad (2.19)$$

где I_{color} – цветовая характеристика области текста, I_{size} – информация о размере области текста, $I_{relations}$ – информация о положении области текста относительно других.

Функция F_e (T , E , F_{info}) сравнения признаков области текста и эталонных признаков. Для расчета вводится матрица эталонных признаков характеристик (включающих цветовую и геометрическую информацию об области текста) E :

$$E = \begin{pmatrix} e_{0,0} & \cdots & e_{n,0} \\ \vdots & \ddots & \vdots \\ e_{0,q} & \cdots & e_{n,q} \end{pmatrix}, \quad (2.20)$$

где n – количество признаков, q – количество типов характеристик, e – эталонное значение признака.

Тогда значение F_e выражается как:

$$F_e(T, E, F_{info}(T, I_{text})) = \min \left(\text{dist}(T, F_{info}(T, I_{text}), E) \right), \quad (2.21)$$

где dist – функция расчета расстояния между эталонными признаками и признаками текста.

Функция $F_{out}(F_e, T)$ формирования результата создает вектор результата R , представляемый в виде:

$$R = (d, x, k), \quad (2.22)$$

где d – тип характеристики, x – текст, описывающий характеристику объекта, k – вектор критериев оценки, содержащий вычисленные расстояния между эталонными признаками и признаками текста.

Таким образом, математическая модель обработки изображения M_{result} представлена как:

$$M_{result} = M_{class}(M_{text}(M_{obj}(M_{roi}(I_{src}))))). \quad (2.23)$$

Новизна математической модели заключается в интеграции моделей поиска объектов, предобработки, распознавания текста и классификации характеристик в единую структуру, позволяя получить характеристики объекта, содержащего символьную маркировку.

2.2. Метод обработки изображений обработки изображений, содержащих символьную информацию

На основе разработанной математической модели разработан метод обработки изображений обработки изображений, содержащих символьную информацию для распознавания символьной маркировки (характеристик объекта) [2, 5, 6], состоящий из следующих этапов:

1. Прием кадра(ов) изображения и поиск двумерного объекта-контейнера (продукция, товар) (основан на выражениях и функциях 1-7).
2. Предобработка области изображения с найденным объектом-контейнером (основан на выражениях и функциях 8-14).
3. Поиск и распознавание символьной маркировки внутри объекта-контейнера (основан на выражениях и функциях 15-17).
4. Классификация характеристик объектов и формирование их геометрических, топологических, цветовых и текстовых признаков (основан на выражениях и функциях 18-22).

Первый этап является проверочным, его цель – поиск объекта-контейнера и проверка его геометрической формы на соответствие заданным шаблонам видов продукции (товара). В результате принимается решение о дальнейшей обработке кадра(ов) изображения при условии, что он содержит объект для распознавания. Этап отличает применение двумерных дескрипторов исходного объекта и заданных шаблонов, основанных на редукции количества проверяемых пикселей из матриц изображений. Формирование и сравнение двумерных дескрипторов реализовано по методу ORB (Oriented FAST and Rotated BRIEF) [111], что позволяет определить наличие объекта для распознавания, содержащего символьную маркировку.

Применение ORB в данном контексте обусловлено его высокой скоростью и устойчивостью к изменениям масштаба и поворота объектов. Это позволяет идентифицировать объект-контейнер, даже если его положение или ориентация на изображении меняются. Процесс сравнения дескрипторов основан на редукции количества пикселей, что уменьшает вычислительную нагрузку и повышает общую пропускную способность системы.

При успешной идентификации объекта-контейнера система переходит к следующему этапу обработки. Если объект не соответствует заданным шаблонам, кадры исключаются из дальнейшего анализа, что предотвращает ненужные вычисления и экономит ресурсы. Таким образом, данный этап обеспечивает идентификацию объекта-контейнера, что позволяет выполнять обработку только релевантных изображений, увеличивая пропускную способность системы.

Второй этап является подготовительным, он осуществляет предобработку изображений, включая компенсацию аффинных искажений и искажений перспективы на изображениях области объекта. В результате определяются прямоугольные блоки изображения, содержащие текст для последующего распознавания. Этап отличает применение метода Гаусса для удаления шумов и выделение границ прямоугольных блоков изображений на основе комбинации методов Канни и Рамера-Дугласа-Пекера [112, 113].

На данном этапе проводится компенсация аффинных искажений и искажений перспективы, которые могут возникнуть при съемке. Данные искажения могут значительно затруднить процесс распознавания текста, поэтому их коррекция необходима для обеспечения точности дальнейшего анализа. Преобразования выполняемые на данном этапе включают в себя такие трансформации, как масштабирование, поворот и исправление перспективы, которые выполняются с помощью соответствующих матричных преобразований.

Далее осуществляется сегментация изображения для выделения прямоугольного блока, содержащего текст. На данном этапе производится удаление шумов с помощью статистического метода, основанного на алгоритме Гаусса. Этот метод позволяет сгладить изображение, устранив случайные шумы, которые могут помешать корректной идентификации границ текстового блока на изображении.

После удаления шумов, границы прямоугольного блока, содержащего текст (объект для распознавания), выделяются с использованием комбинации методов Канни и Рамера-Дугласа-Пекера. Метод Канни применяется для детектирования границ на изображении. Алгоритм Рамера-Дугласа-Пекера, в свою очередь, используется для аппроксимации этих контуров, что позволяет выделить прямоугольные области, содержащие текст.

Выходными данными этого этапа, являются определенные области изображения, содержащие текст (объект для распознавания). Этот блок впоследствии используется на следующем этапе для распознавания текста. Точная сегментация и предобработка изображений на данном этапе важны для минимизации ошибок распознавания на следующих этапах.

Третий этап является основным, он состоит в поиске и выделении из матрицы пикселей требуемых областей текста и их последующем посимвольном распознавании. Области текста выделяются с помощью нейронной сети с архитектурой EAST [70], что позволяет сегментировать изображение, содержащие области текста произвольной структуры. Разделение текста на символы реализовано методом поиска базовой линии строк, что сокращает ошибки

распознавания первого рода за счет уменьшения влияния межсимвольной дистанции и перекрытия символов на точность разбиения для изображений, содержащих различный шрифт и стиль текста, а также курсивное начертание. Посимвольное распознавание символов реализовано нейронной сетью с архитектурой сверточной рекуррентной сети CRNN.

На этом этапе области текста выделяются с помощью нейронной сети с архитектурой EAST (Efficient and Accurate Scene Text Detector). Применение данной архитектуры позволяет сегментировать изображение и выделить текстовые области произвольной структуры. EAST эффективно справляется с задачей локализации текста, даже если текст представлен в различных шрифтах, размерах и стилях. Это достигается за счет анализа изображения на нескольких уровнях и комбинирования результатов для формирования точных границ текстовых областей.

После выделения текстовых областей, изображение разделяется на отдельные символы. Для этого используется метод поиска базовой линии строк. Этот метод позволяет определить линию, вдоль которой располагаются символы текста. Такой подход существенно снижает ошибки распознавания первого рода, так как уменьшается влияние межсимвольной дистанции и перекрытия символов на точность разбиения, для изображений, содержащих текст с различным шрифтом, стилем и курсивным начертанием. Посимвольное распознавание символов реализуется сверточной рекуррентной нейронной сетью (CRNN) [75], обеспечивающей высокую точность распознавания даже при наличии шумов и искажений.

Особенность этапа – потенциальный параллелизм за счет независимой посимвольной обработки нескольких кадров. Выходные данные третьего этапа – символьная последовательность.

Четвертый этап является завершающим. Он необходим для классификации характеристик исходных объектов по их геометрическим, топологическим и цветовым характеристикам, что позволяет получить текстовую информацию о характеристиках объекта. Конструктивный вид функции классификации

представлен количественно-разностной функцией невязки между эталонными и выделенными признаками.

Основная задача этого этапа – классификация исходных объектов по их геометрическим, топологическим, цветовым и текстовым характеристикам. Эти признаки позволяют идентифицировать объекты и определить их состояние, тип и другие важные параметры.

Геометрические признаки включают в себя размеры, форму и пропорции объекта. Они определяются на основе анализа контуров и границ объекта, выделенных на предыдущих этапах. Топологические характеристики связаны с внутренней структурой и взаимосвязями между элементами объекта. Эти сведения позволяют понимать, как различные части объекта связаны друг с другом, что используется при распознавании сложных символов или текстов, содержащих взаимосвязанные элементы. Топологические признаки помогают в улучшении точности распознавания и классификации характеристик объектов.

Цветовые характеристики включают информацию о цветовом распределении и интенсивности на изображении. Цветовые признаки могут быть использованы для различения объектов с похожей формой, но разными цветами, а также для выделения текста на фоне различных цветов. Это помогает повысить точность распознавания, особенно в условиях низкого контраста или сложных цветовых комбинаций. Например, цвет может быть использован для определения различных маркировок, которые имеют стандартизированные цветовые схемы.

Текстовые признаки включают в себя распознанную символьную информацию, которая была получена на третьем этапе. Эти данные являются значениями характеристик идентификации объекта. Текстовая информация может содержать коды, номера, названия, описания и другие данные, которые необходимы для определения принадлежности характеристики к определенному типу информации, описывающей объект.

Для классификации объектов используется количественно-разностная функция невязки между эталонными и выделенными признаками. Эта функция сравнивает текущие признаки объекта с эталонными значениями, настраиваемыми

для конкретных распознаваемых объектов. Эталонные значения представляют собой заранее определенные характеристики для различных типов объектов. На основе этого сравнения принимается решение о классификации характеристик объекта и формируется итоговый результат.

Пример изображения маркировки, при распознавании которой сведения геометрического, топологического, цветового характера могут быть использованы для классификации характеристик объекта представлен на рисунке 2.1. На изображении представлен модуль контроля запуска полетного двигателя. В данном случае на основании информации о цвете, разнице в размерах шрифта и взаимном расположении можно сделать вывод о принадлежности областей текста к определенной характеристике объекта.



Рисунок 2.1 – Пример объекта, где сведения геометрического, топологического, цветового характера могут быть использованы для классификации характеристик

На данном изображении текст представлен в синем и черном цветах на белом фоне, что создает высокий контраст и облегчает выделение и распознавание текста. Синий цвет используется для выделения названия продукта «SmartStart» и логотипа производителя «TCW Technologies», что делает возможным однозначно идентифицировать данные элементы.

На данном примере, крупный шрифт используется для названия продукта «SmartStart». Средний шрифт используется для указания модели «SS-12v», а шрифт меньшего размера – для технических характеристик и информации о производителе. Эти различия в размере шрифта помогают классифицировать характеристики объектов, представленных на изображении.

В данном примере, текст «SmartStart» располагается в верхней части изображения. Под ним расположен текст «MODEL: SS-12v», а еще ниже – технические характеристики и информация о производителе. Такое расположение текстовых блоков помогает логически структурировать информацию, обеспечивая возможность использования топологических эталонных признаков для классификации характеристик объекта.

Применение этого метода позволяет получить высокую точность классификации характеристик за счет комплексного анализа множества признаков. Таким образом, четвертый этап метода обработки изображений завершает процесс анализа, обеспечивая классификацию объектов и формирование их геометрических, топологических, цветовых и текстовых признаков. Это позволяет получить полную и точную информацию о каждом объекте.

Новизна метода заключается в применении двумерных дескрипторов и выделения объекта на основе комбинации методов Канни и Рамера-Дугласа-Пекера, а также применением рекуррентной сверточной нейронной сети, позволяющей вести комплексную обработку пространственно-цветовых характеристик областей текста, таким образом формируя завершённый цикл обработки изображений объектов с символьной маркировкой.

Вывод по второй главе

Разработана математическая модель обработки изображения, отличающаяся использованием последовательности частных математических моделей для выделения объекта, предобработки изображения объекта, распознавания текста и классификации характеристик объекта, в совокупности позволяющая получить характеристики объекта, содержащего символьную маркировку.

Новизна разработанного метода заключается в применении двумерных дескрипторов и выделения объекта на основе комбинации методов Канни и Рамера-Дугласа-Пекера, а также применением рекуррентной сверточной нейронной сети, что позволяет сформировать завершённый цикл обработки изображений объектов с символьной маркировкой и интегрировать ССКЗ в распределённые вычислительные системы мониторинга и обработки автоизображений, информационно-аналитические видеосистемы, массивные видеоподсистемы в составе ГПС. Данный метод применим для организации арифметической, логической и символьной обработки данных, хранения и ввода-вывода информации.

ГЛАВА 3. КОНВЕЙЕРНО-ПАРАЛЛЕЛЬНЫЙ АЛГОРИТМ ОБРАБОТКИ ИЗОБРАЖЕНИЙ, СОДЕРЖАЩИХ СИМВОЛЬНУЮ ИНФОРМАЦИЮ

3.1. Общая структура алгоритма

Конвейерно-параллельный алгоритм обработки изображений, содержащих СИ, предназначен для распознавания символьной маркировки продукции. Алгоритм основан на разработанной модели и методе, обеспечивая возможность распараллеливания процесса обработки, настройку к различным видам характеристик и устойчивость к искажениям, что обеспечивает повышенную точность распознавания [3, 7, 10, 11]. Блок-схема алгоритма представлена на рисунках 3.1, 3.2, 3.3 и 3.4.

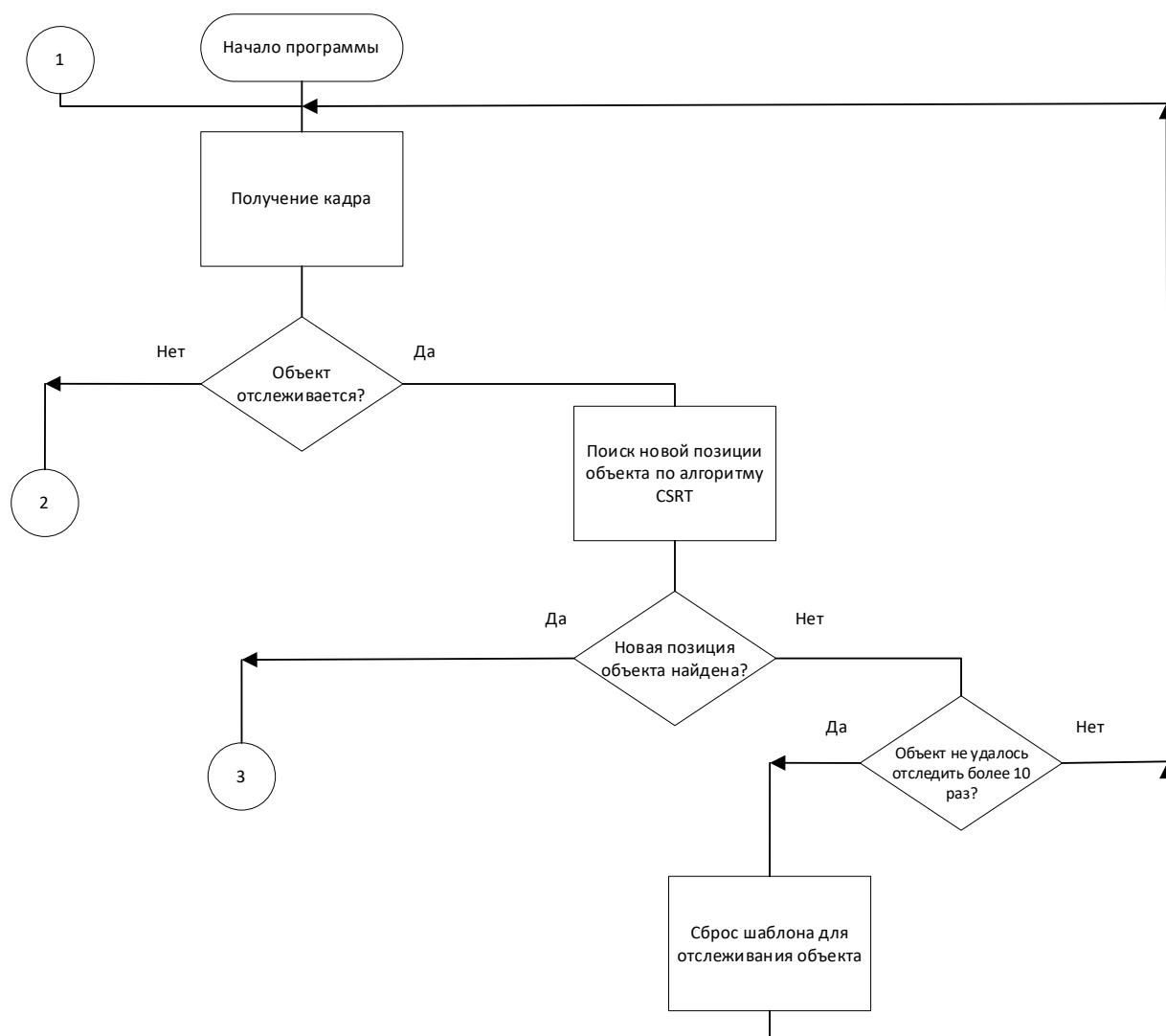


Рисунок 3.1 – Блок-схема конвейерно-параллельного алгоритма обработки изображений, содержащих СИ

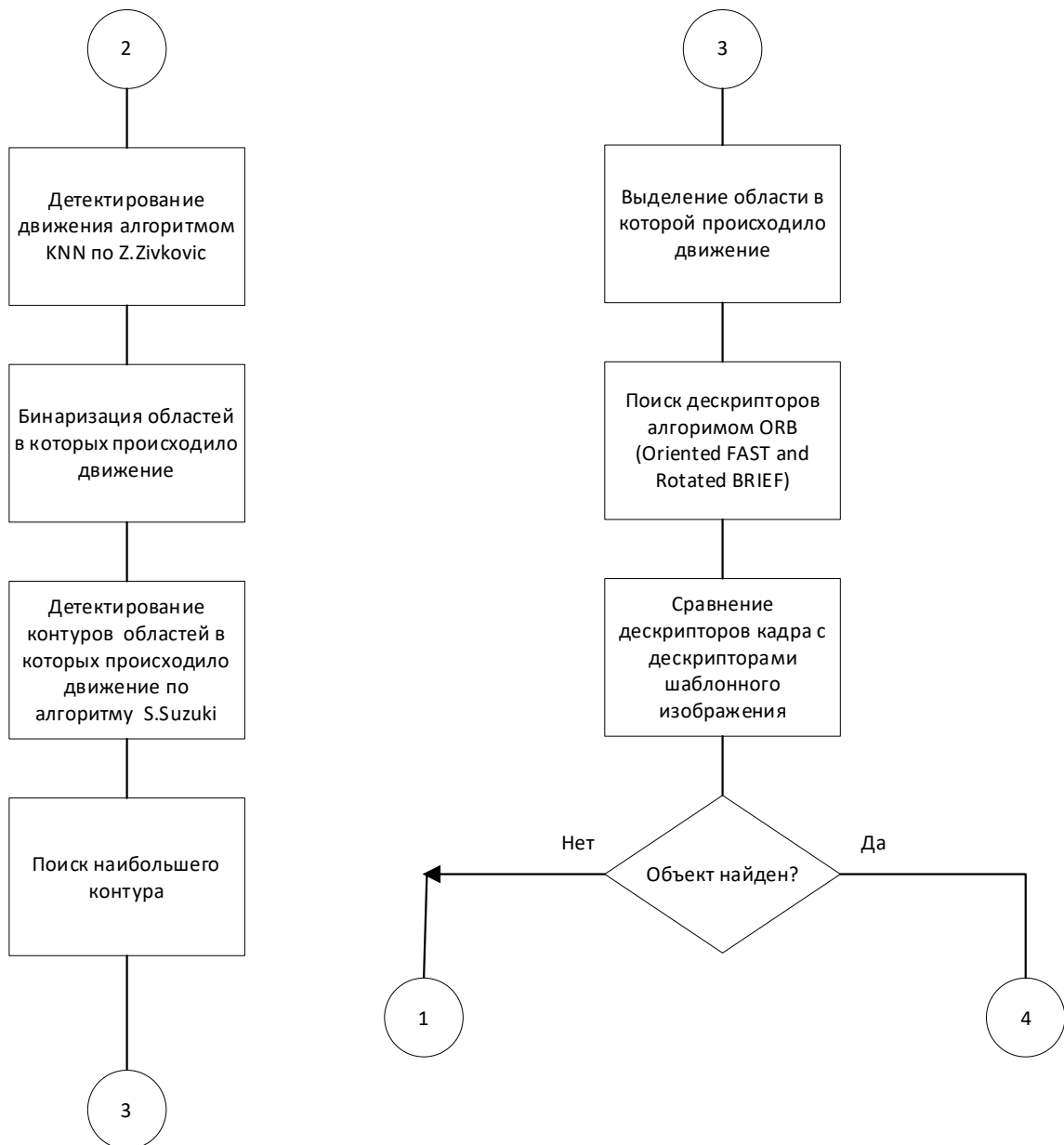


Рисунок 3.2 – Блок-схема конвейерно-параллельного алгоритма обработки изображений, содержащих СИ (продолжение)

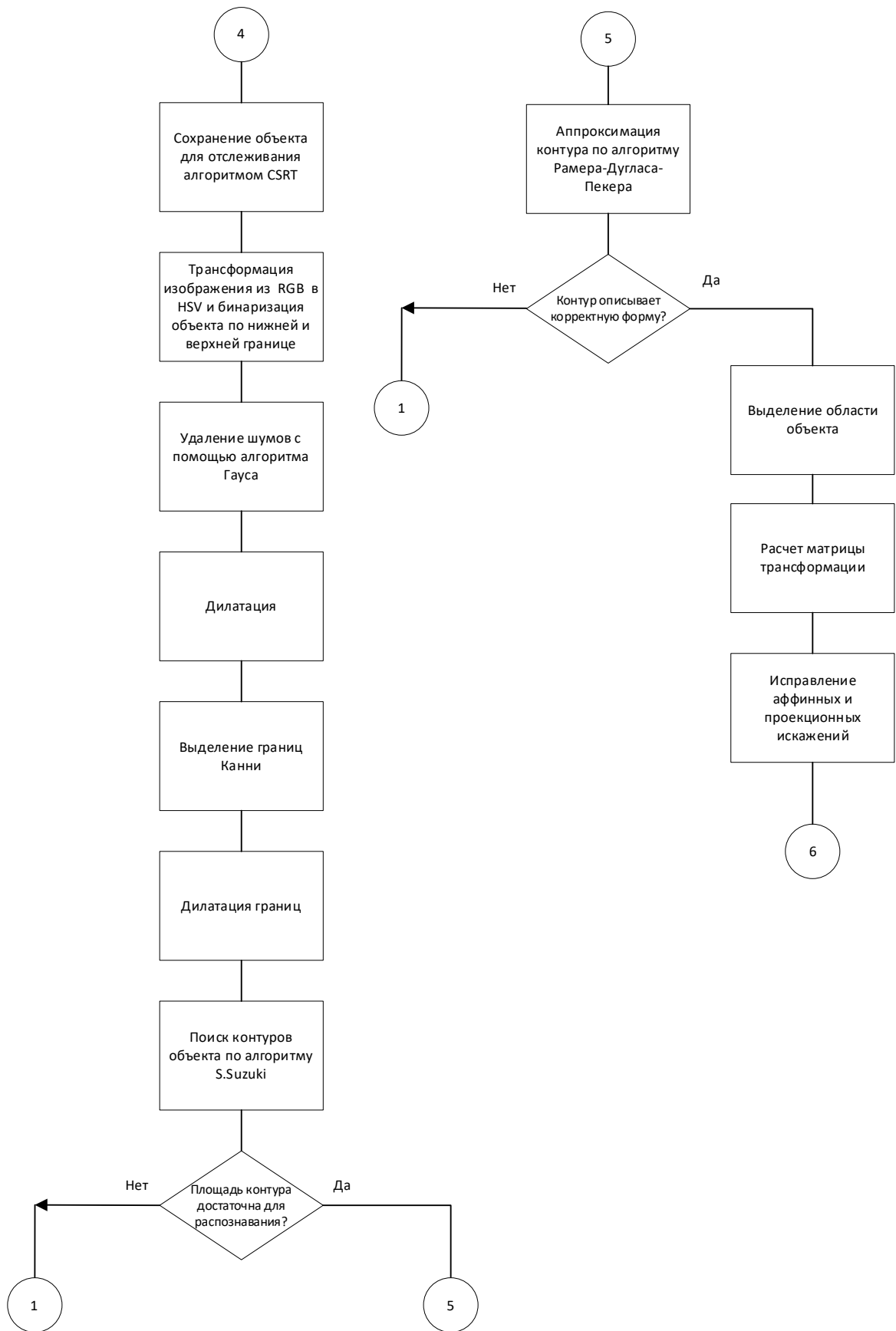


Рисунок 3.3 – Блок-схема конвейерно-параллельного алгоритма обработки изображений, содержащих СИ (продолжение)

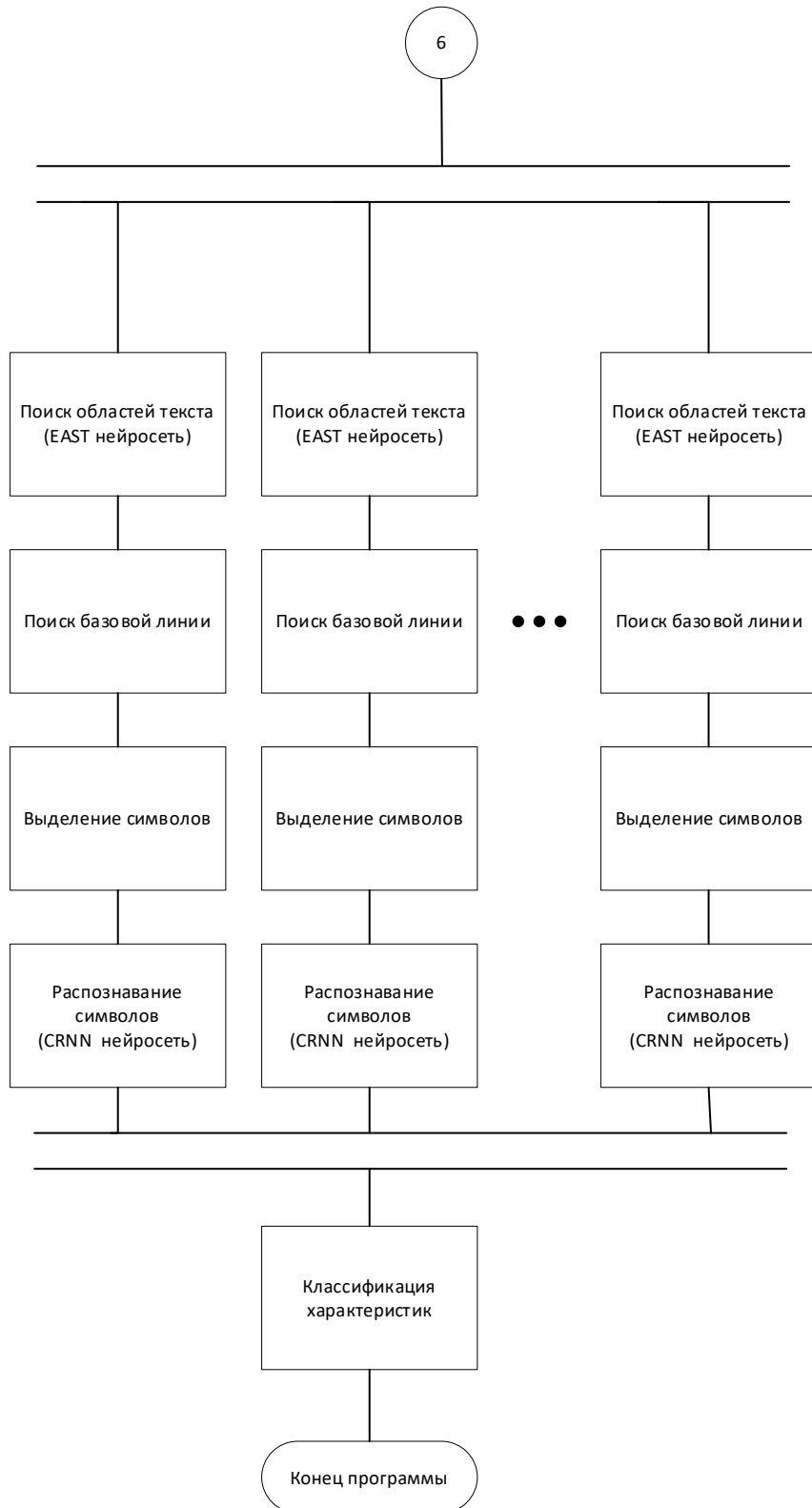


Рисунок 3.4 – Блок-схема конвейерно-параллельного алгоритма обработки изображений, содержащих СИ (продолжение)

Операции выполняемые в процессе работы алгоритма могут быть разбиты на 6 основных групп, каждая из которых выполняет свои задачи и взаимодействует с другими в рамках конвейерно-параллельной обработки.

Первой группой операций алгоритма является **поиск объекта для распознавания** на входных изображениях. Этот процесс включает анализ изображения для определения наличия объекта-контейнера, который может содержать символьную маркировку. Используются методы обработки изображений, такие как поиск области изображения, в которой происходило движение (алгоритм KNN, по Z.Zivkovic [60]), выделение контуров (алгоритм S.Suzuki [114]), поиск и сравнение дескрипторов (ORB) [111] и определения потенциальных объектов для дальнейшего анализа.

После определения объекта-контейнера, вторая группа операций включает **отслеживание его положения** в последовательности кадров. Этот процесс критически важен для обеспечения непрерывности обработки в реальном времени, особенно в условиях движения объектов, например, на производственной линии. Алгоритм Channel and Spatial Reliability Tracker (CSRT) [66] используется для точного определения положения объекта в каждом кадре и сигнализации, в случае если объект для распознавания вышел за пределы области обзора камеры, осуществляющей захват кадров.

Третья группа операций включает **выделение объекта-контейнера для распознавания** из фона и других объектов на изображении. На этом этапе используются методы обработки изображения, такие как удаление шумов (алгоритм Гаусса), дилатация, выделение границ (метод Канни), поиск контуров (алгоритм S.Suzuki), аппроксимация границ методом Рамера-Дугласа-Пекера, проверка формы объекта и исправление аффинных и проекционных искажений, для точного выделения границ объекта и компенсации шумов и искажений на исходном изображении. Это позволяет изолировать объект и подготовить его для дальнейшего анализа.

Четвертая группа операций фокусируется на **поиске областей текста** внутри выделенного объекта-контейнера и **выделении отдельных символов**. На данном

этапе используется нейросеть с архитектурой EAST для распознавания паттернов областей изображения, содержащих текстовую информацию [70]. После этого, алгоритмы разделения символов, такие как поиск базовой линии и анализ расстояний между символами, применяются для точного выделения каждого символа.

В пятой группе операций происходит непосредственно **распознавание текста**. Для этого используются современные методы машинного обучения, такие как сверточные рекуррентные нейронные сети (CRNN) [75], которые способны эффективно распознавать текстовые символы даже в условиях шума и искажений. Результатом этого процесса является последовательность символов, представляющая текстовую информацию на объекте.

Завершающая группа операций **включает классификацию характеристик объекта** на основе полученной текстовой информации и других характеристик, таких как геометрические, топологические и цветовые признаки. Используются методы классификации, такие как количественно-разностные функции невязки, для определения типа объекта и его характеристик. Это позволяет получить полную информацию о каждом объекте и сформировать результат работы алгоритма.

Параллельная обработка выполняется на этапах поиска символов и их распознавания. Это позволяет значительно ускорить процесс обработки, так как несколько кадров могут обрабатываться одновременно. Преимущества этого подхода включают сокращение времени обработки и возможность обработки больших объемов данных за единицу времени. Остальной процесс обработки работает по конвейерному принципу. Совмещение конвейерного и параллельного подходов к обработке предоставляет значительные преимущества, такие как повышение общей пропускной способности системы, улучшение эффективности обработки за счет одновременного выполнения различных задач, и возможность последовательного получения информации, соответствующей порядку поступления кадров.

3.2. Поиск объекта для распознавания

Поиск объекта для распознавания является первой из групп операций в конвейерно-параллельном алгоритме обработки изображений, содержащих СИ. Этот процесс включает в себя анализ входных изображений для определения наличия объекта-контейнера, который потенциально содержит символьную маркировку. Для этого используются различные методы обработки изображений, такие как анализ движения, выделение контуров и сравнение дескрипторов.

Для выделения области изображения, в которой произошло движение применяется алгоритм KNN по Z. Zivkovic [60]. Этот алгоритм анализирует последовательность кадров и определяет области, где произошли изменения, что указывает на наличие объекта в движении. Обнаружение движущихся объектов позволяет сосредоточить дальнейшую обработку на релевантных частях изображения, что повышает точность всей системы позволяя обрабатывать ограниченную область исходного кадра, способствуя уменьшению количества ошибок первого рода при выделении и распознавании текста, и повышает пропускную способность системы за счет уменьшения объема обрабатываемой информации. Исходный кадр и пример работы алгоритма KNN представлен на рисунке 3.5.

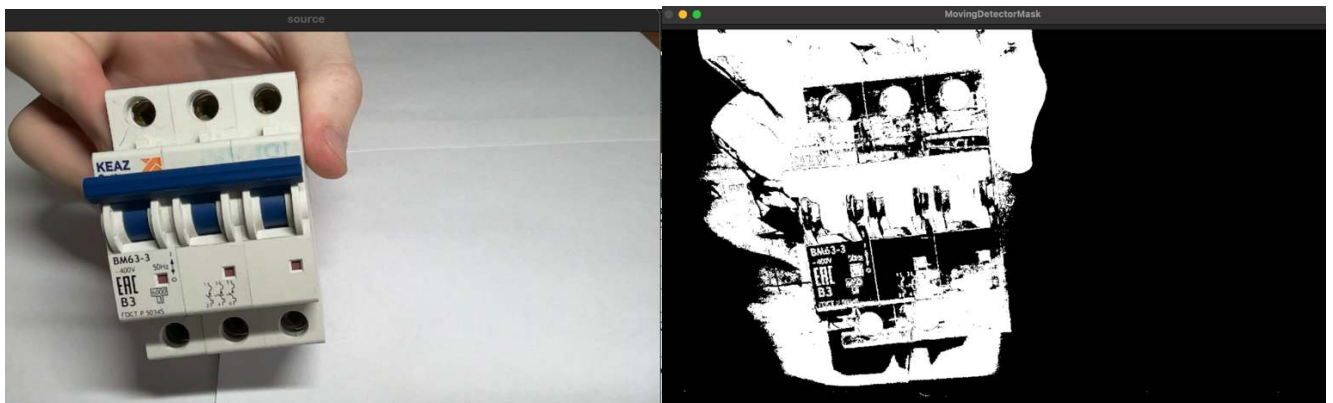


Рисунок 3.5 – Исходный кадр и пример работы алгоритма KNN.

Далее используется алгоритм S. Suzuki для выделения контуров объекта. Этот метод позволяет определить границы движущегося объекта в видеопотоке, что является важным шагом для его последующего выделения и анализа. Данный алгоритм позволяет увеличить точность сравнения дескрипторов движущегося

объекта с шаблонными дескрипторами, за счет выполнения поиска дескрипторов только у релевантной части изображения, без учета фона. Пример работы алгоритма выделения контуров представлен на рисунке 3.6.

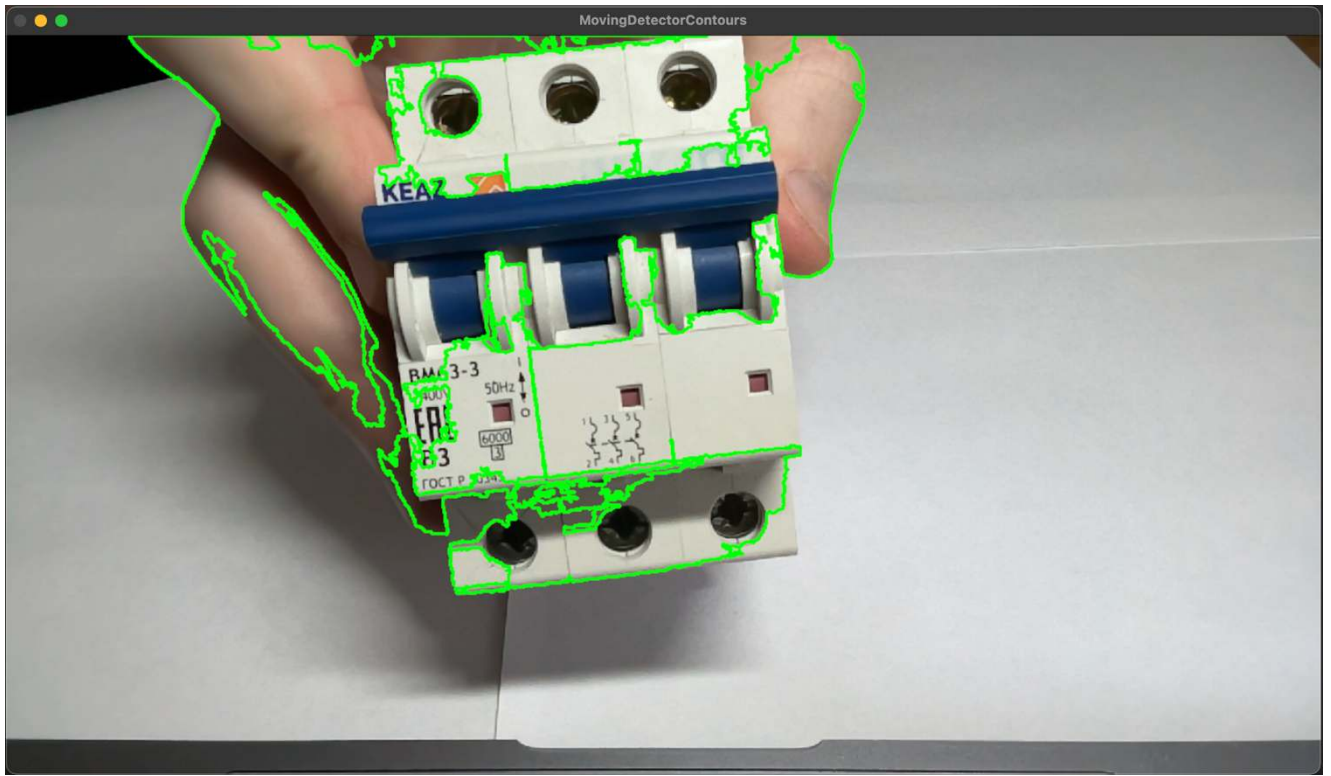


Рисунок 3.6 – Пример работы алгоритма выделения контуров движущегося объекта.

Для точного определения объекта-контейнера применяется метод ORB (Oriented FAST and Rotated BRIEF) [111]. Этот алгоритм позволяет создать двумерные дескрипторы как для исходного объекта, так и для заданных шаблонов, и сравнить их. Дескрипторы формируются путем редукции количества проверяемых пикселей из матриц изображений, что позволяет эффективно и точно определить наличие объекта для распознавания. ORB может надежно идентифицировать объекты, даже если они изменили свой масштаб или ориентацию, что позволяет с высокой точностью определять наличие объекта для распознавания в кадре. Пример работы алгоритма ORB представлен на рисунке 3.7, на нем можно увидеть кадр с движущимся объектом, шаблонное изображение объекта для распознавания и найденные соответствия между дескрипторами данных изображений (представлены зелеными линиями)

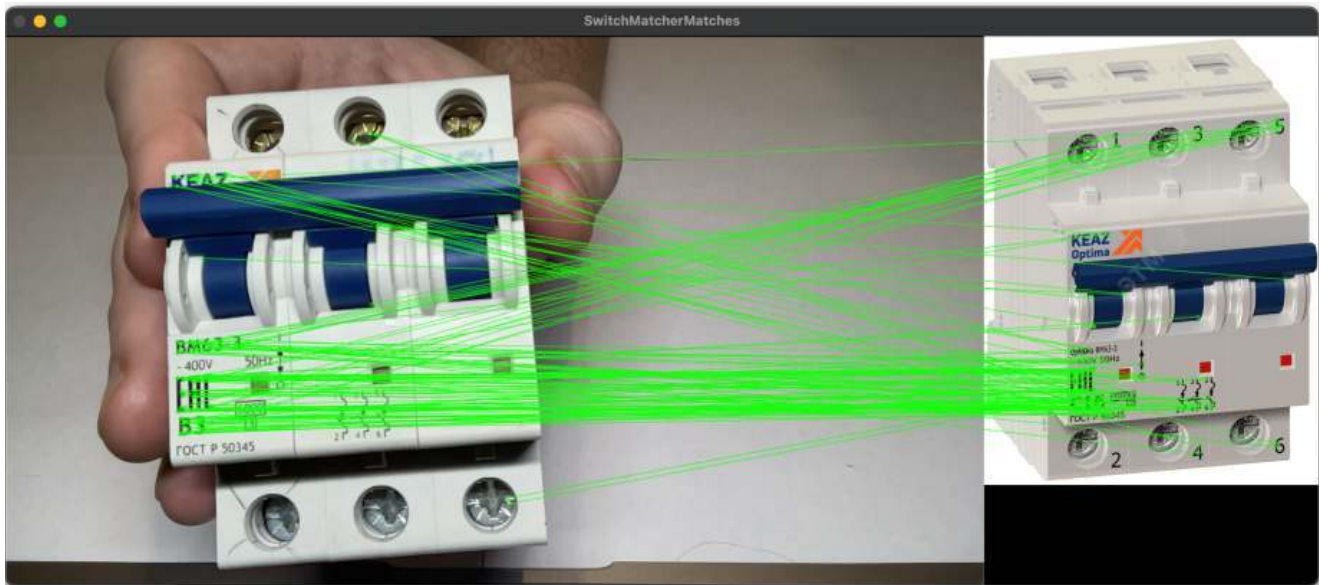


Рисунок 3.7 – Пример работы алгоритма ORB (кадр, в котором происходит поиск объекта слева, шаблонное изображение справа)

Поиск объекта для распознавания является важным шагом в процессе обработки изображений, содержащих СИ. Применение алгоритмов KNN, S. Suzuki и ORB позволяет выделить объект-контейнер из общего фона и подготовить его для дальнейшего анализа. Данный этап позволяет выявить наличие объекта для распознавания в кадре видеопотока, таким образом увеличивая общую пропускную способность системы, за счет определения необходимости выполнения дальнейших этапов обработки, в случае если объект для распознавания был найден, и отсеивания кадров, в которых отсутствует объект для распознавания.

3.3. Отслеживание объекта для распознавания

После того как объект-контейнер был найден на изображении, необходимо обеспечить его непрерывное отслеживание в последовательности кадров. Этот этап важен для обеспечения стабильности процесса обработки, в условиях движения объектов, например, на производственной линии. Для решения этой задачи используется алгоритм отслеживания объектов CSRT (Channel and Spatial Reliability Tracker) [66].

Отслеживание объекта необходимо для поддержания непрерывности процесса обработки. В условиях динамических сцен, где объекты могут двигаться,

изменять свою позицию, масштаб и ориентацию, важно иметь возможность точно определять их положение на каждом кадре. Непрерывное отслеживание объекта гарантирует, что система сохраняет контекст, связанный с данным объектом, от одного кадра к другому. Отслеживание объекта после однократного обнаружения позволяет пропустить этап поиска объекта для дальнейших кадров, что увеличивает общую пропускную способность системы за счет уменьшения количества выполняемых операций.

Алгоритм CSRT выбран для отслеживания объектов благодаря его высокой точности и устойчивости к изменениям в масштабе, вращении и частичным окклюзиям. Он использует информацию о каналах и пространственную надежность для улучшения точности отслеживания.

После того как объект-контейнер был найден и выделен, его положение и размеры используются для инициализации трекера CSRT. На каждом последующем кадре трекер обновляет свою модель объекта и определяет его новое положение. Это включает анализ изменений в масштабе, ориентации и внешнем виде объекта. Трекер проверяет уровень уверенности в положении текущего объекта. Уровень уверенности определяет достоверность определения нового положения движущегося объекта в кадре. В случае снижения уровня уверенности или потери объекта трекер пытается его заново обнаружить в последующих кадрах. Если объект для распознавания выходит за пределы области обзора камеры, система сигнализирует об этом, что позволяет своевременно принять меры для восстановления отслеживания.

В случае, если объект отсутствует или трекер сигнализирует о недостаточном уровне уверенности в положении объекта в 10 последовательных кадрах, алгоритмом принимается решение о том, что объект покинул область обзора камеры и возвращается к выполнению группы операций по поиску объекта. Таким образом, алгоритм сохраняет контекст и учитывает случаи последовательной обработки одного и того же объекта в нескольких кадрах подряд, даже в случаях его перекрытия или частичного выхода за границы обзора камеры, но при этом

выполняет операцию поиска нового объекта повторно, в случае если это необходимо.

Отслеживание объекта для распознавания приносит следующие преимущества: обеспечение непрерывности отслеживания объекта снижает вероятность его потери и ошибок на последующих этапах обработки, позволяет эффективно отслеживать объекты даже в сложных динамических сценах, что важно для использования на производственных линиях и в других приложениях, система концентрирует ресурсы на актуальных областях изображения, повышая общую эффективность и скорость обработки, способность алгоритма адаптироваться к изменениям внешнего вида объекта улучшает его устойчивость и точность в различных условиях.

3.4. Выделение объекта для распознавания

После того как объект-контейнер был успешно обнаружен и отслежен в последовательности кадров, необходимо выделить его из фона и других объектов на изображении. Этот этап включает в себя серию операций по предобработке изображения и улучшению его качества для дальнейшего анализа. Точность выделения объекта-контейнера напрямую влияет на эффективность и точность последующих этапов распознавания.

Выделение объекта для распознавания необходимо для реализации следующих операций.

Изоляции релевантной информации. Удаление фона и других нерелевантных элементов позволяет обрабатывать исключительно объекте-контейнер и содержащуюся на нем символьную информацию, что позволяет повысить комплексную характеристику «точность × пропускная способность», за счет уменьшения ошибок первого рода при распознавании символов и уменьшения размера обрабатываемого изображения.

Улучшения качества изображения. Компенсация искажений и устранение шумов помогают улучшить видимость и четкость символов на объекте, что повышает точность распознавания.

Исправление проекционных и аффинных искажений позволяет получить большую точность при последующей обработке изображения с помощью нейросетей EAST и CRNN.

Процесс выделения объекта для распознавания из фона и других объектов на изображении включает несколько этапов предобработки, обеспечивающих улучшение качества изображения и подготовку его к дальнейшему анализу. Операции на этом этапе включают преобразование цветовой модели изображения в HSV, бинаризацию по верхнему и нижнему порогам, удаление шумов с использованием фильтра Гаусса, дилатацию, выделение границ с помощью метода Канни, нахождение контуров по алгоритму S. Suzuki [114], аппроксимацию контуров методом Рамера-Дугласа-Пекера [113, 112], а также исправление аффинных и проекционных искажений. Дополнительно выполняются проверки площади и формы объекта-контейнера для исключения неподходящих кадров из дальнейшей обработки.

Первоначально изображение преобразуется из цветовой модели RGB в HSV (Hue, Saturation, Value). Эта операция позволяет более выделять цветовые характеристики объекта, так как в HSV модели оттенок, насыщенность и яркость представляются отдельными компонентами. Преобразование в HSV облегчает задачу выделения объекта на изображении, так как позволяет определять и анализировать цветовые характеристики изображения, что особенно важно для дальнейшей бинаризации изображения.

Следующий этап включает бинаризацию изображения по верхнему и нижнему порогам. Бинаризация преобразует изображение в черно-белую форму, где пиксели, попадающие в заданный диапазон интенсивностей, принимают значение белого цвета (значимая информация), а все остальные пиксели – черного (неважная информация). Этот процесс позволяет четко выделить объект-контейнер, отделив его от фона и других незначимых элементов на изображении [46]. Выбор верхнего и нижнего порогов осуществляется на основе анализа гистограмм изображений, сделанных в тех же условиях, в которых будет использоваться алгоритм, что обеспечивает гибкость метода к различным условиям

освещения и контрастности. Пример бинаризации изображения представлен на рисунке 3.8.

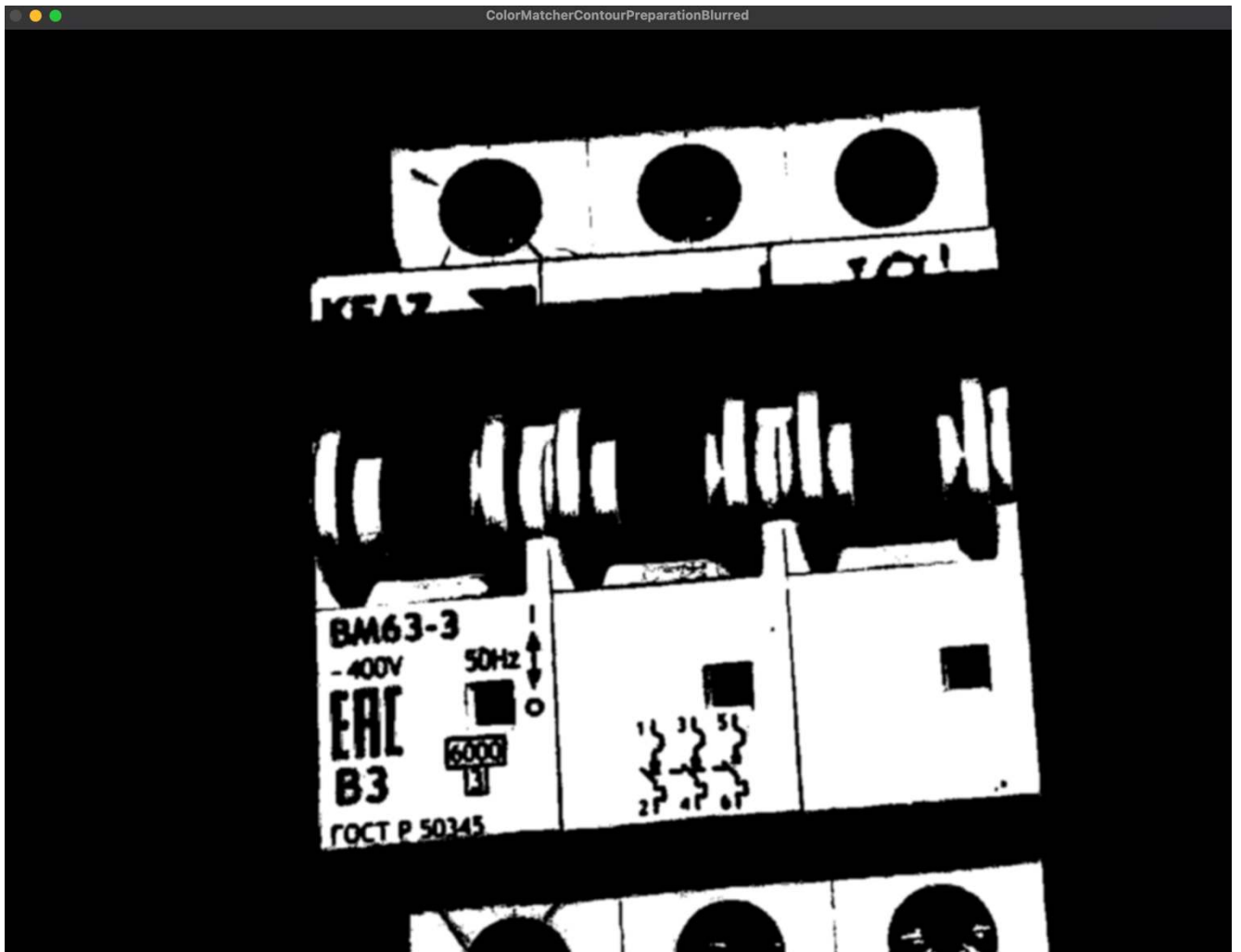


Рисунок 3.8 – Пример работы алгоритма бинаризации

Удаление шумов осуществляется с использованием фильтра Гаусса, который сглаживает изображение и устраняет мелкие артефакты, сохраняя при этом основные контуры объекта. Этот метод основан на применении размытия Гаусса, где каждый пиксель изображения заменяется средним значением интенсивностей соседних пикселей, взвешенных по распределению Гаусса. В результате, фильтр уменьшает влияние шумов, улучшая четкость границ и контуров объектов-контейнеров.

Дилатация расширяет границы объектов на изображении, объединяя разрозненные части объекта-контейнера. Этот процесс помогает восстановить линии границ на объектах со сложной структурой, которые подвержены влиянию шумов и искажений [46]. Дилатация выполняется путем замены каждого пикселя

изображением, заданным структурным элементом, максимальным значением в окрестности этого пикселя. Таким образом, данный метод сглаживает и расширяет границы объектов, улучшая их выделение. Пример работы алгоритма дилатации в комбинации с размытием Гаусса представлен на рисунке 3.9.

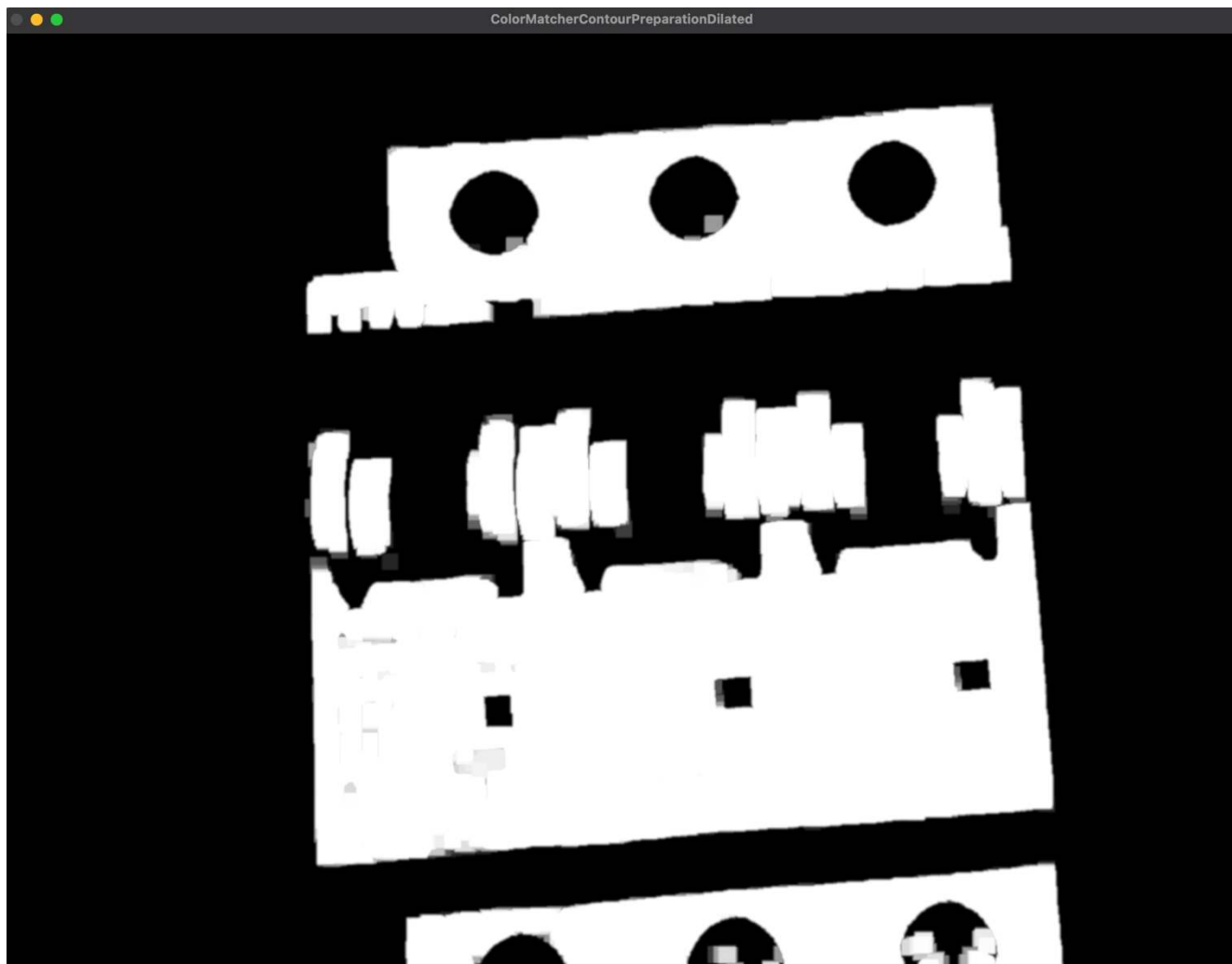


Рисунок 3.9 – Пример работы алгоритма дилатации в комбинации с размытием Гаусса

Выделение границ – это процесс преобразования изображения таким образом, чтобы на нем остались только пиксели, соответствующие границам объектов. Этот метод помогает определить резкие изменения интенсивности пикселей, которые возникают на краях объектов, разделяющих различные области изображения [41]. Метод Канни используется для выделения границ объекта-контейнера. Этот алгоритм включает вычисление градиентов изображения, подавление немаксимумов и двойную пороговую фильтрацию с последующей трассировкой краев. В результате этого процесса получается изображение, на

котором четко видны границы объектов, что позволяет их легко идентифицировать и отделить от фона (пример на рисунке 3.10).

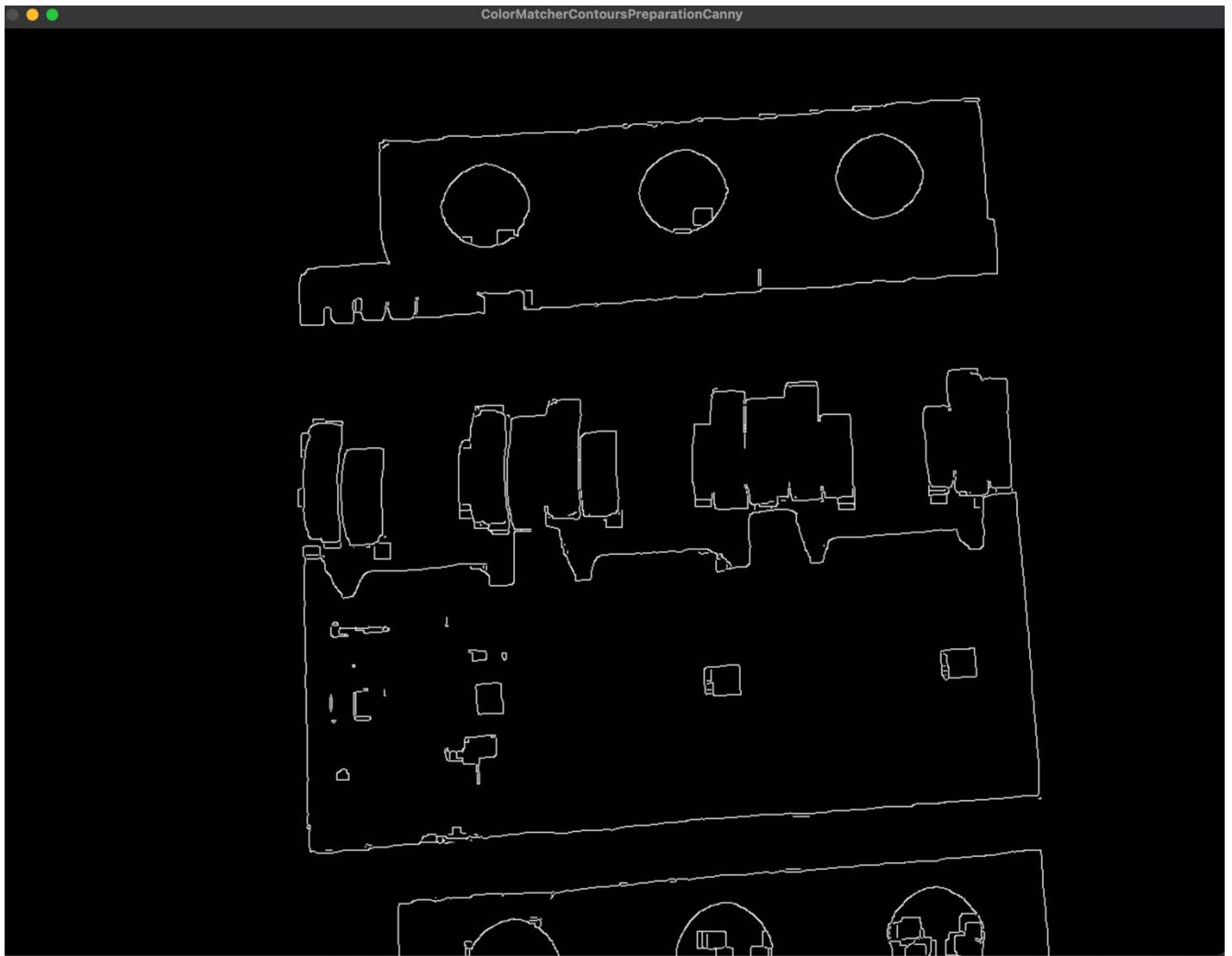


Рисунок 3.10 – Пример работы метода Канни

После использования метода Канни операция дилатации производится повторно (пример на рисунке 3.11). Данная операция необходима для изображений, содержащих шумы или искажения т.к. в случае их появления границы объектов, представленных в кадре, могли быть нарушены. Повторная дилатация позволяет восстановить границы, которые были повреждены или разорваны.

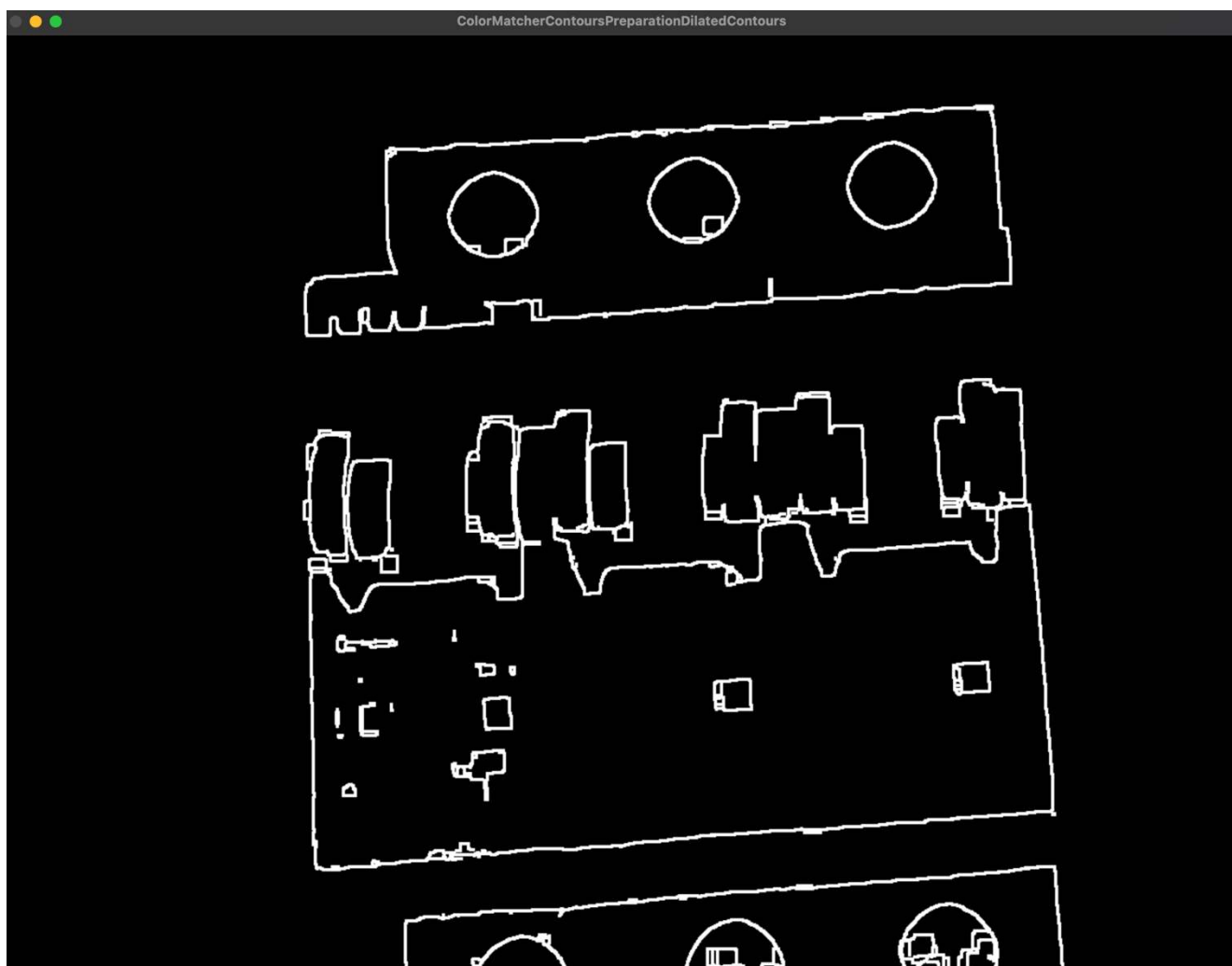


Рисунок 3.11 – Пример дилатации границ, выделенных методом Канни

Алгоритм S. Suzuki применяется для нахождения контуров объектов на изображении, что помогает выделить область, содержащую объект-контейнер. Метод Рамера-Дугласа-Пекера используется для аппроксимации найденных контуров, что позволяет упростить их и улучшить точность выделения объекта. Этот алгоритм сокращает количество точек, описывающих контур, сохраняя его форму с определенной точностью. Упрощение контуров позволяет определить ориентацию объекта и так же наличие проекционных искажений, для последующего исправления. Кроме того, аппроксимация контура позволяет провести проверку формы объекта, что позволяет убедиться в том, что объект может быть распознан последующими этапами алгоритма. Пример выделения контура, аппроксимации границ и прямоугольная область изображения, которые впоследствии будут передан для исправления аффинных и проекционных искажений представлен на рисунке 3.12.

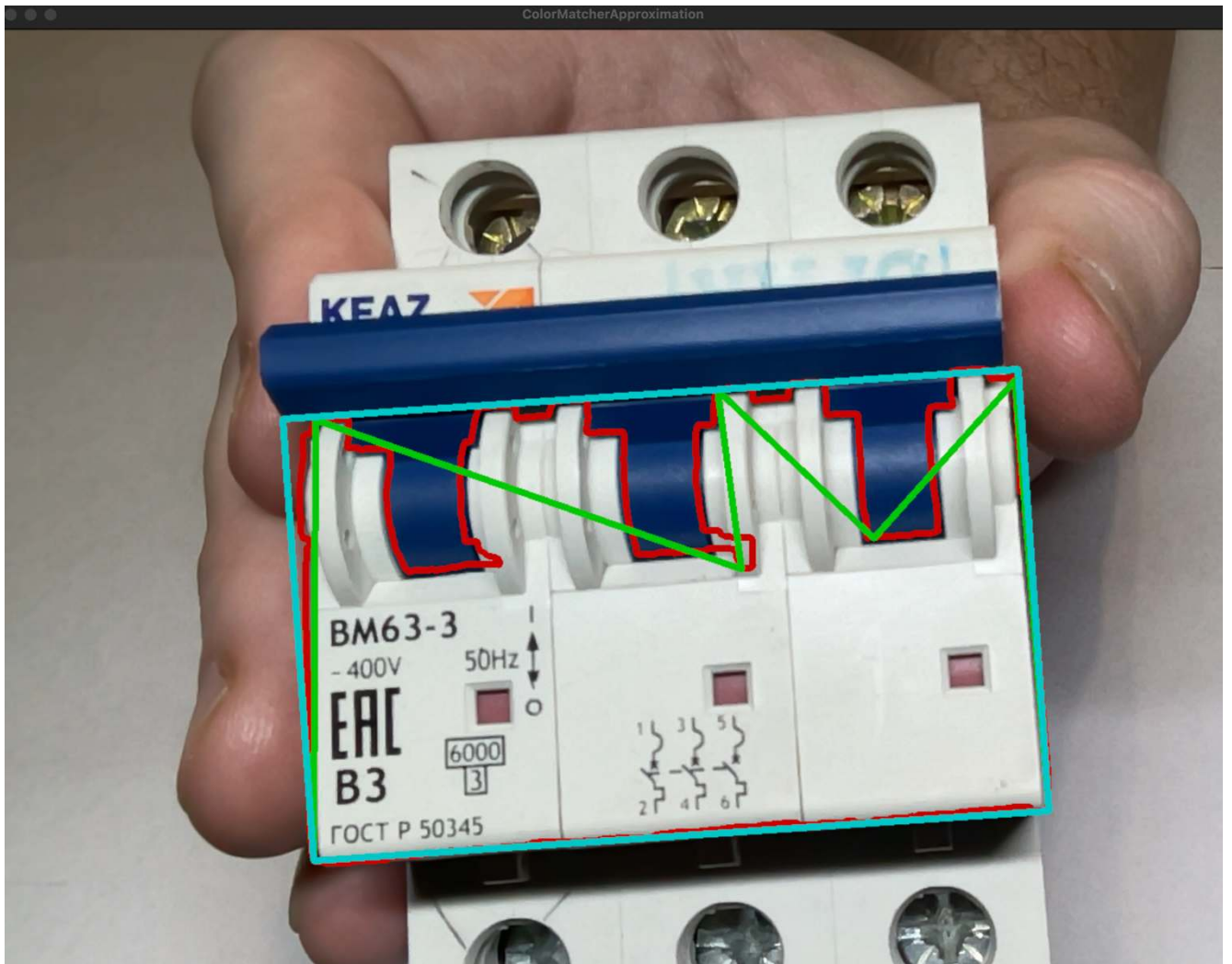


Рисунок 3.12 – Пример выделения контура, аппроксимации границ и выделения прямоугольной области изображения, для исправления аффинных и проекционных искажений

Исправление аффинных и проекционных искажений необходимо для приведения объекта-контейнера к стандартной форме и размерам, что упрощает дальнейший анализ. Данные исправления производятся на выделенном предыдущими операциями алгоритма участке исходного изображения. Аффинные преобразования включают масштабирование, повороты и сдвиги, в то время как проекционные искажения учитывают перспективные изменения. Исправление этих искажений приводит к нормализации изображения объекта, что упрощает дальнейший анализ и распознавание символов. Участок исходного изображения, содержащий СИ, с исправленными аффинными и проекционными искажениями,

который впоследствии будет передан дальнейшим операциям алгоритма для распознавания СИ представлен на рисунке 3.13.

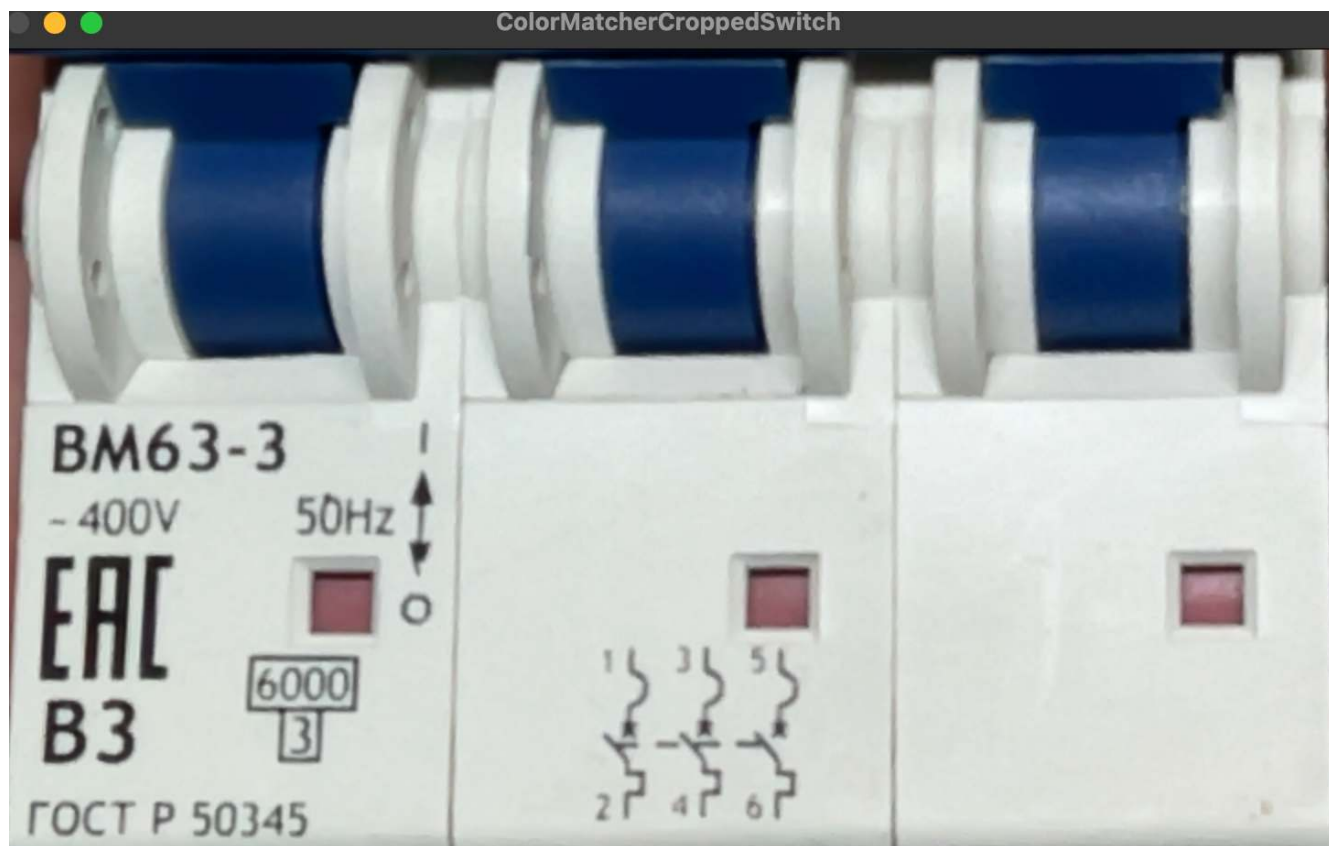


Рисунок 3.13 – Участок исходного изображения, содержащий СИ, с исправленными аффинными и проекционными искажениями, который впоследствии будет передан дальнейшим операциям алгоритма для распознавания СИ

В процессе выполнения данной группы методов выполняются операции проверки площади выделенного объекта и формы объекта-контейнера. Проверка площади позволяет убедиться в том, что объект имеет достаточной размер, чтобы СИ могла быть распознана нейросетевыми методами, и позволяет исключить из процесса обработки кадры, на которых объект продукции находится слишком далеко от камеры. Проверка формы объекта-контейнера позволяет убедиться в том, что изображение объекта не содержит окклюзий и маркировка объекта полностью доступна для выделения и распознавания, также исключая из процесса обработки кадры, которые заведомо не могут быть корректно обработаны на дальнейших этапах.

Таким образом, этап выделения объекта для распознавания конвейерно-параллельного алгоритма обработки изображений, обеспечивает повышение комплексной характеристики «точность × пропускная способность» выполняя предобработку данных для дальнейшего анализа символьной информации.

3.5. Поиск областей текста и выделение символов

Процесс начинается с поиска областей, содержащих текстовую информацию, внутри выделенного объекта-контейнера. Для этого используется нейронная сеть с архитектурой EAST (Efficient and Accurate Scene Text Detector), которая обеспечивает распознавание текстовых областей на изображении. EAST выполняет анализ изображения и определяет области, которые с высокой вероятностью содержат текст. EAST способна справляться с текстом различной структуры и ориентации, обеспечивая точное выделение областей текста. Произведено обучение нейросетей для распознавания текста. Для обучения использовались публичная выборка COCO-text дополненная реальными изображениями объектов, содержащих маркировку, с нанесенными искажениями.

Архитектура EAST [70] использует несколько уровней в своей сети для улучшения обнаружения текста разного размера и формы на изображениях. Эти уровни являются частью концепции пирамидальных признаков, объединяющих признаки, извлеченные на разных уровнях нейронной сети, таким образом, чтобы использовать как низкоуровневые, так и высокоуровневые представления данных. Каждый из этих уровней по мере продвижения по сети уменьшает пространственное разрешение входного изображения, но увеличивает глубину карты признаков. На каждом последующем уровне сети карта признаков содержит более абстрактную и высокоуровневую информацию, хотя и с меньшей детализацией в пространстве. В данной работе используется 4 уровня сетевых экстракторов признаков (размерностей 64, 128, 256, 384), которые объединяются 3 уровнями слоями слияния признаков, после чего выполняется предсказание карты оценок, предсказание геометрической карты и нестандартное подавление (NMS). NMS (Non-Maximum Suppression, подавление немаксимальных значений) – это

алгоритм, используемый в задачах компьютерного зрения, с целью уменьшить количество перекрывающихся областей с текстом, оставив только те, которые имеют наивысшую показатель уверенности.

Основные слои и функции активации используемой нейронной сети представлены в таблице 3.1.

Таблица 3.1 Основные слои и функции активации используемой нейронной сети

Этап/Компонент	Слои/Операции	Функции активации	Описание и взаимосвязь
Сетевой экстрактор признаков	Сверточные слои, Пулингговые слои	ReLU	Сверточные и пулингговые слои в сети VGG16 извлекают и снижают размерность признаков изображения.
Слияние признаков	1x1 Conv, 3x3 Conv	ReLU	Карты признаков из разных уровней сети экстрактора объединяются для улучшения детектирования текстов различных размеров.
Предсказание карты оценок	3x3 Conv, 1x1 Conv	Sigmoid	Прогнозирует вероятность того, что пиксель принадлежит текстовому блоку. Выходная карта оценок имеет ту же размерность, что и карта признаков.
Предсказание геометрической карты	3x3 Conv, 1x1 Conv	Linear	Прогнозирует геометрию текстового блока для каждого пикселя (четыре дистанции или параметры повернутой ограничивающей рамки).
Подавление немаксимальных значений (NMS)	Алгоритм подавления немаксимальных значений	—	Применяется для удаления перекрывающихся текстовых блоков, чтобы сохранить только наиболее вероятные.

Пример работы нейронной сети представлен на рисунке 3.14.



Рисунок 3.14 – Пример работы нейронной сети EAST

После того как текстовые области были обнаружены, следующий шаг включает выделение отдельных символов внутри этих областей. Этот процесс требует разделения текстовых строк на отдельные символы, что может быть сложной задачей из-за перекрытий символов, варьирующегося межсимвольного расстояния, различий в шрифтах и стилях текста. Для решения этой задачи применяется метод поиска базовой линии текста, который позволяет определить основную линию, вдоль которой располагаются символы. Используя информацию о базовой линии и анализ расстояний между символами, алгоритм разделяет строку текста на отдельные символы. Этот метод помогает уменьшить ошибки распознавания, особенно для текстов, содержащих сложные шрифты или курсивное начертание.

Кроме того, метод разделения текста на символы учитывает потенциальный параллелизм в обработке. Это означает, что несколько текстовых областей могут обрабатываться одновременно, что значительно ускоряет весь процесс распознавания. Такая параллельная обработка применима в условиях, где

необходимо быстро и точно распознавать текстовую информацию на множестве объектов.

В результате работы данной группы операций на выходе формируется последовательность изображений символов, готовая для дальнейшего распознавания. Эти символы затем передаются на следующий этап обработки, где осуществляется их распознавание с использованием нейросетевых методов.

3.6. Распознавание текста

Процесс распознавания текста, является ключевым из операций конвейерно-параллельного алгоритма обработки изображений, содержащих СИ. На данном этапе происходит извлечение текстовых данных из выделенных символов, что необходимо для последующей классификации и анализа характеристик объекта.

Процесс распознавания текста осуществляется с использованием сверточных рекуррентных нейронных сетей (CRNN [75]). Она состоит из трех основных компонентов: сверточной части, рекуррентной части и окончательного классификационного слоя.

Первый компонент, сверточная нейронная сеть (CNN), отвечает за извлечение пространственных признаков из входных данных, таких как изображения. В этой части происходит последовательное применение сверток и пулинга, что позволяет выделить иерархически значимые признаки, начиная от простых элементов, таких как грани и углы, до более сложных структур, таких как объекты и их части.

Второй компонент, рекуррентная нейронная сеть (RNN), используется для обработки временных зависимостей между извлеченными признаками. Рекуррентная сеть принимает на вход последовательность признаков, полученных из сверточной сети, и моделирует временные или последовательные зависимости в данных, такие как последовательность признаков символов в тексте. Для этой цели используются сети с долгой краткосрочной памятью (LSTM). Последний компонент, классификационный слой, преобразует выходные данные рекуррентной сети в итоговые предсказания. Он представляет собой полносвязный

слой с функцией активации, который сопоставляет выходные данные к соответствующим классам или меткам. Это позволяет CRNN распознавать символы, учитывая их взаимные характеристики и последовательность признаков символов, что позволяет показывать высокую точность в процессе работы с текстами, искаженными шумами или имеющими нестандартные шрифты.

Пример распознавания текста представлен на рисунке 3.15.

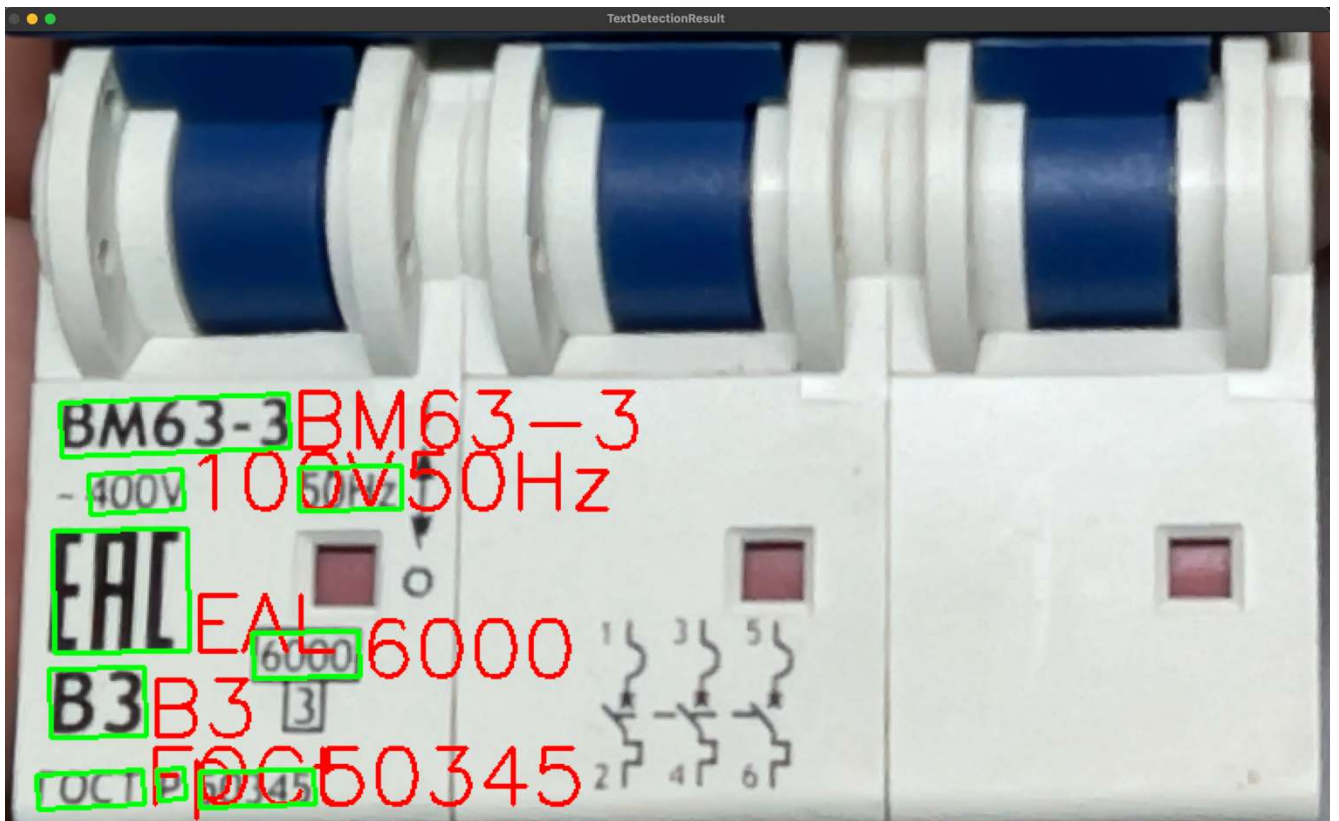


Рисунок 3.15 – Пример распознавания текста

Для обучения нейронной сети распознавания текста использовались публичные наборы данных SynthText для русского языка и MJsynth для английского языка. Данные выборки были дополнены реальными изображениями объектов, содержащих маркировку. Для увеличения точности работы сети и увеличения устойчивости к различным искажениям первоначальная выборка была дополнена изображениями, содержащими следующие изменения: дополнительные искажающие элементы; наклон текста; изменение цветовой характеристики изображения (как для фона, так и для символов); изменение геометрии изображения; изменение размеров символов в рамках одной строки; добавление

шумов на исходное изображение. Слои и функции активации CRNN представлены в таблице 3.2.

Таблица 3.2 – Слои и функции активации CRNN

Слой	Функция	Функция активации	Взаимосвязи
Входной слой	Принимает исходные данные.	Нет	Передает данные в сверточные слои.
Сверточные слои (Conv)	Извлечение пространственных признаков, применяя свертки.	ReLU	Последовательно связаны, передают признаки в слои пулинга или следующие свертки.
Пулинговые слои (Pooling)	Уменьшение размерности данных, выделение наиболее значимых признаков.	Нет	Следуют после сверточных слоев, передают уменьшенные признаки в следующие слои.
Рекуррентные слои (LSTM/GRU)	Моделирование зависимостей между признаками.	Sigmoid, Tanh	Принимают последовательности признаков из сверточной части, передают результаты в классификационный слой.
Полносвязный слой (Fully Connected)	Сопоставление признаков с конечными классами (значениями символов).	Softmax	Получает данные от рекуррентных слоев, производит итоговые предсказания.

Результатом работы CRNN является текстовая строка, представляющая распознанные символы объекта. Эта строка служит основой для последующих операций классификации и анализа, что позволяет получить точные данные о маркировке объекта.

Таким образом, этап распознавания текста обеспечивает получение необходимой символьной информации для дальнейшей обработки и классификации характеристик распознаваемого объекта, в составе алгоритма распознавания СИ.

3.7. Классификация характеристик

Операции классификации характеристик предназначены для систематизации и анализа признаков, присущих объектам распознавания, с целью их последующей классификации. Данные операции используются для определения ключевых признаков, а также определяют соответствие областей текста заданным категориям характеристик. В процессе классификации выполняется анализ и сопоставлению признаков, таких как цветовые, геометрические, топологические и символные, с заранее определенными эталонными значениями.

Эталонные признаки задаются в виде матрицы, где столбцы представляют собой различные характеристики, такие как серия, производитель и другие параметры продукции, а строки содержат конкретные признаки для каждой характеристики. В результате каждая ячейка этой матрицы содержит эталонное значение признака для соответствующей характеристики. Такой подход позволяет формализовать процесс классификации и обеспечить соответствие между распознаваемыми объектами и эталонными признаками.

Для определения степени соответствия выделенных признаков эталонным используется количественно-разностная функция невязки. Эта функция вычисляет расхождение между значениями признаков, выделенных из анализируемых объектов, и их эталонными значениями, представленными в матрице. Сравнение проводится по каждому признаку отдельно, что позволяет учесть различия на уровне цветовых характеристик, геометрических форм, топологических структур и символьных элементов. Невязка, вычисляемая с помощью данной функции, отражает степень отклонения обрабатываемого значения от эталона служит основой для принятия решения о его классификации. Чем меньше невязка, тем большая степень уверенности в том, что данная область текста является заданной характеристикой.

Пример работы операции классификации характеристик представлен на рисунке 3.16.



Рисунок 3.16 – Пример работы операции классификации характеристик

Таким образом, каждая область текста классифицируется как та характеристика, которая демонстрирует наименьшее отклонение от эталонных признаков, определённых в матрице. Этот процесс обеспечивает высокую точность распознавания и классификации, поскольку каждая область текста объекта сравнивается с строго определёнными эталонными характеристиками, зафиксированными в матричной форме.

3.8. Допущения и ограничения

В данном разделе рассматриваются допущения и ограничения, которые были приняты при разработке и тестировании алгоритма и вычислительного модуля обработки символьной информации. Для обеспечения стабильной и точной работы системы предполагается, что входные изображения имеют разрешение 1920x1080 пикселей. Освещенность сцены, в которой проводится съемка, должна находиться в диапазоне от 150 до 1500 люкс, чтобы минимизировать влияние недостаточного или чрезмерного освещения на качество распознавания. Кроме того, пиковое

соотношение сигнала к шуму (PSNR) для входных данных должно быть не ниже 10 dB, что гарантирует уровень шумовых искажений, с которыми система способна справиться. При этом объект для распознавания в кадре должен занимать не менее 8.5% от площади кадра.

В рамках проведенного исследования для различных алгоритмов были экспериментально определены константы и диапазоны значений, при которых достигается наилучшая точность в различных условиях.

Для алгоритма поиска движения в кадре на основе метода ближайших соседей (KNN) было установлено, что оптимальная точность детектирования движения достигается при использовании следующих параметров: история (количество кадров запоминаемых алгоритмом и используемых для определения фона и движущегося объекта [48]) из 15 кадров, пороговое значение среднеквадратичного отклонения для определения совпадения значения пикселя текущего кадра с предыдущим кадром [48] равно 400.

Алгоритм поиска контуров S.Suzuki [98] используется с модификацией, осуществляющей компрессию вертикальных, горизонтальных и диагональных линий, позволяющей уменьшить количество точек необходимых для представления контура, таким образом ускоряя последующие этапы обработки.

Сравнение дескрипторов, найденных алгоритмом ORB [99] для текущего кадра и шаблонных изображений, происходит с пороговым значением T_{desc} (формула 2.6) равным 0.5.

Алгоритма отслеживания движения CSRT используется со стандартными параметрами [54], кроме меры PSR (пороговое значение для определения выхода объекта из зоны обзора), которая составляет 0.045.

Верхний и нижние пороги для бинаризации являются настраиваемыми и для различных типов распознаваемых объектов могут отличаться. Применение фильтрации Гаусса выполняется с маской размерности 5x5 пикселей, с использованием экстраполяции граничных пикселей методом свертки [34]. Операции дилатации выполняется с маской размерности 5x5 пикселей представленной в формуле 2.12. Выделение контуров алгоритмом Канни

используется с пороговыми значениями гистерезиса 25 и 75 и размером оператора Собела 3×3 [29]. Аппроксимация контуров методом Рамера-Дугласа-Пекера выполняется с параметром максимальной дистанции между оригинальным контуром и его аппроксимацией [101, 100] составляющей 4% от периметра оригинального контура. Расчёт матриц трансформаций для исправления аффинных и проекционных искажений происходит с использованием метода Гаусса с использованием оптимального поворотного элемента [29].

Найденные экспериментальным путем оптимальные значения для данных параметров позволяет значительно улучшить точность обработки, при этом обеспечивая эффективную реализацию аппаратно-программных систем компьютерного зрения за счет оптимизации применяемого встраиваемого программного обеспечения.

3.9. Исследование точности выделения объекта и распознавания символов

В данном разделе рассматривается тестирование точности распознавания предложенного алгоритма распознавания символьной маркировки продукции. Целью тестирования являлась оценка точность выделения объекта и распознавания символов разработанного алгоритма в условиях различного уровня шума, распределяемого по нормальному закону. В качестве меры уровня шума использовался показатель PSNR (Peak Signal-to-Noise Ratio), который характеризует отношение максимального значения сигнала к уровню шума. Меньшие значения PSNR указывают на наличие большего количества шума на изображении. Были проведены эксперименты, в ходе которых оценивалась точность выделения объекта и точность распознавания символов на изображениях, подвергнутых различным степеням шумового воздействия. Полученные статистические данные представлены в виде таблиц 3.3 и 3.4 соответственно, содержащих результаты оценки точности алгоритма при различных уровнях шума. Доверительный интервал рассчитывался по формуле доверительного интервала для среднего с уровнем достоверности 0.99.

Таблица 3.3 – Статистические данные по точности обнаружения объекта в кадре при различных искажениях

Уровень шума (PSNR)	Площадь перекрытия объекта	Точность обнаружение объекта в кадре
От 20 dB	До 5%	98% ± 0.3%
15-20 dB	До 15%	81% ± 0.7%
10-15 dB	До 30%	21% ± 0.9%

Таблица 3.4 – Точность распознавания символов разработанного алгоритма при различном уровне шума

Уровень шума (PSNR)	Точность распознавания	Ошибки первого рода (Распознанные символы, не присутствовавшие на изображении)	Ошибки второго рода (Некорректно распознанные символы, присутствовавшие на изображении)
От 20 dB	97.7% ± 0.5%	2%	0.4%
15-20 dB	91.8% ± 0.7%	6.6%	1.7%
10-15 dB	80% ± 0.8%	17.2%	2.9%

Испытания показали, что предложенное решение обладает высокой степенью устойчивости к разнообразным видам искажений, что позволяет достигать значительных результатов при распознавании символов даже в условиях присутствия высокого уровня шума. Точность распознавания варьируется от 97% до 80% в зависимости от уровня PSNR, при уменьшении этого значения точность

снижается. Эти результаты указывают на то, что алгоритм способен эффективно распознавать символьную маркировку в различных условиях, демонстрируя высокую устойчивость к воздействию шумовых помех.

Основным вызовом при распознавании изображений с низким уровнем PSNR стало некорректное определение участков изображения как символов. Это проявилось в значительном увеличении ошибок первого рода (ложноположительных срабатываний), которые достигли 17,15% при наименьших значениях PSNR, что указывает на наибольший уровень шума. Данный результат свидетельствует о том, что при сильных искажениях алгоритм может ошибочно распознавать фрагменты изображения как символы.

Таким образом, тестирование показало, что предложенное решение обладает высокой точностью распознавания при широком диапазоне значений PSNR.

Вывод по третьей главе

Разработан конвейерно-параллельный алгоритм обработки изображений, содержащих СИ, осуществляющий распознавание символьной маркировки продукции, основанный на разработанной модели и методе, отличающийся возможностью распараллеливания процесса обработки, настройкой к различным видам характеристик и устойчивостью к искажениям, что обеспечивает повышенную точность распознавания. Данный алгоритм применим для организации арифметической, логической и символьной обработки данных, хранения и ввода-вывода информации.

ГЛАВА 4. КОНВЕЙЕРНО-ПАРАЛЛЕЛЬНЫЙ ВЫЧИСЛИТЕЛЬНЫЙ МОДУЛЬ ОБРАБОТКИ СИМВОЛЬНОЙ ИНФОРМАЦИИ

4.1. Структурная организация вычислительного модуля

В данном разделе представлена структурная организация вычислительного модуля обработки символьной информации, предназначенного для распознавания и классификации символов на изображениях [9]. Модуль разрабатывался с целью повышения комплексной характеристики «точность × пропускная способность» распознавания, адаптации к различным типам символьных маркировок и устойчивости к шумовым искажениям. Структура модуля включает в себя несколько блоков, каждый из которых выполняет специфические задачи, обеспечивая комплексную обработку и распознавание характеристик объектов на изображении с использованием СИ.

Вычислительный модуль обработки символьной информации включает несколько блоков, каждый из которых выполняет определённую функцию в цепочке обработки данных:

1. Блок получения изображения отвечает за ввод исходного изображения в систему. Полученное изображение становится исходными данными для всех последующих этапов обработки (основан на выражениях и функциях 2.1-2.7).

2. Блок предобработки изображения, выполняющий этапы анализа характеристик объекта, удаления шумов и поиска областей текста, включает:

- 2.1. Блок анализа цветовой информации выполняет анализ цветовых характеристик изображения, извлекая дополнительную информацию, которая может быть использована на последующих этапах обработки (основан на выражениях и функциях 2.9-2.10).

- 2.2. Блок удаления шумов обеспечивает фильтрацию изображения, удаляя шумы и улучшая качество изображения, что облегчает дальнейшее распознавание символов (основан на выражениях и функциях 2.11-2.12).

2.3. Блок поиска области объекта выделяет области, содержащие объект на изображении, которые будут переданы для последующей классификации СИ (основан на выражениях и функция 2.13-2.14)

3. Блоки классификации символьной информации используются параллельно, с целью увеличения пропускной способности и содержат:

3.1. Блок поиска слов осуществляет поиск и выделение отдельных слов в областях объектов, выделенных на предыдущем этапе (основан на выражении 2.16).

3.2. Блок поиска символов определяет конкретные символы в пределах каждого найденного слова (основан на выражении 2.17).

3.3. Блок классификации символов проводит классификацию и распознавание каждого символа, формируя итоговый распознанный текст.

4. Блок формирования выходного значения служит для классификации характеристик на основании данных, полученных из блоков классификации и содержит (основан на выражениях и функциях 2.18-2.22):

4.1. Блок классификации характеристик сопоставляет распознанные символы с характеристиками объекта, такими как серия или производитель, и определяет соответствие с использованием количественно-разностной функции невязки.

4.2. Блок формирования результата распознавания генерирует окончательный список характеристик объекта на основе распознанного текста и классифицированных признаков.

Структурная организация вычислительного модуля представлена на рисунке 4.1.

Получено положительное решение о выдаче патента на устройство для распознавания символьной информации маркировки продукции, разработанное на основе данной структурной организации вычислительного модуля [12].

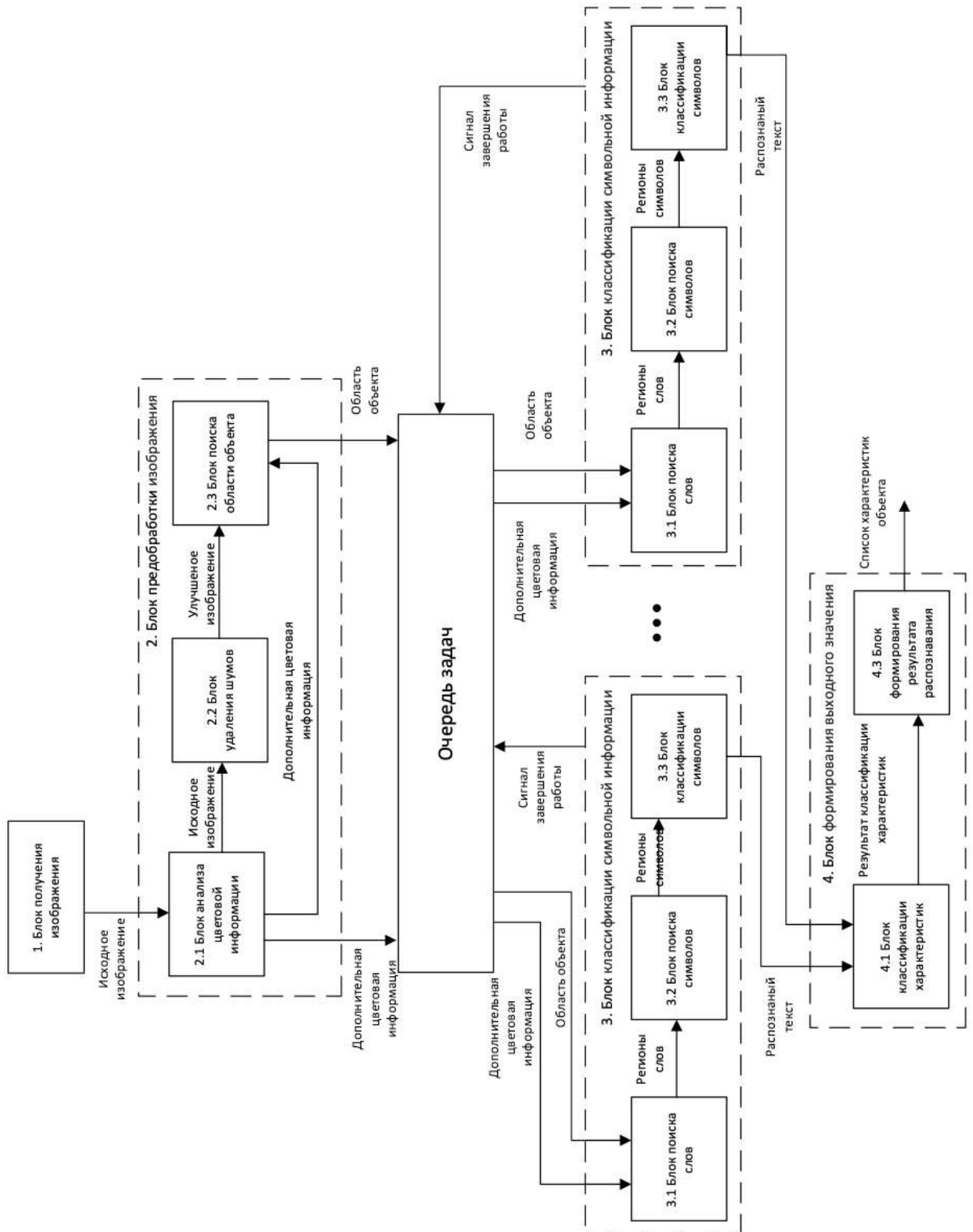


Рисунок 4.1 – Структурная организация вычислительного модуля

Диаграмма потоков данных, описывающая процесс передачи данных между различными операциями, выполняемыми в блоках представлена на рисунке 4.2, где

I_n и I_m представляют собой обрабатываемые кадры изображения с соответствующими номерами.

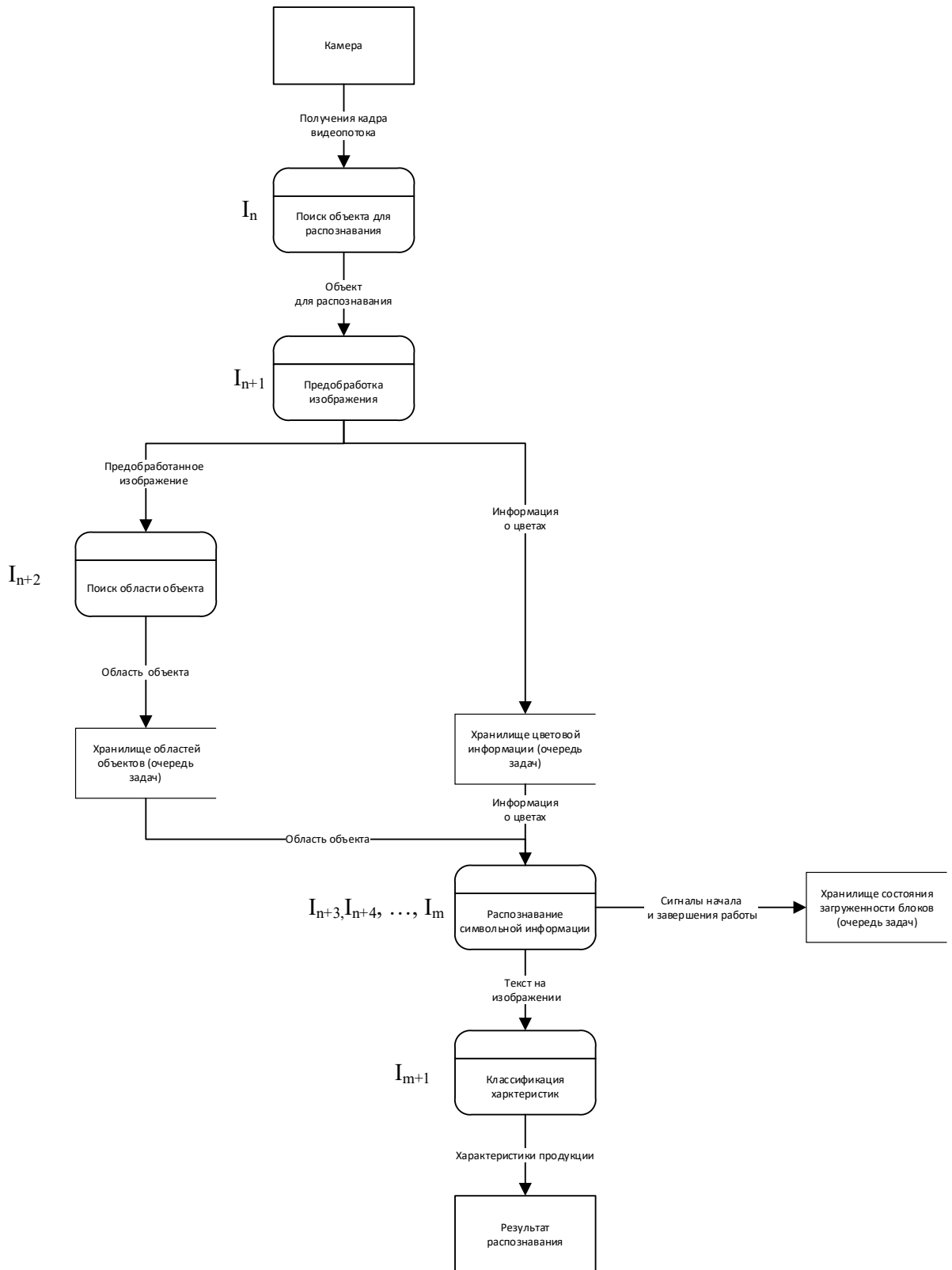


Рисунок 4.2 – Диаграмма потоков данных в вычислительном модуле обработки СИ

Управляющим элементом всего модуля является **очередь задач**, которая организует параллельное выполнение операций, принимая данные из блока предобработки и распределяя их между блоками классификации СИ. Она обеспечивает передачу данных основываясь на информации о загруженности каждого из блоков классификации.

Главный управляющий поток, представленный очередью задач, получает на вход данные, такие как области объектов, выделенные в процессе предобработки. Эти данные разбиваются на отдельные задачи, каждая из которых будет передана на обработку в один из блоков 3. Когда очередь задач получает новую область объекта, она передает данную область для обработки в первый доступный блок 3.

Каждый блок 3 обрабатывает полученную задачу (распознает слова или символы в заданной области). Важно отметить, что блоки работают независимо друг от друга, что позволяет обрабатывать несколько задач одновременно, значительно повышая пропускную способность всего модуля.

После завершения обработки задачи, блок 3 отправляет полученный результат напрямую в блок 4 (блок формирования выходного значения), где данные проходят дальнейшую классификацию и структурирование, включая сопоставление распознанного текста и его признаков с эталонными признаками для получения характеристик объекта и формирование окончательных результатов.

После передачи результата в блок 4, блок 3 отправляет сигнал очереди задач о том, что он завершил обработку текущей задачи и готов принять следующую. Этот сигнал позволяет очереди задач немедленно назначить новому освободившемуся блоку следующую задачу из очереди. Получив сигнал от одного из блоков 3 о завершении работы и готовности к новой задаче, очередь задач распределяет следующую задачу на этот блок. Таким образом, процесс продолжает работать циклически, обеспечивая непрерывный поток задач.

Параллельная работа блока 3 позволяет обрабатывать различные области объектов, выделенные на этапе предобработки, несколькими экземплярами блоков 3.1 (поиска слов), 3.2 (поиска символов) и 3.3 (классификации символов). Параллельное выполнение процессов ускоряет распознавание символов, поскольку

обработка отдельных областей объектов не зависит друг от друга и может выполняться одновременно на нескольких графических процессорах.

Параллельная работа блоков позволяет использовать более ресурсоемкие нейросетевые методы выделения текста распознавания, которые являются слишком ресурсоемкими для последовательной обработки. Таким образом, система может достичь более высокой характеристики «точность × пропускная способность» распознавания символов за счет использования более точных методов без ущерба для скорости обработки.

В структурной организации вычислительного модуля, происходит распределение задач между центральным процессором (CPU) и графическими процессорами работающими параллельно (GPU), что позволяет повысить характеристику «точность × пропускная способность».

В первую очередь, блоки, требующие интенсивной обработки данных или способные работать в параллельном режиме, выполняются на GPU. Например, блок предобработки изображения, ответственный за удаление шумов, отслеживание объекта, поиск контуров, исправление искажений. Данный процесс включает в себя сложные вычисления, которые могут выполняться графическим процессором. GPU оптимально справляется с такими задачами, особенно при обработке изображений высокого разрешения, где требуется значительное количество параллельных операций.

На этапе классификации символьной информации работа также распределена на GPU. В частности, блоки поиска слов и поиска символов, которые ответственны за распознавание и выделение отдельных слов и символов в тексте, используют нейросетевые алгоритмы. Эти алгоритмы предполагают выполнение большого числа операций с плавающей запятой, и поэтому их реализация на GPU обеспечивает существенное повышение производительности. GPU позволяет одновременно обрабатывать несколько областей текста, что повышает общую пропускную способность модуля.

Классификация символов, следующий этап обработки, также выполняется на графическом процессоре. Данный процесс включает в себя нейронные сети,

которые требуют значительных вычислительных ресурсов. Применение GPU позволяет параллельно выполнять множество операций, необходимых для точного распознавания символов, для повышения пропускной способности обработки видеопотока.

В то же время блоки, не требующие интенсивных вычислительных ресурсов, такие как блок получения изображения и блок анализа цветовой информации, выполняются на центральном процессоре. Получение изображения, представляющее собой ввод данных в систему, не требует значительных вычислительных мощностей и реализуется конвейерно, поэтому эффективно реализуется на CPU. Анализ цветовой информации включает операции, такие как преобразование цветных пространств и вычисление статистических характеристик изображения. Эти задачи являются менее ресурсоемкими и не требуют параллельной обработки, что делает CPU подходящей платформой для их выполнения.

Завершающая часть обработки – блок формирования выходного значения, также выполняется на CPU. Данный блок включает в себя задачи сопоставления распознанных символов с эталонными данными и формирования итогового списка характеристик объекта. Эти задачи не требуют значительных вычислительных ресурсов и могут эффективно выполняться конвейерно, что делает CPU оптимальной платформой для их реализации.

В структуре вычислительного модуля очередь задач играет ключевую роль в координации обработки данных, и её реализация на центральном процессоре (CPU) обеспечивает упорядоченность и последовательность всего процесса. Очередь задач, функционируя на CPU, выполняет задачи управления распределением рабочих нагрузок между блоками модуля, которые задействуют графические процессоры работающие параллельно. Это позволяет эффективно организовать обработку данных в параллельных вычислениях, обеспечивая последовательный порядок выполнения задач распознавания и получение итогового результата.

Одной из важнейших функций очереди задач является сохранение порядка обработки изображений. В процессе работы модуля изображения поступают на

вход в последовательности, которая должна быть сохранена до самого конца обработки. Поскольку различные блоки модуля, особенно те, которые работают на GPU, могут обрабатывать данные в параллельном режиме, существует вероятность несоответствия между порядком поступления и завершения обработки изображений. Реализованная на CPU очередь задач предотвращает это, гарантируя, что обработанные данные будут возвращены в том же порядке, в котором они поступили на вход системы.

Кроме того, очередь задач, реализованная на CPU, обеспечивает эффективное управление ресурсами модуля. CPU контролирует загрузку параллельно подключенных графических процессоров, распределяя задачи таким образом, чтобы избежать перегрузки отдельных блоков и обеспечить максимальную производительность системы. Это достигается за счёт динамического распределения задач между блоками, что позволяет оптимально использовать вычислительные мощности GPU. Очередь задач анализирует текущую загрузенность каждого блока и направляет задачи на выполнение в первый доступный блок, что позволяет максимально использовать параллелизм и повышает общую пропускную способность модуля.

Ещё одним важным аспектом является синхронизация финальных этапов обработки. После того как GPU завершает обработку задач, результаты возвращаются на CPU, который отвечает за их дальнейшую передачу в блоки классификации и формирования итоговых значений. Благодаря реализации очереди задач на CPU, обеспечивается синхронизация всех этапов обработки и исключается вероятность расхождения между порядком выполнения и формированием конечных данных.

Таким образом, основная особенность структурной организации вычислительного модуля заключается в разделении задач между CPU и GPU с учётом их вычислительной сложности и необходимости в конвейерно-параллельной обработке. В вычислительном модуле конвейерный подход применяется на этапе предобработки изображения и в блоке формирования выходного значения. На этих этапах обработка данных происходит конвейерно:

каждый блок передаёт результат своей работы на вход следующего блока, таким образом, обеспечивая непрерывный поток данных. Например, после получения изображения, оно проходит через блоки анализа цветовой информации, удаления шумов, а затем результат этой предобработки передаётся на дальнейшую классификацию и распознавание. Конвейерная обработка позволяет обеспечить структурированное и упорядоченное выполнение всех необходимых операций, сохраняя при этом последовательность данных, что особенно важно для корректного формирования итогового результата.

В реализации вычислительного модуля обработки символьной информации параллельный подход используется на этапах поиска текста и классификации символьной информации. Блоки поиска слов, поиска символов и классификации символов работают независимо друг от друга, что позволяет выполнять несколько задач одновременно. Таким образом получается достичь увеличения пропускной способности модуля, когда разные области текста могут обрабатываться параллельно на нескольких графических процессорах (GPU). Параллельная обработка значительно увеличивает пропускную способность модуля, поскольку позволяет одновременно распознавать символы в различных частях изображения, не дожидаясь завершения обработки предыдущих задач. Это не только увеличивает пропускную способность, но и позволяет использовать более сложные и ресурсоемкие алгоритмы распознавания, обеспечивая высокую точность без ущерба для скорости работы системы.

Допустима реализация подобной структурной организации вычислительного модуля на программируемых логических интегральных схемах (ПЛИС). ПЛИС представляет собой платформу, которая позволяет проектировать аппаратные схемы с гибкой архитектурой, полностью настраиваемой под конкретные задачи, такие как распознавание символов на изображениях или видеопотоке. Это даёт уникальное преимущество в создании конвейерно-параллельных архитектур, оптимизированных под требования системы, где важна как высокая точность, так и пропускная способность обработки данных.

Одним из ключевых аспектов использования ПЛИС для таких задач является возможность параллелизации вычислений на аппаратном уровне. ПЛИС предлагает возможность создания множества специализированных блоков, которые могут быть настроены для выполнения отдельных операций одновременно. Для обработки символьной информации, распознавание текстов и маркировок на изображениях, различные блоки модуля, реализованные с аппаратной поддержкой, для различных этапов обработки, работающих параллельно, ускоряют пропускную способность модуля.

Особенностью ПЛИС является то, что её архитектура может быть адаптирована для выполнения ресурсоёмких вычислений, таких как нейросетевые алгоритмы, применяемые для распознавания символов или объектов. На этапе поиска и распознавания текста на изображениях, например, использование сверточных рекуррентных нейронных сетей (CRNN) требует больших объёмов вычислений. В отличие от стандартных процессоров, которые ограничены скоростью последовательных операций, ПЛИС позволяет выполнять вычислительные задачи одновременно, распределяя их между различными аппаратными блоками. Это позволяет достичь более высокой производительности при обработке изображений с использованием более сложных и ресурсоемких алгоритмов, где необходимо производить большее количество вычислений за короткий промежуток времени.

Конвейерная архитектура позволяет настроить аппаратные блоки таким образом, чтобы данные поступали от одного блока к другому с минимальной задержкой, обеспечивая непрерывный поток информации. В подобной архитектуре система управления задачами, которая координирует работу всех блоков, ПЛИС позволяет реализовать очередь задач, которая распределяет рабочую нагрузку между параллельно работающими блоками, обеспечивая при этом сохранение порядка поступления данных.

Такая структура позволяет организовать эффективную обработку символьной информации, начиная с получения изображения и заканчивая формированием структурированной информации об объекте.

4.2. Тестирование комплексной характеристики «точность × пропускная способность» вычислительного модуля

Для оценки эффективности разработанного вычислительного модуля обработки символьной информации было проведено компьютерное моделирование, в ходе которого тестировалась комплексная характеристика «точность × пропускная способность» модуля. Моделирование позволило провести анализ работы модуля в различных условиях, имитируя разнообразные сценарии использования и влияния различных факторов, таких как уровень шумовых искажений, типы символьных маркировок, размер объектов для распознавания.

Комплексная характеристика «точность × пропускная способность» оценивалась в количестве корректно распознанных кадров в секунду (крк/с), вычисляемых как $t = a \times b$, где a – точность распознавания символов в диапазоне $[0..1]$, b – пропускная способность.

Было создано несколько тестовых наборов видеопотоков с различными типами символьных маркировок, различным уровнем качества изображения (включая наличие шумов, распределяемых по нормальному закону) и различными условиями съемки (разные углы обзора, освещенность). При этом уровень шума измерялся как пиковое отношение сигнала к шуму (PSNR), меньшее значение данного показателя указывает на большую зашумленность изображения. Оценка производилась в трех интервалах «от 20 dB», «15-20 dB» и «10-15 dB». Компьютерное моделирование для проведения тестирования проводилось с использованием процессора с частотой 2.3GHz, графических процессоров с частотой 1290MHz и объемом памяти 4GB.

Для оценки пропускной способности модуль подвергался моделированию нагрузки путем подачи видеопотока на входной блок и измерения количества кадров, распознанных системой в секунду. Для оценки точности работы модуля проводилось сравнение распознанных символов с эталонными значениями, содержащимися в тестовых наборах данных. Подобное тестирование так же проводилось для TesseractOCR и Keras (CRNN)), что позволило выполнить

сравнение с аналогами. Так же было произведено сравнение с устройствами распознавания символов, выполненными в трех вариациях: на основе ПЛИС, с использованием графического процессора и программной реализацией [89].

Сравнение с аналогами приведено в таблице 4.1.

Таблица 4.1 – Сравнение с аналогами

Критерий	Поддержка программной реализации	Поддержка аппаратной реализации	Поддержка ускорения на графических процессорах	Возможность параллельной работы	Сохранение порядка обработки кадров
Разработанный вычислительный модуль	Да	Да	Да	Да	Да
Keras (CRNN)	Да	Нет	Да	Да	Нет
TesseractOCR	Да	Нет	Нет	Нет	Нет
Устройство распознавания символов (на основе ПЛИС)	Нет	Да	Нет	Да	Нет
Устройство распознавания символов (программная реализация)	Да	Нет	Нет	Да	Нет
Устройство распознавания символов (с графическим процессором)	Нет	Да	Да	Да	Нет

Сравнение пропускной способности разработанного решения и аналогов представлено на рисунке 4.3.

Для оптимизации пропускной способности системы замерялось время, затрачиваемое на обработку каждого кадров, содержащихся в тестовых видеопотоках. Проведение данного тестирования позволило выявить оптимальное количество параллельных блоков 3 (распознавания символьной информации), для обеспечения максимальной пропускной способности разработанного модуля.

При этом пропускная способность обусловлена следующими показателями:

1. Блоки 1 и 2 (получения изображения и предобработки изображения) обрабатывают 1 кадр за 35мс, что позволяет достичь пропускной способности до 33.3 кадров в секунду

2. Блок 3 (распознавания символьной информации) обрабатывает один кадр за 400мс, что при использовании 14 параллельных блоков позволяет достичь пропускной способности до 35 кадров в секунду

3. Блок 4 (формирования выходного значения) обрабатывает 1 кадр за 5мс.

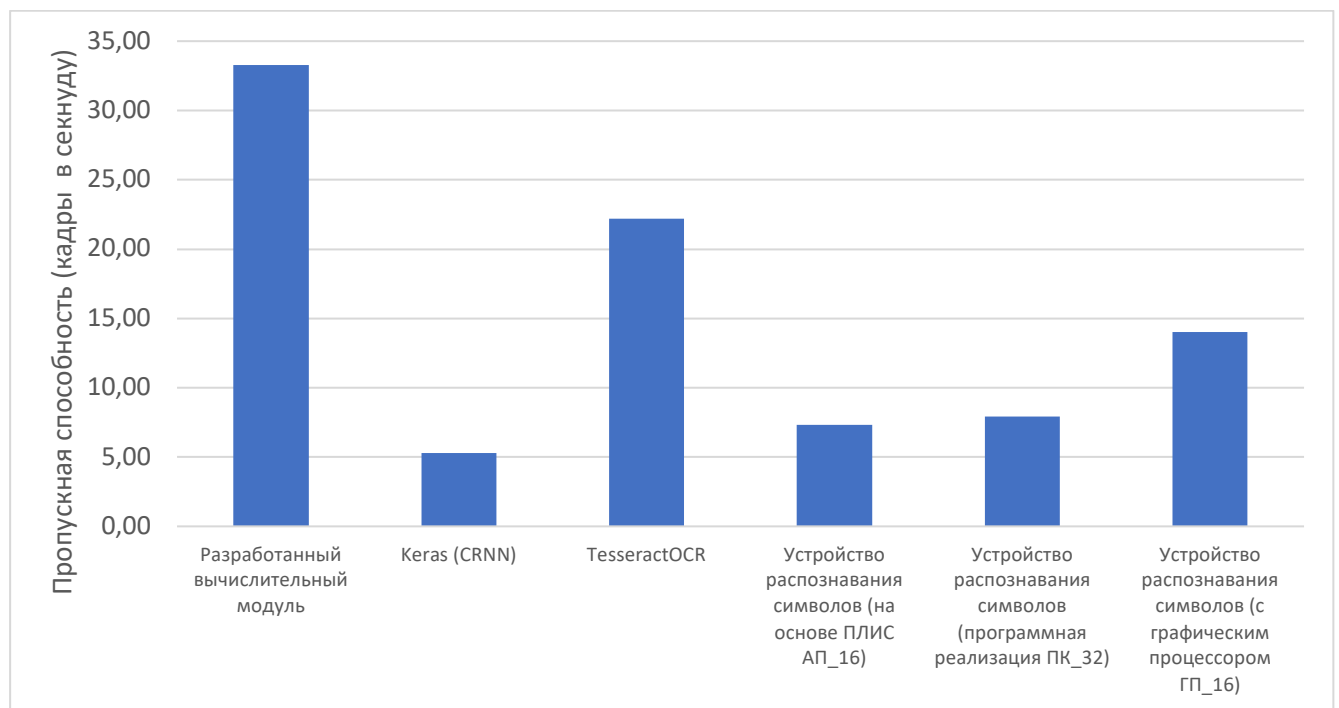


Рисунок 4.3 – Сравнение пропускной способности разработанного решения и аналогов

Результаты тестирования комплексной характеристики «точность × пропускная способность» при различном уровне шума, распределяемом по нормальному закону, приведены в таблице 4.2 и на рисунке 4.4.

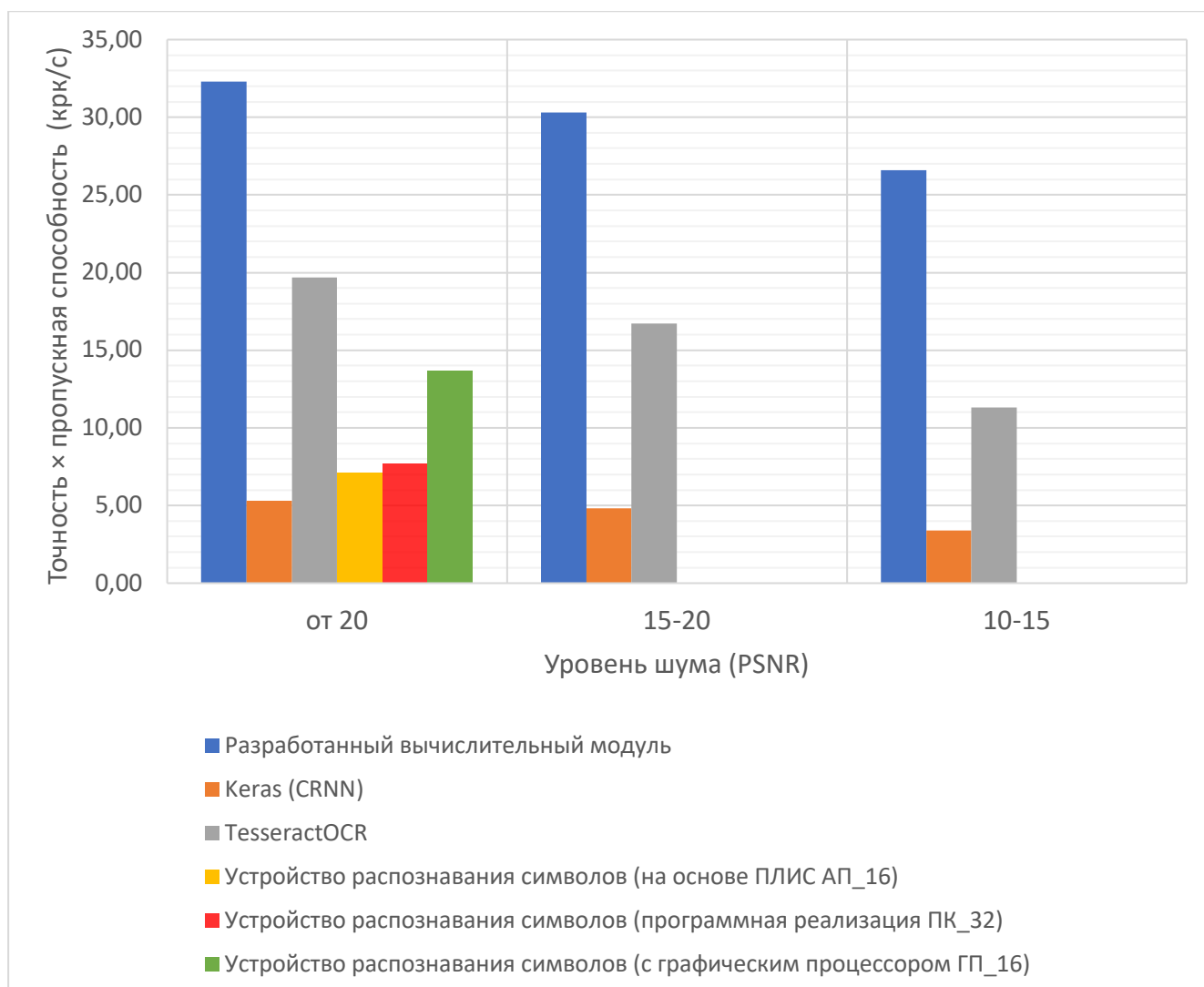


Рисунок 4.4 – Результаты тестирования по комплексной характеристике «точность × пропускная способность»

Таблица 4.2 – Результаты тестирования по комплексной характеристике «точность × пропускная способность»

Решение	Точность × пропускная способность при PSNR от 20 dB	Точность × пропускная способность при PSNR 15-20 dB	Точность × пропускная способность при PSNR 10-15 dB
Разработанный вычислительный модуль	$0.97 \times 33.3 = 32.3$ крк/с	$0.91 \times 33.3 = 30.3$ крк/с	$0.80 \times 33.3 = 26.6$ крк/с
Keras (CRNN)	$0.93 \times 5.7 = 5.3$ крк/с	$0.84 \times 5.7 = 4.8$ крк/с	$0.6 \times 5.7 = 3.4$ крк/с
TesseractOCR	$0.89 \times 22.2 = 19.7$ крк/с	$0.75 \times 22.2 = 16.7$ крк/с	$0.51 \times 22.2 = 11.3$ крк/с
Устройство распознавания символов (на основе ПЛИС АП_16)	$0.98 \times 7.3 = 7.1$ крк/с	-	-
Устройство распознавания символов (программная реализация ПК_32)	$0.98 \times 7.9 = 7.7$ крк/с	-	-
Устройство распознавания символов (с графическим процессором ГП_16)	$0.98 \times 14 = 13.7$ крк/с	-	-

Разработанная организация вычислительного модуля, который учитывает особенности созданного алгоритма распознавания и обеспечивает его реализацию на современных вычислительных устройствах, позволяет обрабатывать символьную информацию на маркировке продукции и отличающаяся применением конвейерно-параллельной обработки.

Вывод по четвертой главе

Разработка архитектурных и структурных решений, лежащих в основе вычислительного модуля, обеспечивает эффективную техническую реализацию аппаратно-программных систем с учетом оптимизации применяемой электронной компонентной базы и встраиваемого программного обеспечения, позволяет вычислительному модулю превосходить аналоги по показателю «точность × пропускная способность», на величину от 39% до 57% в зависимости от уровня шума.

ОСНОВНЫЕ РЕЗУЛЬТАТЫ РАБОТЫ

Диссертационная работа посвящена решению научно-технической задачи, связанной с обработкой символьной информации на маркировке продукции, в условиях работы с символами различных алфавитов и наличия шумов и искажений на изображениях продукции. В ходе исследования получены следующие основные результаты:

1. Проведен анализ существующих моделей, методов, алгоритмов и вычислительных модулей для обработки изображений и распознавания символов. Выделена необходимость создания моделей, методов, алгоритмов и вычислительных модулей, направленных на ликвидацию проблем выполнения распознавания маркировки продукции, возникающих при наличии шумов на изображении и работе с несколькими алфавитами.

2. Разработана математическая модель обработки изображений, отличающаяся использованием последовательности разработанных частных математических моделей для выделения объекта, предобработки изображения объекта, распознавания текста и классификации характеристик объекта, в совокупности позволяющая получить характеристики объекта, содержащего символьную маркировку.

3. Метод обработки изображений, содержащих символьную информацию, включающий последовательность проверочного, подготовительного, основного (распознавание) и классификационного этапов и отличающийся применением двумерных дескрипторов и выделения объекта на основе комбинации методов Канни и Рамера-Дугласа-Пекера, а также применением рекуррентной сверточной нейронной сети, позволяющий вести комплексную обработку областей текста в условиях геометрической, информационной, цветовой неопределенностей для последующей классификации характеристик объектов.

4. Разработан конвейерно-параллельный алгоритм обработки изображений, содержащих СИ, который реализует распознавание символьной

маркировки продукции в соответствии с созданной математической моделью и методом обработки изображений, содержащих СИ, и отличающийся совмещением предварительной обработки изображений и параллельной классификации символьной информации.

5. Разработана структурная организация вычислительного модуля обработки символьной информации, основанная на конвейерно-параллельном алгоритме обработки изображений, содержащих СИ. Спроектированная структурная организация вычислительного модуля позволяет повысить комплексный показатель «точность × пропускная способность», в сравнении с аналогами до 39% (для изображений с уровнем шума (PSNR) от 20dB), до 44% (PSNR – 15-20dB) и до 57% (PSNR – 10-15dB).

Список использованной литературы

1. Хомяков О.О. Распознавание символьной информации для автоматизации производственных процессов / О.О. Хомяков, В.С. Панищев, М.И. Труфанов, О.Г. Добросердов. // Известия Юго-Западного государственного университета, № 1, 2021. С. 122-137.

2. Метод распознавания текстовых данных на изображениях / О.О. Хомяков, В.С. Панищев, Д.В. Титов, С.И. Егоров. // Известия ЮФУ. Технические науки, No. 4 (234), 2023. pp. 57-65.

3. Математическая модель и параллельный алгоритм обработки изображений, содержащих символьную информацию / О.О. Хомяков, В.С. Панищев, В.С. Титов, Э.И. Ватутин. // Труды МАИ, № 137, 2024. С. 25.

4. Хомяков О.О. Распознавание символьной информации на маркировке автоматических выключателей / О.О. Хомяков, В.С. Панищев. // Информационные технологии и математическое моделирование систем, 2019. С. 106-109.

5. Хомяков О.О. Улучшение качества распознавания маркировки автоматических выключателей для частных случаев / О.О. Хомяков, В.С. Панищев. // Лучший исследовательский проект, 2020. С. 169-174.

6. Сравнение алгоритмов обработки данных в задаче распознавания символьной маркировки / О.О. Хомяков, В.С. Панищев, М.Р. Мбуйи, М.Н.К. Л. // Национальный суперкомпьютерный форум, 2021. С. 2.

7. Определение производителя продукции с помощью технологии оптического распознавания символов на примере молочной продукции / О.О. Хомяков, В.С. Панищев, Н.К. Мухамадсадилов, К.С. Калугин. // Всероссийская научная конференция Регионы России, 2022. С. 334-335.

8. Хомяков О.О. Развитие искусственного интеллекта в цифровом мире: философские аспекты // Исторические, философские, методологические проблемы современной науки, 2022. С. 221-227.

9. Хомяков О.О. Система распознавания характеристик продукции с использованием текстовой маркировки / О.О. Хомяков, А.И. Безмельцев, А.А.

Савоськин. // Актуальные вопросы современной науки: теория и практика научных исследований, 2022. С. 62-64.

10.Хомяков О.О. Увеличение точности классификации характеристик продукции с использованием алгоритмов анализа структуры документа / О.О. Хомяков, В.С. Панищев. // Оптико-электронные приборы и устройства в системах распознавания образов и обработки изображений, 2023. С. 181-182.

11.Хомяков О.О. Увеличение пропускной способности алгоритма распознавания символов в видеопотоке при условии наличия искажений // Наука и инновации – современные концепции, 2024. С. 129-132.

12.Устройство для распознавания символьной информации маркировки продукции / Хомяков О.О., Панищев В.С., Савоськин А. А., Безмельцев А. И. / заявка № 2024113679 от 21.05.2024, решение о выдаче патента от 26.09.2024.

13.Провоторов А. Методика поэтапного обнаружения маркировки слябов / А. Провоторов, А. Орлов. // Современные проблемы науки и образования, № 6, 2012. С. 89.

14.Shangbang L. Scene Text Detection and Recognition: The Deep Learning Era / L. Shangbang, H. Xin, Y. Cong. // International Journal of Computer Vision, 2020. P. 24.

15.Chowdhury A.M. Extracting and segmenting container name from container images / A.M. Chowdhury, K. Deb. // International Journal of Computer Applications, No. 74 (19), 2013. pp. 18-22.

16.Zhu Y. Scene text detection and recognition: recent advances and future trends / Y. Zhu, C. Yao, X. Bai. // Frontiers of Computer Science, No. 10 (1), 2015. pp. 19-36.

17.Robot navigation using human cues: A robot navigation system for symbolic goal-directed exploration / R. Schulz, B. Talbot, O. Lam, F. Dayoub, P. Corke, B. Upcroft, G. Wyeth. // Proceedings of the 2015 IEEE International Conference on Robotics and Automation , 2015. pp. 1100-1105.

18.Photoocr: Reading text in uncontrolled conditions / A. Bissacco, M. Cummins, Y. Netzer, H. Neven. // Proceedings of the IEEE International Conference on Computer Vision, 2013. pp. 785–792.

19.Середа С.В. Применение машинного зрения в логистике // Sciences of Europe, № 65, 2021. С. 45-50.

20.Bulatov K.B. A method to reduce errors of string recognition based on combination of several recognition results with per-character alternatives // Вестник Южно-Уральского государственного университета. Серия: Математическое моделирование и программирование, Т. 12, № 3, 2019. С. 74-88.

21.Detecting figures and part labels in patents: competition-based development of graphics recognition algorithms / C. Riedl, R. Zanibbi, M. Hearst, S. Zhu, M. Menietti, J. Crusan, I. Metelsky, K. Lakhani. // International Journal on Document Analysis and Recognition, Vol. 19, 2016. pp. 155-172.

22.Hayashi N. Constructing an automatic object-recognition algorithm using labeling information for efficient recycling of WEEE / N. Hayashi, S. Koyanaka, T. Oki. // Waste Management, 2019. pp. 337-346.

23.Metal stamping character recognition algorithm based on multi-directional illumination image fusion enhancement technology / X. Zhong, Z. You, M. Qian, J. Zhang, X. Hu. // EURASIP Journal on Image and Video Processing, 2018. pp. 1-11.

24.Данилов А.Д. Интеллектуализация управления гибким автоматизированным производством // Вестник ВГТУ, № 5, 2016. С. 60-64.

25. Automatic Recognition of Container Serial Code / C. Hsu, Y. Yang, A. Chang, S.M.S. Morsalin, G. Shen, L. Shiu. // 2023 International Conference on Consumer Electronics, 2023. pp. 257-258.

26.Орлов А.А. Применение машины экстремального обучения нейронной сети для распознавания символов маркировки / А.А. Орлов, Е.С. Абрамова, А.В. Мареев. // Оптико-электронные приборы и устройства в системах распознавания образов и обработки изображений, 2023. С. 176-178.

27.An efficient industrial system for vehicle tyre (tire) detection and text recognition using deep learning / W. Kazmi, I. Nabney, G. Vogiatzis, P. Rose, A. Codd. // IEEE Transactions on Intelligent Transportation Systems, 2020. pp. 1-12.

28. Копылов И.В. Идентификация маркировок рулонов металлопроката / И.В. Копылов, А.В. Казаков, Л.Л. Малыгин. // Вестник череповецкого государственного университета, № 5 (74), 2016. С. 12-15.

29. Mahajan S. DELIGHT-Net: DEep and LIGHTweight network to segment Indian text at word level from wild scenic images / S. Mahajan, R. Rani, K. Trehan. // International Journal of Multimedia Information Retrieval, No. 12(2), 2023. pp. 29-37.

30. Марков В.В. Решение задачи распознавания символов на изображениях товарных ценников // Евразийский союз ученых, № 5-3 (74), 2020. С. 63-65.

31. Пыж С.В. Мобильная автоматизированная система инвентаризации металлургической продукции и обеспечения безопасности в складских помещениях / С.В. Пыж, О.Г. Ганичева. // Научно-технический прогресс в чёрной металлургии, 2013. С. 266-271.

32. Маслов И.А. Оптическое распознавание символов в информационных системах и проблемы внедрения // E-Scio, № 3 (78), 2023. С. 7.

33. Gurevich I.B. Image formalization space: Formulation of tasks, structural properties, and elements / I.B. Gurevich, V.V. Yashina. // Pattern Recognition and Image Analysis. Advances in Mathematical Theory and Applications., Vol. 21, No. 2, 2011. pp. 134-139.

34. Gurevich I.B. Comparative analysis and classification of features for image models / I.B. Gurevich, I.V. Koryabkina. // Pattern Recognition and Image Analysis. Advances in Mathematical Theory and Applications, Vol. 16, No. 3, 2006. pp. 265-297.

35. FOTS: Fast Oriented Text Spotting with a Unified Network / X. Liu, D. Liang, S. Yan, D. Chen, Y. Qiao, J. Yan. // Computer Vision and Pattern Recognition, 2018. pp. 5676-5685.

36. Liu Y. Convolutional Recurrent Neural-Network-Based Machine Learning for Scene Text Recognition Application / Y. Liu, Y. Wang, H. Shi. // Symmetry, No. 15 (4), 2023. P. 849.

37. Papangkorn I. Recognition of similar characters using gradient features of discriminative regions / I. Papangkorn, B. Jakramate, M. Sanparith. // Expert Systems with Applications, Vol. 134, 2019. pp. 120-137.

38. Smith R.W. An Overview of the Tesseract OCR Engine // Ninth International Conference on Document Analysis and Recognition (ICDAR 2007), 2007. pp. 629-633.
39. Islam N. A Survey on Optical Character Recognition System // ITB Journal of Information and Communication Technology, No. 10, 2016. P. 4.
40. Сойфер В.А., Гашников М.В. Методы компьютерной обработки изображений. Москва: ФИЗМАТЛИТ, 2003. 784 с.
41. Pratt W. Digital Image Processing. New York: Wiley, 2001. 750 pp.
42. Yazdi M. New Trends on Moving Object Detection in Video Images Captured by a moving Camera: A Survey / M. Yazdi, T. Bouwmans. // Computer Science Review, No. 28, 2018. pp. 157-177.
43. Щеглов Д.С. Оптимизированный подбор фильтров при оптическом распознавании символьной информации / Д.С. Щеглов, Б.С. Горячкин, С.А. Пьянзин. // StudNet, № 4, 2020. С. 258-264.
44. Soua M. GPU parallel implementation of the new hybrid binarization based on Kmeans method (НБК) / M. Soua, R. Kachouri, M. Akil. // Journal of Real-Time Image Processing, No. 14, 2014. pp. 1-15.
45. Gatos B. Improved document image binarization by using a combination of multiple binarization techniques and adapted edge information / B. Gatos, I. Pratikakis, S. Perantonis. // Pattern Recognition 2009, 2009. P. 4.
46. Гонсалес Р., Вудс Р. Цифровая обработка изображений. Москва: Техносфера, 2012. 1104 pp.
47. TextDragon: An End-to-End Framework for Arbitrary Shaped Text Spotting / W. Feng, W. He, F. Yin, X. Zhang, C. Liu. // International Conference on Computer Vision (ICCV), 2019. pp. 9075-9084.
48. An End-to-End TextSpotter with Explicit Alignment and Attention / T. He, Z. Tian, W. Huang, C. Shen, Y. Qiao, C. Sun. // IEEE Conference on Computer Vision and Pattern Recognition, 2018. pp. 5020-5029.
49. Duda R.O. Pattern Classification 2nd Edition with Computer Manual. Wiley-Interscience, 2004. 816 pp.

50.Duda R.O., Hart P.E., Stork D.G. Pattern Classification (2nd edition). New York: John Wiley & Sons, 2001. 680 pp.

51.Optical character recognition on heterogeneous SoC for HD automatic number plate recognition system. / A. Farhat, O. Hommos, A. Al-Zawqari, A. Al-Qahtani, F. Bensaali, A. Amira, X. Zhai. // EURASIP Journal on Image and Video Processing, 2018. P. 17.

52.Ферцев А.А. Реализация нейронной сети для распознавания изображений с помощью технологии Nvidia CUDA // Прикладная информатика, № 6 (36), 2011. С. 102-110.

53.Распознавание символов на основе вейвлет-преобразования, метода главных компонент и нейронных сетей / В.Г. Спицын, Ю.А. Болотова, Н.Х. Фан, Т.Т.Ч. Буй. // Компьютерная оптика, № 2, 2016. С. 249-257.

54.Хаустов П.А. Алгоритмы распознавания рукописных символов на основе построения структурных моделей // Компьютерная оптика, № 1, 2017. С. 67-78.

55.Leveraging Transfer Learning and GAN Models for OCR from Engineering Documents / W. Khallouli, R. Pamie-George, S. Kovacic, A. Sousa-Poza, M. Canan, J. Li. // IEEE APOT, 2022. P. 7.

56.Журавлев Ю.И. Задачи построения линейных и нелинейных классификаторов в случае многих классов / Ю.И. Журавлев, Ю.П. Лаптин, А.П. Виноградов. // Математические методы распознавания образов, Т. 15, № 1, 2011. С. 72-74.

57.Журавлев Ю.И. О решении задач распознавания по прецедентам при большом числе классов / Ю.И. Журавлев, В.В. Рязанов. // Доклады Академии наук, Vol. 476, No. 5, 2017. pp. 489-491.

58.Журавлев Ю.И. Логические корректоры в задаче классификации по прецедентам / Ю.И. Журавлев, Е.В. Дюкова, П.А. Прокофьев. // Журнал вычислительной математики и математической физики, Т. 57, № 11, 2017. С. 1906-1927.

59. Real time rectangular document detection on mobile devices / N. Skoryukina, D. Nikolaev, A. Sheshkus, D. Polevoy. // Proceedings of SPIE - The International Society for Optical Engineering, 2015. P. 6.

60. Zivkovic Z. Approximate Bayesian methods for kernel-based object tracking / Z. Zivkovic, A. Taylan Cemgil, B. Kröse. // Computer Vision and Image Understanding, 2009. pp. 743-749.

61. Олькина Д.С. Алгоритм семантической сегментации изображений для решения задачи позиционирования летательного аппарата на земной поверхности // Труды МАИ, № 130, 2023. С. 16.

62. Исследование методов сегментации изображений текстовых блоков документов с помощью алгоритмов структурного анализа и машинного обучения / Т.С. Чернов, Д.А. Ильин, П.В. Безматерных, И.А. Фараджев, С.М. Карпенко. // Вестник российского фонда фундаментальных исследований, 2013. С. 55-71.

63. Горемыкин И.В. Дескрипторы цветовых регионов в задаче обеспечения цветовой согласованности изображения / И.В. Горемыкин, Л.Н. Бутенко. // Перспективы развития информационных технологий, № 27, 2015. С. 6-17.

64. Захаров А.А. Метод нахождения соответствий на изображениях с использованием структур дескрипторов / А.А. Захаров, А.Л. Жизняков, В.С. Титов. // Компьютерная оптика, Vol. 43, No. 5, 2019. pp. 810-817.

65. Tek S. GPU Accelerated Real-Time Object Detection on High Resolution Videos using Modified Census Transform / S. Tek, M. Gökmen. // VISAPP 2012 - Proceedings of the International Conference on Computer Vision Theory and Applications, 2012. P. 4.

66. Discriminative correlation filter tracker with channel and spatial reliability / A. Lukezic, T. Vojir, L. Cehovin Zajc, J. Matas, M. Kristan. // International Journal of Computer Vision, 2018. pp. 6309-6318.

67. Zhao Y. Deep Learning Neural Network-Based Detection of Wafer Marking Character Recognition in Complex Backgrounds / Y. Zhao, J. Xie, P. He. // Electronics, No. 12 (20). P. 17.

68. Deep-CNN TL: Text Localization from Natural Scene Images Using Deep Convolution Neural Network with Transfer Learning / M. Chaitra, D. Ramegowda, M. Gopalakrishna, B. Prakash. // *Arabian Journal for Science and Engineering*, No. 47 (9), 2021. pp. 9629-9640.

69. Zhao W. End-to-End Scene Text Recognition with Character Centroid Prediction / W. Zhao, J. Ma. // *International Conference on Neural Information Processing*, Vol. 1, 2017. pp. 291-299.

70. EAST: An Efficient and Accurate Scene Text Detector / X. Zhou, C. Yao, H. Wen, Y. Wang, S. Zhou, W. He, J. Liang. // *Computer Vision and Pattern Recognition (CVPR)*, 2017. pp. 5551-5560.

71. Силаков Н.В., Т.К.Л. Обзор алгоритмов обнаружения текстовых областей на кадрах видеопотока / Н.В., Т.К.Л. Силаков. // *Вестник РГГУ*, № 2, 2020. С. 27-45.

72. Hinami R. Bidirectional extraction and recognition of scene text with layout consistency / R. Hinami, X. Liu, N. Chiba. // *International Journal on Document Analysis and Recognition*, Vol. 19, 2016. pp. 83-98.

73. Орлов А.А. Алгоритм локализации изображений маркировок промышленных изделий на основе двумерного рекуррентного поиска усредненного максимума / А.А. Орлов, Д.Г. Привезенцев, А.В. Астафьев. // *СВЧ-техника и телекоммуникационные технологии*, Vol. 1, 2015. pp. 397-398.

74. Busta M. Deep TextSpotter: An End-to-End Trainable Scene Text Localization and Recognition Framework / M. Busta, L. Neumann, J. Matas. // *2017 IEEE International Conference on Computer Vision (ICCV)*, 2017. pp. 2223-2231.

75. Shi B. An End-to-End Trainable Neural Network for Image-Based Sequence Recognition and Its Application to Scene Text Recognition / B. Shi, X. Bai, C. Yao. // *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2015. P. 9.

76. El-Sawy A. Characters Recognition using Convolutional Neural Network. / A. El-Sawy, M. Benha. // *WSEAS TRANSACTIONS on COMPUTER RESEARCH*, 2017. pp. 11-19.

77. Text recognition in multimedia documents: a study of two neural-based OCRs using and avoiding character segmentation / K. Elagouni, C. Garcia, F. Mamalet, P.

Sébillot. // *International Journal on Document Analysis and Recognition*, Vol. 17, 2013. pp. 19-31.

78.Хомоненко А.Д. Обоснование архитектуры сверточной нейронной сети для автономного распознавания объектов на изображениях бортовой вычислительной системой / А.Д. Хомоненко, Е.Л. Яковлев. // *Научные технологии в космических исследованиях Земли*, № 6, 2018. С. 86-93.

79.Implementation of an intelligent decision support system to accompany the manufacturing process / В.А. Головкин, А.А. Крошечко, М.В. Ковалев, В.В. Таберко, Д.С. Иванюк. // *Открытые семантические технологии проектирования интеллектуальных систем*, № 4, 2020. С. 175-182.

80.Применение методов машинного обучения для классификации радиосигналов / И.В. Малыгин, С.А. Бельков, А.Д. Тарасов, М.Р. Усвяцов. // *Труды МАИ*, № 96, 2017. С. 15.

81.Ревякин А.М. Подходы к разработке системы распознавания для решения задачи определения контента цифровых изображений / А.М. Ревякин, А.В. Скурнович. // *Вестник евразийской науки*, № 4 (35), 2016. С. 14.

82.Макарова А.И. Сравнительный анализ реализаций метода опорных векторов для обучения двухклассовому распознаванию в условиях большого числа объектов / А.И. Макарова, В.В. Сулимова. // *Известия ТулГУ. Технические науки*, № 10, 2018. С. 12.

83.Прототипирование высокоскоростной нейронной сети в ПЛИС для классификации изображений видеопотока / Р.А. Соловьев, А.Г. Кустов, Д.В. Тельпухов, В.С. Рухлов. // *Cloud of science*, № 4, 2018. С. 680-703.

84.Сидорова Е.А. Подход к моделированию процесса извлечения информации из текста на основе онтологии // *Онтология проектирования*, No. 1 (27), 2018. pp. 134-151.

85.Коршунов А. Тематическое моделирование текстов на естественном языке / А. Коршунов, А. Гомзин. // *Труды ИСП РАН*, 2012. С. 215-242.

86.Хорошевский В.Г. Алгоритмы распределения ветвей параллельных программ по процессорным ядрам вычислительных систем / В.Г. Хорошевский, М.Г. Курносков. // Автометрия, Т. 44, № 2, 2008. С. 56-67.

87.Хорошевский В.Г. Распределённые вычислительные системы с программируемой структурой // Вестник СибГУТИ, No. 2 (10), 2010. pp. 3-41.

88.Современные и перспективные высокопроизводительные вычислительные системы с реконфигурируемой архитектурой / И.А. Каляев, И.И. Левин, А.И. Дордопуло, Ю.И. Доронченко, М.К. Раскладкин. // Вестник Южно-Уральского государственного университета. Серия: Вычислительная математика и информатика, Vol. 4, No. 3, 2015. pp. 24-39.

89.Устройство на основе ПЛИС для распознавания рукописных цифр на изображениях / И.В. Зоев, А.П. Береснев, Н.Г. Марков, А.Н. Мальчуков. // Компьютерная оптика, 2017. С. 938-949.

90.Корнеев В.В., Киселев А.В. Современные микропроцессоры. Санкт-Петербург: БХВ-Петербург, 2003. 448 с.

91.An efficient FPGA implementation of Optical Character Recognition for License Plate Recognition / J. Yuan, B. Youssefi, M. Mirhassani, R. Muscedere. // IEEE 30th Canadian Conference on Electrical and Computer Engineering (CCECE), 2017. pp. 1-4.

92.Efficient License Plate Recognition via Parallel Position-Aware Attention / T. Wang, W. Wang, C. Li, J. Tang. // Pattern Recognition and Computer Vision, 2022. pp. 346-360.

93.Параллельно-конвейерная форма программы для программирования вычислительных систем гибридного типа / И.А. Каляев, А.И. Дордопуло, И.И. Левин, В.А. Гудков, А.А. Гуленок, Г.А. Евстафьев. // Вестник Уфимского государственного авиационного технического университета, Т. 20, № 3 (73), 2016. С. 122-128.

94.A Fast Feature Extraction in Object Recognition Using Parallel processing on CPU and GPU / J. Kim, E. Park, X. Cui, H. Kim, W. Gruver. // IEEE International Conference on Systems, Man and Cybernetics, 2009. pp. 3842 - 3847.

95.Chung I. Parallel Deep Neural Network Training for Big Data on Blue GeneQ. / I. Chung, T. Sainath, B. Ramabhadran, M. Pichen, J. Gunnels, V. Austel, U. Chauhari, B. Kingsbury. // International Conference for High Performance Computing, Networking, Storage and Analysis, SC, 2015. pp. 745-753.

96.Halvorsen M. Hardware Acceleration of Convolutional Neural Networks // Norwegian University of Science and Technology, 2015. P. 65.

97.A Parallel Neuromorphic Text Recognition System and Its Implementation on a Heterogeneous High-Performance Computing Cluster. / Q. Qiu, Q. Wu, P.R.L.R. Bishop M. // Computers, IEEE Transactions, No. 62, 2013. pp. 886-899.

98.GPU-based and Streaming-enabled Implementation of Pre-processing Flow towards Enhancing Optical Character Recognition Accuracy and Efficiency / Y.S. Gener, P. Dattilo, D. Gajaria, A. Fusco, A. Akoglu. // Dept. of Electrical and Computer Engineering, University of Arizona, Tucson, 2023. P. 12.

99.Yu K. Memory-Tree Based Design of Optical Character Recognition in FPGA / K. Yu, M. Kim, J. Choi. // Electronics, Vol. 12, No. 754, 2023. P. 16.

100.Патент № 2 662 632 Российская Федерация, МПК G06F 17/21 (2006.01) G06F 17/24 (2006.01) G06F 17/30 (2006.01) G06F 3/0488 (2013.01) Представление документов фиксированного формата в формате с измененной компоновкой : N 2016105659 : заявлено 19.08.2014 : опубл. 26.02.2015 / Цзэн В., Аткинсон С.Р., Нараянан Р., Лиу М., Гил А.

101.Патент № 2 585 972 Российская Федерация, МПК G06K 9/34 (2006.01) Модуль обнаружения формул : N 2014130243/08 : заявлено 23.01.2012 : опубл. 10.02.2016 / Лазаревиц М., Расковиц М., Обулен А., Петровиц Т.В.

102.Патент № 2 733 115 Российская Федерация, МПК G06K 9/20 (2006.01) G06K 9/36 (2006.01) G06K 9/62 (2006.01) B42D 25/00 (2014.01) Способ и устройство для верифицирования сертификатов и идентичностей : N 2019113157 : заявлено 10.08.2018 : опубл. 25.04.2019 / Чжэн Д., Сюй В., Ли Л.

103.Патент № 2 682 007 Российская Федерация, МПК G06K 9/62 (2006.01) Способ и система для идентификации номера и типа вагона и способ и система для

инспекции безопасности : N 2017140638 : заявлено 09.08.2016 : опубл. 06.07.2017 / Сюй Я., Юй В.

104. Патент № 2 621 626 Российская Федерация, МПК G06F 9/44 (2006.01) G06F 3/041 (2006.01) Способ и устройство управления расписанием с использованием оптического устройства чтения символов : N 2014107560 : заявлено 30.08.2012 : опубл. 07.03.2013 / Парк М.Д., Ким Т.Й., Ахн Ю.М., Янг Г.Х.

105. Орлов А.А. Реализация технологии автоматической идентификации промышленных изделий в условиях производства трубопроводной продукции / А.А. Орлов, А.В. Провоторов, А.В. Астафьев. // Труды международного симпозиума "Надежность и качество", Т. 2, 2011. С. 197.

106. Автоматизация склада производственного предприятия на основе технологии штрих-кодирования / Е.С. Болтунова, А.А. Калашников, Д.Б. Смирнов, Г.И. Юрковская. // Актуальные проблемы авиации и космонавтики, 2022. С. 877-879.

107. Орлов А.А. Система автоматического контроля движения промышленных изделий с использованием радиочастотных идентификационных меток / А.А. Орлов, Д.В. Титов, А.В. Астафьев. // Известия высших учебных заведений. Приборостроение, Т. 64, № 3, 2021. С. 176-182.

108. Ташметов Т. Применение штрих-кодирования на складах вагонного депо для учета товарно-материальных ценностей / Т. Ташметов, О. Тулкин, М. Фозил. // АВРИИТТ-2021, 2021. С. 43-47.

109. Фомин В.В. параллельные вычисления при реализации web-инструментария распознавания образов на основе методов прецедентов / В.В. Фомин, И.В. Александров. // Программные продукты и системы, № 2, 2017. С. 207-216.

110. Geng Y. A Parallel Computing Method for Entity Recognition based on MapReduce. / Y. Geng, P. Li, J. Zhao. // 2nd International Conference on Electronics, Network and Computer Engineering (ICENCE 2016), 2016. pp. 648-653.

111.Orb: an efficient alternative to sift or surf / E. Rublee, V. Rabaud, K. Konolige, G. Bradski. // Computer Vision (ICCV), 2011 IEEE International Conference, 2011. pp. 2564-2571.

112.Douglas D.H. Algorithms for the reduction of the number of points required to represent a digitized line or its caricature / D.H. Douglas, T.K. Peucker. // Geovisualization. University of Toronto Press, 1973. pp. 112-122.

113.Ramer U. An iterative procedure for the polygonal approximation of plane curves // Computer Graphics and Image Processing, 1972. pp. 244-256.

114.Suzuki S. topological structural analysis of digitized binary images by border following / S. Suzuki, K. be. // Computer Vision, Graphics, and Image Processing, Vol. 30, 1985. pp. 32-46.

115.Алексеев В.В. Аналитическая модель обработки графических изображений в системах жизнеобеспечения города для выявления поврежденных объектов / В.В. Алексеев, Д.В. Лакомов. // Труды МАИ, № 103, 2018. С. 24.

116.Воронов В.И. исследование параллельных структур нейронных сетей для использования в задачах по семантической классификации текста на русском языке в условиях ограничения вычислительных ресурсов (на примере оперативных сводок в системе МВД России) / В.И. Воронов, Э.В. Мартыненко. // Экономика и качество систем связи, № 3 (9), 2018. С. 52-60.

117.Васильева А.С. Ускорение процесса распознавания с помощью нейронной сети и графического процессора // Наука, техника и образование, № 4 (45), 2018. С. 55-56.

ПРИЛОЖЕНИЕ

Устройство для распознавания символьной информации маркировки продукции / Хомяков О.О., Панищев В.С., Савоськин А. А., Безмельцев А. И. / заявка № 2024113679 от 21.05.2024, решение о выдаче патента от 26.09.2024

РОССИЙСКАЯ ФЕДЕРАЦИЯ



ФЕДЕРАЛЬНАЯ СЛУЖБА
ПО ИНТЕЛЛЕКТУАЛЬНОЙ СОБСТВЕННОСТИ

(19) **RU** ⁽¹¹⁾ **2024 113 679** ⁽¹³⁾ **A****(12) ЗАЯВКА НА ИЗОБРЕТЕНИЕ**

Состояние делопроизводства: Экспертиза по существу завершена (последнее изменение статуса: 27.09.2024)

(21)(22) Заявка: **2024113679**, 21.05.2024

Делопроизводство

Исходящая корреспонденция		Входящая корреспонденция	
Решение о выдаче патента	26.09.2024		
Отчет об информационном поиске	26.09.2024		
Уведомление об удовлетворении ходатайства	01.07.2024	Ходатайство о проведении экспертизы заявки по существу	28.06.2024
Уведомление об удовлетворении ходатайства	01.07.2024	Ходатайство об освобождении от уплаты пошлин или уменьшении размера	28.06.2024
Уведомление о положительном результате формальной экспертизы	10.06.2024		
Уведомление об удовлетворении ходатайства	10.06.2024	Ходатайство об освобождении от уплаты пошлин или уменьшении размера	21.05.2024
Уведомление о зачете пошлины	10.06.2024	Платежный документ	21.05.2024
Уведомление о поступлении документов заявки	21.05.2024		

**ФЕДЕРАЛЬНАЯ СЛУЖБА ПО ИНТЕЛЛЕКТУАЛЬНОЙ СОБСТВЕННОСТИ
(РОСПАТЕНТ)**

Бережковская наб., 30, корп. 1, Москва, Г-59, ГСП-3, 125993. Телефон (8-499) 240- 60- 15. Факс (8-495) 531- 63- 18

На № - от -

Наш № 2024113679/28(030748)

При переписке просим ссылаться на номер заявки

Исходящая корреспонденция от

26.09.2024

ФГБОУ ВО "Юго-Западный государственный университет", Пахомова Екатерина Геннадиевна
ул. 50 лет Октября, 94
г. Курск
305040

Р Е Ш Е Н И Е
о выдаче патента на изобретение

(21) Заявка № 2024113679/28(030748)

(22) Дата подачи заявки 21.05.2024

В результате экспертизы заявки на изобретение по существу установлено, что заявленное изобретение относится к объектам патентных прав, соответствует условиям патентоспособности, сущность заявленного изобретения (изобретений) в документах заявки раскрыта с полнотой, достаточной для осуществления изобретения (изобретений)*, в связи с чем принято решение о выдаче патента на изобретение.

Заключение по результатам экспертизы прилагается.

Приложение: на 6 л. в 1 экз.

Начальник Управления
организации
предоставления
государственных услуг

Документ подписан электронной подписью
Сведения о сертификате ЭП
Сертификат
04285D8800C7B115AF49B2C98AE6B64B15
Владелец Травников
Дмитрий Владимирович
Срок действия с 09.08.2024 по 09.08.2025

Д. В. Травников



*Проверка достаточности раскрытия сущности заявленного изобретения проводится по заявкам на изобретения, поданным после 01.10.2014.

ЗАКЛЮЧЕНИЕ ПО РЕЗУЛЬТАТАМ ЭКСПЕРТИЗЫ

(21) Заявка № 2024113679/28(030748)

(22) Дата подачи заявки 21.05.2024

(24) Дата начала отсчета срока действия патента 21.05.2024

ПРИОРИТЕТ УСТАНОВЛЕН ПО ДАТЕ

(22) подачи заявки 21.05.2024

(72) Автор(ы) Хомяков Олег Олегович, Панищев Владимир Славиевич, Савоськин Алексей Александрович, Безмельцев Александр Игоревич, RU

(73) Патентообладатель(и) Федеральное государственное бюджетное образовательное учреждение высшего образования "Юго-Западный государственный университет", RU

(54) Название изобретения Устройство для распознавания символьной информации маркировки продукции

Диплом конкурса инновационных проектов «Инновация и изобретение года»
(номинация «Инновационная идея»), Курск 2021



АДМИНИСТРАЦИЯ КУРСКОЙ ОБЛАСТИ

ДИПЛОМ

НАГРАЖДАЕТСЯ

**ХОМЯКОВ
ОЛЕГ ОЛЕГОВИЧ**

за участие в ежегодном областном
конкурсе инновационных проектов

«ИННОВАЦИЯ И ИЗОБРЕТЕНИЕ ГОДА»

В НОМИНАЦИИ

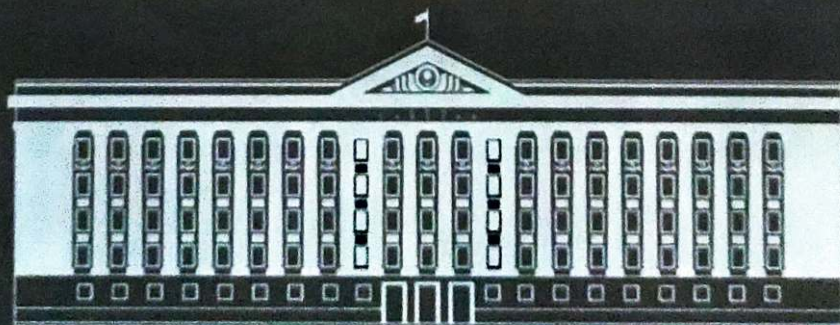
«ИННОВАЦИОННАЯ ИДЕЯ»

Проект: «Распознавание символьной информации
для автоматизации производственных процессов»

ГУБЕРНАТОР
КУРСКОЙ ОБЛАСТИ



Р. СТАРОВОЙТ



Г. КУРСК 2021

Акт об использовании результатов диссертационной работы в образовательном процессе

УТВЕРЖДАЮ
Проректор по учебной
работе ЮЗГУ
О.Г. Локтицова

«08» 04 2024 г.

АКТ об использовании результатов диссертационной работы Хомякова Олега Олеговича в образовательном процессе

Комиссия в составе начальника учебно-методического управления к.х.н., доцента В.В. Протасова, д.т.н., доцента, зав. кафедрой вычислительной техники И.Е. Чернецкой, д.т.н., профессора, профессора кафедры вычислительной техники Титова В. С. составила настоящий акт о том, что результаты диссертационной работы Хомякова Олега Олеговича, внедрены в образовательный процесс кафедры вычислительной техники, а именно:

- при проведении занятий по дисциплине «Цифровая обработка и анализ изображений в информационных системах» по направлению подготовки 09.04.01 «Информатика и вычислительная техника» используются разделы диссертационной работы, связанные с описанием алгоритма обработки изображений, содержащих символьную информацию;

- при проведении занятий по дисциплине «Архитектура систем обработки, анализа и интерпретации данных» по направлению подготовки 09.04.01 «Информатика и вычислительная техника» используются разделы диссертационной работы, связанные с описанием структурно-функциональной организации вычислительного модуля обработки символьной информации.

Начальник учебно-методического управления
к.х.н., доцент

В.В. Протасов

Зав. кафедрой ВТ, д.т.н., доцент
Профессор кафедры ВТ, д.т.н., профессор

И.Е. Чернецкая
В.С. Титов



Акт об использовании результатов диссертационной работы


 «УТВЕРЖДАЮ»
 генеральный директор ООО «Ценозавр» П.И. Умеренков
 «05» _____ 2024 г.

АКТ

об использовании результатов диссертационной работы
 Хомякова Олега Олеговича

Научно-техническая комиссия в составе:

председатель комиссии: генеральный директор ООО «Ценозавр»
 П.И. Умеренков

и члены комиссии: архитектор программного обеспечения К.С.
 Калугин
 программист В.А. Зенков

составили настоящий акт о том, что результаты диссертационной работы Хомякова Олега Олеговича, а именно:

- алгоритм обработки изображений, содержащих символьную информацию, осуществляющий распознавание символьной маркировки продукции;
- решения по программной реализации разработанного алгоритма обработки

использованы в процессе анализа символьной информации, на изображениях товаров торговых сетей, что позволило повысить точность распознавания текстового описания продукции на 4-20%, в зависимости от качества входных изображений.

Данный акт не может служить основанием для финансовых расчетов между организациями.

Председатель комиссии:

Члены комиссии:

 П.И. Умеренков
 К.С. Калугин
 В.А. Зенков